bioRχiv

Enhancer binding kinetics explain transcription factor hub formation

Samantha Fallacaro,[1,2,3] Manya Kapoor,[1,3,4] Lillian Encarnation,[1,3,5] Apratim Mukherjee,[2,3] Meghan A. Turner,[6,7] Hernan G. Garcia,[6,8,9,10,11] and Mustafa Mir[1,3,12,13*]

[1]Department of Cell and Developmental Biology, Perelman School of Medicine, University of Pennsylvania; Philadelphia, PA, USA
[2]Developmental, Stem Cell, and Regenerative Biology Graduate Group, Perelman School of Medicine; Philadelphia, PA, USA
[3]Center for Computational and Genomic Medicine, Children's Hospital of Philadelphia, Philadelphia, PA, USA
[4]Department of Bioengineering, University of Pennsylvania; Philadelphia, PA, USA
[5]Department of Biology, University of Pennsylvania; Philadelphia, PA, USA
[6]Biophysics Graduate Group, University of California at Berkeley, Berkeley, CA, USA
[7]Current Address: Allen Institute for Brain Science, Seattle, WA, USA
[8]Department of Molecular and Cell Biology, University of California, Berkeley, CA, USA
[9]Department of Physics, University of California, Berkeley, CA, USA
[10]Institute for Quantitative Biosciences-QB3, University of California, Berkeley, CA, USA
[11]Chan Zuckerberg Biohub – San Francisco, San Francisco, CA, USA
[12]Howard Hughes Medical Institute, Children's Hospital of Philadelphia, Philadelphia, PA, USA
[13]Epigenetics Institute, University of Pennsylvania Perelman School of Medicine; Philadelphia, PA, USA
*Correspondence should be addressed to M.M. (mustafa.mir@pennmedicine.upenn.edu)

Transcription factors (TFs) form dynamic, high-concentration clusters, or "hubs", proposed to increase TF binding frequency at target enhancers. However, how enhancer sequence shapes hub properties remains unclear. We developed a live-imaging based framework to quantify the spatiotemporal relationship between TF hubs and actively transcribed genes in live Drosophila embryos. Examining hubs formed by the TF Dorsal across enhancers with defined binding-site composition, we find that hub enrichment and persistence scale with the number of Dorsal binding motifs. However, these hub properties do not predict transcriptional burst behavior for a given enhancer. Combining quantitative imaging with computational modeling, we show that Dorsal hubs arise from TF–DNA binding kinetics rather than actively increasing target occupancy. These findings support a model in which TF hubs are emergent features of enhancer-encoded TF–DNA interactions encoded by enhancer sequence rather than higher-order regulatory assemblies.

Introduction

The spatiotemporal regulation of gene expression during embryonic development depends on the ability of sequence-specific transcription factors to find and bind their target sites within the crowded nuclear environment. Live-imaging and single-molecule tracking experiments have shown that transcription factors (TFs) bind their genomic targets transiently, with residence times on the order of tens of seconds (*1, 2*). The local accumulation of TFs around targets, through the formation of hubs or condensates, has been proposed as a mechanism to increase binding frequency and maintain target occupancy despite short residence times (*1–7*). These high-local concentration clusters are further thought to catalyze the recruitment of co-factors and transcriptional machinery to amplify gene activation (*8–22*). Here, we sought to address the central question of this model: whether hubs drive TF target occupancy by increasing their binding frequency or whether they are an emergent feature of underlying binding kinetics.

Clusters of transcription factors are referred to as condensates when they exhibit non-stochiometric accumulations and their biophysical properties are consistent with formation via phase separation mechanisms (*23–25*). Under the phase separation model, transcriptional condensate formation depends on the nuclear concentration and sequence properties of condensate constituents. Although substantial work has focused on how intrinsic protein features and concentration thresholds govern condensate formation (*7*), far less attention has been directed toward how genomic context may tune hub properties at individual loci in an in vivo context.

Hub localization to genomic sites is dynamically influenced by a combination of targeting through a TF's protein-protein and protein-DNA interactions (*11, 16, 21, 26–30*). A TF's locus-specific network of interaction partners also plays a role in generating compositionally and functionally distinct hubs (*27, 30, 31*). The nuclear concentration of a TF and the number of TF binding sites at a locus have a direct impact on TF occupancy and transcriptional output (*11, 32*). Recent work has suggested a spatio-temporal relationship where the proximity of a hub to its regulatory elements can increase transcriptional bursting (*16, 33–35*). If hubs differentially influence TF occupancy at genes, their properties should be highly gene-dependent, rather than globally dictated by transcription factor identity and nuclear concentration. Based on these observations we hypothesized that the biophysical properties attributed to hubs, such as their size

and stability, are not emergent properties of higher-order phenomena such as phase-separation. We reasoned that hub properties instead reflect the molecular scale binding and diffusion kinetics of hub constituents in an enhancer-specific manner.

To investigate whether hubs drive or reflect binding kinetics, we leverage the Drosophila dorsoventral patterning system. In this system, the transcription factor Dorsal forms a nuclear concentration gradient and activates distinct sets of genes in a spatially-dependent manner along the embryo's dorsoventral axis (*36, 37*). Due to its morphogenic nature, Dorsal-regulated gene expression provides a natural framework to determine whether hubs function uniformly or differentially at target genes that are expressed at different Dorsal concentration levels. Previous work in the Drosophila embryo has shown that TF hubs containing Bicoid, Zelda, or Dorsal are recruited to specific genomic loci even at low concentrations, and suggests that morphogen hubs play an active role in fine-tuning transcriptional responses to spatial gradients (*8, 9, 21, 32, 38, 39*). Dorsal clusters have recently been proposed to form in a concentration-threshold-dependent manner via liquid-liquid phase separation mechanisms (*40*). However, how these clusters associate with target sites and if their properties are tuned in a site-specific manner has not previously been explored.

Here, we use high-resolution lattice light-sheet imaging to analyze how Dorsal hubs form at and interact with transcriptionally active loci. To enable these insights, we developed a live imaging-based framework that quantifies the spatiotemporal relationship between TF hubs and active transcription sites *in vivo*. Using this approach, we measured Dorsal hub properties including local Dorsal enrichment and the persistence of hubs at enhancers of *sog* and *snail*, and at synthetic enhancers with varying numbers of Dorsal binding sites. We further used our live-imaging data, along with single-molecule tracking experiments, to parameterize a molecular kinetics model that recapitulates Dorsal hub formation at the enhancers we investigated. Together, our results suggest that TF clustering, or hub formation, primarily reflects enhancer-dependent binding kinetics rather than a distinct regulatory layer that increases binding frequency.

Results

Dorsal hub persistence and density scale with nuclear concentration

To investigate the biophysical properties of Dorsal hubs, we performed lattice light-sheet imaging on blastoderm stage Drosophila embryos expressing endogenously-tagged Dorsal-mNeonGreen in the 13th and 14th syncytial nuclear cycles (nc13 and nc14) (*41*). Dorsal forms a nuclear concentration gradient and is found at high concentrations in ventral nuclei, moderate to low concentrations in lateral nuclei, and is largely excluded from dorsal nuclei (Fig. 1A, Movie S1). Dorsal's nuclear concentration gradient is not persistent and is re-established after each nuclear division. At the onset of mitosis, the cytoplasmic-nuclear concentration ratio flattens, and after nuclear membrane reformation the gradient is re-established through a balance of nuclear import and export during interphase (*37*). We observed that Dorsal hubs are visible as soon as Dorsal accumulates in the nucleus after mitosis, and remain visible throughout interphase at nuclei in both ventral and lateral regions (Fig. 1B-C). Further, hubs form even as the nucleoplasmic Dorsal concentration is diminished in the dorsal-lateral region (Fig. S1A,B, Movie S2), which demonstrates that hub formation does not require high nuclear protein levels.

To understand how nuclear concentrations influence hub formation, we compared ventral nuclei along the presumptive ventral furrow with nuclei at the lateral surface (Fig. 1B). Here, we define lateral nuclei as the region of the embryo where Dorsal nuclear concentration diminishes to be equivalent to the cytoplasmic concentration (Fig. S1A,B). Nuclear Dorsal intensity in both ventral and lateral nuclei rises for the first 4.5 minutes of interphase and then falls for the remainder of interphase (Fig. 1D). To compare Dorsal hubs in these nuclei, we developed a custom analysis pipeline to detect hubs and quantify their intensities, spatial distributions, and persistence (Fig. S1C, Movie S3). While the total number of detected hubs per nucleus increases during interphase (Fig. 1E-F), this increase is correlated with the expansion of nuclear volume in this time period (Fig. S1D-E). Thus, despite the dynamic change in nuclear concentration during interphase, the density of Dorsal hubs (number of hubs per unit volume) remains relatively constant during the interphases of nc13 and nc14 (Figs. 1G-H). We suspect that the apparent increase in hub number is due to tightly spaced accumulations at binding sites becoming resolvable as the nuclear volume expands rather than de novo formation of additional hubs.

Although the density of hubs is uncoupled with nuclear concentration, the average hub intensity is strongly correlated with nuclear Dorsal concentration (Figs. 1I-J, S1H). To further assess hub variability using a segmentation-independent metric, we calculated the coefficient of variation (CV) of nuclear Dorsal intensity (Fig. S1F-G). The CV exhibits an intermediate correlation with nuclear concentration likely since the CV reflects intensity dispersion and not just morphological differences. Based on these data we hypothesized that hubs form rapidly at dorsal target sites at the start of interphase, and then increasing amounts of Dorsal interact with these sites as nuclear concentrations increase.

Dorsal hubs preferentially enrich and persist at target sites

To determine whether Dorsal hubs preferentially form at transcriptionally active target sites, we performed two-color imaging of Dorsal-mNeonGreen with MCP-mCherry to visualize the site of transcription from MS2 reporters. We characterized hub formation at reporters of *snail* (*sna*), a ventrally expressed Dorsal target, *short gastrulation* (*sog*), a laterally expressed target, and at *hunchback* (*hb*) as a non-Dorsal target control in both ventral (*hb(v)*) and lateral (*hb(l)*) nuclei (Figs. 2A-B, S2, Movies S4, S5). These reporters differ substantially in the number and arrangement of Dorsal binding sites with *sna* containing the highest site

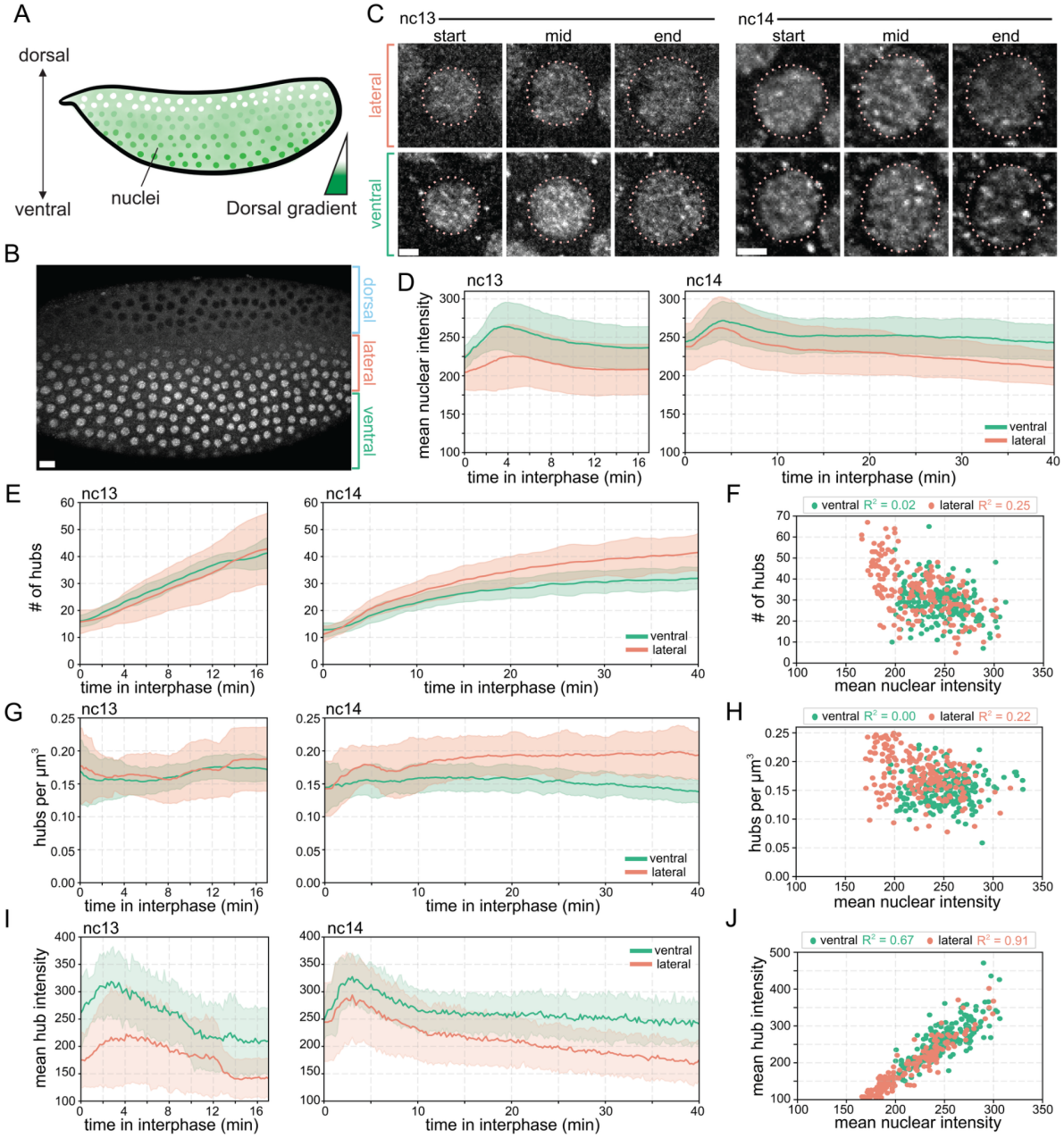
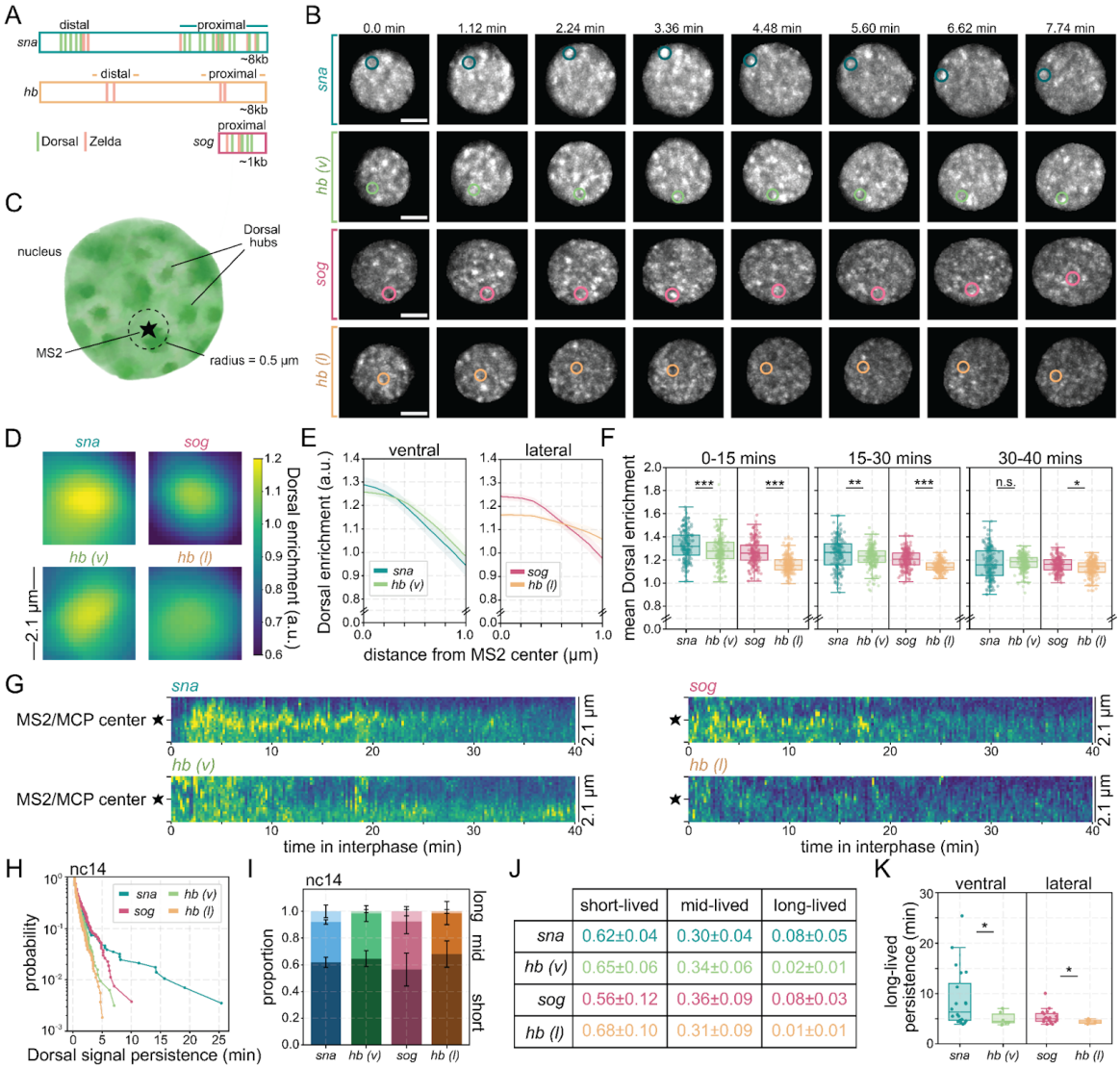


Figure 1: Dorsal hub densities are constant in interphase, but hub enrichment depends on Dorsal nuclear concentration. (A) Schematic of the Dorsal gradient in *Drosophila* embryos. Dorsal translocates into ventral surface nuclei at higher rates creating a nuclear concentration gradient along the dorsoventral axis. **(B)** Maximum intensity projection image showing nuclear gradient of Dorsal-mNeonGreen in a *Drosophila* embryo in nc13. Scale bar is 15 microns. **(C)** Representative single z-slice images of Dorsal-mNeonGreen in nuclei at the lateral (top) and ventral (bottom) embryonic surfaces at the start, middle, and end of nc13 (left) and nc14 (right). Scale bars are 2 microns. **(D)** Average nuclear mean intensity in ventral (green) or lateral (orange) nuclei. Shading shows standard deviation between embryo replicates. N = 5 nuclei per embryo, 3 embryo replicates for each axis position. **(E)** Average number of hubs per nucleus in ventral (green) or lateral (orange) nuclei. Shading shows standard deviation between embryo replicates. **(F)** Scatter plot of number of hubs vs. mean nuclear intensities. R^2 value is Pearson's correlation (N = 5814 ventral hubs, 6928 lateral hubs from N = 3 embryos per axis position). **(G)** Average hub density (number of hubs per cubic micron) in ventral (green) or lateral (orange) nuclei. Shading shows standard deviation between embryo replicates. **(H)** Scatter plot of hub densities vs. mean nuclear intensities. R^2 value is Pearson's correlation. **(I)** Mean intensities of Dorsal hubs in ventral (green) or lateral (orange) nuclei. Solid line shows average intensity and shading shows standard deviation between embryo replicates. **(J)** Scatter plot showing hub mean intensity and mean nuclear intensity. R^2 is Pearson's correlation.

density (~16 sites) compared to *sog* (~4 sites) and hb (no Dorsal binding sites) (Figs. 2A, S2). Importantly, all constructs were integrated into the same genomic docking site on chromosome III, allowing us to examine enhancer grammar as the primary variable. Visually, we found that Dorsal hubs are most stably associated with *sna* and slightly less so at *sog* compared to non-specific interactions at *hb(v)* and *hb(l)* (Movies S4, S5).

