Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Curr Biol. Author manuscript; available in PMC 2022 November 22.

Published in final edited form as:

Curr Biol. 2021 November 22; 31(22): 5102–5110.e5. doi:10.1016/j.cub.2021.09.027.

Spatial organization of transcribing loci during early genome activation in Drosophila

Shao-Kuei Huang[1], **Peter H. Whitney**[1], **Sayantan Dutta**[2,3], **Stanislav Y. Shvartsman**[2,3,4], **Christine A. Rushlow**[1,+]

[1]Department of Biology, New York University, New York, NY 10003, USA.

[2]The Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544, USA

[3]Department of Chemical and Biological Engineering, Princeton University, Princeton, NJ 08544, USA

[4]Center for Computational Biology, Flatiron Research Institute, New York, NY 10010, USA

Summary

The early *Drosophila* embryo provides unique experimental advantages for addressing fundamental questions of gene regulation at multiple levels of organization, from individual gene loci to the entire genome. Using 1.5-hour old *Drosophila* embryos undergoing the first wave of genome activation[1], we detected ~110 discrete “speckles” of RNA Polymerase II (Pol-II) per nucleus, two of which were larger and localized to the histone locus bodies (HLBs)[2,3]. In the absence of the primary driver of *Drosophila* genome activation, the pioneer factor Zelda (Zld)[1,4,5], 70% fewer speckles were present, however, the HLBs tended to be larger than wild-type (*wt*) HLBs, indicating that Pol-II accumulates at the HLBs in the absence of robust early-gene transcription. We observed a uniform distribution of distances between active genes in the nuclei of both *wt* and *zld* mutant embryos, indicating that early co-regulated genes do not cluster into nuclear sub-domains. However, in instances whereby transcribing genes did come into close 3D proximity (within 400 nm), they were found to have distinct Pol-II speckles. In contrast to the emerging model whereby active genes are clustered to facilitate co-regulation and sharing of transcriptional resources, our data support an “individualist” model of gene control at early genome activation in *Drosophila*. This model is in contrast to a “collectivist” model where active genes are spatially clustered and share transcriptional resources, motivating rigorous tests of both models in other experimental systems.

[+]Lead contact and corresponding author (chris.rushlow@nyu.edu).

Author Contributions

All authors contributed to the design of the study, preparation of figures, and writing of the manuscript. SKH carried out the imaging experiments. PW and SKH performed the image analysis. SD carried out the Pol-II spot distribution *in silico*/simulation analysis.

Declaration of Interests

The authors declare no competing interests.

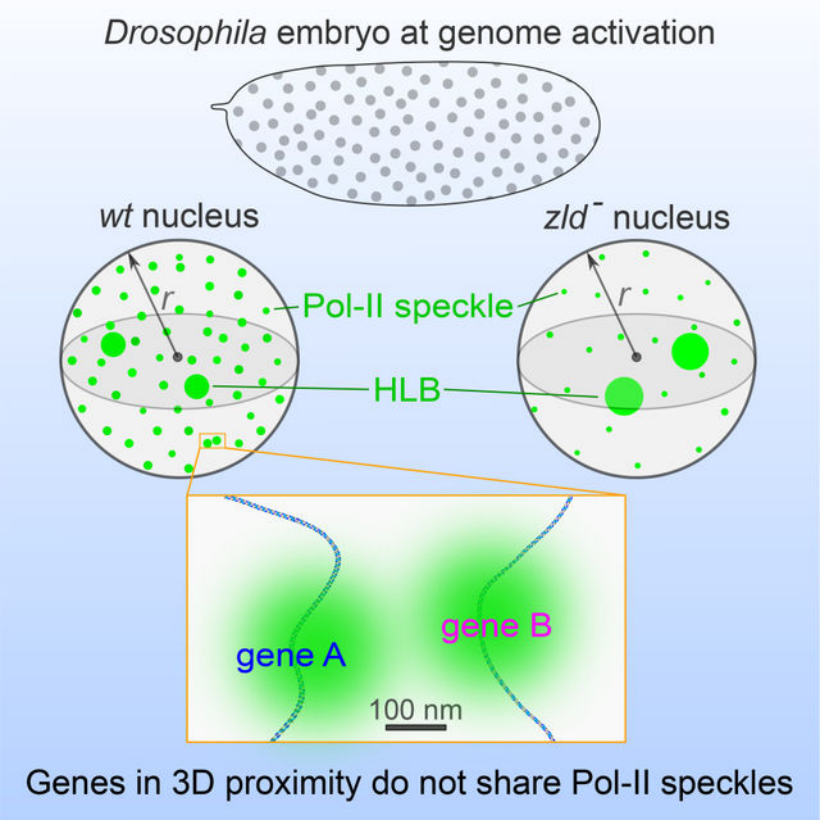
Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Blurb

Genome activation in Drosophila begins at one hour of development. Huang *et al.* demonstrate that RNA Polymerase II accumulates in discrete foci, or speckles, whose number and size depend on the pioneer factor Zelda. Transcribing genes that come into close 3D proximity (<400 nm) do not share a speckle, suggesting that co-regulated genes do not share RNA Polymerase II factories.

Graphical Abstract



Keywords

RNA polymerase II; speckles; Zelda; genome activation; histone locus body

Results and Discussion

RNA Polymerase II appears in discrete transcriptional foci during zygotic genome activation

Zygotic genome activation in *Drosophila* begins one hour after fertilization with a minor wave of transcription (tens of genes), followed by a major wave of activation an hour later (thousands of genes)[1,6–8]. The limited number of the early expressed genes provided a unique opportunity to globally visualize RNA Polymerase II (Pol-II) at sites of nascent transcription at the single-nucleus level, thus lending a complementary view to Pol-II activity seen in ChIP profiles[7,8].

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

High resolution microscopy of wild-type (*wt*) nuclear cycle 12 (nc12; ~1.5 hours old) embryos incubated with antibodies that recognize the RPB1 subunit of Pol-II revealed distinct Pol-II spot foci (Figure 1A, Pol-II in green), reminiscent of the membraneless condensates of Pol-II seen at active genes in mammalian cells, also called "speckles"[9–14]. Interestingly, two of the Pol-II foci in each nucleus were typically larger than others, and turned out to be the histone locus bodies (HLBs), multiprotein-RNA complexes involved in the transcription and processing of histone RNAs, as they stained for Multi sex combs (Mxc) protein, a component of the HLB[15,16], and *histone H3* (*His3*) RNAs (Figure 1A, Mxc in yellow, *His3* in magenta; see Methods).

Since Zld is a pioneer transcription factor required for genome activation[1,4,17], embryos devoid of *zld* RNA and protein (Figure S1B), herein referred to as *zld*$^{-}$, were stained for Pol-II. There appeared to be fewer speckles in *zld*$^{-}$, but surprisingly, the HLBs seemed larger (Figure 1A–B). These observations were confirmed by quantifying Pol-II spot fluorescence intensity using Imaris software (see Methods). First, to ensure that any observed Pol-II intensity differences between genotypes was not due to cell-cycle dependent accumulation of Pol-II, each nc12 embryo was staged within interphase. Since nuclei change shape as interphase proceeds, appearing less round (see example Hoechst stainings in Figure 1C, and His2av-RFP in Figure S1C), we used the feature of "nuclear sphericity" as a proxy for time (Figure S1D), i.e, pseudo-time, and compared the total nuclear Pol-II signal intensity in *wt* and *zld*$^{-}$ nuclei over time. As nuclei progressed through interphase, Pol-II levels increased in both genotypes alike (Figure 1C), suggesting that Pol-II is continually imported into the nucleus during interphase, and there is no difference between genotypes.

We next measured Pol-II intensity at the HLBs. Although HLBs in both *wt* and *zld*$^{-}$ nuclei acquired more Pol-II signal over time, which is consistent with the observation that HLBs accumulate Pol-II and grow larger over time[18], *zld*$^{-}$ HLBs tended to acquire Pol-II "earlier", suggesting an increase in the rate of accumulation of Pol-II at HLBs, and maintain it higher than in *wt* (Figure 1D). To assess whether this Pol-II accumulation had a functional consequence on transcription at the HLBs, we examined the output of *His3* mRNA via single molecule (sm) FISH over the same pseudo-time axis (Figure 1D). *His3* intensity levels showed a strikingly similar trajectory to Pol-II, indicating that the reallocation of Pol-II to the HLBs in *zld*$^{-}$ embryos indeed affects transcription. We next compared *wt* and *zld*$^{-}$ HLB total volumes (see Methods) and found that *zld*$^{-}$ HLBs tended to be larger (Figure 1E), as their Pol-II and *His3* signal intensities trended higher (see Figure 1F–H histograms and HLB metaspots; data shown for early interphase only). For all measurements, we found the greatest differences in "early" embryos, suggesting that a key limiting factor for transcription initiation at the HLBs is simply the ability to recruit Pol-II from the nucleoplasm. In the absence of the wide scale genome activation typically seen in *wt* embryos, the available pool of Pol-II is higher in *zld*$^{-}$, thus leading to higher transcriptional output of the core histone genes.

Pol-II accumulation at transcriptional foci depends on Zld.

We next sought to gain insight into Zld's effect on genome activation by examining Pol-II staining not found at the HLBs. We suspected that single Pol-II spots represent single

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

transcribing genes, as reported in mammalian cells[11,19–21]. To confirm, we performed single molecule FISH (smFISH) on a reporter transgene (*sog-lacZ,* described below) and co-stained for Pol-II (Figure 2A). We demonstrated that the smFISH staining produced linear signal amplification by rank-ordering small extra-nuclear spots to observe a "stair-step" pattern, showing a defined amount of fluorescence is associated with a signal molecule, which we estimated to be ~1650 AU (Figure 2B–C). Next, we examined sites of nascent transcription, high intensity foci of smFISH staining inside each nucleus, and estimated the number of nascent transcripts based on the fluorescence of a single molecule. Since there was a wide range of expression levels across the reporter expression domain, this analysis was able to demonstrate that the number of nascent transcripts scales with brightness of the corresponding Pol-II spot (Figure 2G–I). Additionally, because we see relatively few deviations from our line of best fit (Figure 2C), we are confident that the spots of Pol-II do indeed represent single genes.

