# Global maps of transcription factor properties reveal threshold-based formation of DNAbound and mobile clusters

Sadia Siddika Dima<sup>1</sup>, Meghan A. Turner<sup>2,3</sup>, Hernan G. Garcia<sup>2,4,5,6,7</sup>, Gregory T. Reeves<sup>1,8\*</sup>

<sup>1</sup>Department of Chemical Engineering, Texas A&M University, College Station, TX 77843

<sup>2</sup>Biophysics Graduate Group, University of California at Berkeley, Berkeley, USA

<sup>3</sup>Current address: Allen Institute for Brain Science, 615 Westlake Ave N, Seattle, WA, USA 98109

<sup>4</sup>Department of Molecular and Cell Biology, University of California, Berkeley, CA, USA

<sup>5</sup>Department of Physics, University of California, Berkeley, CA, USA

<sup>6</sup>Institute for Quantitative Biosciences-QB3, University of California, Berkeley, CA, USA

<sup>7</sup>Chan Zuckerberg Biohub – San Francisco, San Francisco, CA, USA

<sup>8</sup> Faculty in Genetics and Genomics, Texas A&M University, College Station, TX 77843

\* To whom correspondence should be addressed:
100 Spence Street
Texas A&M University, Chemical Engineering Dept
College Station, Texas 77843 MS 3122
979-862-4021
gtreeves@tamu.edu

## Abstract

The relationship between bulk transcription factor concentration and DNA binding has been a central question in gene regulation for decades. Recent studies propose that DNA-bound transcription factor hubs, or clusters, aid in fast and precise transcriptional interpretation. Using live imaging techniques, we quantify the concentration, binding, and mobility of the morphogen Dorsal (Dl), both in bulk and in clusters, in *Drosophila* blastoderm embryo. Our experiments encompass multiple length and time scales, allowing us to obtain a nucleus-wide view of the mechanism connecting hub formation to bulk Dl concentration. Our results show that previously unobserved, slowly-moving clusters of Dl are present, in addition to the expected populations of freely mobile and DNA-bound Dl. Furthermore, both mobile clusters and DNA-bound Dl appear only once a threshold concentration in the nucleus is surpassed, a behavior consistent with liquid-liquid phase separation. Thus, our work elucidates how bulk transcription factor concentration factor concentration and spatiotemporal changes of different populations needed for gene regulation.

## Introduction

Cell fate determination and maintenance in metazoans depend on the accurate spatiotemporal regulation of gene expression. The binding of TFs to regulatory motifs in the DNA in eukaryotes, and thereby the target gene expression, is dependent on a plethora of factors including its concentration, binding affinity, chromatin accessibility, and protein-protein interactions (1-4) Therefore, bulk concentration, a parameter commonly estimated from the fluorescence intensity of fixed or live imaging, cannot serve as a direct proxy to TF activity; instead, measurements of TF-DNA binding, in relationship to the bulk concentration, are required. Recent studies also suggest TFs form DNA-bound clusters (or hubs) of elevated concentration (reviewed in (5, 6)), potentially mediated by intrinsically disordered regions (7-9). It has been proposed that these clusters represent liquid-liquid phase-separated droplets; although that suggestion remains a matter of debate (5, 6). In *Drosophila*, the formation of TF hubs has been found to facilitate binding to low-affinity sites by increasing the local concentration of TFs (10-12). These hubs aid in transcriptional burst induction (9) and function in a gene-specific, or enhancer-specific manner (13). However, the mechanisms connecting bulk transcription factor concentration, on the

micron scale, to DNA binding and hub formation, on the nanometer scale, remain poorly understood.