We built an analysis pipeline to localize and track sites of nascent transcription from the MS2/MCP signal over time and the kinetics of Dorsal interactions at these sites (Fig. 2C, Movie S6). First, to quantify site-specific enrichment independently of hub segmentation, we measured the average Dorsal intensity centered on the MS2/MCP spot over time (Fig. 2D–E). Due to the higher nuclear background in ventral nuclei, the hub enrichment is higher at *hb(v)* than *sog*. However when comparing *hb(v)* with *sna*, and *hb(l)* with

sog, our analysis shows that Dorsal is enriched at actively transcribing target genes relative to non-target controls. The radial profiles demonstrate that Dorsal enrichment at target loci decays substantially within ~0.5 μm of the MS2/MCP site with sog enrichment approaching *hb(l)* levels at this distance (Fig. 2E). We therefore used a 0.5 μm sphere average to capture specific locus-associated signals while reducing noise and accounting for slight spatial offsets of hubs relative to the MS2/MCP spot (Fig. 2F). The mean enrichment in this sphere closely matches the trends of enrichment measured directly at the MS2/MCP center. Consistent with the nucleus wide trends (Fig. 1), Dorsal enrichment at MS2/MCP correlates strongly linearly with nuclear Dorsal concentration (Figs. S3A), suggesting that hubs persistently form at target genes and the concentration of Dorsal within them scales with available nuclear Dorsal. Towards the end of nc14 as the Dorsal nuclear concentration



	short-lived	mid-lived	long-lived
sna	0.62±0.04	0.30±0.04	0.08±0.05
hb (v)	0.65±0.06	0.34±0.06	0.02±0.01
sog	0.56±0.12	0.36±0.09	0.08±0.03
hb (l)	0.68±0.10	0.31±0.09	0.01±0.01

Figure 2: Dorsal accumulations are persistent at target genes and transient at non-targets. (A) Schematic of *sna*, *sog*, and *hb* enhancers with Dorsal and Zelda (a cofactor common to all three enhancers) binding sites shown. Each enhancer was positioned directly upstream of its promoter, followed by 24 MS2 sequences, and the *yellow* coding sequence (see Table S1 for genotypes of embryo imaged). Dorsal target genes, *snail* and *sog,* are expressed along the ventral and lateral surfaces respectively while *hunchback* is expressed anteriorly. *hunchback* is not a Dorsal target gene and is used as a negative control for both the anterior-ventral surface (*hb(v)*) to compare to *snail* or on the anterior-lateral surface (*hb(l)*) to compare to *sog*. **(B)** Representative maximum intensity projections of single nuclei with Dorsal-mNeonGreen over ~8 mins of nc14. The MS2/MCP reporter location is circled in each image. The cytoplasmic Dorsal signal is masked, and contrast was adjusted independently for each row to aid visualization. Scale bars are 2 microns. **(C)** Schematic of analysis, a single nucleus with Dorsal hubs is depicted,the MS2/MCP spot is denoted as a star. A 0.5 micron radius sphere is considered around the MS2/MCP spot for quantification of Dorsal signal at the reporter. **(D)** Images show average Dorsal enrichment centered at the MS2/MCP site. Images are 2.1 μm x 2.1 μm. The color bar shows the level of Dorsal enrichment in the image (Dorsal intensity over nuclear mean intensity). **(E)** Average radial profile with standard error centered at MS2/MCP spots shows Dorsal enrichment as a function of distance from MS2/MCP spot center. **(F)** Boxplots showing the distribution of mean Dorsal enrichment within 0.5 micron of the MS2/MCP spot during the beginning, middle, and end of nc14 at each reporter. A Mann-Whitney U-test was performed to determine significance and following p-values were used: *p<0.05, **p<0.01, and ***p<0.001 **(G)** Representative kymographs (one nucleus per reporter is shown) shows the Dorsal enrichment centered at the MS2/MCP spot across time in nc14. **(H)** Survival plots showing the length of time any hub is consistently present within the interaction sphere for each reporter for nc14. **(I)** A Gaussian Mixture Model was used to determine short-, mid-, or long-lived persistent hub interaction frequencies. The proportion in each bin is shown with error bars designating standard deviation among embryo replicates. **(J)** Table of mean persistence and standard deviations for each population. **(K)** Box plots show the average duration of persistent hub presence within the long-lived category. N = 5 nuclei per embryo and 3 embryos replicates for each reporter. A Mann-Whitney U-test was performed to determine significance and following p-values were used: *p<0.05, **p<0.01, and ***p<0.001.

drops, so does hub intensity, and the distinction between target and non-target loci is diminished.

Examination of kymographs centered on MS2/MCP spots (Fig. 2G) show that target loci exhibit extended periods of sustained enrichment compared to non-target loci. To characterize the temporal stability of hub presence at a locus, we defined a persistent enrichment event as a continuous time interval when the mean Dorsal intensity within the interaction sphere exceeds the nuclear mean intensity. Survival probability analysis of periods of enrichment persistence revealed a long-lived population at target genes that is significantly higher than at non-targets (Fig. 2H). To further characterize persistence, we fit the distribution enrichment-segment durations to a three-component Gaussian mixture model (Fig. S3B) and classified segments as short-, mid-, and long-lived. Both *sna* and *sog* exhibit rare but pronounced long-lived enrichment segments (~8% of events in nc14) (Fig. 2I–J, Fig. S3C), consistent with prior literature reporting infrequent but stable hub associations to genomic sites (*16, 22*). However, the long-lived population is significantly more persistent at *sna* than at *sog* suggesting site-specific or nuclear concentration dependent effects (Fig. 2K).We repeated these analyses using endogenous MS2 knock-in *sna* and *sog* fly lines (Fig. S3D–K, Movie S7). While absolute enrichment values differed slightly from transgenic reporters, potentially reflecting chromatin context, specific enrichment and elevated persistence were preserved. Notably, endogenous *sna* exhibited a markedly higher proportion of long-lived enrichment segments (~40%) (Fig. S3K).

To assess the effects of Dorsal nuclear concentration, we imaged the *sog* reporter near the dorsal–lateral transition point where Dorsal concentration is relatively low (as in Fig. S1A–B) and compared them to more ventrally positioned nuclei that remain within the *sog* domain but have higher nuclear concentrations. We refer to these populations as $sog^{lateral}$ and sog^{medial} respectively (Fig. S4A–B, Movie S8). We observed clear differences in Dorsal enrichment relative to the nuclear mean between these two populations (Fig. S4C–E). In medial nuclei where nucleoplasmic Dorsal concentration is higher, the relative enrichment at the locus appears reduced. This reduction reflects an increased background intensity, since as nuclear background increases, the contrast between locus-associated enrichment and the surrounding nucleoplasm diminishes. Persistence analysis revealed a reduced long-lived enrichment population in sog^{medial} nuclei compared to $sog^{lateral}$ nuclei (Fig. S4F–I). This reduction in persistence does not necessarily indicate reduced binding-site occupancy. In fact, previous work has shown that for another Dorsal target gene, *zen*, Dorsal maintains similar levels of occupancy in lateral and ventral regions despites changes in nuclear concentration (*42*). Therefore, the change in persistence at *sog* can be attributed to elevated background, where enrichment

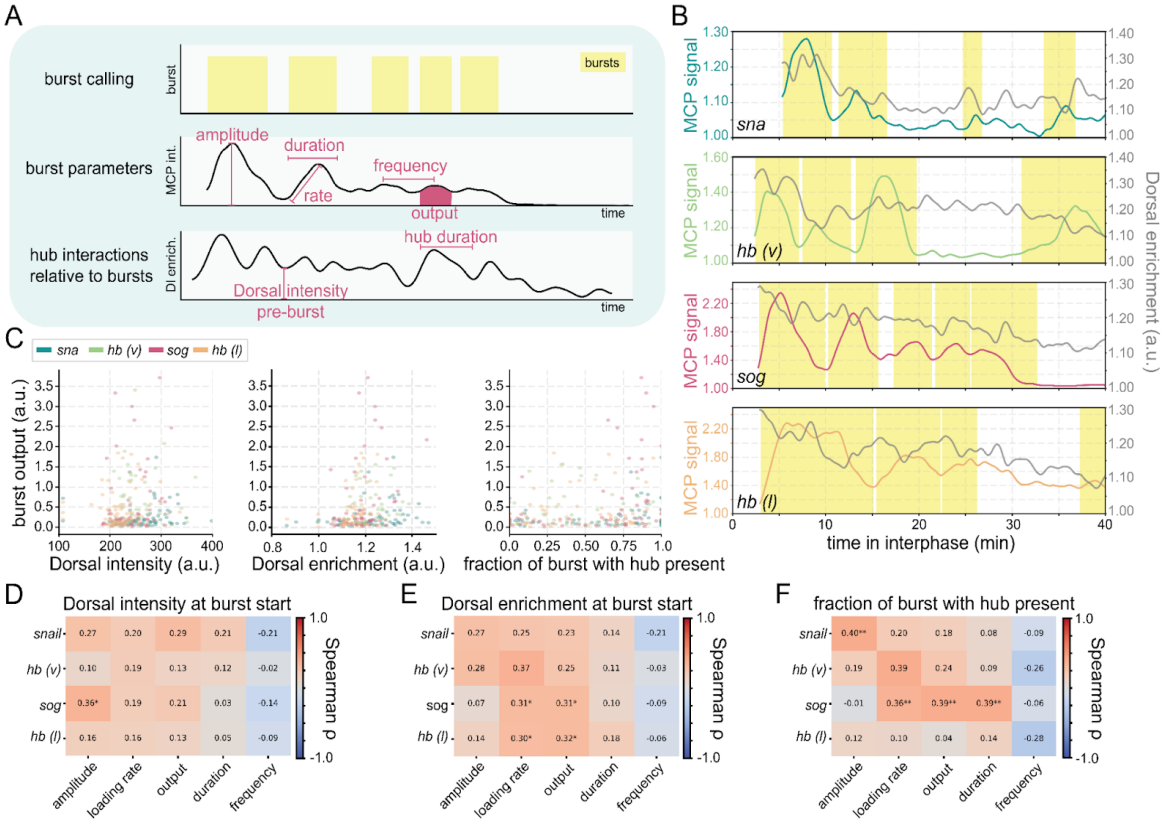


Figure 3: Dorsal hubs show minimal correlation with transcriptional burst kinetics. (A) Schematic showing the burst calling, burst parameters extracted from the MS2/MCP intensity, and the hub interaction parameters quantified from the Dorsal hub overlap and intensity. Bursts are identified in yellow. **(B)** Representative traces of MS2/MCP intensity (colored) and Dorsal intensity (grey) at each reporter gene in nc14. **(C)** Scatter plots showing burst output with the mean Dorsal intensity before the burst (left), the mean Dorsal enrichment before the burst (center), or the fraction of time during a burst when a hub is persisting near the MS2/MCP site (right). Marker colors correspond to each reporter analyzed. **(D-F)** Heat maps show the Spearman's rank correlation coefficient (ρ) between the Dorsal intensity at burst start **(D)**, Dorsal enrichment at burst start **(E)**, and fraction of burst with hub present **(F)** for each gene and each burst parameter. The asterisks indicate p-values: *p<0.05, **p<0.01, and ***p<0.001.

segments are more likely to fall below the nuclear mean threshold, shortening the apparent persistence of above-background enrichment events. Our persistence measurements thus represent a conservative estimate of enhancer-specific differences in hub stability.

Together, these data indicate that Dorsal enrichment and persistence is site-specific. While Dorsal hubs transiently appear to associate with non-target loci, these interactions are likely incidental, and stable hub enrichment is restricted to target genes.

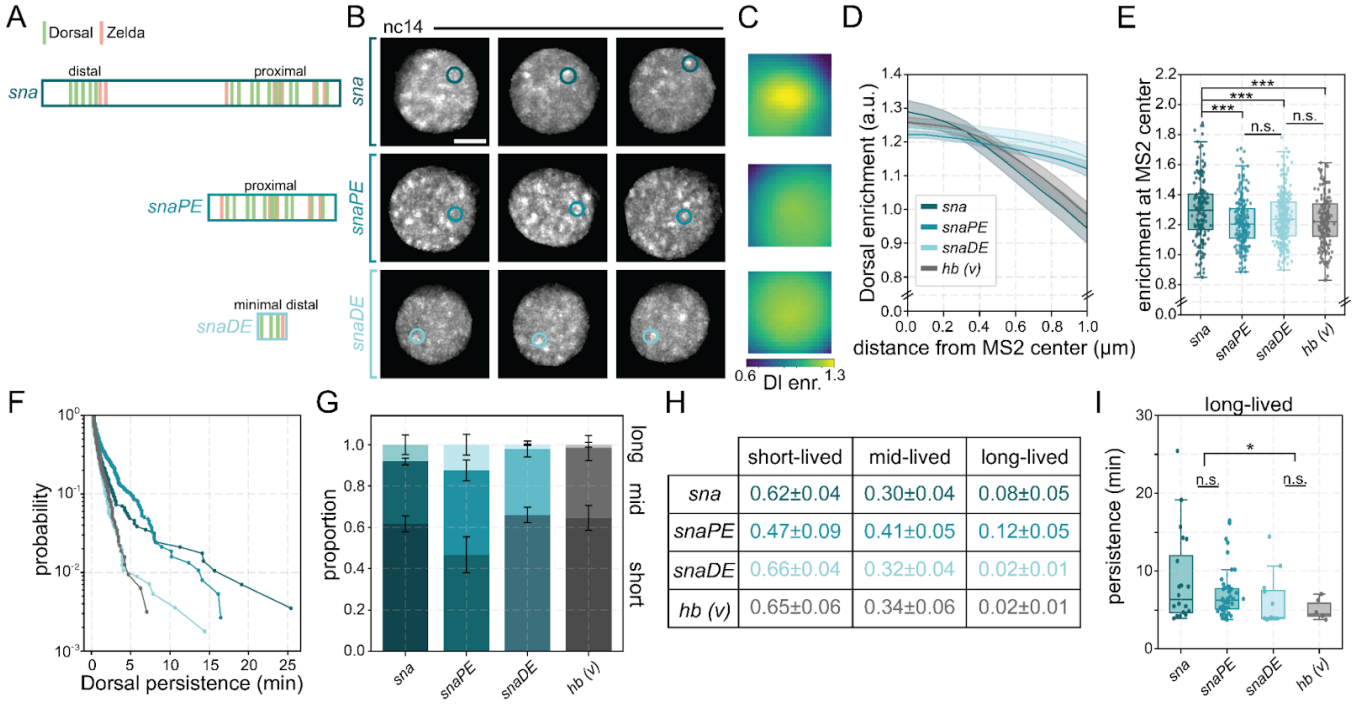
Hub enrichment and persistence do not strongly predict burst parameters

To assess whether Dorsal hub interactions are quantitatively linked to transcriptional kinetics, we analyzed the relationship between hub properties and transcriptional burst kinetics (Fig. 3A). Transcription occurs in bursts corresponding to transitions between inactive and active promoter states (*43*). We segmented MS2/MCP fluorescence traces normalized to MCP background levels into individual bursts by identifying changes in the rate of increase of MS2/MCP signal (Fig. 3B). For each reporter, we extracted burst parameters including amplitude (amount of transcripts), loading rate (slope of rise), total output (integrated intensity), duration of bursts, and frequency (interval between bursts) (*44*). We then quantified three hub-related features at each locus: (1) Dorsal intensity and (2) Dorsal enrichment at the MS2/MCP site at the burst onset, and (3) the fraction of the burst duration with persistent Dorsal enrichment (Figs. 3C, S5).

Across target and non-target reporters correlations between hub metrics and burst parameters were minimal (Fig. 3D–F). Because the MS2/MCP signal is detectable only after transcription initiation, these analyses are restricted to periods of active transcription and do not capture hub behavior preceding burst onset. These results indicate that while Dorsal hubs are enriched at target loci, variation in hub intensity or persistence at a given enhancer during active transcription does not robustly predict burst parameters.