To compare Pol-II spots in *wt* and *zld*$^{-}$ embryos, we used the Imaris "spots" finding function, testing several different thresholds to effectively eliminate background noise before calling the number of Pol-II spots in *wt* versus *zld*$^{-}$ nuclei (see Methods and Figure S2). As expected, there was a dramatic reduction in the total number of Pol-II spots in *zld*$^{-}$ nuclei (early 33 ± 13.8, late 37 ± 14.2) compared to *wt* (early 95 ± 13.0, late 109 ± 17.6) (Figure 2D), and many of the remaining spots had less Pol-II signal (Figure 2E–F), both consistent with past observations on Zld's role in transcription[1,4,22–24]. Most genes expressed at nc12 are completely dependent on Zld, and are thus absent in *zld*$^{-}$, hence the reduced number of Pol-II spots. However, some genes are only partially dependent on Zld, as they are also regulated by the patterning factors Bicoid (Bcd) and Dorsal (Dl)[4,25,26], with Zld potentiating their activity by increasing binding-site accessibility[27–29] and their local concentrations at target enhancers[24,26,30]. Without Zld, there is reduced transcriptional output of these genes, which could give rise to less intense Pol-II spots.

To test the prediction that Pol-II spot intensity levels can be modulated by Zld, we used two transgenes containing enhancer-reporter constructs with or without Zld binding sites, rather than *zld*$^{-}$ mutants, to eliminate potential effects of Pol-II availability in *zld*$^{-}$ nuclei since Pol-II moves to the HLBs. We previously showed that the *short gastrulation* (*sog*) shadow enhancer with three Zld binding sites (3TAG) drives higher and more robust reporter expression than an enhancer without Zld sites (0TAG), especially in nuclei with low-level Dl[22,24]. Thus, we compared Pol-II signal intensity at the 3TAG and 0TAG transgene loci, in nuclei containing high or low levels of Dl. To further control for Pol-II variability between embryos, 3TAG and 0TAG were expressed in the same nucleus as heterozygotes (see Methods). To distinguish between the transgenes, each contained a different reporter, either *yellow* (*y*) or *lacZ* (Figure 2G, 3TAG-*y* in yellow, 0TAG-*lacZ* in magenta). We found that Pol-II signal intensity was significantly lower at the 0TAG locus compared to the 3TAG locus, and as we predicted, the effect was more pronounced in the neuroectoderm where Dl morphogen levels are at their lowest (Figure 2H–I). These results are consistent with the known effects of Zld on Dl target genes[24,26,27], and demonstrate that the reduced transcription in the absence of Zld binding is a consequence of reduced Pol-II levels at the site of transcription.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Pol-II speckles are distributed evenly and randomly throughout the nucleus

Now confident that our Pol-II spots accurately reflect actively transcribed genes in nc12, we analyzed the spatial distribution of these spots to gain insight on the 3-dimensional (3D) distribution of transcribed genes during early genome activation. Multiple studies have implicated the 3D spatial conformation of genes as an important regulator of gene expression[31–33]. Critically, we wanted to know if there was any evidence of non-uniformity in the locations of active transcription. For example, there could be non-uniformity along the radius, i.e., a radial gradient, or there could be local clustering (Figure 3A), the latter indicative of "transcriptional hubs" where active loci share Pol-II machinery[20,34–36]. First, we asked if there is a systematic inhomogeneity in the density of spots as we scan radially outwards from the center of the nucleus. To address this question, we divided the nuclear regions into concentric 3D shells from the center of mass along the radius (like onion layers). We calculated the local density of spots for each shell, ρ(shell), then divided by the average total nuclear density (ρ). Figure 3B shows that the ρ(shell)/ρ ratio for each shell along the radius was ~1 in both *wt* and *zld*$^{-}$, indicating that Pol-II spots are evenly distributed along the radius, and this is not influenced by Zld.

Next, we asked whether the Pol-II speckles could be locally clustered (see Figure 3A, right) by assessing the distribution of the distances between spots and their neighbors. First, we took the 3D coordinates of the spots and constructed the Delaunay triangulation of the set of points using MATLAB (see Methods). Next, we identified all pairs of points in each tetrahedron to be each other's nearest (Voronoi) neighbors[37] (Figure 3C shows a Delaunay tetrahedron with a spot and its neighbors highlighted in red), then calculated the probability distribution of distance between each pair of nearest neighbors $V(r)$ (see Methods). Figure 3D shows these functions calculated for a uniformly-dispersed versus a clustered distribution of points generated *in silico*. If distributed uniformly, spot distances will fit a Gaussian distribution with a single peak (Figure 3D, red dashed line). If however spots are clustered, we would expect neighbor distances for spots within the same cluster to be small, while the distances between spots in different clusters to be large, therefore the distribution would have two peaks (Figure 3D, black dashed line). *wt* spots fit a Gaussian distribution with one peak (Figure 3D, solid blue line). Moreover, removing Zld does not affect the distribution of spots, but instead, the average distance between neighboring spots, which is increased as a consequence of the reduced number of Pol-II spots observed in *zld*$^{-}$ (Figure 3E, coral curve shifts right). Taken together these analyses show that Pol-II speckles are uniformly distributed throughout the nucleus in both *wt* and *zld*$^{-}$ nuclei, and suggest that although Zld affects the relative enrichment of Pol-II molecules in speckles, it does not control their 3D organization.

Genes in close proximity do not share a Pol-II speckle

The above analysis assessed clustering of Pol-II speckles from a global perspective, i.e., a bird's eye view of the nucleus, however, it was not able to address whether two, or a few, genes could come together in 3D space, particularly if co-regulated by the same transcription factor. 3D interactions between early genes have been observed in Hi-C studies[38–40], which revealed the emergence of TADs at nc12 as well as low-frequency interactions between TAD boundaries[38], many positioned at Zld target loci (see Hi-C contact map in Figure S3B). We

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

speculated that these interactions could reflect Zld target genes coming together in 3D space to form a Pol-II hub.

Since it is not possible to know which genes are specifically being transcribed at each of the Pol-II spots, we instead looked for instances where transcribing genes came into close proximity in 3D, and then asked if they shared a Pol-II spot. Focusing on a 5 megabase (Mb) region of chromosome 2L comprising several highly expressed genes (see Pol-II ChIP-seq track[8] in Figure 4A), we performed dual color RNA FISH of genes in pairwise combinations to first assess if their nascent transcription foci ever appeared close together despite lying far apart on the chromosome (Figure 4B–C). For each of six pairs of genes, we observed the two FISH signal centers come within less than 400 nm of each other at least 10% of the time (Figure S3C); 400 nm was used as the cutoff to define "close proximity" since promoter-enhancer interactions have been observed at this distance[41,42]. For example, *CG15382* and *CG14014* are separated by 3.4 Mb, but their transcriptional foci were seen as close as 230 nm. In a control experiment, two probes from the *slam* gene, 87 bp apart, showed 3D distances of 172±46 nm (Figure S3C; Figure 4D), thus we considered 218 nm the operational resolution limit in our imaging experiments. In contrast, two probes for genes on different chromosomes (*slam* on 2 and *hb* on 3) never came into close proximity (Figure S3C; Figure 4E).

Combining FISH with Pol-II antibody staining, we then asked whether two genes transcribing in close proximity could occupy a single Pol-II spot, indicating a shared transcriptional hub. Figure 4F shows images of *Bsg25D* and *CG14014* foci with their associated Pol-II spots, and in each case, the Pol-II spots could be resolved as discrete spots (if above the resolution limit of ~218 nm). Similar results were seen for FISH pairs, *slam-elba1*, *slam-elba2*, and *CG14014-CG15382* (Figure 4G–I, respectively). A striking example of loci in close proximity was observed at the *elba1* locus where two sister chromatids were actively transcribing, each in its own discrete Pol-II spot, which measured 223 nm apart (Figure 4J). These results indicate that genes in 3D close proximity do not share Pol-II hubs.

Transcriptional output is independent of 3D distance between loci

Lastly, we asked whether the transcriptional output of genes transcribing in 3D close proximity was increased compared to when they were distant, as it has been proposed that sharing of transcriptional machinery could lead to a greater transcriptional efficiency and output[20,43]. We used smFISH to quantify nascent transcript levels of *CG14014* and *Bsg25D* when the two loci were in close proximity (<400 nm) versus not close (500nm-2300nm apart) in single nuclei (see example image in Figure S4A). For each nucleus we generated a close/distant smFISH signal intensity ratio for each gene probe, which would be greater than 1 (>1) if there was higher transcriptional output when the two genes were in close proximity, however, the ratio on average was ~1 for both *CG14014* and *Bsg25D* (Figure 4K). These ratios were not significantly different from those of the control groups, which comprised the remaining pairs of foci where both pairs of *CG14014* and *Bsg25D* foci were in close proximity or were distant (see p-values in Figure 4K and control box plot in Figure S4C). Furthermore, subdividing the close-proximity distances into three bins of increasing distances (<200nm, 200–300nm, and 300–400nm), also gave ratios close to 1

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(Figure S4D). The same analysis was performed for two additional gene pairs, *Bsg25D-slam* and *CG14014-slam*, yielding the same results (Figures 4K and S4C, E–F), indicating that there was no 'benefit' in them being transcribed in close proximity. These results are in agreement with recent studies demonstrating that 3D proximity is uncoupled from transcriptional status[39,44].

Taken together, our results support a role for Zld in recruiting and maintaining Pol-II at gene loci in nc12 embryos when genome activation is initially underway. Zld co-regulated genes do not appear to transcribe in shared Pol-II factories/hubs even when within 400 nm of each other, and neither is their transcriptional output increased. Although we cannot rule out the possibility that on occasion two or more loci might share a Pol-II hub, at nc12 or in later cycles as hundreds of genes become active, our results support an "individualist" model, whereby each active gene transcribes in its own discrete Pol-II entity. The individualist model is in contrast to a "collectivist" model where active genes are spatially clustered and share transcriptional resources, motivating rigorous tests of both models in other experimental systems.

STAR Methods

Resource availability

Lead contact—Requests for further information, resources, and reagents should be directed to and will be fulfilled by the Lead Contact, Christine Rushlow (chris.rushlow@nyu.edu).

Materials availability—This study generated new unique reagent: Zld antibody available upon request.