Reducing Dorsal motifs within the *snail* enhancer decreases hub enrichment and persistence

To test whether enhancer binding-site number directly modulates Dorsal hub properties, we leveraged the dual-enhancer structure of the *snail* locus. We focused on this locus to remove the concentration effects between enhancers active in different concentration regions (like *sog* and *sna*). The full snail reporter (*sna*) contains both the proximal and distal enhancers, so we compared this



	short-lived	mid-lived	long-lived
sna	0.62±0.04	0.30±0.04	0.08±0.05
snaPE	0.47±0.09	0.41±0.05	0.12±0.05
snaDE	0.66±0.04	0.32±0.04	0.02±0.01
hb (v)	0.65±0.06	0.34±0.06	0.02±0.01

Figure 4: Reducing Dorsal motifs decreases hub enrichment and persistence. (A) Schematic depicting enhancers with Dorsal and Zelda binding sites for each *snail* reporter with either both enhancers (*sna*), the proximal enhancer (*snaPE*), or the distal enhancer (*snaDE*). Each enhancer was positioned just upstream of its promoter, 24 MS2 sequences, and the *yellow* coding sequence (see Table S1 for genotypes of embryo imaged). Each of these *snail* reporters were imaged along the ventral surface in nc14. **(B)** Representative maximum intensity projections of single nuclei with Dorsal-mNeonGreen in nc14. The MS2/MCP spot is circled in each image. The cytoplasmic Dorsal signal is masked, and contrast was adjusted independently for each row to aid visualization. Scale bars are 2 microns. **(C)** Images show average Dorsal enrichment centered at the MS2 /MCPsite. Images are 2.1 μm x 2.1 μm. The color bar on the bottom shows the level of Dorsal enrichment in the image (Dorsal intensity over nuclear mean intensity). **(D)** Average radial profile with standard error centered at the MS2/MCP spot shows Dorsal enrichment as a function of distance from MS2/MCP spot center. **(E)** Boxplots showing the distribution of the Dorsal enrichment at the center of MS2/MCP site (radius = 0.0 μm in (D)). A Mann-Whitney U-test was performed to determine significance and following p-values were used: *p<0.05, **p<0.01, and ***p<0.001. **(F)** Survival plots showing the length of time any hub is consistently present within the interaction sphere for each reporter for nc14. **(G)** The proportion of long-, mid, -and short-lived persistent hub formation in each bin is shown with error bars designating standard deviation among embryo replicates. **(H)** The proportion of short-, mid-, and long-lived hub persistence events are listed with their mean and standard deviation in the table. **(I)** Box plots showing the average duration of consistent hub interactions within the short-, mid-, and long-lived populations. N = *sna:* 18 nuclei in 3 embryos; *snaPE:* 22 nuclei in 3 embryos; *snaDE:* 25 nuclei in 4 embryos; and *hb(v):* 18 nuclei in 4 embryos. A Mann-Whitney U-test was performed to determine significance and following p-values were used: *p<0.05, **p<0.01, and ***p<0.001.

construct to reporters containing only the proximal enhancer (*snaPE*) or only the minimal distal enhancer (*snaDE*), progressively reducing the total number of Dorsal binding sites (Fig. 4A–B, Movie S9). We included *hunchback* imaged on the ventral surface (*hb(v)*) as a negative control. Quantification of Dorsal enrichment at the MS2/MCP site revealed a reduction in enrichment from *sna* to either *snaPE* or *snaDE* (Fig. 4C–E). Notably, enrichment of both *snaPE* and *snaDE* approached levels indistinguishable from *hb(v)* suggesting that a minimum density of binding sites may be required to generate enrichment detectable above nuclear background. Although absolute levels remained near background, both *sna* and *snaPE* displayed significantly higher proportions of long-lived enrichment segments and longer persistence durations compared to *snaDE* and *hb(v)* (Fig. 4F–I). In contrast, persistence at *snaDE* was statistically indistinguishable from *hb(v)*. These results indicate that increasing Dorsal binding-site number enhances both

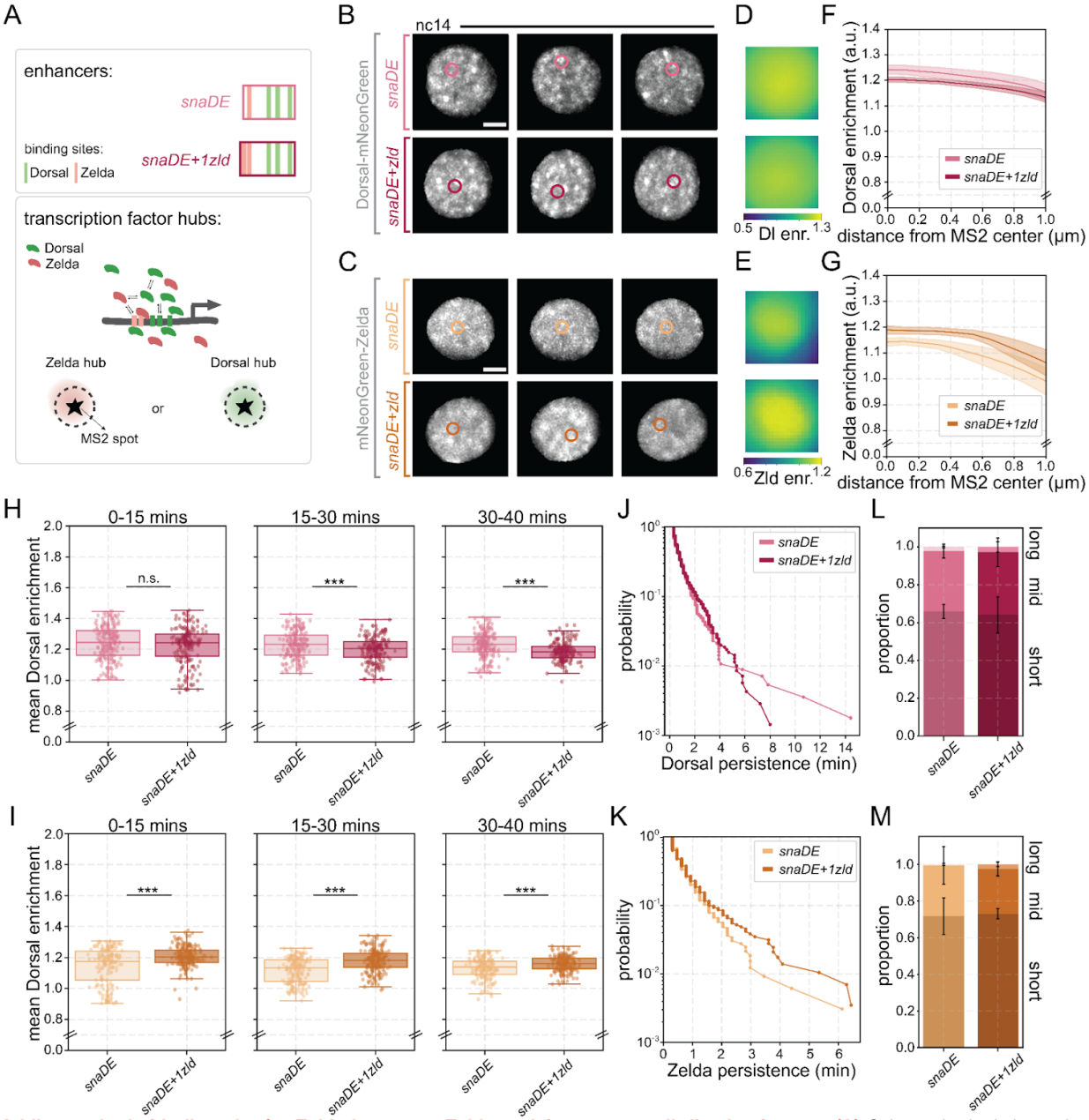


Figure 5: Adding a single binding site for Zelda increases Zelda enrichment at snail distal enhancer. (A) Schematic depicting enhancers with Dorsal and Zelda binding sites for *snail* distal enhancer reporter without (*snaDE*) or with (*snaDE+1zld*). Each enhancer was positioned just upstream of its promoter, 24 MS2 stem loops, and the *yellow* coding sequence (see Table S1 for exact genotypes of embryo imaged). Each *snail* reporter was imaged at the ventral surface in nc14. These reporters were imaged with either Dorsal-mNeonGreen or Zelda-mNeonGreen to visualize hubs formed by each TF. **(B-C)** Representative maximum intensity projections of single nuclei with Dorsal-mNeonGreen **(B)** or mNeonGreen-Zelda **(C)** in nc14. The MS2/MCP spot is circled in each image. The cytoplasmic Dorsal signal is masked, and contrast was adjusted independently for each row to aid visualization. Scale bars are 2 microns. **(D-E)** Images show average Dorsal enrichment centered at the MS2/MCP site. Images are 2.1 μm x 2.1 μm. The color bar on the bottom shows the level of Dorsal **(D)** or Zelda **(E)** enrichment in the image (local intensity over nuclear mean intensity). **(F-G)** Average radial profile with standard error centered at the MS2/MCP spot shows Dorsal **(F)** or Zelda **(G)** enrichment as a function of distance from MS2/MCP spot center. **(H, I)** Boxplots showing the distribution of mean Dorsal **(H)** or Zelda **(I)** intensities during the beginning, middle, and end of nc14 at each gene. A Mann-Whitney U-test was performed to determine significance and following p-values were used: *p<0.05, **p<0.01, and ***p<0.001. **(J, K)** Survival plots showing the length of time a Dorsal **(J)** or Zelda **(K)** hub is consistently present within 0.5 μm of MS2/MCP center for nc14. **(L, M)** Using a Gaussian Mixture Model, each Dorsal **(L)** or Zelda **(M)** hub persistence event was categorized into short-, mid-, or long-lived bins. The proportion in each bin is shown with error bars designating standard deviation among embryo replicates. N = *snaDE* (Dorsal): 25 nuclei in 4 embryos; *snaDE+1zld* (Dorsal): 36 nuclei in 4 embryos; *snaDE* (Zelda: 24 nuclei in 3 embryos; and *snaDE+1zld* (Zelda): 19 nuclei in 3 embryos.

enrichment magnitude and persistence, and differences between three binding sites and the non-specific control are indistinguishable in our assay.

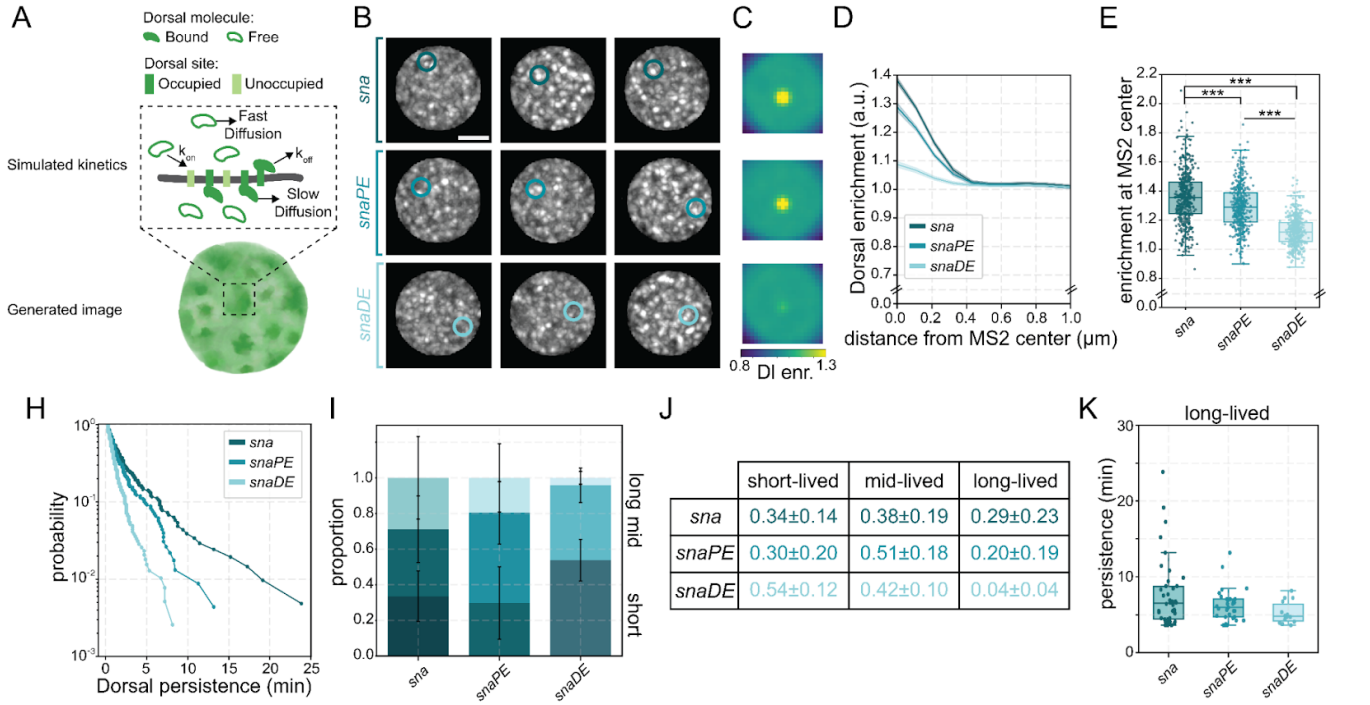
To further probe the sensitivity limits of our detection, we imaged additional reporters in which one or two Dorsal binding sites were removed from the distal enhancer (*snaDE-1xdl* and *snaDE-2xdl*; Fig. S6A–B, Movie S10). Across these constructs, neither enrichment nor persistence differed significantly from the unmodified *snaDE* reporter or from *hb(v)* (Fig. S6C–H). These data indicate that further reductions in binding-site number within the distal enhancer fall below the threshold required for detectable changes in hub properties in our assay.

We next tested whether altering non-Dorsal binding sites within the enhancer influences hub behavior. Zelda is a pioneer transcription factor known to regulate early embryonic genes and to form its own nuclear hubs (*9, 39*). We compared the distal enhancer reporter to a modified reporter containing an additional Zelda binding site (*snaDE+1zld*) (Fig. 5A–B, Movie S11). Unexpectedly, addition of a single Zelda binding site resulted in a small but significant decrease in Dorsal enrichment at the MS2/MCP site (Fig. 5D, F, H) without altering Dorsal persistence (Fig. 5J, L; Fig. S7A–D). We then examined Zelda localization at these reporters (Fig. 5C, Movie S12). Zelda enrichment increases significantly at *snaDE+1zld* relative to *snaDE* (Fig. 5E, G, I) although Zelda persistence remains unchanged (Fig. 5K, M; Fig. S7C–D).

Together, these results demonstrate that enhancer binding site composition tunes transcription factor enrichment and persistence at actively transcribing reporter sites. Dorsal hub enrichment and persistence decrease when lowering binding-site number but these measurements are constrained by detection limits imposed by nuclear background. Moreover, addition of a single Zelda binding site is sufficient to measurably increase Zelda enrichment while modestly altering Dorsal enrichment, suggesting that enhancer composition can differentially shape factor-specific hub properties

Enhancer-dependent Dorsal hub properties are dictated by molecular binding kinetics

We hypothesized that if Dorsal hub properties simply reflect binding kinetics at enhancer-encoded binding sites, then a kinetic binding model should be sufficient to recapitulate our experimental results on hub enrichment and persistence. To test this, we incorporated single-molecule tracking–derived kinetic parameters into a quantitative model of Dorsal binding. To parameterize the simulations, we performed SMT experiments in live Drosophila embryos as previously shown (*8, 9, 20, 21, 41, 45*). We used a fly line containing endogenously-tagged Dorsal-mEos4a to enable sparse photoconversion for single molecule detection at



	short-lived	mid-lived	long-lived
sna	0.34±0.14	0.38±0.19	0.29±0.23
snaPE	0.30±0.20	0.51±0.18	0.20±0.19
snaDE	0.54±0.12	0.42±0.10	0.04±0.04

Figure 6: Molecular binding kinetics simulations recapitulate hub enrichment and persistence. (A) Schematic showing the generation of synthetic nucleus images by modeling Dorsal single-molecule kinetics and binding (k_{on} = binding probability; k_{off} = disassociation rate). **(B)** Representative maximum intensity projections of simulated single nuclei with Dorsal signal in grey. The binding site cluster is circled in each image. Contrast is scaled equally for all conditions. Scale bars are 2 microns. **(C)** Images show simulated average Dorsal enrichment centered at the binding site clusters. Images are 2.1 μm x 2.1 μm. The color bar on the bottom shows the level of Dorsal enrichment in the image (Dorsal intensity over nuclear mean intensity). **(D)** Average radial profile with standard error centered at the binding site cluster shows Dorsal enrichment as a function of distance from spot center. **(E)** Boxplots showing the distribution of the Dorsal enrichment at the center of the binding site cluster (radius = 0.0 μm). A Mann-Whitney U-test was performed to determine significance and following p-values were used: *p<0.05, **p<0.01, and ***p<0.001. **(H)** Survival plots showing the length of time any hub is consistently present within the interaction sphere for each reporter. **(I)** The proportion of long-, mid, -and short-lived persistent hub formation in each bin is shown with error bars designating standard deviation among embryo replicates. **(J)** The proportion of short-, mid-, and long-lived hub persistence events are listed with their mean and standard deviation in the table. **(K)** Box plots showing the average duration of consistent hub interactions within the long-lived populations. N = 15 nuclei per simulated enhancer. A Mann-Whitney U-test was performed to determine significance and following p-values were used: *p<0.05, **p<0.01, and ***p<0.001.

both ventral and lateral surfaces of the embryo (*41*). Using long exposure times (500 msec) to blur out the unbound population of molecules and reduce the effects of photobleaching, we measured the residence time of Dorsal to be approximately four seconds in both regions (Fig. S8A–B). Using short exposure times (10 msec) to track both bound and unbound molecules, we estimated diffusion coefficients, and the immobile and free fractions (*46*) and found them to be similar (~58%) in both ventral and lateral nuclei (Fig. S8C–D).

Using our SMT data (Fig. S8) we simulated the diffusion and binding kinetics of Dorsal molecules in nc14 nuclei. We inferred the number of Dorsal binding sites (*47*) and Dorsal concentrations (*48*) from previously published data (for a full list of parameters, see Table S2). We simulated diffraction-limited microscope images by convolving the simulated molecule positions with the microscope's point spread function, mimicking the signal recorded by our camera sensor. These images were generated in 3D to obtain z-stacks over time, taking into account variables such as slice spacing and exposure time, to match our experimental data. We inserted binding site arrays to match the *snail* (Fig. 6) and *sog* enhancers (Fig. S9) we examined experimentally. Using the same analysis pipelines we applied to our experimental data, we characterized hub enrichment and persistence in our simulated data.

We found that our model is able to recapitulate the differences we see in hub properties at different *snail* enhancers (full, proximal, or distal) (Fig. 6). Hub enrichment scales with the number of Dorsal binding sites as in experimental data (Fig. 6C–E), however, the enrichment magnitude is increased relative to our real data. This increase is likely due to two factors: lower background noise in the simulated images; and spread of enrichment signal in our experimental data due to our reliance on using the MS2/MCP spot (i.e. the distance of the MCP-mCherry signal plus the stem loops from the enhancer) for estimating the position of the binding sites as opposed to using the precise position of binding sites in our simulations. The persistence of Dorsal signals at different reporters also follows the same trend as our experimental data, with persistence scaling with the number of binding sites (Fig. 6H–K). Accounting for the known cooperativity of Dorsal binding (*49–52*) was necessary in our simulations and non-cooperative binding failed to scale Dorsal persistence with the number of binding sites (Fig. S10). Additionally, we recapitulated the hub properties at sog^{medial} and $sog^{lateral}$ and found that the increased concentration in the 'medial' embryonic region did lead to lower persistence consistent with our data (Fig. S9). Therefore, our simulations demonstrate that molecular binding kinetics, and binding site number, are sufficient to explain enhancer-specific hub properties.