Data and code availability—All data reported in this paper will be shared by the lead contact upon request. All original code has been deposited at Figshare and is publicly available as of the date of publication. DOIs are listed in the key resources table. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Experimental Model and Subject Details

All flies were grown on standard fly (*Drosophila melanogaster*) cornmeal-molasses-yeast media. Embryos were depleted of maternal and early zygotic *zld* using the Maternal-Gal4-shRNA system[45]. *MTD-Gal4 / UAS-shRNA-zld* females were crossed to w^{1118} males as described in Sun et al. 2015[27]; the resulting embryos were devoid of *zld* RNAs and referred to as zld^{-}. zld^{-} embryos were compared to wild-type (*wt*) embryos derived from *MTD-GAL4 /* w^{1118} flies (for Pol-II spot analysis experiment) or *y[1] w[*]* flies (for *in silico* analysis and proximity experiments). Strains *y[1] w[*]; sog 3TAG*-MS2v7-*lacZ* and *y[1] w[*]; sog 0TAG*-MS2v7-*lacZ,* referred to as 3TAG and 0TAG, respectively, are described in Yamada et al., 2019[24]. 3TAG and 0TAG transgenes with the *yellow* (*y+*) gene reporter in place of *lacZ* (*y[1] w[*]; sog 3TAG*-MS2v5-*y* and *y[1] w[*]; sog 0TAG*-MS2v5-*y*) were constructed in the pbPHi-eveprMS2 vector[46], and integrated on chromosome 2 (VK18) by

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Best Gene, Inc. (Chino Hills, CA; https://www.thebestgene.com). Two sets of heterozygotes were generated: 3TAG-*lacZ*/0TAG-*y* and 3TAG-*y,* 0TAG-*lacZ* to control for differences in reporter sequences.

Methods details

Embryo Staging—Embryo staging was performed by defining the "sphericity" of nuclei, a quality that changes over the course of a nuclear cycle. Sphericity is the volume divided by the surface area of an object. An index can be created by comparing the sphericity of any given object to that of a sphere with the same volume. Because a sphere has the minimum possible surface area to volume ratio, dividing the s surface area by the sphere's surface area, a unitless index between 0 and 1 is generated, with 1 being perfectly spherical. We observed in live imaging experiments that the average sphericity index of Drosophila nuclei decreases as the nuclei approach mitosis (Fig. S1C), therefore we employed sphericity as a proxy for time in our fixed imaging experiments.

Immunofluorescence (IF)—Embryos were collected 1 to 2.5 hours after egg laying at room temperature, dechorionated with Clorox, fixed in 4% formaldehyde/heptane, and devitellinized with methanol. Fixed embryos were rehydrated and stained with primary antibodies overnight, washed in PBT (1X PBS, 0.1% Tween 20), and then stained with secondary antibodies conjugated with Alexa Flour 488, 555, or 647 (Invitrogen and ThermoFisher Scientific) for 2 hours at room temperature. After Hoechst 33342 (Tocris Bioscience) staining (1μg/ml) for 15 minutes, embryos were washed in PBT and mounted onto glass microscope slides (Fisher Scientific) using Aqua-Poly/Mount (Polysciences) and Number 1.5 glass coverslips (Fisher Scientific). Primary antibodies included: anti-RBP1 conjugated with alexa488 (CTD4H8, Sigma-Aldrich, 05–623-AF488), which detects all forms of RNA polymerase II (Pol-II; 1:100 dilution); guinea pig anti-Multi sex combs (Mxc) to detect the histone locus body[16], a gift from Dr. Robert Duronio (1:5000 dilution); rabbit anti-Zld antibodies raised against the N-terminal domain of Zld, amino acids 1–517 (Pocono Rabbit firm, Canadensis, PA, see Figure S1A,B). For the Pol-II foci clustering *in silico* analysis, embryos were stained with the IF protocol described above except the Aqua-Poly/Mount (Polysciences) was diluted to 70% with PBS/Tween to prevent embryo drifting. Embryos were re-suspended in mounting media and placed on a 35mm glass-bottom dish (Bioptechs) without a coverslip to minimize deformation of nuclei.

Fluorescent *in situ* hybridization (FISH)—Hybridization of fixed embryos was done following a standard RNA FISH protocol[47]. The template DNA fragments were generated by PCR with anti-sense cDNA primers also containing the T7 promoter sequence. DIG or Fluorescein-labeled (Roche) anti-sense probes of *CG15382*, *elba1*, *Bsg25D*, *CG14014*, *hunchback* (*hb*), *CG15876*, *CG13712*, *slam* and *slam*#2 (two different regions of the *slam* gene) were generated by *in vitro* transcription with RNA labeling kits (Roche). Alexa Flour 488, 555, or 647 conjugated secondary antibodies were used to detect DIG or Fluorescein (FL) antibodies. In the experiments combining IF and FISH, IF was done prior to FISH.

IF and FISH with oligonucleotide probes targeting *His3*—To detect *His3* RNAs, a 21nt DNA probe (ACTTCACGTTTGAAAACACAA; Integrated DNA Technologies)

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

targeting the 5' UTR of the *His3* transcript was directly conjugated with ATTO 633 to ensure a linear signal to transcript ratio. Targeting the 5'UTR in this manner allowed us to only label transcripts from the main *His3* gene and no other histone variants. Because the HLB genes are massively amplified in *Drosophila*, only a single oligo was needed to achieve a high signal to noise ratio when visualizing nascent transcripts. Briefly, fixed embryos were hybridized first with the *His3* probe using the single molecule FISH (smFISH) Stellaris (LGC Biosearch Technologies) "*Drosophila* Embryo" protocol and reagents since a small oligo probe is compatible with the Stellaris protocol. After overnight hybridization and subsequent wash steps, embryos were then stained using anti-MXC (1:5000) and anti-RNA Pol-II conjugated to Alexa 488 (1:50) diluted in PBT for 1.5 hours at room temperature. Embryos were then washed in PBT and stained with anti-Guinea pig Cy3 (1:500) diluted in PBT. Finally, embryos were washed in PBT and stained with Hoechst 33342 (1μg/ml). Embryos were mounted on glass slides (Fisher Scientific) using ProLong Antifade Mountant (ThermoFisher) and High Precision Deckgläser No. 1.5H coverslips (Thor Labs) and allowed to cure for 24 hours.

Single molecule fluorescent *in situ* hybridization (smFISH)—To quantify the output of 3TAG and 0TAG transgenes, we generated Stellaris (LGC Biosearch Technologies) smFISH probes targeting the *lacZ*, *yellow*, and *MS2* (spacer sequences in MS2.v7) reporter genes using the Stellaris probe designer. Each reporter gene was labeled with 48 20nt DNA probes directly conjugated to either Quasar 670 (*lacZ*) or Quasar 570 (*MS2* and *yellow*) to achieve linear signal amplification. smFISH probes against *Bsg25D*, *CG14014*, and *slam* were generated following the Gaspar et al. (2017)[48] protocol with minor modifications. Probes comprising 33 (*Bsg25D*), 40 (*CG14014*), and 40 (*slam*) oligonucleotides (20 bases each; Integrated DNA Technologies) with gaps of at least two bases were NH2-dd-UTP conjugated using terminal deoxy-nucleotidyl transferase (NEB) and labeled with succinimidyl (NHS)-ester conjugated Alexa fluor 647 or Alexa flour 594 dyes (Invitrogen). smFISH was done following the Gaspar et al. (2017)[48] protocol. Two different fluorophores were used for experiments comparing two different genes. Stellaris reagents and protocols were used for smFISH. IF was done before smFISH when combining IF and smFISH.

High resolution image acquisition—Images of IF, FISH, and smFISH were acquired with the Zeiss confocal microscope, LSM 880, utilizing the Airyscan Detector and a Plan-Apochromat 100x/1.46 Oil DIC M27 Elyra objective. Each embryo was imaged in 3 positions when possible. Each nc12 image contained at least 14 nuclei. The following parameters were used for imaging: Resolution = 24.526 pixels per micron, Voxel size = 0.041 × 0.041 × 0.150 micron^3, Bits per voxel = 16, Laser Power #1(633) = 0.005, Laser Power #2(561) = 0.020, Laser Power #3(205) = 0.029, Laser Power #4(488) = 0.028. Each channel used appropriate filter sets to minimize bleed through, and channels cycled each frame to minimize z-drift. For the distance-measuring experiments, fluorescent beads (FocalCheck thin-ring fluorescent microspheres kit; Molecular Probes) were used to ensure that lasers were aligned within 50nm tolerance. Confocal settings were: 1580×1580 (Figures 1 and 2) or 1156×1156 pixel, 25–30 z-stacks (except for the Pol-II distribution experiment, which was scanned in 70 z-stacks) 0.15μm apart, 16 bit. All images were

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

processed using the Zeiss Zen software with “Airyscan Processing” using the automatic values for “strength”. Mxc signals were processed in FIJI with functions of remove outliers (radius=6, threshold=50) and median filter (radius=4). 3D images of Pol-II, dual-FISH pairs, and Pol-II with dual-FISH pairs were snapshots of Imaris (Bitplane, Oxford Instruments, http://www.bitplane.com/imaris/imaris) 3D views.

ChIP-seq tracks of Pol-II and Zld and Hi-C contact map—The ChIP-seq tracks of Pol-II[8] and Zld[17] of a 2 MB region of chromosome 2L (shown in Figure 4A and Figure S3A) were generated using the Integrated Genome Browser[49] (version 9.1.4, https://www.bioviz.org/); Y-axis, normalized sequencing reads. Genome-wide, we estimated the total number of highly bound genes to be 143 (using a cutoff of 30), 111 (using a cutoff of 40), or 83 (using a cutoff of 50). Figure S3B shows a Hi-C contact heatmap of the corresponding region at nc12 was generated from Hi-C data[38] using Hi-glass software[50] (https://higlass.io/).

QUANTIFICATION AND STATISTICAL ANALYSIS

Nuclear Volume measurements—Volumes of nuclei were defined using the Imaris “surfaces” function suing the following parameters: Enable Smooth = true, Surface Grain Size = 0.200μm, Enable Eliminate Background = true, Diameter Of Largest Sphere = 4.00μm, Enable Automatic Threshold = true. Data was exported to Rstudio (for all scripts see (https://rushlowlab.bio.nyu.edu/), and each image was assigned a sphericity score by taking the median of sphericity values. For any analysis that binned nuclei into stages, “Early” images were defined as having a sphericity index greater than 0.75, and “Late” images were defined as having a sphericity index less than 0.75. The positions of nuclear centers measured by the function were used to assign spots to the nearest nucleus.

HLB Measurements—HLBs were defined and quantified by finding the center of HLBs of all pictures using the Imaris “spots” function using the following parameters: Estimated XY Diameter = 0.408μm, Estimated Z Diameter = 0.815μm, Background Subtraction = true, “Quality” above 47.6. Centers were found using the *His3* probe signal due to the high signal to noise ratio of this channel. All position data was extracted from Imaris as .csv files and subsequent analysis was done in Rstudio. Channel alignment for Pol-II and MXC channels was performed using a custom content-aware alignment method in Rstudio named “findcenters,” which performed a lateral transformation to align the highest intensity points of all channels. After alignment, single 2μm × 2μm slices were taken of each HLB, and voxel values at each position in the slice were averaged together for all channels to generate the final metaplots. Intensity measurements were taken using the same procedures listed above, but instead of single slices, a sphere with a 0.5μm radius was defined at each HLB, and all voxel values within that sphere were averaged to create a measurement of intensity.