Discussion

Our findings demonstrate that transcription factor hubs are not homogeneous entities, but instead reflect binding kinetics at specific enhancer sequences. We show that across synthetic and endogenous reporters, Dorsal hub persistence and enrichment, are enhancer specific. These properties scale with the number of binding-sites and nuclear concentration but only weakly correlate with transcriptional burst kinetics (Fig. 7). These observations suggest that hubs reflect the kinetics of TF binding at enhancers rather than constituting an independent regulatory layer that determines instantaneous transcriptional output.

We found that the nuclear density of Dorsal hubs does not depend on nuclear concentration, while hub enrichment does. The decoupling of density from concentration is inconsistent with models where hubs arise from concentration-threshold dependent assembly of condensates. Although Dorsal clustering has been suggested to be concentration dependent and arising from thermodynamic phase separation (*40*), our results indicate that hub density is not dictated by bulk nuclear levels. Instead, nuclear concentration appears to tune factor enrichment at existing genomic binding sites rather than creating new assemblies. This is further supported by our computational modeling which recapitulates hub properties such as enrichment and persistence through binding kinetics alone.

By using reporters integrated in a common genomic docking site, we show that increasing Dorsal binding site density increases enrichment and persistence. Conversely, reducing binding sites progressively diminishes detectable enrichment and long-lived interactions, eventually reaching levels indistinguishable from non-specific control sites. These effects extend beyond Dorsal itself. The addition of a single Zelda binding site measurably increases Zelda

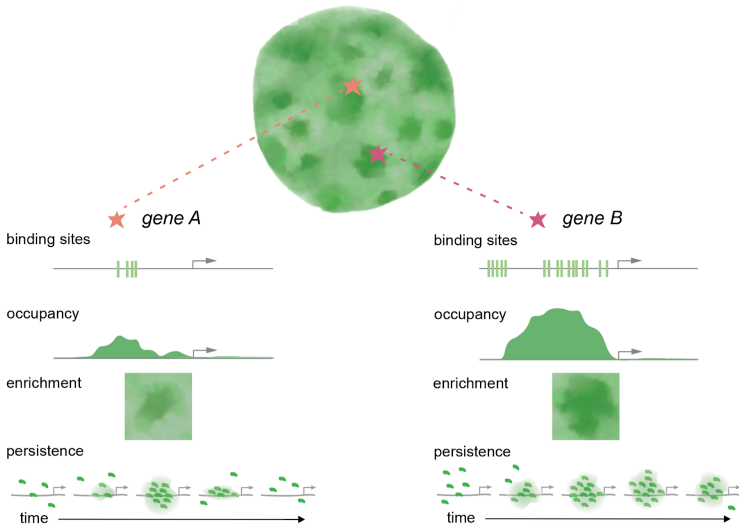


Figure 7: Transcription factor hub properties reflect enhancer binding site number. (A, B) Binding sites at *cis*-regulatory regions mediate changes in Dorsal gene occupancy which is reflected in hub properties apparent at the locus. As a result, different target sites with lower **(A)** or higher **(B)** number of binding sites have differing enrichment and persistence of Dorsal hubs.

enrichment while modestly reducing Dorsal enrichment, highlighting that enhancer composition differentially tunes factor-specific hub properties. Cis-regulatory element control of transcription factor hubs is not completely unprecedented. Prior work demonstrated that removing Zelda binding sites at *sog* reduced both Dorsal clustering and transcription, particularly in lateral regions where Dorsal concentration is lower (*32*). Other work also demonstrated that high-affinity Bicoid binding sites did not result in increased Bicoid clustering intensity, suggesting that clustering is not simply dictated by affinity (*38*). However, this prior study did not account for differences in number of binding sites or hub interaction times, which may be a crucial distinguishing factor, akin to the differences observed between *snail* and *sog* in our data. Overall, our data suggests that enhancer grammar may play a key role in determining hub properties and function. A key implication of our results is that persistence and enrichment can be interpreted as real-time occupancy readouts. Genomic approaches such as ChIP-seq have shown that motif number and affinity correlate with transcription factor occupancy. Here, our time-averaged enrichment measurements correlate with binding-site number across genes in vivo, suggesting that hub enrichment reflects stoichiometric binding at enhancer sites.

Importantly, our data reveal that hub properties are only weakly predictive of transcription burst kinetics. In line with our data, previous studies on Gal4 clusters found no difference in burst size when the clusters were positioned closer to their target site and in vivo single molecule tracking of Gal4 at its target site suggests that burst duration and size are primarily determined by transcription factor dwell time and binding affinity (*35*). Although stable enrichment is more prevalent at target genes, variation in enrichment or persistence during active transcription does not strongly scale with burst kinetics. Thus while binding-site composition tunes occupancy, transcriptional bursting is likely influenced by steps downstream to TF recruitment.

As has been shown previously, the presence of a TF hub is not sufficient for transcription initiation (*33*). Indeed, we previously examined the effect of enhancer placement relative to the promoter (*53*). Downstream placement of the enhancer significantly dampened transcriptional output, yet we detected no change in local Dorsal enrichment between upstream and downstream configurations. This dissociation between enrichment and transcription further supports the conclusion that hub formation and enrichment are dictated by binding-site composition rather than act as transcriptional modulators at a given enhancer. Although hub properties at burst onset do not strongly predict burst parameters at a single locus, this does not rule out hub contributions prior to activation or between bursts as previously reported (*22, 33, 54*). Additionally, recent work emphasizes that cooperative TF binding and exchange can stabilize high-occupancy states that underlie transcriptional bursting (*34, 55*). We similarly find in our modeling that cooperativity is required to reproduce the measured hub properties. Still, the differences in burst kinetics and transcriptional output between genes are likely dependent on the level of TF occupancy which is reflected in the larger differences in hub properties between reporters.

Our findings also highlight a critical practical consideration: hub detection is inherently constrained by nuclear background, and the contrast of the imaging modality being used. As nuclear Dorsal concentration increases, the background intensity is elevated, reducing contrast between locus-associated enrichment and the surrounding nucleoplasm without necessarily altering underlying binding kinetics at a locus. This limitation in detection could easily lead to misleading conclusions that condensates or hubs are discrete regulatory structures that highly concentrate molecules. Indeed, our recent work on RNA polymerase II (Pol II) clustering similarly demonstrates that Pol II clusters at active genes can be explained by transcription-dependent engagement and recruitment kinetics and are thus only visible at genes that sufficiently accumulate Pol II (*20*).

Together our results suggest the clustering of Dorsal, and other TFs, observed *in vivo* may be interpretable through the lens of binding kinetics and occupancy rather than phase separation. Second, they underscore the importance of cooperative binding and enhancer features in shaping apparent hub behavior. Third, they highlight how detection thresholds and nuclear background can influence conclusions about condensate formation. Finally, by demonstrating that enhancer perturbations produce quantifiable changes in enrichment and persistence, our study establishes that hub properties are encoded within DNA sequence. We suggest that the many reports of transcription factors and regulatory proteins forming condensates or hubs through higher-order mechanisms should be examined through the lens of molecular binding kinetics. Instead, rapid search processes may facilitate a high on-rate and frequent exchange of TF molecules at their targets, leading to the appearance of persistent hubs at gene loci despite short residence times. Future work combining enhancer dissection, kinetic modeling, and hub molecular composition analysis will further clarify how generalizable this binding and search process centric framework is across transcription factors and developmental systems.

References

1. A. P. Boka, A. Mukherjee, M. Mir, Single-molecule tracking technologies for quantifying the dynamics of gene regulation in cells, tissue and embryos. *Development* **148** (2021).
2. F. Lu, T. Lionnet, Transcription factor dynamics. *Cold Spring Harb. Perspect. Biol.* **13**, a040949 (2021).
3. G. Hayward-Lara, M. D. Fischer, M. Mir, Dynamic microenvironments shape nuclear organization and gene expression. *Curr. Opin. Genet. Dev.* **86**, 102177 (2024).
4. K. Wagh, D. A. Stavreva, A. Upadhyaya, G. L. Hager, Transcription factor dynamics: One molecule at a time. *Annu. Rev. Cell Dev. Biol.* **39**, 277–305 (2023).
5. A. Tsai, A. K. Muthusamy, M. R. P. Alves, L. D. Lavis, R. H. Singer, D. L. Stern, J. Crocker, Nuclear microenvironments modulate transcription from low-affinity enhancers. *Elife* **6** (2017).

6. A. S. Hansen, A. Amitai, C. Cattoglio, R. Tjian, X. Darzacq, Guided nuclear exploration increases CTCF target search efficiency. *Nat. Chem. Biol.* **16**, 257–266 (2020).
7. A. Navalkar, M. Eppert, B. R. Sabari, T. Mittag, Density transitions in the regulation of transcription. *Mol. Cell*, doi: 10.1016/j.molcel.2026.01.009 (2026).
8. M. Mir, A. Reimer, J. E. Haines, X.-Y. Li, M. Stadler, H. Garcia, M. B. Eisen, X. Darzacq, Dense Bicoid hubs accentuate binding along the morphogen gradient. *Genes Dev.* **31**, 1784–1794 (2017).
9. M. Mir, M. R. Stadler, S. A. Ortiz, C. E. Hannon, M. M. Harrison, X. Darzacq, M. B. Eisen, Dynamic multifactor hubs interact transiently with sites of active transcription in Drosophila embryos. *Elife* **7** (2018).
10. B. R. Sabari, A. Dall'Agnese, A. Boija, I. A. Klein, E. L. Coffey, K. Shrinivas, B. J. Abraham, N. M. Hannett, A. V. Zamudio, J. C. Manteiga, C. H. Li, Y. E. Guo, D. S. Day, J. Schuijers, E. Vasile, S. Malik, D. Hnisz, T. I. Lee, I. I. Cisse, R. G. Roeder, P. A. Sharp, A. K. Chakraborty, R. A. Young, Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361**, eaar3958 (2018).
11. K. Kawasaki, T. Fukaya, Functional coordination between transcription factor clustering and gene activity. *Mol. Cell* **83**, 1605–1622.e9 (2023).
12. W.-K. Cho, J.-H. Spille, M. Hecht, C. Lee, C. Li, V. Grube, I. I. Cisse, Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science* **361**, 412–415 (2018).
13. A. Boija, I. A. Klein, B. R. Sabari, A. Dall'Agnese, E. L. Coffey, A. V. Zamudio, C. H. Li, K. Shrinivas, J. C. Manteiga, N. M. Hannett, B. J. Abraham, L. K. Afeyan, Y. E. Guo, J. K. Rimel, C. B. Fant, J. Schuijers, T. I. Lee, D. J. Taatjes, R. A. Young, Transcription factors activate genes through the phase-separation capacity of their activation domains. *Cell* **175**, 1842–1855.e16 (2018).
14. I. I. Cisse, I. Izeddin, S. Z. Causse, L. Boudarene, A. Senecal, L. Muresan, C. Dugast-Darzacq, B. Hajj, M. Dahan, X. Darzacq, Real-time dynamics of RNA polymerase II clustering in live human cells. *Science* **341**, 664–667 (2013).
15. M.-T. Wei, Y.-C. Chang, S. F. Shimobayashi, Y. Shin, A. R. Strom, C. P. Brangwynne, Nucleated transcriptional condensates amplify gene expression. *Nat. Cell Biol.* **22**, 1187–1196 (2020).
16. M. Du, S. H. Stitzinger, J.-H. Spille, W.-K. Cho, C. Lee, M. Hijaz, A. Quintana, I. I. Cissé, Direct observation of a condensate effect on super-enhancer controlled gene bursting. *Cell* **187**, 2595–2598 (2024).
17. A. Tsai, M. R. Alves, J. Crocker, Multi-enhancer transcriptional hubs confer phenotypic robustness. *Elife* **8** (2019).
18. S. Chong, C. Dugast-Darzacq, Z. Liu, P. Dong, G. M. Dailey, C. Cattoglio, A. Heckert, S. Banala, L. Lavis, X. Darzacq, R. Tjian, Imaging dynamic and selective low-complexity domain interactions that control gene transcription. *Science* **361** (2018).
19. M. Mazzocca, A. Loffreda, E. Colombo, T. Fillot, D. Gnani, P. Falletta, E. Monteleone, S. Capozi, E. Bertrand, G. Legube, Z. Lavagnino, C. Tacchetti, D. Mazza, Chromatin organization drives the search mechanism of nuclear factors. *Nat. Commun.* **14**, 6433 (2023).
20. A. Mukherjee, M. Kapoor, K. Shankta, S. Fallacaro, R. D. Carter, G. Hayward-Lara, P. Ratchasanmuang, Y. I. Haloush, M. Mir, A single cluster of RNA Polymerase II molecules is stably associated with active genes, *bioRxivorg* (2025). https://doi.org/10.1101/2025.02.10.637507.
21. S. Fallacaro, A. Mukherjee, P. Ratchasanmuang, J. Zinski, Y. I. Haloush, K. Shankta, M. Mir, A fine kinetic balance of interactions directs transcription factor hubs to genes, *bioRxivorg* (2024). https://doi.org/10.1101/2024.04.16.589811.
22. C. Santa-Cruz Mateos, C. Roussos, F. J. de Haro Arbona, J. Falo-Sanjuan, S. Bray, Direct observation of Notch signalling induced transcription hubs mediating gene-expression responses, *bioRxiv* (2025). https://doi.org/10.1101/2025.07.04.663121.
23. S. F. Banani, H. O. Lee, A. A. Hyman, M. K. Rosen, Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **18**, 285–298 (2017).
24. A. Musacchio, On the role of phase separation in the biogenesis of membraneless compartments. *EMBO J.* **41**, e109952 (2022).
25. G. Pei, H. Lyons, P. Li, B. R. Sabari, Transcription regulation by biomolecular condensates. *Nat. Rev. Mol. Cell Biol.* **26**, 213–236 (2025).
26. M. M. Gaskill, I. V. Soluri, A. E. Branks, A. P. Boka, M. R. Stadler, K. Vietor, H.-Y. S. Huang, T. J. Gibson, A. Mukherjee, M. Mir, S. A. Blythe, M. M. Harrison, Localization of the Drosophila pioneer factor GAF to subnuclear foci is driven by DNA binding and required to silence satellite repeat expression. *Dev. Cell* **58**, 1610–1624.e8 (2023).
27. A. Patil, A. R. Strom, J. A. Paulo, C. K. Collings, K. M. Ruff, M. K. Shinn, A. Sankar, K. S. Cervantes, T. Wauer, J. D. St Laurent, G. Xu, L. A. Becker, S. P. Gygi, R. V. Pappu, C. P. Brangwynne, C. Kadoch, A disordered region controls cBAF activity via condensation and partner recruitment. *Cell* **186**, 4936–4955.e26 (2023).
28. N. M. Chabot, R. Purkanti, A. Del Panta Ridolfi, D. D. Nogare, H. Oda, H. Kimura, F. Jug, A. Dal Co, N. L. Vastenhouw, Local DNA compaction creates TF-DNA clusters that enable transcription, *bioRxiv* (2024). https://doi.org/10.1101/2024.07.25.605169.
29. M. Ugolini, M. A. Kerlin, K. Kuznetsova, H. Oda, H. Kimura, N. L. Vastenhouw, Transcription bodies regulate gene expression by sequestering CDK9. *Nat. Cell Biol.* **26**, 604–612 (2024).
30. S. Dongre, N. L. Vastenhouw, Specificity in clustering of gene-specific transcription factors is encoded in the genome. *Nucleic Acids Res.* **53** (2025).
31. T. Mittag, R. V. Pappu, A conceptual framework for understanding phase separation and addressing open questions and challenges. *Mol. Cell* **82**, 2201–2214 (2022).
32. S. Yamada, P. H. Whitney, S.-K. Huang, E. C. Eck, H. G. Garcia, C. A. Rushlow, The Drosophila pioneer factor Zelda modulates the nuclear microenvironment of a dorsal target enhancer to potentiate transcriptional output. *Curr. Biol.* **29**, 1387–1393.e5 (2019).
33. F. J. DeHaro-Arbona, C. Roussos, S. Baloul, J. Townson, M. J. Gómez Lamarca, S. Bray, Dynamic modes of Notch transcription hubs conferring memory and stochastic activation revealed by live imaging the co-activator Mastermind. *Elife* **12** (2024).
34. W. Pomp, J. V. W. Meeussen, T. L. Lenstra, Transcription factor exchange enables prolonged transcriptional bursts. *Mol. Cell* **84**, 1036–1048.e9 (2024).
35. B. T. Donovan, A. Huynh, D. A. Ball, H. P. Patel, M. G. Poirier, D. R. Larson, M. L. Ferguson, T. L. Lenstra, Live-cell imaging reveals the interplay between transcription factors, nucleosomes, and bursting. *EMBO J.* **38**, e100809 (2019).
36. C. A. Rushlow, K. Han, J. L. Manley, M. Levine, The graded distribution of the dorsal morphogen is initiated by selective nuclear transport in Drosophila. *Cell* **59**, 1165–1177 (1989).
37. A. E. Schloop, P. U. Bandodkar, G. T. Reeves, Formation, interpretation, and regulation of the Drosophila Dorsal/NF-κB gradient. *Curr. Top. Dev. Biol.* **137**, 143–191 (2020).
38. R. Munshi, J. Ling, S. Ryabichko, E. F. Wieschaus, T. Gregor, Transcription factor clusters as information transfer agents. *Sci. Adv.* **11**, eadp3251 (2025).
39. J. Dufourt, A. Trullo, J. Hunter, C. Fernandez, J. Lazaro, M. Dejean, L. Morales, S. Nait-Amer, K. N. Schulz, M. M. Harrison, C. Favard, O. Radulescu, M. Lagha, Temporal control of gene expression by the pioneer factor Zelda through transient interactions in hubs. *Nat. Commun.* **9**, 5194 (2018).
40. S. S. Dima, M. A. Turner, H. G. Garcia, G. T. Reeves, Global maps of transcription factor properties reveal threshold-based formation of DNA-bound and mobile clusters. *Sci. Adv.* **12** (2026).
41. M. A. Turner, N. M. Gravina, B. Moretti, S. Dima, G. Martini, G. Reeves, H. G. Garcia, Novel fluorescent and photoconvertible fusions reveal dorsal activator dynamics, *bioRxivorg* (2025). https://doi.org/10.1101/2025.05.12.653543.
42. P.-H. Holmqvist, A. Boija, P. Philip, F. Crona, P. Stenberg, M. Mannervik, Preferential genome targeting of the CBP co-activator by Rel and Smad proteins in early Drosophila melanogaster embryos. *PLoS Genet.* **8**, e1002769 (2012).
43. E. Tunnacliffe, J. R. Chubb, What is a transcriptional burst? *Trends Genet.* **36**, 288–297 (2020).
44. J. V. W. Meeussen, T. L. Lenstra, Time will tell: comparing timescales to gain insight into transcriptional bursting. *Trends Genet.* **40**, 160–174 (2024).
45. M. Mir, A. Reimer, M. Stadler, A. Tangara, A. S. Hansen, D. Hockemeyer, M. B. Eisen, H. Garcia, X. Darzacq, Single molecule imaging in live embryos using lattice light-sheet microscopy. *Methods Mol. Biol.* **1814**, 541–559 (2018).