Pol-II HLB volumes were calculated using a custom R Script “findvolumes” to deal specifically with the difficulty of defining an accurate threshold value in images where background intensity is known to be variable based on biological factors. The method works by defining a 1μm × 1μm × 1μm cube centered on each HLB position, then adjusting a threshold value from the maximum to minimum voxel intensity value within the cube and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

counting the number of contiguous voxels that satisfy this threshold by flood-filling from a central seed point of the cube. To define an optimal threshold, the method looks for a characteristic spike in voxels included for a given threshold. This spike is seen clearly by taking the 3rd derivative (commonly referred to as "jerk") of a loess smoothed line generated by plotting the thresholds used against the number of voxels that threshold generates. This behavior is indicative of the threshold value reaching the background intensity values outside of the HLB, which rapidly changes the rate at which new voxels are added to the flood-fill in subsequent threshold values, and hence produces a spike in the 3rd derivative.

Pol-II spot calling (non HLB)—Visually identified nc12 embryos were acquired and loaded into Imaris software (Oxford Instruments). Images were first smoothed using a Gausian distribution to call the background level, and that was subtracted from each image by selecting the "background subtraction" icon. To "count" the number of Pol-II spots in each nucleus for comparison between *wt* and *zld*$^{-}$ genotypes, the following protocol was used. 1) The Imaris "spots" function, with the method of standard deviation between center-to-surrounding signal, was used in all 300 nm diameter spheres detected in the image. 300 nm was chosen because it included visually determined spots that were missed using the 250 nm or 400 nm window. Several different thresholds (T10-T110) in the Imaris "spots" function were tested and compared for spot calling (see example images in Figure S2A–B); the lower the threshold, the more spots were called (Figure S2C). 2) We chose a specific threshold that best represented the Pol-II spots visually identifiable by eye, i.e. spots stronger than diffuse nuclear "background" signal. There was no significant difference between the specific thresholds used for images of *wt* early (77±9.7), *wt* late (79±15.6), *zld*$^{-}$ early (87±13.0), and *zld*$^{-}$ late (85±4.5) (one way ANOVA on Rank, p=0.51) (Figure S2D). 3) To further exclude "background noise" that remained after thresholding (spots still undetectable by eye), we created a filter based on a specific Pol-II nuclear intensity level similar to diffuse nuclear signal. This approach using Pol-II signal to estimate nuclear space is similar to that used by Cisse et al., 2013[51]. A second filter based on Hoechst staining, which defined nuclear spaces, was also used in order to eliminate extra-nuclear spots so they would not be counted in the nuclear spot counts. By using these filters, both "weak" background spots inside the nucleus, and all extra-nuclear spots, were excluded in counts (white spots in Fig. S2A, B). 4) To obtain the final number of spots per nucleus, spots first had to be assigned to specific nuclei, by first identifying all nuclei in the image using the Imaris "surfaces" function, which can distinguish when Pol-II signal intensity drops sharply (see image in Figure S2E) and define homogenous masses. Then the centers of each homogeneous mass were calculated to obtain nuclear centers, and each Pol-II spot was then assigned to the nucleus whose center of mass was closest. Spots per nucleus was then calculated and an average obtained to compare number of spots in *wt* versus *zld*$^{-}$ nuclei.

Spot calling for *in silico* analysis—Because the Pol-II spot distribution analysis used a different mounting method to avoid flattening of embryos, the signal varied between embryos even on the same slide, thus thresholds needed adjusting for each embryo. However, there was no significant difference between thresholds for *yw* (105 ± 29.5) and *zld*$^{-}$ (100 ± 28.7) (t-test, p=0.25). Diffuse Pol-II spots lying outside of the nucleus and large spots comprising the HLBs as well as the centrosomes were removed.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Number density of Pol-II spots: To calculate number density of Pol-II spots (ρ), we drew an imaginary sphere of radius D_0, the average diameter of nucleus (4.76 μm) centered at center of mass of each nucleus and counted the number of spots inside it and divided it by the volume of the sphere to obtain the number density of Pol-II spots in each nucleus. The mean and standard deviation is calculated for all nuclei in *wt* and the *zld*$^-$ embryos. We averaged over 300 spheres for both cases and densities for *wt* and *zld*$^-$.

Radial distribution of density: To quantify how the density of spots varies with the distance from centers of mass of the nuclei, we drew spherical shells of thickness 0.1 μm centered at the center of mass of each nucleus and calculated the number of Pol-II spots in each shell. We defined $\rho(r)$, local density at distance r from the center of mass as the ratio of the mean number of particles in the spherical shell at distance r (averaged over all *wt* and *zld*$^-$ embryos separately) and the volume of the spherical shell at distance r. Figure 3B suggests that local density at all distances from the center is close to the global density of dots in both *wt* and *zld*$^-$ nuclei.

Distribution of distances between neighbors: To check if there is any correlation between the position of Pol-II spots within nuclei, we calculated $V(r)$, the probability distribution function of distance between nearest neighbor spots. To find the nearest neighbor of each spot, we found the Delaunay triangulation of the point pattern (where each point represents a Pol-II spot) using the *delaunayTriangulation* function in MATLAB. Next, we allocated all pairs of points in each Delaunay tetrahedron to be each other's nearest neighbors. Then, we calculated the probability distribution of distance between each pair of nearest neighbors.

Generation of synthetic point patterns: To check whether the Pol-II spots within the nucleus are uniformly distributed or clustered, we generated synthetic point patterns, uniformly distributed within a sphere, and distributed in a few clusters, with number densities similar to Pol-II spots in the *wt* and *zld*$^-$ nuclei. To generate a uniform (Poisson) distribution of particles, we took a sphere of diameter D and generated $\rho\pi D^3/6$ particles within the sphere using the random number generator in MATLAB. To generate clustered distribution of particles, we took a sphere of diameter D and divided $\rho\pi D^3/6$ particles into m clusters. First, we placed m points within the sphere randomly as the center of each cluster. Next, we drew a sphere of diameter $f(D/m)^{1/3}$, where f<1, around each center. Finally, we distributed $\rho\pi D^3/6m$ particles randomly into each of the m clusters.

Dual-color FISH 3D distance analysis—After stacks of images were loaded into Imaris, the positions of foci were determined with the Imaris "spots" function to detect up to 8 foci per nucleus. Duplicated sister chromatids were excluded from further analysis. Samples were mounted with coverslips to flatten embryos in a traditional manner so the measured distance may be larger than actual length. Distances between foci of different genes were measured with the Imaris "measure" function by assigning start and end points from the center of defined foci. Distances between FISH foci for two different genes were measured for the following pairs of genes (1D distances listed, starting with closest; also see Figure S3C): *slam* and *slam*#2 (*slam*-RC: 1500 to 3427 and 425 to 1412, respectively, 87 bp apart; *CR44504-RA* (+144 to +900) and *hiiragi*-RD (+1273 to +3038), 4 kb apart; *CG13712*-RA (−111 to +660) and *CG15876*-RA (+29 to +478), 14.4 kb apart; *BSG25D*-RD

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(+10 to +2380) and *CG14014*-RB (+58 to +1034), 272.5 kb apart; *CG14014*-RB (+58 to +1034) and *slam*-RC (+1500 to +3427), 824.2 kb apart, *elba1*-RA (+120 to +1095) and, *slam*-RC (+1500 to +3427), 1690.7 kb apart, *CG14014*-RB (+58 to +1034) and *CG15382-RA* (+113 to +979), 3393.1 kb apart, *hb*-RA (+164to +2440) and *slam*-RC (+1500~+3427) on different chromosomes; *hb* probe was a gift from Dr. Steve Small. Gene foci pairs in a nucleus were assigned in the following manner: a focus of "gene A" was assigned to the closest "gene B" focus. Up to two pairs of foci can be assigned per nucleus (representing the two genes on the two homologous chromosomes). Using the Mann-Whitney Rank Sum Test for any two of the groups, pairs with distances between 1700 kb ~ 15 kb have a significant positive correlation between 1D distance and 3D distance (Figure S3C). No significant differences were noted between 3.4 Mb and 1.7 Mb pairs, or 15 kb and 4 kb pairs. A control pair of *hb* and *slam*, two genes on different chromosomes, had the farthest 3D distance (2186±909 nm) and no close proximity was ever observed, as expected. Two probes against different regions of the *slam* gene (*slam* and *slam#2*, 87 bp apart) had an average distance of 185±33 nm, thus we established our resolution limit as 218 nm (mean plus one standard deviation covers 84% of data point).

smFISH 3D data analysis—3TAG and 0TAG transgene transcriptional output (Figure 2) was quantified in a similar manner to HLBs, by first defining the centers of nascent transcripts using the Imaris "spots" function using the following parameters: Estimated XY Diameter = 0.330μm, Estimated Z Diameter = 0.660μm, Background Subtraction = true, "Quality" above 13.5. Intensity of Pol-II channel and Pol-II signal metaplots were generated by the methods detailed above. In order to spatially bin nuclei as either "Dorsal" or "Ventral", the Dl gradient as defined by the Dl antibody stain was visually inspected.

smFISH for gene pairs (Figure 4K and S4) were called the same way as Pol-II spots to detect up to eight foci per nucleus, and their position and intensity was recorded. Duplicated sister chromatids were excluded from further analysis. Distances between foci of different genes were measured with the Imaris "measure" function by assigning start and end points from the center of defined foci. To compare smFISH signal of foci transcribed in close proximity versus distant, we chose nuclei containing one pair of close foci (<400 nm) and another pair of distant foci (*Bsg25D-CG14014* pair: 500nm-2300nm with an average of 828nm ± 346nm, *CG14014-slam* pair: 500nm-2500nm with an average of 1008nm ±414 nm, *Bsg25D-slam* pair: 500nm-3300nm with an average of 1057nm ± 438nm) and generated their signal intensity ratio (close/distant). These ratios were compared to the control group, comprising the remaining pairs of foci where both pairs of foci were either close or were distant. Those that did not have two pairs were omitted from the analysis. The Mann-Whitney Rank Sum Test was used to compare experimental groups with their control groups. The One-way ANOVA on ranks test was used to compare control groups.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Acknowledgements

The authors thank Robert Duronio for the Mxc antibody and Sevinc Ercan for her insightful ideas over the course of this study. The authors also thank Bikhyat Shrestha, Nan Lin, Herman Liu, and Sammy Almeida for their help with embryo collections and Imaris image analyses, and Shigehiro Yamada for making the 3TAG and 0TAG plasmids. The research was supported by National Institute of Health (NIH) research grants: RO1GM63024 to CAR, RO1GM134204 to SYS, and T32HD7520 Training Program in Developmental Genetics to PHW.