46. A. S. Hansen, M. Woringer, J. B. Grimm, L. D. Lavis, R. Tjian, X. Darzacq, Robust model-based analysis of single-particle tracking experiments with Spot-On. *Elife* **7** (2018).
47. Y. Sun, C.-Y. Nien, K. Chen, H.-Y. Liu, J. Johnston, J. Zeitlinger, C. Rushlow, Zelda overcomes the high intrinsic nucleosome barrier at enhancers during Drosophila zygotic genome activation. *Genome Res.* **25**, 1703–1714 (2015).
48. S. Alamos, A. Reimer, C. Westrum, M. A. Turner, P. Talledo, J. Zhao, E. Luu, H. G. Garcia, Minimal synthetic enhancers reveal control of the probability of transcriptional engagement and its timing by a morphogen gradient. *Cell Syst.* **14**, 220–236.e3 (2023).
49. L. Dunipace, J. M. McGehee, J. Irizarry, A. Stathopoulos, The proximal enhancer of the snail gene mediates negative autoregulatory feedback in Drosophila melanogaster. *Genetics* **230** (2025).
50. J. P. Bothma, H. G. Garcia, S. Ng, M. W. Perry, T. Gregor, M. Levine, Enhancer additivity and non-additivity are determined by enhancer strength in the Drosophila embryo. *Elife* **4** (2015).
51. J. Irizarry, A. Stathopoulos, Dynamic patterning by morphogens illuminated by cis-regulatory studies. *Development* **148**, dev196113 (2021).
52. R. Sayal, J. M. Dresch, I. Pushel, B. R. Taylor, D. N. Arnosti, Quantitative perturbation-based analysis of gene expression predicts enhancer activity in early Drosophila embryo. *Elife* **5** (2016).
53. E. A. Leyes Porello, J. Wu, S. Fallacaro, E. Dreispiel Juan, R. T. Trudeau, J. A. Vidal, C. Nieto, A. Singh, M. Mir, B. Lim, Enhancer placement impacts transcriptional dynamics in Drosophila embryos, *bioRxivorg* (2025). https://doi.org/10.1101/2025.08.14.670187.
54. C.-Y. Cho, P. H. O'Farrell, Stepwise modifications of transcriptional hubs link pioneer factor activity to a burst of transcription. *Nat. Commun.* **14**, 4848 (2023).
55. C. Fountas, T. L. Lenstra, Better together: how cooperativity influences transcriptional bursting. *Curr. Opin. Genet. Dev.* **89**, 102274 (2024).
56. T.-M. Fu, G. Liu, D. E. Milkie, X. Ruan, F. Görlitz, Y. Shi, V. Ferro, N. S. Divekar, W. Wang, H. M. York, V. Kilic, M. Mueller, Y. Liang, T. A. Daugird, M. J. Gacha-Garay, K. A. Larkin, R. C. Adikes, N. Harrison, C. Shirazinejad, S. Williams, J. L. Nourse, S.-H. Sheu, L. Gao, T. Li, C. Mondal, K. Achour, W. Hercule, D. Stabley, K. Emmerich, P. Dong, D. Drubin, Z. J. Liu, D. Clapham, J. S. Mumm, M. Koyama, A. Killilea, J. J. Bravo-Cordero, C. D. Keene, L. Luo, T. Kirchhausen, M. M. Pathak, S. Arumugam, J. K. Nunez, R. Gao, D. Q. Matus, B. L. Martin, I. A. Swinburne, E. Betzig, W. R. Legant, S. Upadhyayula, A multimodal adaptive optical microscope for in vivo imaging from molecules to organisms, *bioRxivorg* (2025). https://doi.org/10.1101/2025.06.02.657494.
57. T.-L. Liu, S. Upadhyayula, D. E. Milkie, V. Singh, K. Wang, I. A. Swinburne, K. R. Mosaliganti, Z. M. Collins, T. W. Hiscock, J. Shea, A. Q. Kohrman, T. N. Medwig, D. Dambournet, R. Forster, B. Cunniff, Y. Ruan, H. Yashiro, S. Scholpp, E. M. Meyerowitz, D. Hockemeyer, D. G. Drubin, B. L. Martin, D. Q. Matus, M. Koyama, S. G. Megason, T. Kirchhausen, E. Betzig, Observing the cell in its native state: Imaging subcellular dynamics in multicellular organisms. *Science* **360** (2018).
58. *Cudadecon: GPU Accelerated 3D Image Deconvolution Using CUDA* (Github; https://github.com/scopetools/cudadecon).
59. X. Ruan, M. Mueller, G. Liu, F. Görlitz, T.-M. Fu, D. E. Milkie, J. L. Lillvis, A. Kuhn, J. Gan Chong, J. L. Hong, C. Y. A. Herr, W. Hercule, M. Nienhaus, A. N. Killilea, E. Betzig, S. Upadhyayula, Image processing tools for petabyte-scale light sheet microscopy data. *Nat. Methods* **21**, 2342–2352 (2024).
60. M. Pachitariu, C. Stringer, Cellpose 2.0: how to train your own model. *Nat. Methods* **19**, 1634–1641 (2022).
61. A. Archit, L. Freckmann, S. Nair, N. Khalid, P. Hilt, V. Rajashekar, M. Freitag, C. Teuber, M. Spitzner, C. Tapia Contreras, G. Buckley, S. von Haaren, S. Gupta, M. Grade, M. Wirth, G. Schneider, A. Dengel, S. Ahmed, C. Pape, Segment Anything for microscopy. *Nat. Methods* **22**, 579–591 (2025).
62. S. Fallacaro, A. Mukherjee, E. Gamauf, M. Mir, Quantifying transcription factor enrichment at target loci in live embryos. *Methods Mol. Biol.* **2923**, 119–141 (2025).
63. M. Pachitariu, M. Rariden, C. Stringer, Cellpose-SAM: superhuman generalization for cellular segmentation, *bioRxiv* (2025). https://doi.org/10.1101/2025.04.28.651001.
64. A. Heckert, *Quot: A Simple Single Molecule Tracking Pipeline with a Graphic User Interface for Quality Control* (Github; https://github.com/alecheckert/quot).
65. A. S. Hansen, I. Pustova, C. Cattoglio, R. Tjian, X. Darzacq, CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *Elife* **6** (2017).
66. K. J. Brennan, M. Weilert, S. Krueger, A. Pampari, H.-Y. Liu, A. W. H. Yang, J. A. Morrison, T. R. Hughes, C. A. Rushlow, A. Kundaje, J. Zeitlinger, Chromatin accessibility in the Drosophila embryo is determined by transcription factor pioneering and enhancer activation. *Dev. Cell* **58**, 1898–1916.e9 (2023).
67. E. Esposito, B. Lim, G. Guessous, H. Falahati, M. Levine, Mitosis-associated repression in development. *Genes Dev.* **30**, 1503–1508 (2016).
68. H. Deng, B. Lim, Shared transcriptional machinery at homologous alleles leads to reduced transcription in early Drosophila embryos. *Front. Cell Dev. Biol.* **10**, 912838 (2022).
69. S. Syed, Y. Duan, B. Lim, Modulation of protein-DNA binding reveals mechanisms of spatiotemporal gene control in early Drosophila embryos. *Elife* **12** (2023).

Acknowledgements

We thank Bomyi Lim and Mounia Lagha for fly lines and helpful discussions about this work. We also thank all members of the Mir Lab for feedback throughout this project. This paper was typeset with the bioRxiv word template by @Chrelli: www.github.com/chrelli/bioRxiv-word-template

Funding

National Institutes of Health: DP2HD108775 (MM), R01GM139913 (HGG) and R01GM152815 (HGG)
Margaret Q Landenberger Foundation (MM)
Howard Hughes Medical Institute, Freeman Hrabowski Scholars Program (MM)
National Science Foundation Graduate Research Fellowship grant DGE-2236662 (SF) and DGE-1752814 (MAT)
Koret-UC Berkeley-Tel Aviv University Initiative in Computational Biology and Bioinformatics (HGG)
Winkler Scholar Faculty Award (HGG)
Chan Zuckerberg Initiative Grant CZIF2024-010479 (HGG)

Author contributions

Conceptualization: MM, SF
Methodology: SF, MM, AM
Software: SF, MK
Validation: SF, MM, MK
Formal Analysis: SF, MM, MK, LE
Investigation: SF, MM, LE
Resources: MAT, HGG, AM
Data Curation: SF, MM, LE
Writing – Original Draft: SF, MM, MK, LE
Writing – Review & Editing: SF, MM, AM, MAT, HGG, MK, LE
Visualization: SF, MK, LE
Supervision: MM
Project Administration: MM
Funding Acquisition: MM, SF, HGG, MAT

Competing interest statement

Authors declare they have no competing interests.

Materials and Methods

Embryo preparation for live imaging

For each MS2 experiment, Drosophila embryos were generated by crossing virgin female flies heterozygous for MCP-mCherry (III) and Dorsal-mNeonGreen to males containing the MS2 reporter of interest. The resulting embryos were collected for live imaging. Only embryos positive for all three (Dorsal-mNeonGreen, MCP-mCherry, and MS2-gene-of-interest) were imaged. Details of all lines used for imaging and where they appear in this work are listed in Table S1. For single molecule tracking experiments, virgin female Dorsal-mNeonGreen; MCP-mCherry flies were crossed to male Dorsal-mEos4a flies. Flies positive for all three markers were used for imaging. Dorsal-mNeonGreen intensity was further used for assessing the position of the embryos along the dorsoventral axis; MCP-mCherry was used for

assessing nuclear boundaries; and Dorsal-mEos4a was sparsely activated for single molecule detection. Embryos were collected over a 90-minute window and manually dechorionated by gently rolling them on double-sided tape using a blunt needle until the chorion was removed. Dechorionated embryos were then arranged on an agar pad before being transferred onto an adhesive heptane-acrylic spot, prepared by dissolving double-sided tape in heptane on a 25 mm diameter glass coverslip. Embryos were manually oriented on the coverslip either ventral side up (facing the objective) or gently rotated to position the dorsal–lateral transition region within the imaging plane (Fig. S1A,B).

Microscope calibration and imaging

All imaging was performed on a Multimodal Optical Scope with Adaptive Imaging Correction (MOSAIC) (*56*). For volumetric experiments, the lattice light-sheet mode was used with the 488 nm and 589 nm laser lines. These laser beams were first expanded to a combined 2 mm diameter, passed through a half-wave plate to adjust polarization, and relayed into an acousto-optic tunable filter (AOTF) (Quanta-Tech, AA Opto Electronic) to selectively control wavelength and power modulation. The collimated beams were then expanded in one dimension using a Powell lens (Laserline Optics, Canada), and their width was adjusted and collimated using a pair of cylindrical lenses. The beam was subsequently relayed onto a grayscale spatial light modulator (Meadowlark Optics, AVR Optics) after passing through a second half-wave plate, where it was diffracted and projected onto a custom annular mask to define the minimum and maximum numerical aperture (NA) of the light-sheet while blocking undiffracted light. The light from the annulus was demagnified and relayed onto a resonant galvanometer scanner (Cambridge Technology, Novanta Photonics), conjugate to the sample plane, to mitigate shadowing artifacts and improve uniformity by introducing a slight wobble in the excitation angle. Finally, the light was projected onto a pair of galvanometer scanning mirrors (Cambridge Technology, Novanta Photonics) conjugate to the pupil plane, allowing scanning of the light-sheet along the x and z optical axes before being focused onto the sample using a Thorlabs TL20X-MPL excitation objective. Fluorescence emission was collected orthogonally by a Zeiss 20×, 1.0 NA detection objective and relayed onto a deformable mirror (ALPAO) positioned in the detection path to correct optical aberrations, as previously described (57). The emitted signal was then split using a Semrock Di03-R561-t3-25x36 dichroic beam splitter and directed onto two sCMOS detectors (Hamamatsu ORCA Fusion). The green emission channel (for 488 nm excitation) was filtered using a Semrock FF03-525/50-25 emission filter and a Chroma ZET488NF notch filter to reject laser light, while the red emission channel (for 589 nm excitation) was filtered using a Semrock FF01-615/24 emission filter and a Chroma ZET594NF notch filter. All volumetric imaging imaging used a multi-bessel lattice sheet with maximum numerical aperture to minimum numerical aperture ratio of 0.4/0.3 488 nm (used to excite mNeonGreen) and 589 nm (used to excite mCherry). Two color channels were acquired sequentially for a volume of 18.9 μm sampling z-planes spaced 0.3 μm apart with an exposure time of 30 msec. A laser power of 0.124 mW was used for the 488 nm laser and 1.344 mW for the 589 nm laser. Each volume was acquired every four seconds during acquisition for a total time of ~9.56 seconds in between stacks.

For single-molecule imaging, a Gaussian light sheet was used on the MOSAIC. Briefly, after passing through the AOTF, the beam was expanded using a Powell lens then relayed onto a custom mask for filtering. The filtered sheet was then projected onto a resonant galvanometer (Cambridge Technology, Novanta Photonics). Lastly, an excitation objective (Thorlabs, TL20X-MPL) was used to focus the light-sheet onto the sample. The detection path for the emitted light is the same as described above for the lattice light-sheet configuration. For this light path, the SLM was bypassed to maximize laser power at the excitation objective. For single molecule imaging, the red emission channel (for 561 excitation) was filtered using a Semrock FF01-585/29 emission filter and a Chroma ZET561NF Notch filter. The 405 nm laser line was kept on constantly during the acquisition period for photoswitching. Data was acquired at 10 msec and 500 msec exposure times for the fast and slow single-molecule imaging experiments respectively. The excitation laser power was optimized empirically for each exposure time to achieve sufficient contrast for single-molecule tracking. The powers of the photoswitching laser were also optimized empirically to achieve low enough detection densities to enable tracking. For fast single-molecule imaging experiments, 8000 frames were acquired with 24.1 mW of the 560 laser, 123 μW of the 488 laser, and 7.3 μW of the photoswitching laser for a total acquisition time of 80 sec. For slow single-molecule imaging experiments, 400 frames were acquired with 0.66 mW of the 560 laser. 123 μW of the 488 laser. and 0.46 μW of the photoswitching laser for a total acquisition time of 200 sec. Powers were measured at the back focal plane of the excitation objective

Volumetric image analysis

All volumetric imaging datasets were pre-processed before downstream analysis using GPU-accelerated 3D deconvolution via CUDA (*58*). Each dataset was deconvolved using a Richardson-Lucy based algorithm with 8 iterations, utilizing point spread functions (PSFs) captured from bead images acquired on the day of imaging to ensure accurate correction. For visualization and rendering, we employed a custom Imaris converter, leveraging fast TIFF and Zarr file readers to efficiently generate Imaris files (*59*). Following deconvolution, images underwent nuclear segmentation using a custom-trained model in Cellpose 2.0 (*60*). Ground truth training data was generated through a combination of micro-SAM (*61*) and manual curation on MCP-mCherry images. The trained Cellpose model was then applied to segment individual slices across each dataset. To improve segmentation accuracy, a post-processing pipeline was developed to stitch segmented slices together and interpolate missing nuclear structures. We then implemented a nearest-neighbor tracking algorithm to follow nuclei across interphase in each nuclear cycle.

To quantify hub properties, we used a custom image analysis pipeline (*20, 21*) . First, nuclear intensity was normalized to its mean value to assess local transcription factor enrichments above background levels. Hubs were segmented by first applying a median filter to reduce noise, followed by image erosion and reconstruction to enhance feature contrast. The reconstructed image was then subtracted from the median-filtered image to generate a binary mask identifying high-density regions. To resolve closely associated hubs, local maxima peaks were identified and used as markers for watershed-based segmentation, ensuring separation of fused or amorphous structures. Finally, region properties were extracted to quantify hub features, including integrated intensity, mean intensity, and size.

After segmenting nuclei and hubs, we manually select nuclei for further analysis under the criteria of checking if anomalies are present in the segmentation and to check if the MS2 spot is closer than 0.2 μm to the edge of the nucleus which may skew analysis. We then generate 4D images of cropped single nuclei. These selected nuclei are then subjected to another segmentation regime this time to detect the MS2/MCP spot. This regime uses a median filter, difference-of-gaussians filter, and then a percentile threshold to segment high intensity spots. The exact sigma values and percentile used were adjusted nucleus-to-nucleus to ensure the best possible MS2/MCP segmentation. Then we identify the spot inside the nucleus with the highest intensities to be the likely MS2/MCP spot. We then used a nearest neighbour algorithm to track the MS2/MCP spot in time while also manually correcting any tracking errors in each nucleus. Due to the bursty nature of transcription, we interpolate the MS2/MCP position in between bursts of visible MS2/MCP spots and extrapolate the MS2/MCP position 10 frames before the first appearance of the MS2/MCP signal while correcting for the motion and size of the nucleus.

We calculated enrichment (Dorsal intensity at a locus over the nuclear background) as previously described in (*62*). However, instead of solely using spot detection, we used the calculated MS2/MCP centroid positions tracked over time as the center position for our cropped images. We then quantified enrichment segments for our persistence measurements at the MS2/MCP spot by assessing the overlap between the hub labels and a 0.5 μm radius sphere centered at the MS2/MCP weighted centroid coordinates as we have done previously (*21*). We could then monitor the duration, intensity, and volume overlap at each locus regardless of transcriptional activity. Hub interaction durations were fit using a three-component Gaussian mixture model in order to establish short-, mid-, and long-lived populations. Three components were used based on the elbow method.

We also called bursts to assess transcriptional activity by applying a Gaussian smoothing filter to the MS2/MCP intensity and calculated the derivative to capture periods where the transcriptional activity peaks. We then quantified parameters including amplitude, duration, output, loading rate, and frequency based on these bursts. Based on the start times of these bursts we could then assess if a hub was present at the start of a burst, the arrival time of that hub overlapping with the MS2/MCP region and then overlap intensity and volume. To assess if a hub was present we used two different metrics: we first assessed if any hub voxels overlapped with the

MS2/MCP region and then assessed if a hub was present and was over a threshold intensity value (the nuclear mean intensity).