References

1. Liang HL, Nien CY, Liu HY, Metzstein MM, Kirov N, and Rushlow C (2008). The zinc-finger protein Zelda is a key activator of the early zygotic genome in Drosophila. Nature 456, 400–403. 10.1038/nature07388. [PubMed: 18931655]
2. White AE, Leslie ME, Calvi BR, Marzluff WF, and Duronio RJ (2007). Developmental and cell cycle regulation of the Drosophila histone locus body. Mol Biol Cell 18, 2491–2502. 10.1091/mbc.e06-11-1033. [PubMed: 17442888]
3. Nizami ZF, Deryusheva S, and Gall JG (2010). Cajal bodies and histone locus bodies in Drosophila and Xenopus. Cold Spring Harb Symp Quant Biol 75, 313–320. 10.1101/sqb.2010.75.005. [PubMed: 21047905]
4. Nien CY, Liang HL, Butcher S, Sun Y, Fu S, Gocha T, Kirov N, Manak JR, and Rushlow C (2011). Temporal coordination of gene networks by Zelda in the early Drosophila embryo. PLoS Genet 7, e1002339. 10.1371/journal.pgen.1002339. [PubMed: 22028675]
5. Harrison MM, Botchan MR, and Cline TW (2010). Grainyhead and Zelda compete for binding to the promoters of the earliest-expressed Drosophila genes. Dev Biol 345, 248–255. 10.1016/j.ydbio.2010.06.026. [PubMed: 20599892]
6. De Renzis S, Elemento O, Tavazoie S, and Wieschaus EF (2007). Unmasking activation of the zygotic genome using chromosomal deletions in the Drosophila embryo. PLoS Biol 5, e117. 10.1371/journal.pbio.0050117. [PubMed: 17456005]
7. Chen K, Johnston J, Shao W, Meier S, Staber C, and Zeitlinger J (2013). A global change in RNA polymerase II pausing during the Drosophila midblastula transition. Elife 2, e00861. 10.7554/eLife.00861. [PubMed: 23951546]
8. Blythe SA, and Wieschaus EF (2015). Zygotic genome activation triggers the DNA replication checkpoint at the midblastula transition. Cell 160, 1169–1181. 10.1016/j.cell.2015.01.050. [PubMed: 25748651]
9. Cho WK, Spille JH, Hecht M, Lee C, Li C, Grube V, and Cisse II (2018). Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. Science 361, 412–415. 10.1126/science.aar4199. [PubMed: 29930094]
10. Boehning M, Dugast-Darzacq C, Rankovic M, Hansen AS, Yu T, Marie-Nelly H, McSwiggen DT, Kokic G, Dailey GM, Cramer P, et al. (2018). RNA polymerase II clustering through carboxy-terminal domain phase separation. Nat Struct Mol Biol 25, 833–840. 10.1038/s41594-018-0112-y. [PubMed: 30127355]
11. Iborra FJ, Pombo A, Jackson DA, and Cook PR (1996). Active RNA polymerases are localized within discrete transcription “factories’ in human nuclei. J Cell Sci 109 (Pt 6), 1427–1436. [PubMed: 8799830]
12. Carmo-Fonseca M, Tollervey D, Pepperkok R, Barabino SM, Merdes A, Brunner C, Zamore PD, Green MR, Hurt E, and Lamond AI (1991). Mammalian nuclei contain foci which are highly enriched in components of the pre-mRNA splicing machinery. EMBO J 10, 195–206. [PubMed: 1824936]
13. Faro-Trindade I, and Cook PR (2006). Transcription factories: structures conserved during differentiation and evolution. Biochem Soc Trans 34, 1133–1137. 10.1042/BST0341133. [PubMed: 17073768]
14. Spector DL (1993). Macromolecular domains within the cell nucleus. Annu Rev Cell Biol 9, 265–315. 10.1146/annurev.cb.09.110193.001405. [PubMed: 8280462]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

15. Guglielmi B, La Rochelle N, and Tjian R (2013). Gene-specific transcriptional mechanisms at the histone gene cluster revealed by single-cell imaging. Mol Cell 51, 480–492. 10.1016/j.molcel.2013.08.009. [PubMed: 23973376]

16. Terzo EA, Lyons SM, Poulton JS, Temple BR, Marzluff WF, and 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. 10.1091/mbc.E14-10-1445. [PubMed: 25694448]

17. Harrison MM, Li XY, Kaplan T, Botchan MR, and Eisen MB (2011). Zelda binding in the early Drosophila melanogaster embryo marks regions subsequently activated at the maternal-to-zygotic transition. PLoS Genet 7, e1002266. 10.1371/journal.pgen.1002266. [PubMed: 22028662]

18. Hur W, Kemp JP Jr., Tarzia M, Deneke VE, Marzluff WF, Duronio RJ, and Di Talia S (2020). CDK-Regulated Phase Separation Seeded by Histone Genes Ensures Precise Growth and Function of Histone Locus Bodies. Dev Cell 54, 379–394 e376. 10.1016/j.devcel.2020.06.003. [PubMed: 32579968]

19. Li J, Dong A, Saydaminova K, Chang H, Wang G, Ochiai H, Yamamoto T, and Pertsinidis A (2019). Single-Molecule Nanoscopy Elucidates RNA Polymerase II Transcription at Single Genes in Live Cells. Cell 178, 491–506 e428. 10.1016/j.cell.2019.05.029. [PubMed: 31155237]

20. Sexton T, Umlauf D, Kurukuti S, and Fraser P (2007). The role of transcription factories in large-scale structure and dynamics of interphase chromatin. Semin Cell Dev Biol 18, 691–697. 10.1016/j.semcdb.2007.08.008. [PubMed: 17950637]

21. Eskiw CH, and Fraser P (2011). Ultrastructural study of transcription factories in mouse erythroblasts. J Cell Sci 124, 3676–3683. 10.1242/jcs.087981. [PubMed: 22045738]

22. Foo SM, Sun Y, Lim B, Ziukaite R, O'Brien K, Nien CY, Kirov N, Shvartsman SY, and Rushlow CA (2014). Zelda potentiates morphogen activity by increasing chromatin accessibility. Curr Biol 24, 1341–1346. 10.1016/j.cub.2014.04.032. [PubMed: 24909324]

23. Xu Z, Chen H, Ling J, Yu D, Struffi P, and Small S (2014). Impacts of the ubiquitous factor Zelda on Bicoid-dependent DNA binding and transcription in Drosophila. Genes Dev 28, 608–621. 10.1101/gad.234534.113. [PubMed: 24637116]

24. Yamada S, Whitney PH, Huang SK, Eck EC, Garcia HG, and Rushlow CA (2019). The Drosophila Pioneer Factor Zelda Modulates the Nuclear Microenvironment of a Dorsal Target Enhancer to Potentiate Transcriptional Output. Curr Biol 29, 1387–1393 e1385. 10.1016/j.cub.2019.03.019. [PubMed: 30982648]

25. Kanodia JS, Liang HL, Kim Y, Lim B, Zhan M, Lu H, Rushlow CA, and Shvartsman SY (2012). Pattern formation by graded and uniform signals in the early Drosophila embryo. Biophys J 102, 427–433. 10.1016/j.bpj.2011.12.042. [PubMed: 22325264]

26. Dufourt J, Trullo A, Hunter J, Fernandez C, Lazaro J, Dejean M, Morales L, Nait-Amer S, Schulz KN, Harrison MM, et al. (2018). Temporal control of gene expression by the pioneer factor Zelda through transient interactions in hubs. Nat Commun 9, 5194. 10.1038/s41467-018-07613-z. [PubMed: 30518940]

27. Sun Y, Nien CY, Chen K, Liu HY, Johnston J, Zeitlinger J, and Rushlow C (2015). Zelda overcomes the high intrinsic nucleosome barrier at enhancers during Drosophila zygotic genome activation. Genome Res 25, 1703–1714. 10.1101/gr.192542.115. [PubMed: 26335633]

28. Schulz KN, Bondra ER, Moshe A, Villalta JE, Lieb JD, Kaplan T, McKay DJ, and Harrison MM (2015). Zelda is differentially required for chromatin accessibility, transcription factor binding, and gene expression in the early Drosophila embryo. Genome Res 25, 1715–1726. 10.1101/gr.192682.115. [PubMed: 26335634]

29. Yanez-Cuna JO, Dinh HQ, Kvon EZ, Shlyueva D, and Stark A (2012). Uncovering cis-regulatory sequence requirements for context-specific transcription factor binding. Genome Res 22, 2018–2030. 10.1101/gr.132811.111. [PubMed: 22534400]

30. Mir M, Stadler MR, Ortiz SA, Hannon CE, Harrison MM, Darzacq X, and Eisen MB (2018). Dynamic multifactor hubs interact transiently with sites of active transcription in Drosophila embryos. Elife 7. 10.7554/eLife.40497.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