Fast single molecule tracking (SMT) analysis

To process the fast single-molecule tracking (SMT) data, we first removed the initial ~1,000 frames affected by fluorophore bleaching. Maximum projections were generated and input into the Cellpose-SAM model (*63*) to obtain nuclear masks. These masks were manually edited to ensure localizations along the nuclear border were excluded. Localization and tracking of the single molecule trajectories was performed using the open source software package *quot* (*64*). For the 10 msec data the following key parameters on *quot* were used: (i) for detection - method = "identity", Spot detection settings = "llr", k = 1.3, w = 15, t = 18, (ii) for localization - method = "ls_int_gaussian", window size = 9, sigma = 1.0, ridge = 0.0001, maximum iterations = 10, damp = 0.3, camera gain = 1, camera background = 108 (iii) for tracking - method = "euclidean", pixel size (in μm) = 0.108, search radius = 0.7 μm, maximum blinking = 0 frames. We then filtered the list of tracked molecules output by quot using the nuclear masks as described earlier to ensure we analyzed only trajectories localized in the nucleus.

To analyze the fast single molecule trajectories to infer the diffusion kinetics we used Spot-On (*46*). Spot-On is a publicly available package that fits a kinetic model to the jump length distribution of the trajectories and estimates the fraction bound and diffusion coefficients for either a two state or a three state model, correcting for the probability of molecules diffusing out of focus. Trajectories from all videos were concatenated into one file per gradient position. For the analysis, we used the following parameters altered from default:

TimeGap = 12
dZ = 0.700
GapsAllowed = 0
TimePoints = 4
UseEntireTraj = 1
JumpsToConsider = 0
NumberOfStates = 2
FitIterations = 2
LocError = 0.030
UseWeights = 1
D_Free_2State = [0.2656 25]
D_Bound_2State = [0.0001 0.2656]

Single molecule trajectories were taken from a minimum of three independent embryos for all positions and imaging conditions. Detailed trajectory number and the number of videos from which the trajectories were compiled can be found in each figure caption.

Slow single molecule tracking (SMT) analysis

To process the slow SMT data, the same methodology was followed as described above for fast SMT, with the exception of two *quot* parameters: k =1.5 and t=15.5. Residence times were calculated from single-molecule data acquired at 500 msec exposure times as described previously (*9, 65*). In brief, imaging at 500msec exposure times effectively blurs out fast moving molecules, leaving the slowly diffusing and bound molecules for tracking. The longer exposure time enables lower excitation laser power, leading to less photobleaching and longer trajectories. The persistence of trajectories over time is used to calculate a survival probability, which is 1- cumulative distribution function of trajectory lengths (in number of frames), providing an estimate of a genome average residence time. An objective threshold of two seconds (*9*) was applied to the minimum trajectory length to remove the effects of tracking errors and slowly diffusing molecules. Probabilities below 10-3 were not considered for fitting to avoid using the sparsely populated distribution tails.

The survival probability curves were then fit to a double exponential model of the form $F * \exp(-k_{ns} * t) + (1 - F) * \exp(-k_s * t)$ using the curve_fit function in Python where k_s is the slower off-rate and k_{ns} is the faster/nonspecific off-rate. An exponential weighting function was used to ensure proper estimation of the slower off-rate so that the fit was not dominated by the more numerous early-time point trajectories. As the inferred off-rate ks is biased by photobleaching and chromatin motion, bias correction was performed as previously described: $k_{s,true}=k_s-k_{bias}$, where k_{bias} is the slower-off rate estimate from fitting the survival probabilities from H2B-mEos4a trajectories imaged in the same conditions. The genome wide residence times were then calculated as $1/k_{s,true}$. Errors reported were estimated from the standard deviations of the fit parameters using standard propagation of error methods.

Simulations

Simulations modeling Dorsal diffusion and binding were developed in Python. Nuclei in nc14 were modeled as spheres of 5 μm and Dorsal molecules were initialized at random positions within the spheres. The concentration of Dorsal along the gradient was estimated using the known ventral concentration, 500 nM (*48*), and nuclear mean intensity ratios from our live imaging data. Dorsal binding sites are also distributed in clusters of varying densities. Each binding site is assigned a residence time with probabilities broadly based on the survival curves in Fig. S8A, with a greater proportion of sites having shorter residence times. A three-state model was used with Dorsal molecules being either freely-diffusing, specifically bound or non-specifically bound. Free diffusion was modeled by Brownian motion with a simulation time step of 10 msec to mimic our single-molecule tracking. At each time step, molecules within a 50 nm capture radius of an unoccupied binding site can bind the site with a certain probability k_{on}, reflecting the accessibility of the site. Once bound, the molecule diffuses with the binding site, which undergoes fractional Brownian motion to mimic the sub-diffusive motion of DNA (Rouse polymer model). This is done via the Davies-Harte algorithm, on top of which a Gaussian jiggle is added to represent local fluctuations. For memory efficiency, fractional Brownian motion increments are generated per display frame, and then bridged together such that subsequent frames are conditioned on previous ones. After the residence time of the binding site, the molecule unbinds and returns to the freely-diffusing state. When cooperativity is enabled, the presence of a bound molecule within a cluster of binding sites increases the k_{on} for subsequent molecules based on a Hill coefficient.

Simulated z-stacks were recorded at intervals of 10 sec with a camera exposure time of 30 msec. The molecules intensities at each simulated step within the exposure time are added together and projected on a 3D voxel grid, with an xy-spacing of 108 nm (corresponding to the pixel size on the MOSAIC camera) and a z-spacing of 300 nm (corresponding to the width of our z-stacks). Each molecule is convolved with a PSF, an anisotropic Gaussian with lateral full-width at half maximum (FWHM) of 250 nm and axial FWHM of 600 nm, the estimated post-deconvolution blur obtained by deconvolving a fiducial image from the MOSAIC with a theoretical PSF.

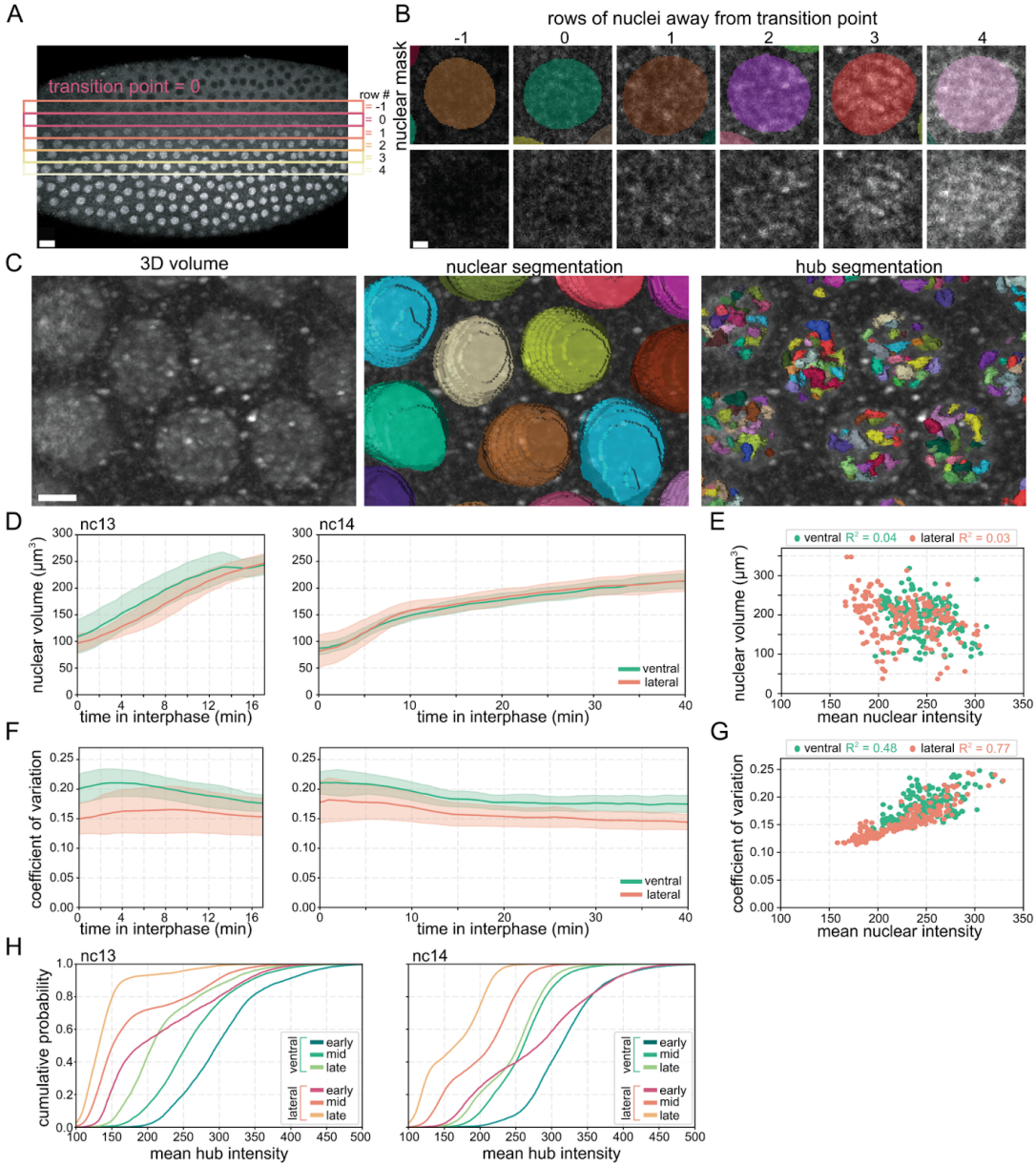


Figure S1: Nuclear segmentation and hub detection across the dorsoventral axis. (A, B) Images showing where nuclei were selected at the lateral surface within a max projection of a full field of view **(A)** and for a single slice for each nucleus. **(B)** Nuclei labelled as lateral were imaged within five nuclear rows of the transition point (row 0 of nuclei where the cytoplasm and nuclear signals are indistinguishable). Each row has a varying nuclear to cytoplasmic ratio of signal yet all have hubs within the nucleus. Scale bar in B is 15 microns and 1 micron in C. **(C)** Images showing results of nuclear and hub segmentation. Left image shows a representative 3D rendering of Dorsal-mNeonGreen in ventral nuclei. Scale bar is 2 microns. Center image shows a 3D rendering of nuclear segmentation results. Right image shows segmentation of Dorsal hubs. **(D)** Average nuclear volume in ventral (green) or lateral (orange) nuclei. Shading shows standard deviation between embryo replicates. N = 5 nuclei per embryo, 3 embryo replicates for each surface. **(E)** Scatter plot showing nuclear volume and mean nuclear intensity. R^2 is Pearson's correlation. **(F)** Coefficient of variation per nucleus in ventral (green) or lateral (orange) nuclei. **(G)** Scatter plot showing coefficient of variation and mean nuclear intensity. R^2 is Pearson's correlation. **(H)** Cumulative probability distributions of mean hub intensity for early, mid, and late nc13 (left) and nc14 (right).

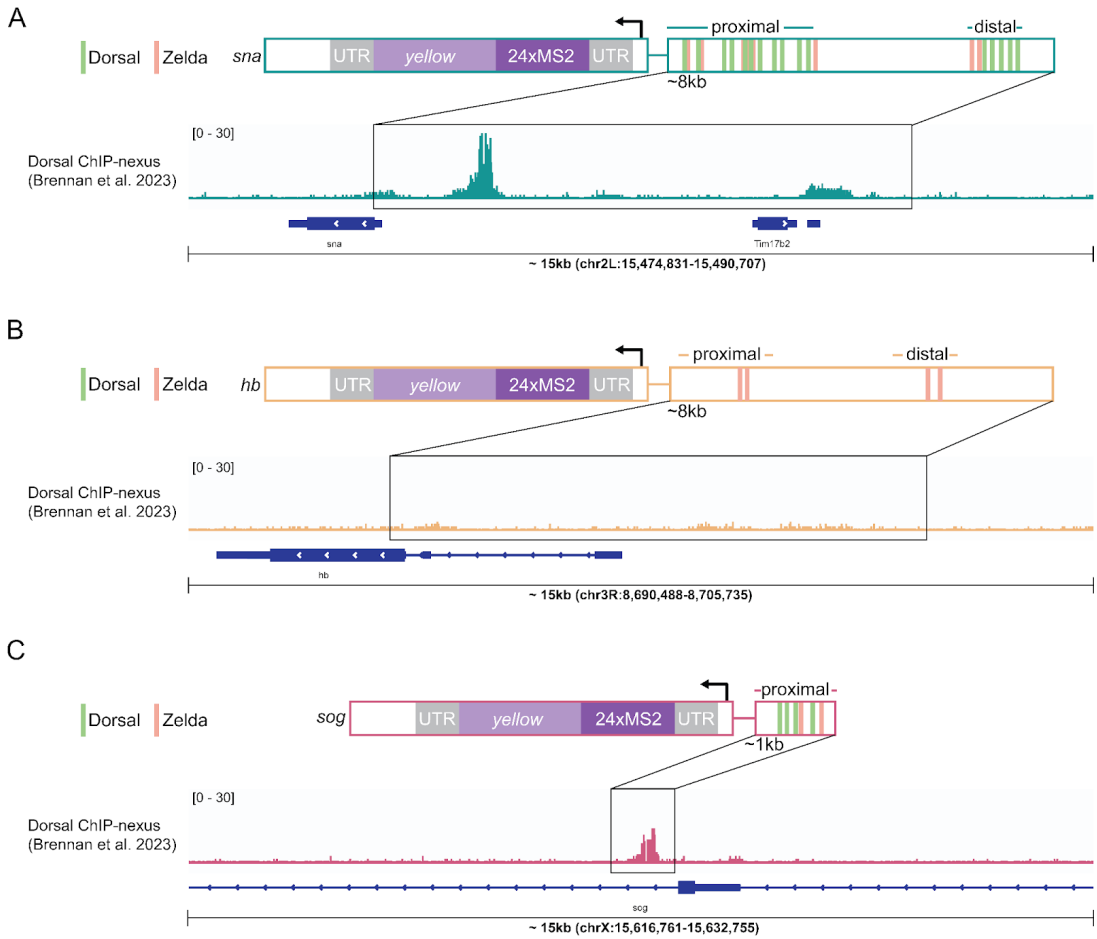
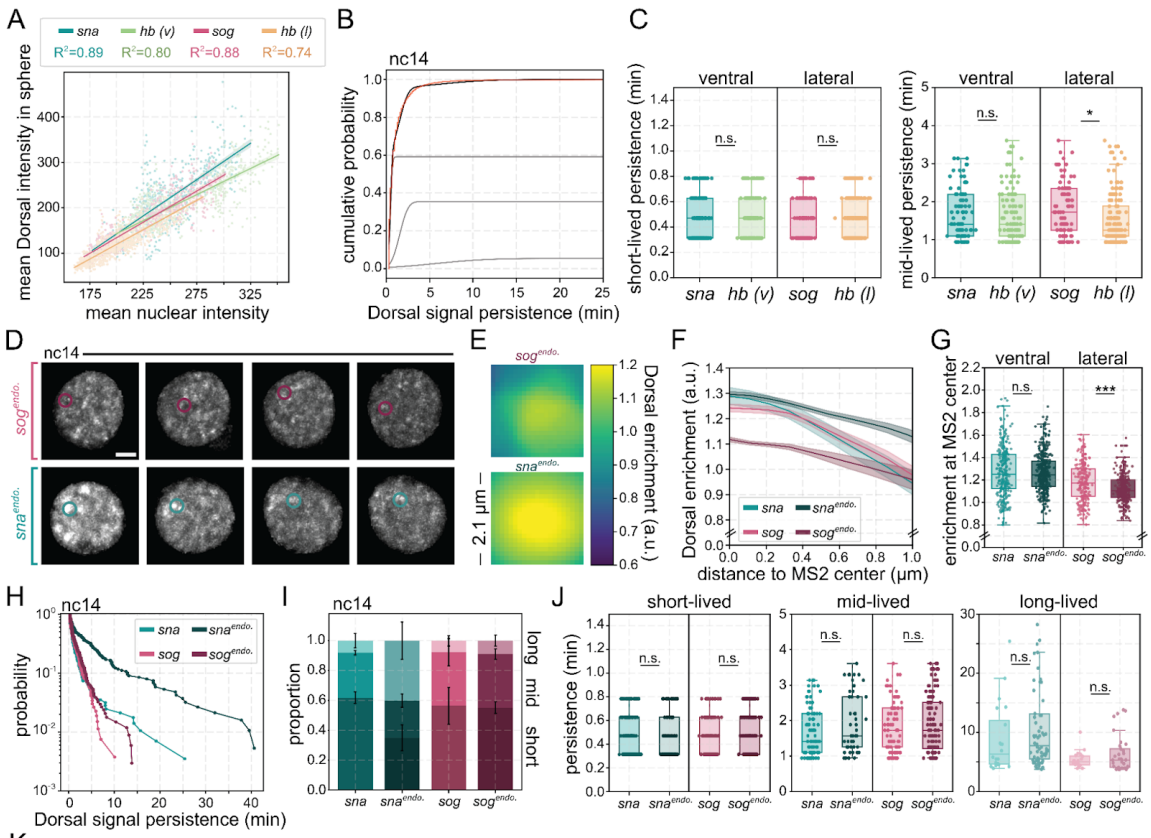


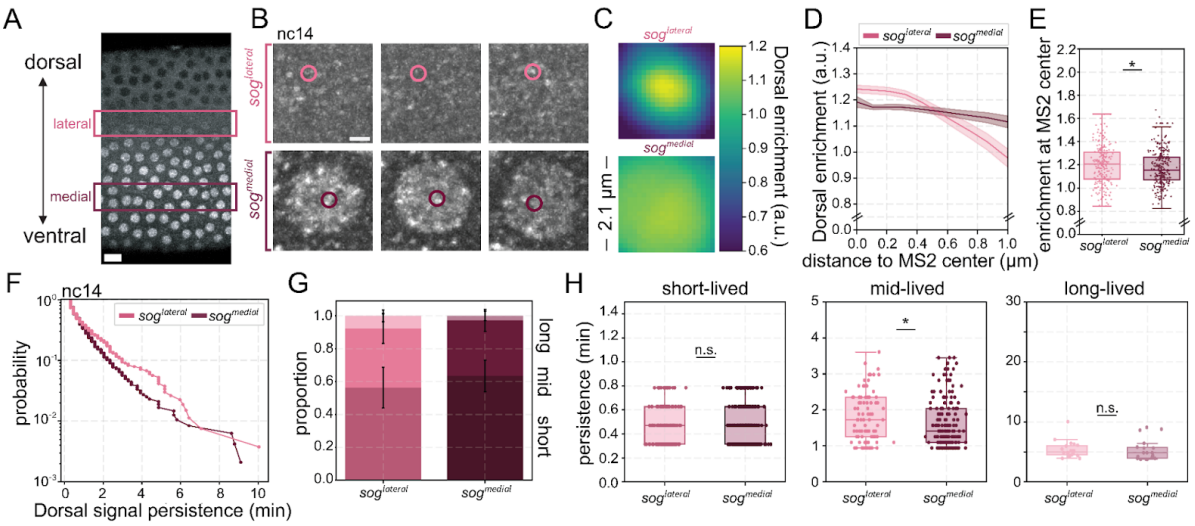
Figure S2: Design of MS2 reporters and Dorsal occupancy in enhancer regions. (A) Design of *snail-MS2* from (*15*) along with Dorsal occupancy (Dorsal ChIP-nexus from (*66*)) at the endogenous enhancer region corresponding to the region in the reporter. **(B)** Design of *hunchback-MS2* from (*15*) along with Dorsal occupancy (Dorsal ChIP-nexus from (*66*)) at the endogenous enhancer region corresponding to the region in the reporter. **(C)** Design of *sog-MS2* from (*67*) along with Dorsal occupancy (Dorsal ChIP-nexus from (*66*)) at the endogenous enhancer region corresponding to the region in the reporter.