31. Andrey G, Montavon T, Mascrez B, Gonzalez F, Noordermeer D, Leleu M, Trono D, Spitz F, and Duboule D (2013). A switch between topological domains underlies HoxD genes collinearity in mouse limbs. Science 340, 1234167. 10.1126/science.1234167. [PubMed: 23744951]
32. Simonis M, Klous P, Splinter E, Moshkin Y, Willemsen R, de Wit E, van Steensel B, and de Laat W (2006). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). Nat Genet 38, 1348–1354. 10.1038/ng1896. [PubMed: 17033623]
33. Szabo Q, Bantignies F, and Cavalli G (2019). Principles of genome folding into topologically associating domains. Sci Adv 5, eaaw1668. 10.1126/sciadv.aaw1668. [PubMed: 30989119]
34. Rieder D, Ploner C, Krogsdam AM, Stocker G, Fischer M, Scheideler M, Dani C, Amri EZ, Muller WG, McNally JG, and Trajanoski Z (2014). Co-expressed genes prepositioned in spatial neighborhoods stochastically associate with SC35 speckles and RNA polymerase II factories. Cell Mol Life Sci 71, 1741–1759. 10.1007/s00018-013-1465-3. [PubMed: 24026398]
35. Chakalova L, and Fraser P (2010). Organization of transcription. Cold Spring Harb Perspect Biol 2, a000729. 10.1101/cshperspect.a000729. [PubMed: 20668006]
36. Hnisz D, Shrinivas K, Young RA, Chakraborty AK, and Sharp PA (2017). A Phase Separation Model for Transcriptional Control. Cell 169, 13–23. 10.1016/j.cell.2017.02.007. [PubMed: 28340338]
37. Kumar VS, and Kumaran V (2005). Voronoi neighbor statistics of hard-disks and hard-spheres. J Chem Phys 123, 074502. 10.1063/1.2000233. [PubMed: 16229596]
38. Hug CB, Grimaldi AG, Kruse K, and Vaquerizas JM (2017). Chromatin Architecture Emerges during Zygotic Genome Activation Independent of Transcription. Cell 169, 216–228 e219. 10.1016/j.cell.2017.03.024. [PubMed: 28388407]
39. Ing-Simmons E, Vaid R, Bing XY, Levine M, Mannervik M, and Vaquerizas JM (2021). Independence of chromatin conformation and gene regulation during Drosophila dorsoventral patterning. Nat Genet 53, 487–499. 10.1038/s41588-021-00799-x. [PubMed: 33795866]
40. Stadler MR, Haines JE, and Eisen MB (2017). Convergence of topological domain boundaries, insulators, and polytene interbands revealed by high-resolution mapping of chromatin contacts in the early Drosophila melanogaster embryo. Elife 6. 10.7554/eLife.29550.
41. Chen H, Levo M, Barinov L, Fujioka M, Jaynes JB, and Gregor T (2018). Dynamic interplay between enhancer-promoter topology and gene activity. Nat Genet 50, 1296–1303. 10.1038/s41588-018-0175-z. [PubMed: 30038397]
42. Heist T, Fukaya T, and Levine M (2019). Large distances separate coregulated genes in living Drosophila embryos. Proc Natl Acad Sci U S A 116, 15062–15067. 10.1073/pnas.1908962116. [PubMed: 31285341]
43. Cook PR, and Marenduzzo D (2018). Transcription-driven genome organization: a model for chromosome structure and the regulation of gene expression tested through simulations. Nucleic Acids Res 46, 9895–9906. 10.1093/nar/gky763. [PubMed: 30239812]
44. Espinola SM, Gotz M, Bellec M, Messina O, Fiche JB, Houbron C, Dejean M, Reim I, Cardozo Gizzi AM, Lagha M, and Nollmann M (2021). Cis-regulatory chromatin loops arise before TADs and gene activation, and are independent of cell fate during early Drosophila development. Nat Genet 53, 477–486. 10.1038/s41588-021-00816-z. [PubMed: 33795867]
45. Staller MV, Yan D, Randklev S, Bragdon MD, Wunderlich ZB, Tao R, Perkins LA, Depace AH, and Perrimon N (2013). Depleting gene activities in early Drosophila embryos with the “maternal-Gal4-shRNA” system. Genetics 193, 51–61. 10.1534/genetics.112.144915. [PubMed: 23105012]
46. Bothma JP, Garcia HG, Esposito E, Schlissel G, Gregor T, and Levine M (2014). Dynamic regulation of eve stripe 2 expression reveals transcriptional bursts in living Drosophila embryos. Proc Natl Acad Sci U S A 111, 10598–10603. 10.1073/pnas.1410022111. [PubMed: 24994903]
47. Lecuyer E, Parthasarathy N, and Krause HM (2008). Fluorescent in situ hybridization protocols in Drosophila embryos and tissues. Methods Mol Biol 420, 289–302. 10.1007/978-1-59745-583-1_18. [PubMed: 18641955]
48. Gaspar I, Wippich F, and Ephrussi A (2017). Enzymatic production of single-molecule FISH and RNA capture probes. RNA 23, 1582–1591. 10.1261/rna.061184.117. [PubMed: 28698239]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

49. Freese NH, Norris DC, and Loraine AE (2016). Integrated genome browser: visual analytics platform for genomics. Bioinformatics 32, 2089–2095. 10.1093/bioinformatics/btw069. [PubMed: 27153568]

50. Kerpedjiev P, Abdennur N, Lekschas F, McCallum C, Dinkla K, Strobelt H, Luber JM, Ouellette SB, Azhir A, Kumar N, et al. (2018). HiGlass: web-based visual exploration and analysis of genome interaction maps. Genome Biol 19, 125. 10.1186/s13059-018-1486-1. [PubMed: 30143029]

51. Cisse II, Izeddin I, Causse SZ, Boudarene L, Senecal A, Muresan L, Dugast-Darzacq C, Hajj B, Dahan M, and Darzacq X (2013). Real-time dynamics of RNA polymerase II clustering in live human cells. Science 341, 664–667. 10.1126/science.1239053. [PubMed: 23828889]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Highlights

- RNA Polymerase II (Pol-II) accumulates into nuclear speckles at genome activation
- Without the pioneer factor Zelda, speckle number and size are greatly reduced
- Pol-II moves instead to the histone locus bodies and histone RNA synthesis increases
- Genes in close 3D proximity do not share a Pol-II speckle

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

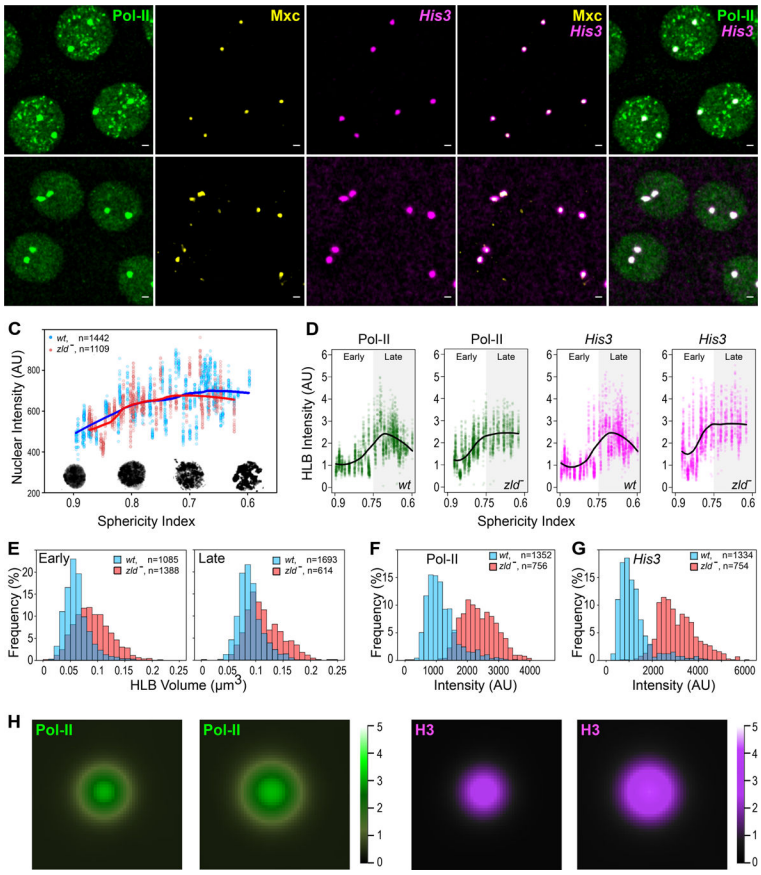


Figure 1. Pol-II accumulates at the HLBs in the absence of early gene transcription.

(A and B) Immunofluorescence of *wt* (**A**) and *zld*¯ (**B**) nc12 embryos using antibodies against RPB1 for Pol-II (green), Multi sex Combs (Mxc, yellow), and *His3* RNA (magenta). Scale bar = 1 μm. **(C)** Total levels of Pol-II fluorescence plotted across a pseudo-time axis. The average voxel intensity of Pol-II was plotted using the median sphericity value of each image. A loess regression line shows similar increases of total Pol-II for both genotypes. Example Hoechst-stained nuclei for the corresponding sphericity values are shown as insets, highlighting the dramatic change in nuclear morphology across the pseudo-time axis. **(D)** Pol-II and *His3* fluorescence at the HLBs plotted in the same manner described in (C). Values were generated by taking the mean voxel value of a sphere with a radius of 500 nm centered on the positions of HLBs. We highlighted distinct “Early” and “Late” phases as images where the median sphericity value was greater or less than 0.75 respectively (see white versus grey shading in Figure 1D panels). Pol-II and *His3* within each genotype agree well across the pseudo-time axis. **(E)** Histograms of volumes of HLBs based on Pol-II signal in both “Early” (left) and “Late” (right) embryos (*wt* shaded in blue, *zld*¯ in coral). Volumes were calculated via a thresholding technique (see Methods) and by counting the number of voxels satisfied by those determined thresholds. Distributions of volumes show significant differences in both Early and Late stages. **(F and G)** All Pol-II (F) and *His3* (G) fluorescence value distributions are from “Early” embryos. Large differences are seen between *wt* (blue) and *zld*¯ (coral). **(H)** HLB metaprofiles (collated 2 μm × 2 μm images with HLBs at the center) showing Pol-II (green) and *His3* fluorescence (magenta) intensities

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

in *wt* (left) and *zld*$^{-}$ (right). Heat maps indicate increasing fluorescence intensity, with white as max intensity and decreasing through green to yellow to black. Note the overall "shelled" organization of *wt* and *zld*$^{-}$ HLBs are similar; the *zld*$^{-}$ HLBs are on average simply larger. Sample numbers are indicated in the upper left (C) or right (E-G) corner of the panel; the same samples were used for the histograms (F-G) and metaprofiles (H).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

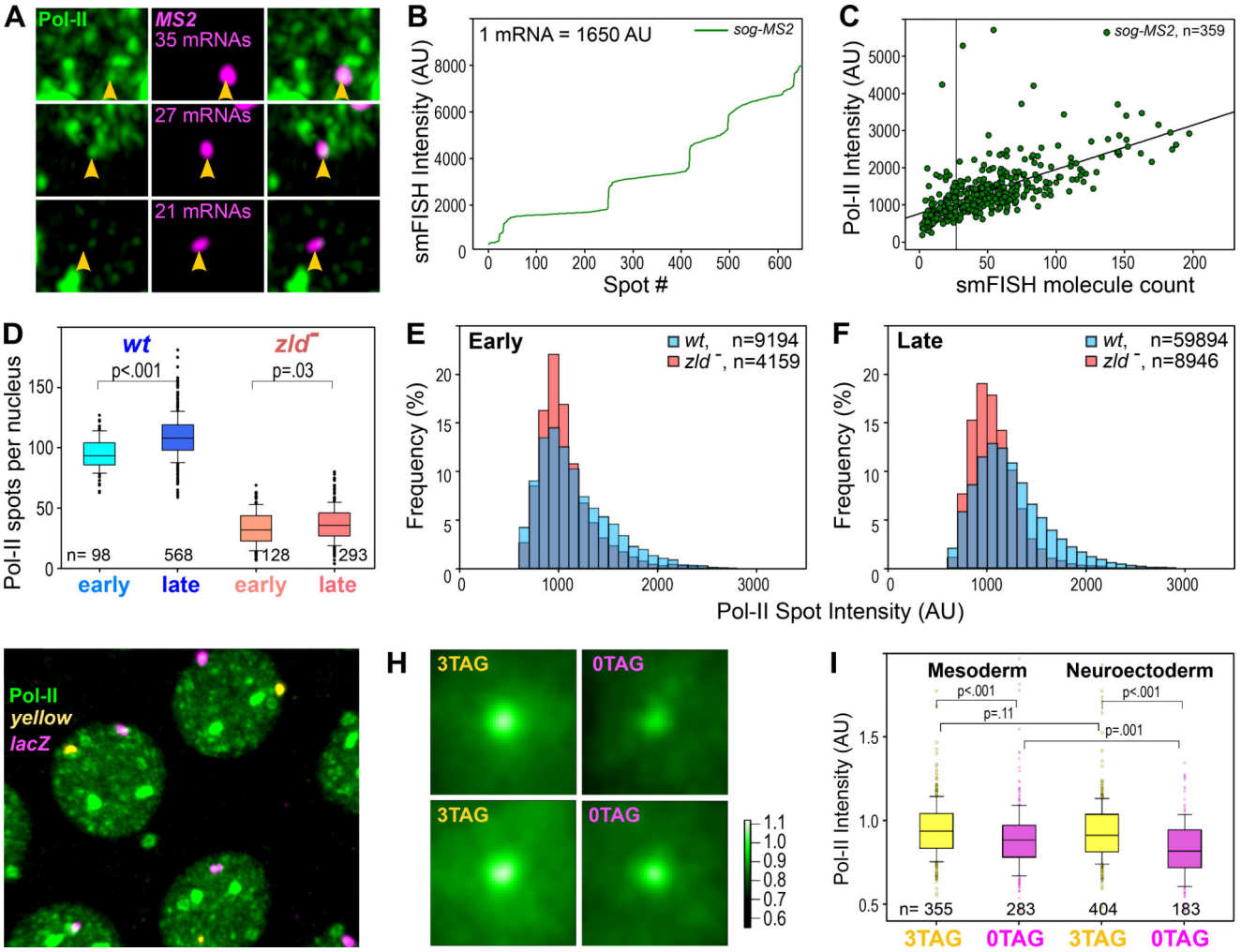


Figure 2. Zld dictates Pol-II speckle number and intensity in early embryos.

(A) Fluorescence imaging of RNA Pol-II and smFISH of the transcripts produced by the *sog-lacZ* (3TAG) enhancer reporter transgene[24]. Spots of RNA Pol-II show overlap with high intensity FISH staining, assumed to be the sites of nascent transcription. The number of nascent transcripts measured scales with brightness of the corresponding RNA Pol-II spot. **(B)** smFISH spot intensities ranked in order of intensity show a "stairstep" pattern, indicating discrete numbers of underlying molecules labeled by the probe set. Each step increases on average by 1650 AU, giving an estimate of the fluorescence produced by a single molecule. **(C)** Number of molecules plotted against the fluorescence of the corresponding RNA Pol-II spot. A strong linear correlation agrees with the assumption that the population of Pol-II within each spot is coupled to nascent transcripts, and therefore represent actively elongating polymerases. **(D)** Box plot showing the distribution of Pol-II spots in *wt* (blue) and *zld*¯ (coral) in early (lighter) versus late (darker) interphase. Sample numbers indicated underneath the plots. *wt* nuclei contain more spots in both early (94.8±13.0) and late (136±26.1) interphase than *zld*¯ early (33.5±13.8) and late (42±17.7). The difference in spot counts between early and late interphase is significant (Mann-Whitney Rank Sum Test) for both genotypes: *wt* (p<.001) and *zld*¯ (p=0.03). **(E-F)**

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Histograms of Pol-II fluorescence intensity (AU) distributions for *wt* (blue) and *zld*$^{-}$ (coral) embryos in early interphase (E), *wt*: 1158±370.8 and *zld*$^{-}$:1071±292.3) and late interphase (F), *wt*: 1225±379.5 and *zld*$^{-}$:1074±241.9). The difference in spot intensities between *wt* and *zld*$^{-}$ are significant (Mann-Whitney Rank Sum Test) for both early ($p<.001$) and late interphase ($p<.001$). **(G)** 3TAG/0TAG heterozygous embryo stained with antibodies against Pol-II (green) and smFISH probes against *yellow* (yellow) and *lacZ* (magenta) to detect nascent transcript foci. Scale bar = 1 μm. **(H)** Metaprofiles (2 μm × 2 μm images) of Pol-II fluorescence intensity at 3TAG-*y* (left) and 0TAG-*lacZ* (right) foci in lateral neuroectoderm (top) and ventral mesoderm (bottom) nuclei where Dl levels are low and high, respectively. **(I)** Box plot distributions of genotypes in (H) showing significant differences in Pol-II intensities between 3TAG and 0TAG in both regions ($p=0.001$), and between 0TAG in the mesoderm versus neuroectoderm ($p=0.001$), but not 3TAG ($p=0.11$). Sample numbers are indicated underneath the plots. In all box plots, the box includes the 25th-75th percentile with the horizontal line marking the median. The lower and upper whiskers reach to 10th and 90th percentiles, respectively, and the outliers are shown as dots.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

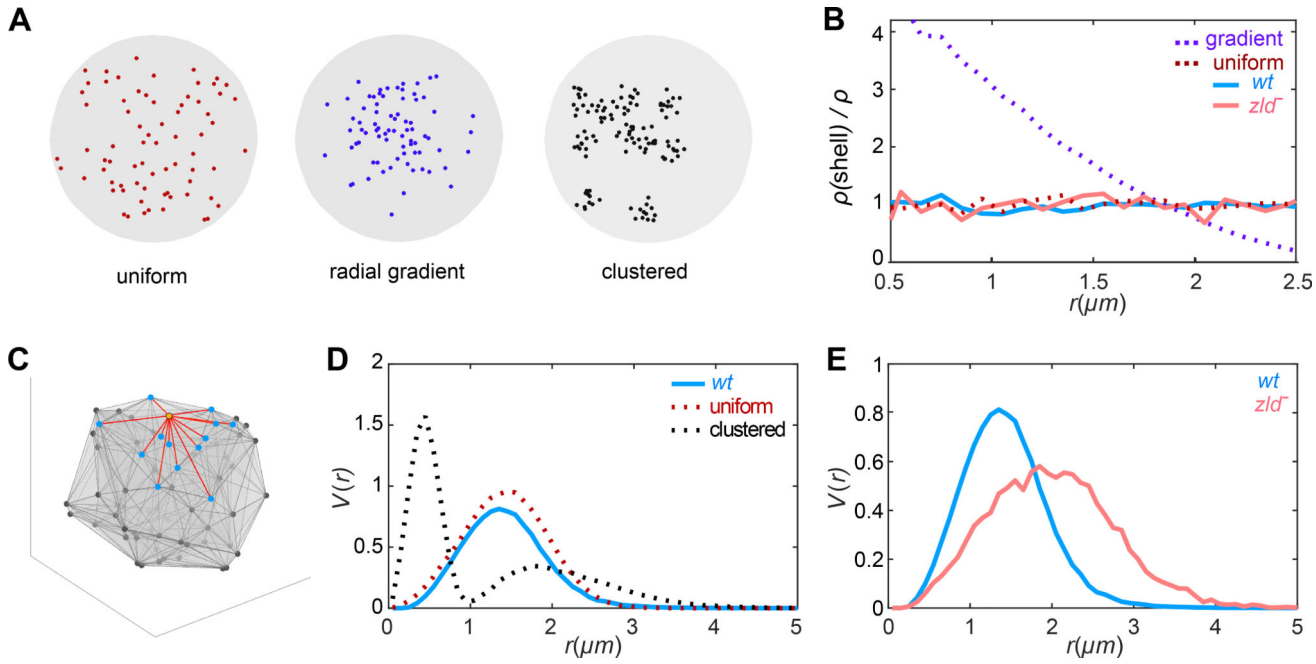


Figure 3. Pol-II speckles are distributed evenly and randomly throughout the nucleus in both *wt* and *zld*⁻.

(A) Synthetic point patterns distributed uniformly (left), in a gradient along the radius (center), or clustered within a nuclear sphere (right). Note that a reverse gradient or extreme clustering (one cluster only) could also occur. **(B)** The ratio of shell density, ρ(shell), of Pol-II spots to global density of spots, ρ, as a function of distance from the center of mass of each nucleus for *wt* (blue) and *zld*⁻ (coral) embryos. Simulated hypothetical ratios are shown as dashed lines, uniform (red) and a center-to-periphery gradient (purple). **(C)** Schematic example of Pol-II speckles in a nucleus. One speckle (orange), its neighbors (blue), and lines between them (red) are highlighted. **(D)** Probability distribution functions of the distance between nearest neighbors ($V(r)$) in *wt* (solid blue line) compared to hypothetical uniform (red dashed line) and clustered (black dashed line) distributions in the same density shown in (A). **(E)** Probability distribution function of distances between nearest neighbors ($V(r)$) in *wt* (blue) and *zld*⁻ (coral) nuclei. The shift of the coral curve to the right reflecting greater distances between spots is due to the reduced total number of spots in *zld*⁻ embryos, while the shift down occurs since Y-axis measures frequency.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