K

	short-lived	mid-lived	long-lived
sna	0.62±0.04	0.30±0.04	0.08±0.05
sna$^{endo.}$	0.35±0.09	0.25±0.04	0.40±0.12
sog	0.56±0.12	0.36±0.09	0.08±0.03
sog$^{endo.}$	0.55±0.04	0.36±0.03	0.09±0.04

Figure S3: Dorsal hubs at transgenic and endogenous enhancers correlate with nuclear intensity and exhibit a range of persistence. (A) Scatterplots show correlation between mean Dorsal intensity at each gene locus and mean nuclear intensity. R^2 is Pearson's Correlation. **(B)** The cumulative probability of Dorsal persistence at a locus is fit to a three-component gaussian mixture model. The fits for each component and all three together are shown. **(C)** Box plots showing the average duration of consistent hub interactions within the short- and mid-lived populations determined by a three-component gaussian-mixture model. A Mann-Whitney U-test was performed to determine significance and following p-values were used: *p<0.05, **p<0.01, and ***p<0.001. **(D)** Representative maximum intensity projections of single nuclei with Dorsal-mNeonGreen in nc14. The MS2/MCP spot for each endogenous knock-in is circled in each image. The cytoplasmic Dorsal signal is masked, and contrast was adjusted independently in each row to aid visualization. Scale bars are 2 microns. **(E)** Images show average Dorsal enrichment centered at the MS2/MCP site. Images are 2.1 μm x 2.1 μm. The color bar on the bottom shows the level of Dorsal enrichment in the image (Dorsal intensity over nuclear mean intensity). **(F)** Average radial profile with standard error centered at the MS2/MCP spot shows Dorsal enrichment as a function of distance from MS2/MCP spot center. Both transgenic enhancers (Fig. 2) and endogenous knock-ins of MS2 loops for each gene are analyzed together (see Table S1 for genotypes of embryo imaged). **(G)** Boxplots showing the distribution of the Dorsal enrichment at the center of MS2/MCP site (radius = 0.0 μm). A Mann-Whitney U-test was performed to determine significance and following p-values were used: *p<0.05, **p<0.01, and ***p<0.001. **(H)** Survival plots showing the length of time any hub is consistently present within the interaction sphere for each reporter for nc14. **(I)** Using a Gaussian Mixture Model, each hub persistence event was categorized into short-, mid-, or long-lived bins. The proportion in each bin is shown with error bars designating difference among embryo replicates. **(J)** Box plots showing the average duration of consistent hub interactions within the short-, mid-, and long-lived populations determined by a three-component gaussian-mixture model. N = *sna*$^{endo.}$*:* 24 nuclei in 3 embryos and *sog*$^{endo.}$*:* 24 nuclei in 3 embryos. A Mann-Whitney U-test was performed to determine significance. **(K)** The proportion of short-, mid-, and long-lived hub persistence events are listed with their mean and standard deviation in the table.



I

	short-lived	mid-lived	long-lived
sog^{lateral}	0.56±0.12	0.36±0.09	0.08±0.03
sog^{medial}	0.63±0.10	0.34±0.07	0.03±0.03

Figure S4: Differences in apparent Dorsal hub properties at *sog* at different Dorsal concentrations is due to nuclear background concentration. (A) Image of Dorsal-mNeonGreen where the 'lateral' and 'medial' *sog* nuclei were imaged (see Table S1 for genotypes of embryo imaged). **(B)** Representative maximum intensity projections of single nuclei with Dorsal-mNeonGreen in nc14. The *sog*-MS2 reporter location is circled in each image. The scale bar is 2 microns. **(C)** Images show average Dorsal enrichment centered at the MS2/MCP site. Images are 2.1 μm x 2.1 μm. The color bar on the bottom shows the level of Dorsal enrichment in the image (Dorsal intensity over nuclear mean intensity). **(D)** Average radial profile with standard error centered at the MS2/MCP spot shows Dorsal enrichment as a function of distance from MS2/MCP spot center. **(E)** Boxplots showing the distribution of the Dorsal enrichment at the center of MS2/MCP site (radius = 0.0 μm). A Mann-Whitney U-test was performed to determine significance and following p-values were used: *p<0.05, **p<0.01, and ***p<0.001. **(F)** Survival plots showing the length of time any hub is consistently present within the interaction sphere for each reporter for nc14. **(G)** Using a Gaussian Mixture Model, each hub persistence event was categorized into short-, mid-, or long-lived bins. The proportion in each bin is shown with error bars designating standard deviation among embryo replicates. **(H)** Box plots showing the average duration of consistent hub interactions within the short-, mid-, and long-lived populations determined by a three-component gaussian-mixture model. N = *sog*^{lateral}*:* 18 nuclei in 3 embryos and *sog*^{medial}*:* 24 nuclei in 3 embryos. A Mann-Whitney U-test was performed to determine significance and following p-values were used: *p<0.05, **p<0.01, and ***p<0.001. **(I)** The proportion of short-, mid-, and long-lived hub persistence events are listed with their mean and standard deviation in the table.

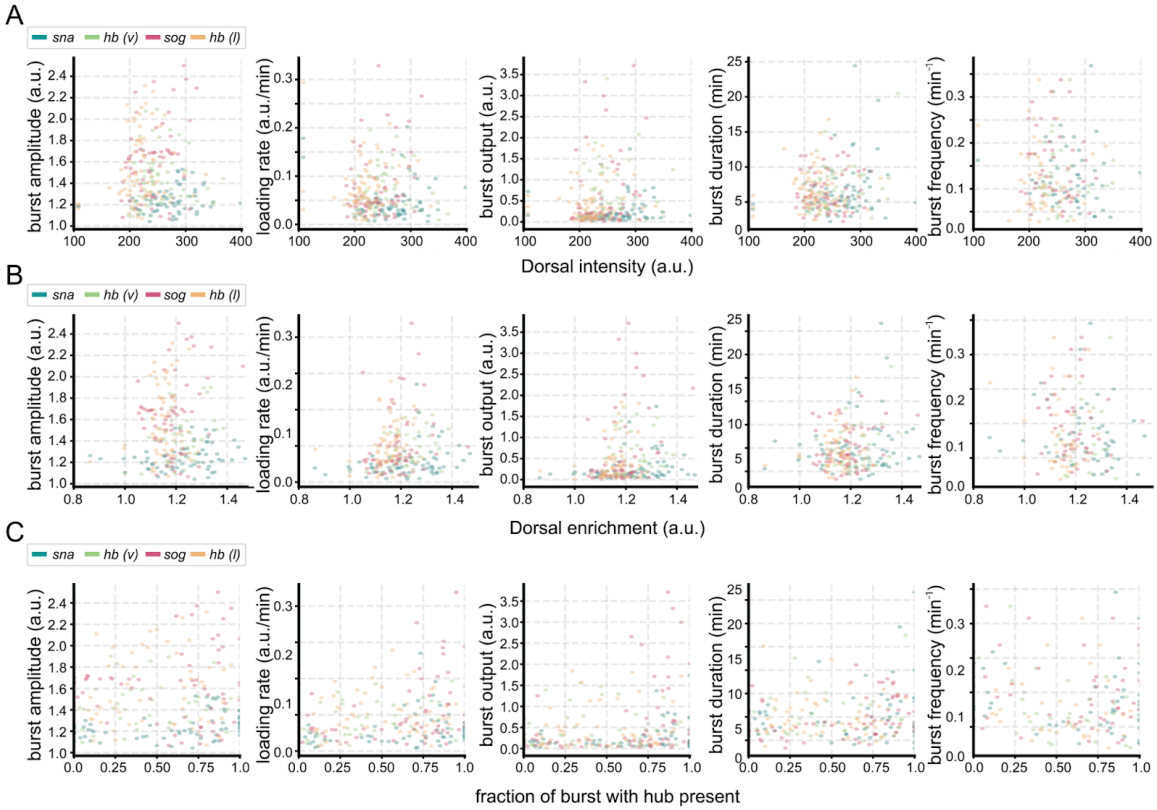


Figure S5: Hub properties and transcription burst kinetics show minimal correlations. (A-C) Scatter plots showing burst metrics (amplitude, loading rate, output, duration, frequency) with the mean Dorsal intensity **(A)**, Dorsal enrichment **(B)**, or fraction of burst with a hub present **(C)** before the burst (left), the mean Dorsal enrichment before the burst (center), or the fraction of time during a burst when a hub is persisting near the MS2/MCP site (right). Marker colors correspond to each reporter analyzed. These data points were used in the heat maps reflecting Spearman's rank correlation in Figure 3D-F.

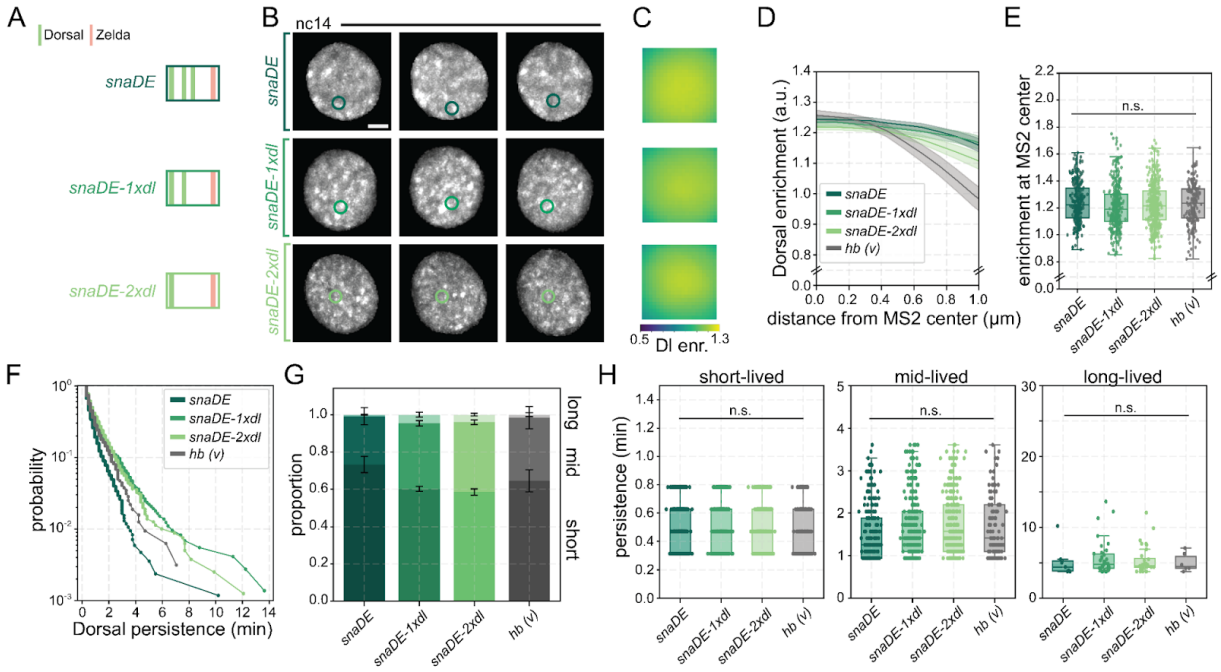
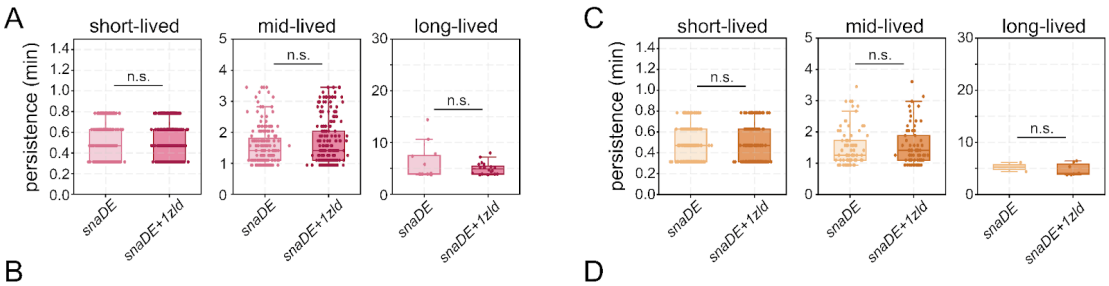


Figure S6: Hub enrichment and signal persistence detection limits. (A) Schematic depicting enhancers with Dorsal and Zelda binding sites for each *snail* reporter with the minimal distal enhancer of *snail* (*snaDE*), minus one Dorsal binding site (*snaDE-1xdl*), or minus two Dorsal binding site (*snaDE-2xdl*). Each enhancer was positioned just upstream of its promoter, 10 MS2 stem loops, and the *yellow* coding sequence (see Table S1 for exact genotypes of embryo imaged). Each of these *snail* reporters were imaged along the ventral surface in nc14. **(B)** Representative maximum intensity projections of single nuclei with Dorsal-mNeonGreen in nc14. The MS2/MCP site is circled in each image. The cytoplasmic Dorsal signal is masked, and contrast was adjusted independently in each row to aid visualization. Scale bars are 2 microns. **(C)** Images show average Dorsal enrichment centered at the MS2/MCP site. Images are 2.1 μm x 2.1 μm. The color bar on the bottom shows the level of Dorsal enrichment in the image (Dorsal intensity over nuclear mean intensity). **(D)** Average radial profile with standard error centered at the MS2/MCP spot shows Dorsal enrichment as a function of distance from MS2/MCP spot center. **(E)** Boxplots showing the distribution of the Dorsal enrichment at the center of MS2/MCP site (radius = 0.0 μm). A Mann-Whitney U-test was performed to determine significance. **(F)** Survival plots showing the length of time any hub is consistently present within the interaction sphere for each reporter for nc14. **(G)** After applying a Gaussian Mixture Model, each hub persistence event was categorized into short-, mid-, or long-lived bins. The proportion in each bin is shown with error bars designating standard deviation among embryo replicates. **(H)** Box plots showing the average duration of consistent hub interactions within the short-, mid-, and long-lived populations determined by a three-component gaussian-mixture model. N = *snaDE:* 28 nuclei in 3 embryos; *snaDE-1xdl:* 33 nuclei in 3 embryos; *snaDE-2xdl:* 44 nuclei in 4 embryos; and *hb (v):* 18 nuclei in 4 embryos. A Mann-Whitney U-test was performed to determine significance and following p-values were used: *p<0.05, **p<0.01, and ***p<0.001.



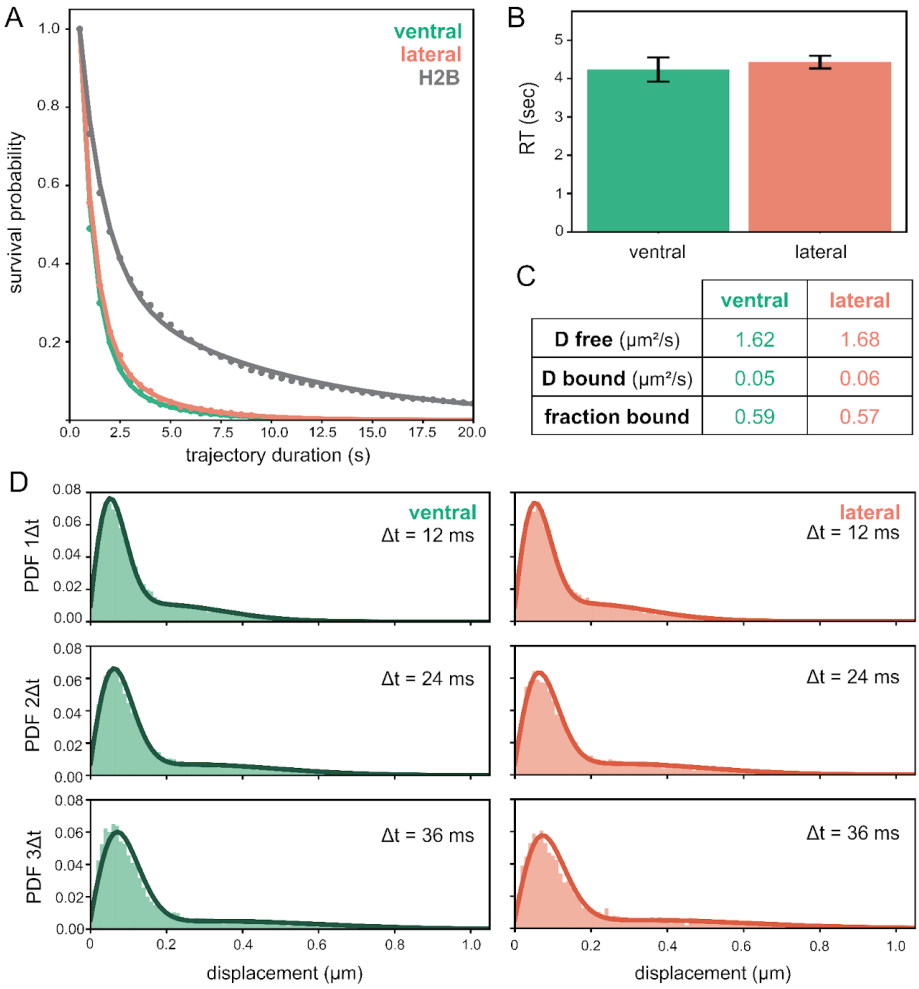
B

Dorsal	short-lived	mid-lived	long-lived
snaDE	0.66±0.04	0.32±0.04	0.02±0.01
snaDE+1zld	0.64±0.10	0.33±0.08	0.03±0.03

D

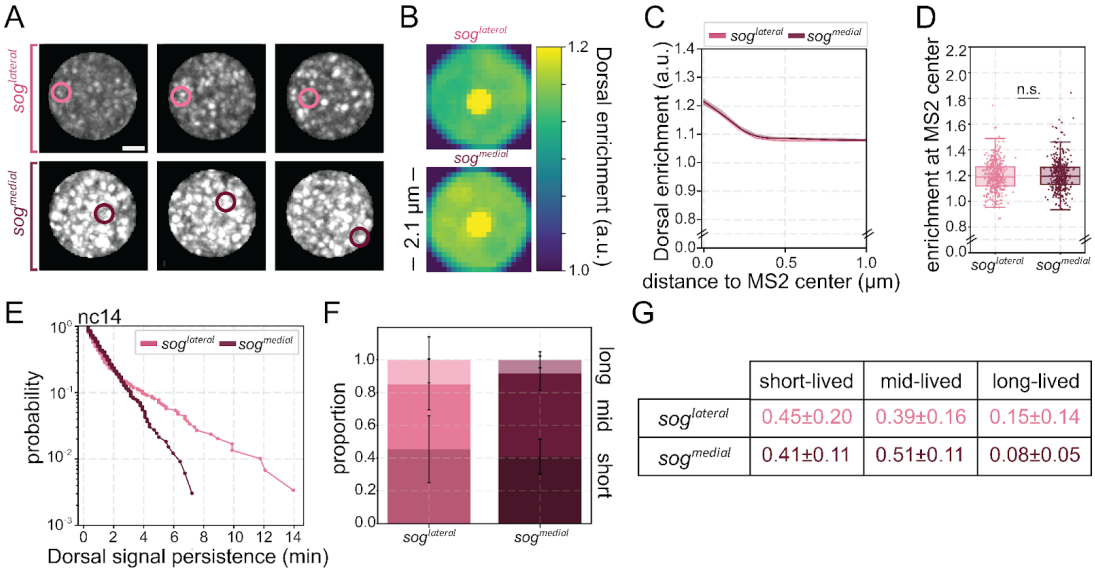
Zelda	short-lived	mid-lived	long-lived
snaDE	0.72±0.10	0.28±0.10	0.01±0.01
snaDE+1zld	0.73±0.03	0.25±0.04	0.02±0.01

Figure S7: Dorsal and Zelda persistence are moderately impacted from a single Zelda binding site addition. (A, C) Box plots showing the average duration of consistent Dorsal **(A)** or Zelda **(C)** hub interactions within the short-, mid-, and long-lived populations determined by a three-component gaussian-mixture model. A Mann-Whitney U-test was performed to determine significance and following p-values were used: *p<0.05, **p<0.01, and ***p<0.001. **(B, D)** The proportion of short-, mid-, and long-lived Dorsal **(B)** or Zelda **(D)** hub persistence events are listed with their means and standard deviations in the table.



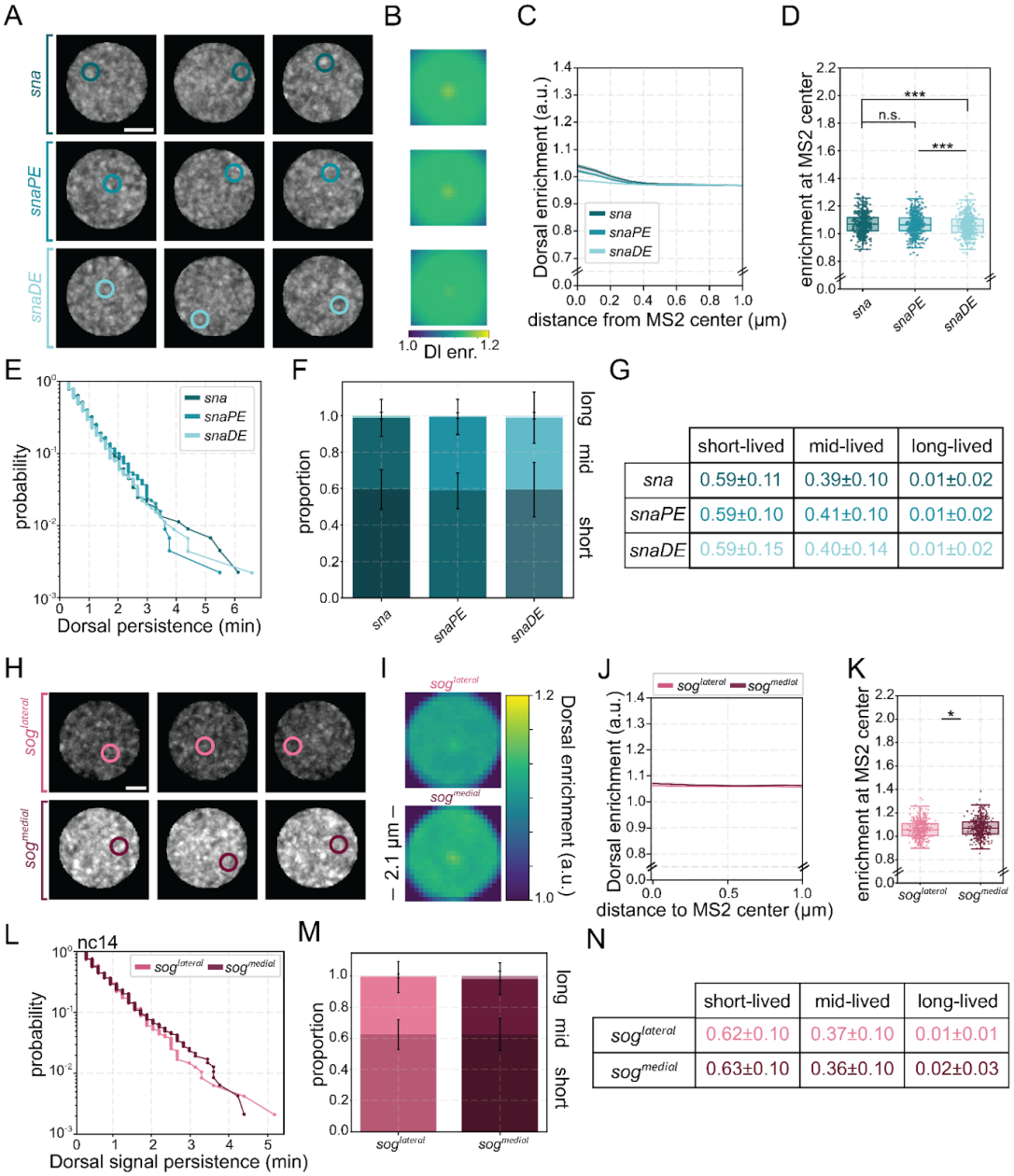
	ventral	lateral
D free (μm²/s)	1.62	1.68
D bound (μm²/s)	0.05	0.06
fraction bound	0.59	0.57

Figure S8: Single molecule tracking (SMT) of Dorsal shows minimal differences on the ventral and lateral regions. (A) Survival probabilities of single-molecule trajectories at 500 msec exposure times for at least three embryos for Dorsal-mEos4a ventrally positioned (N = 978 trajectories, 5 unique embryos), Dorsal-mEos4a laterally positioned (N = 1214 unique trajectories, 4 unique embryos), and H2B-mEos4a (N = 3098 trajectories, 3 unique embryos)**.** **(B)** Dorsal residence times after photobleaching correction for ventral (4.24 ± 0.32 sec) and lateral (4.43 ± 0.17 sec) positions. **(C)** Spot-On diffusion coefficient estimates for free and bound populations. **(D)** Probability density functions (PDFs) of the trajectory displacements of ventral (N = 10968 trajectories, 4 unique embryos) and lateral (N = 5824 trajectories, 11 unique embryos) Dorsal-mEos4a across three time intervals.



	short-lived	mid-lived	long-lived
sog^{lateral}	0.45±0.20	0.39±0.16	0.15±0.14
sog^{medial}	0.41±0.11	0.51±0.11	0.08±0.05

Figure S9: Computational modeling shows the effect of nuclear concentration on persistence measurements. (A) Representative maximum intensity projections of simulated single nuclei with Dorsal signal in grey. The binding site cluster is circled in each image. Contrast is scaled equally for all conditions. Scale bars are 2 microns. **(B)** Images show simulated average Dorsal enrichment centered at the binding site clusters. Images are 2.1 μm x 2.1 μm. The color bar on the bottom shows the level of Dorsal enrichment in the image (Dorsal intensity over nuclear mean intensity). **(C)** Average radial profile with standard error centered at the binding site cluster shows Dorsal enrichment as a function of distance from spot center. **(D)** Boxplots showing the distribution of the Dorsal enrichment at the center of the binding site cluster (radius = 0.0 μm). A Mann-Whitney U-test was performed to determine significance and following p-values were used: *p<0.05, **p<0.01, and ***p<0.001. **(E)** Survival plots showing the length of time any hub is consistently present within the interaction sphere for each reporter. **(F)** The proportion of long-, mid, -and short-lived persistent hub formation in each bin is shown with error bars designating standard deviation among embryo replicates. **(G)** The proportion of short-, mid-, and long-lived hub persistence events are listed with their mean and standard deviation in the table.



	short-lived	mid-lived	long-lived
sna	0.59±0.11	0.39±0.10	0.01±0.02
snaPE	0.59±0.10	0.41±0.10	0.01±0.02
snaDE	0.59±0.15	0.40±0.14	0.01±0.02

	short-lived	mid-lived	long-lived
$sog^{lateral}$	0.62±0.10	0.37±0.10	0.01±0.01
sog^{medial}	0.63±0.10	0.36±0.10	0.02±0.03

Figure S10: Computational modeling without cooperativity fails to recapitulate Dorsal hub properties. (A, H) Representative maximum intensity projections of simulated single nuclei without cooperative binding included in computation (Dorsal signal in grey). The binding site cluster is circled in each image. Contrast is scaled equally for all conditions. Scale bars are 2 microns. **(B, I)** Images show simulated average Dorsal enrichment centered at the binding site clusters. Images are 2.1 μm x 2.1 μm. The color bar on the bottom shows the level of Dorsal enrichment in the image (Dorsal intensity over nuclear mean intensity). **(C, J)** Average radial profile with standard error centered at the binding site cluster shows Dorsal enrichment as a function of distance from spot center. **(D, K)** Boxplots showing the distribution of the Dorsal enrichment at the center of the binding site cluster (radius = 0.0 μm). A Mann-Whitney U-test was performed to determine significance and following p-values were used: *p<0.05, **p<0.01, and ***p<0.001. **(E, L)** Survival plots showing the length of time any hub is consistently present within the interaction sphere for each reporter. **(F, M)** The proportion of long-, mid, -and short-lived persistent hub formation in each bin is shown with error bars designating standard deviation among embryo replicates. **(G, N)** The proportion of short-, mid-, and long-lived hub persistence events are listed with their mean and standard deviation in the table

Table S1: Details of *Drosophila* lines used.

Shorthand	Genotype*	Source	Appears in Figures
Dorsal-mNeonGreen	yw; Dorsal-mNeonGreen/CyO;	(*41*)	1, 2, 3, 4, S1, S3, S4, S5, S6, S7, S8
Dorsal-mEos4a	yw; Dorsal-mEos4a/CyO;	(*41*)	6, S8
mNeonGreen-Zelda	mNeonGreen-Zelda;;	(*9*)	5, S7
MCP-mCherry	nos>MCP-mCherry (III)	(*48*)	2, 3, 4, S3, S4, S5, S6, S7, S8
MCP-mCherry	enosx2>MCP-mCherry (II)	(*41*)	5, S7
sna	sna-24xMS2 (BAC)	(*15*)	2, 3, S3, S5
hb	hb-24xMS2 (BAC)	(*15*)	2, 3, 4, S3, S5
sog, $sog^{lateral}$, sog^{medial}	sog-24xMS2 (BAC)	(*67*)	2, 3, S3, S4, S5
$sna^{endo.}$	sna-24xMS2 (endogenous)	(*68*)	S3
$sog^{endo.}$	sog-24xMS2 (endogenous)	Gift from Bomyi Lim	S3
snaPE	snaPE > snaPr > 24xMS2-yellow	Gift from Bomyi Lim	4
snaDE	snaDE > snaPr > 24xMS2 - yellow	(*39*)	4, 5, S7
snaDE	snaDE > snaPr > 10xMS2 - yellow	(*47*)	S6
snaDE - 1xdl	snaDE del dl1 > snaPr > 10xMS2 - yellow	(*47*)	S6
snaDE - 2xdl	snaDE del dl1/2 > snaPr > 10xMS2 - yellow	(*47*)	S6
snaDE + 1 zld	snaDE + 1 Zld > snaPr > 24xMS2 - yellow	(*39*)	5, S7

* All non-endogenous MS2 reporters inserted into VK33 on Chromosome III

Table S2: Simulation parameters.

Category	Parameter	Value
General Parameters	Dorsal concentration from (*48*)	Ventral = 500 nM
		Medial = 406 nM
		Lateral = 219 nM
	Nucleus diameter	5 μm
	Molecule diameter	15 nm
	Simulation time step	10 ms
	Display time step	10 s
Binding sites	Number (post-replication) from (*10*)	7000
	Endogenous sites per cluster	2 (60%), 10 (25%), 20 (10%), 30 (4%), 40 (1%)
	Endogenous site cluster diameter	3 px = 324 nm
	Diffusion coefficient	0.05 ± 0.03 $\mu m^2/s$ (0.001 μm jiggle)
	Hurst exponent	0.25
	Synthetic sites per cluster (post-replication)	*snail* = 32 (in a 1, 3 or 5 px diameter)
		snailPE = 22 (in a 1, 2 or 4 px diameter)
		snailDE = 6 (in a 1, 2 or 3 px diameter)
		sog = 8 (in a 1, 2 or 3 px diameter)
Three-state diffusion model	Initialization	50% free, 50% non-specifically bound
	Free diffusion coefficient	1.7 ± 0.45 $\mu m^2/s$
	Switching rate (free to non-specific binding)	1 Hz
	Non-specific residence time	0.5 s
	Binding probability k_{on}	0.005
	Capture radius	50 nm
Binding site strength distribution	Base residence time ($1/k_{off}$)	5 ± 1 s
	Endogenous site residence time (multipliers of base residence time)	0.06x (50%), 0.2x (27%), 0.5x (15%), 1x (6%), 2.4x (2%)
	Synthetic site residence time (multipliers of base residence time)	*snail*/*snailPE*/*snailDE* = 0.2x
		sog = 0.5x
	Hill coefficient	1 (without cooperativity) or 100 (with cooperativity)
Visualization	PSF FWHM	Lateral = 250 nm
		Axial = 600 nm
	z-slice spacing	300 nm
	Exposure time	30 ms
	Pixel size	108 nm

Supplemental Movie Captions.

Movie S1: Movies showing Dorsal-mNeonGreen (gray) in nuclei along the lateral (left) or ventral (right) surface of a *Drosophila* embryo in nc13 and nc14. Each movie is a 3D rendering with equivalent contrast limits.

Movie S2: Movie showing Dorsal-mNeonGreen (gray) in 3D rendered nuclei along the lateral transition point of a *Drosophila* embryo in nc13 and nc14.

Movie S3: Example of hub segmentation (colored labels) within ventral nuclei of Dorsal-mNeonGreen (gray). Movie starts on a single z-slice then moves through the whole nucleus in z.

Movie S4: Example 3D rendering of nuclei through time showing Dorsal-mNeonGreen (gray) at MS2 spots (circled) for *sna, hb (v), sog,* and *hb (l)* in nc13. For visualization purposes, cytoplasm has been removed and set to zero, and contrast limits are individually adjusted for ventral or lateral nuclei.

Movie S5: Example 3D rendering of nuclei through time showing Dorsal-mNeonGreen (gray) at MS2 spots (circled) for *sna, hb (v), sog,* and *hb (l)* in nc14. For visualization purposes, cytoplasm has been removed and set to zero, and contrast limits are individually adjusted for ventral or lateral nuclei.

Movie S6: Example of 3D MS2 tracking in time within a single nucleus. Gray channel shows MCP-mCherry distribution in the nucleus. For visualization purposes, cytoplasm has been removed and set to zero. A pink circle marks the spot of nascent transcription of *sog*-MS2. The left shows a top-down (X-Y) view while the right shows a side (Y-Z) view.

Movie S7: Example 3D rendering of nuclei through time showing Dorsal-mNeonGreen (gray) at MS2 spots (circled) for *snail (endogenous)* and *sog (endogenous)* in nc14. For visualization purposes, cytoplasm has been removed and set to zero, and contrast limits are individually adjusted for ventral or lateral nuclei.

Movie S8: Example 3D rendering of nuclei through time showing Dorsal-mNeonGreen (gray) at MS2 spots (circled) for *sog (lateral)* and *sog (medial)* in nc14. Contrast limits are equivalent across conditions.

Movie S9: Example 3D rendering of nuclei through time showing Dorsal-mNeonGreen (gray) at MS2 spots (circled) for *sna, snaPE,* and *snaDE* in nc14. For visualization purposes, cytoplasm has been removed and set to zero. Contrast limits are equivalent across conditions.

Movie S10: Example 3D rendering of nuclei through time showing Dorsal-mNeonGreen (gray) at MS2 spots (circled) for *snaDE, snaDE-1xdl,* and *snaDE-2xdl* in nc14. For visualization purposes, cytoplasm has been removed and set to zero. Contrast limits are equivalent across conditions.

Movie S11: Example 3D rendering of nuclei through time showing Dorsal-mNeonGreen (gray) at MS2 spots (circled) for *snaDE* and *snaDE+1zld* in nc14. For visualization purposes, cytoplasm has been removed and set to zero. Contrast limits are equivalent across conditions.

Movie S12: Example 3D rendering of nuclei through time showing mNeonGreen-Zelda (gray) at MS2 spots (circled) for *snaDE* and *snaDE+1zld* in nc14. For visualization purposes, cytoplasm has been removed and set to zero. Contrast limits are equivalent across conditions.