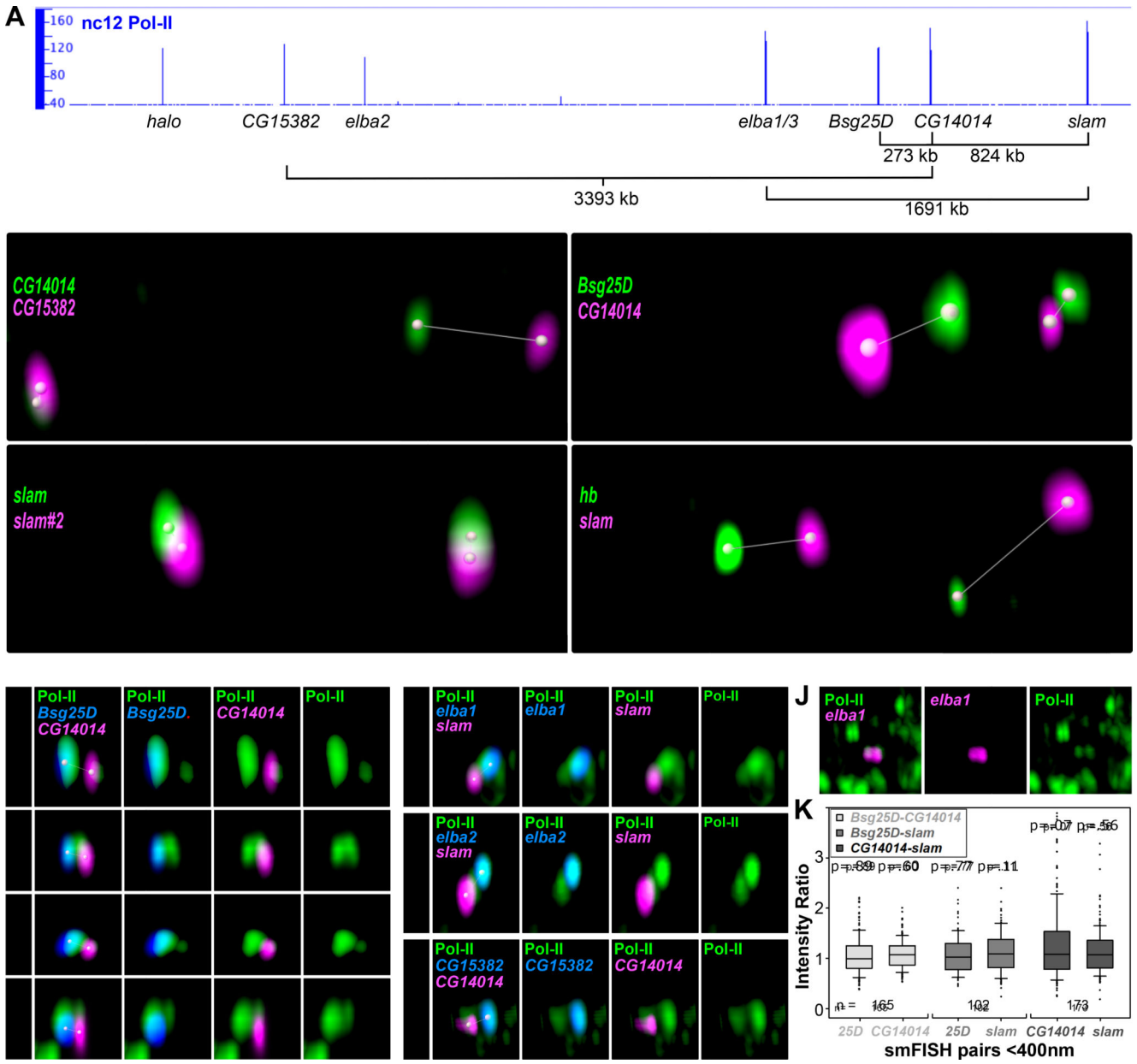


Figure 4. Zld co-regulated genes do not share Pol-II speckles or increased transcriptional output when in close proximity.

A) Integrated Genome Browser view of Pol-II ChIP-seq[8] peaks in a 5Mb region of chromosome 2, with one-dimension (1D) distances labeled for several pairs of active genes. **(B-E)** 3D Imaris views of single nuclei with dual color FISH for the following pairs of genes (color code indicated): *CG15382* and *CG14014* (B), *Bsg25D* and *CG14014* (C), *slam* and *slam#2* (D), and *hb* and *slam* (E). Distances are measured by Imaris software. *slam* and *slam* #2 are two probes for *slam* mRNA, 78bp apart. **(F-I)** Zoom-in 3D images of Pol-II immunofluorescence (IF) (green) and dual color RNA FISH for the following pairs of genes (color code indicated and distances between FISH foci centers shown on the left): *Bsg25D* and *CG14014* (F), *elba1* and *slam* (G), *elba2* and *slam* (H), and *CG15382* and *CG14014*

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(I). Note that Pol-II spots overlap with each RNA FISH signal, and are separable. Scale bar = 500 nm. **(J)** Images of Pol-II IF (green) and *elba1* RNA FISH (magenta) on sister chromatids. Note the *elba1* signals have not separated completely, though the associated Pol-II spots are separable (233nm apart). Scale bar = 500 nm. **(K)** Box plot showing the ratio of smFISH signal intensity (close proximity/distant) for each gene of each pair tested (as indicated); close proximity pairs are <400nm apart, distant pairs are 500nm and up apart (see distribution of distances in Figure S4B). There is no significant difference between each experimental group and its control group (Mann-Whitney Rank Sum Test, p-values inside box plot).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Key Resources Table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse anti-RBP1 conjugated with Alexa Flour488 (CTD4H8)	Sigma-Aldrich	cat# 05-623-AF488 RRID: AB_309852
Guinea pig anti-Multi sex combs (Mxc)	Terzo, Lyons, Poulton, Temple, Marzluff, and Duronio[16]	N/A
Rabbit anti-Zelda antibody	This manuscript	N/A
Donkey anti mouse secondary antibody, Alexa fluor 488	ThermoFisher scientific	Cat# A-21202 RRID: AB_141607
Goat anti guinea pig secondary antibody, Alexa fluor 647	ThermoFisher scientific	Cat# A-21450 RRID:AB_2735091
Sheep anti-DIG-AP antibody	Roche	Cat# 11093274910 RRID:AB_2734716
Rabbit anti-FL antibody	Invitrogen	Cat# A-889 RRID: AB_221561
Donkey anti rabbit secondary antibody, Alexa fluor 555	ThermoFisher scientific	Cat# A-31572 RRID: AB_162543
Donkey anti sheep secondary antibody, Alexa fluor 647	ThermoFisher scientific	Cat# A-21448 RRID:AB_2535865
Chemicals, Peptides, and Recombinant Proteins		
DIG RNA Labeling Kit	Roche	Cat# 11175025910
Fluorescein RNA Labeling Kit	Roche	Cat# 11685619910
smFISH Hybridization Buffer	Biosearch Technologies	Cat# SMF-HB1–10
Terminal Transferase	NEB	Cat# M0315S
Amino-11-ddUTP	Lumiprobe	Cat# A5040
Alexa Fluor™ 555 NHS Ester	Invitrogen	Cat# A-20009
Alexa Fluor™ 594 NHS Ester	Invitrogen	Cat# A-20004
Alexa Fluor™ 647 NHS Ester	Invitrogen	Cat# A-20006
Hoechst 33342	R&D systems	Cat# 5117/50
Experimental Models: Drosophila strains		
y[1] w[1118]	Bloomington Drosophila Stock Center	Cat# 6598
Maternal Triple Driver(MTD)-Gal4: P^{52}; P{Gal4-*nos*.NGT}40; P{*nos*-Gal4-VP16}c	Bloomington Drosophila Stock Center	Cat# 31777, RRID:BDSC_31777
UAS-shRNA-*zld*	Sun, Nien, Chen, Liu, Johnston, Zeitlinger, Rushlow[27]	N/A
y[1] w[*]; sog 3TAG-MS2v7-*lacZ*	Yamada, Whitney, Huang, Eck, Garcia, Rushlow[24]	N/A
y[1] w[*]; sog 0TAG-MS2v7-*lacZ*	Yamada, Whitney, Huang, Eck, Garcia, Rushlow[24]	N/A
Oligonucleotides		
Stellaris *lacZ* Quasar 670 smFISH Probe	Biosearch Technologies	SMF-1065–5
Stellaris *yellow* Quasar 570 smFISH Probe	Biosearch Technologies	SMF-1083–5
Stellaris MS2 Quasar 570 smFISH Probe	Biosearch Technologies	SMF-1063–5
Custom Stellaris smFISH probes (*lacZ, yellow, MS2*)	Biosearch Technologies	Table S1

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Custom probes for smFISH (*histone H3, Bsg25D, CG14014, slam*)	IDT	Table S1
Software and Algorithms		
FIJI (ImageJ)	NIH	http://fiji.sc
Imaris	Bitplane	http://www.bitplane.org
Matlab	The Mathworks Inc.	https://www.mathworks.com
R	The R Foundation	https://www.r-project.org
All the graphing functions in R	This manuscript	https://doi.org/10.6084/m9.figshare.15167187.v1
convert imaris csv file to R data	This manuscript	https://doi.org/10.6084/m9.figshare.15167169.v1
yellow/lacZ quantifier	This manuscript	https://doi.org/10.6084/m9.figshare.15167199.v1
Integrated Genome Browser 9.1.4	Freese, Norris, and Loraine[49]	Bioviz https://www.bioviz.org/genome-dashboard
Hi-Glass	Kerpedjiev et al.[50]	https://higlass.io/
Zen (Black)	Zeiss	https://www.zeiss.com/microscopy/us/products/microscope-software/zen.html
Other		
Confocal microscope	Zeiss	LSM880
Power meter(X-cite)	Lumen Dynamics Group Inc, Canada	Model # XR2100
Aqua-Poly/Mount	Polysciences	Cat# 18606–20
ProLong™ Glass Antifade Mountant	Invitrogen	Cat# P36984
Number 1.5 glass coverslips	Fisher Scientific	Cat# 22266858
High Precision Deckgläser No. 1.5H coverslips	Thor labs	Cat# CG15CH2

1. Terzo, E.A., Lyons, S.M., Poulton, J.S., Temple, B.R., Marzluff, W.F., and Duronio, R.J. (2015). Distinct self-interaction domains promote Multi Sex Combs accumulation in and formation of the Drosophila histone locus body. Mol Biol Cell *26*, 1559–1574. 10.1091/mbc.E14-10-1445.

2. Liang, H.L., Nien, C.Y., Liu, H.Y., Metzstein, M.M., Kirov, N., and Rushlow, C. (2008). The zinc-finger protein Zelda is a key activator of the early zygotic genome in Drosophila. Nature *456*, 400–403. 10.1038/nature07388.

3. Duronio, R.J., and Marzluff, W.F. (2017). Coordinating cell cycle-regulated histone gene expression through assembly and function of the Histone Locus Body. RNA Biol *14*, 726–738. 10.1080/15476286.2016.1265198.

4. Freese, N.H., Norris, D.C., and Loraine, A.E. (2016). Integrated genome browser: visual analytics platform for genomics. Bioinformatics *32*, 2089–2095. 10.1093/bioinformatics/btw069.

5. Kerpedjiev, P., Abdennur, N., Lekschas, F., McCallum, C., Dinkla, K., Strobelt, H., Luber, J.M., Ouellette, S.B., Azhir, A., Kumar, N., et al. (2018). HiGlass: web-based visual exploration and analysis of genome interaction maps. Genome Biol *19*, 125. 10.1186/s13059-018-1486-1.