Dorsal (Dl), a Drosophila homolog of mammalian NF-kB, acts as a morphogen to pattern the dorsal-ventral (DV) axis of the Drosophila blastoderm-stage (1-3 hr old) embryo, regulating hundreds of genes (14-16). DI is one of the earliest known morphogens (17) and one of the best characterized sequence-specific TFs (18). Previous work using fluorescent imaging of live and fixed embryos indicated that the Dl gradient has a Gaussian-like shape with the highest level of Dl present on the ventral side of the embryo (14, 19). The level of Dl oscillates in a progressively increasing saw-tooth pattern during blastoderm nuclear cycles (ncs) 10 to 14 (14, 20). This temporal variation has been proposed to ensure the accurate dynamic expression patterns of its target genes (14). However, modeling efforts that take spatiotemporal variation of the Dl gradient as the only input fail to predict expression pattern and dynamics of all the target gene domains (reviewed in (21, 22)). As such, knowing the bulk nuclear intensity of Dl is not sufficient to explain patterning. Intriguingly, recent groundbreaking work has shown that the activity of many transcription factors, including Dl, is also influenced by the presence of clusters (also referred to as hubs, aggregates, droplets, or condensates), and the enrichment of Dl clusters has been observed at the site of transcription in both fixed and live embryos (13, 23). But how the dynamic Dl gradient dictates the formation of these populations of clusters remains unexplored.

Thus, to understand the transcriptional interpretation of the Dl gradient, knowing the dynamics of Dl bulk nuclear concentration is an important, necessary piece, but it is not sufficient by itself. In the same way, knowing the cluster formation dynamics at specific enhancers is also required for a detailed view of gene regulation (13), but it is again not the full picture. Measurements of different populations of Dl -- such as the freely diffusing population, the DNA-bound population, clusters of Dl -- and, in particular, how they relate to each other in dose/response maps, must be considered. Therefore, we used advanced live imaging and analysis techniques, such as raster image correlation spectroscopy (RICS) (24–29), a type of scanning fluorescence correlation spectroscopy (30-32), and single particle tracking (SPT) of super-resolution images of an endogenously-tagged Dl. These experiments, which span a wide range of length and time scales, indicate the presence of three different populations of Dl: freely moving Dl, DNA-bound Dl, and

a previously unobserved population of slowly-moving clusters of Dl. Interestingly, a recent study similarly observed the formation of mobile clusters of two other blastoderm patterning factors, Bicoid (Bcd) and Capicua (Cic), and proposed these mobile clusters to be an efficient search strategy for regulatory regions (*33*). Thus, mobile cluster formation might be a common mechanism for the accurate transcriptional interpretation of morphogen gradients. Our results further suggest that the total nuclear concentration of Dl must reach a threshold of roughly 30 nM before DNA binding occurs and before the slowly-moving clusters of Dl are formed, potentially indicating an ultrasensitive process or, alternatively, a liquid-liquid phase separation (LLPS) process. Ultimately, our results suggest that the relationships of the different populations of Dl in the nucleus, including DNA-bound as well as slowly-moving clusters, must be taken into account to understand transcriptional regulation on a tissue-wide level. We suggest that the importance of these relationships is also extended to other transcription factor systems.

#### Results

#### The DI nuclear concentration and binding is dynamic throughout the blastoderm stage

Studying the dose/response behavior between the bulk concentration of a TF and its DNA binding interactions, which can provide important insights into gene regulation, requires measurements at a range of concentrations. To do so, we leveraged the fact that the Dl gradient has dynamically varying levels over nc 10-14. We made use of a CRISPR-edited *dl-mNeonGreen* (dl-mNG) construct in which the mNeonGreen was inserted scarlessly in-frame just before the stop codon of the endogenous *dl* locus (referred to as the "wildtype" construct). At two copies, this scarless CRISPR edit is viable and results in similar target gene expression to wildtype. We imaged the ventral-most nuclei (Fig. 1A) from nc 10 to gastrulation in embryos collected from mothers having one copy of the Dl-mNeonGreen wildtype construct and one copy of His2Av-RFP (to mark the DNA). This line is used in the rest of our experiments unless mentioned otherwise. Our time series image acquisitions (Movie S1) were compatible with RICS analysis (see Methods), which allowed us to quantify the temporal dynamics of the total nuclear concentration and binding of Dl. We found the absolute nuclear concentration of Dl increases from one nuclear cycle (nc) to the next from ~ 60 nM in nc 10 to ~120 nM in nc 14 (blue "wt" data in Fig. 1B), consistent with the increase in Dl gradient amplitude over time observed previously (14, 19, 20, 34). During each nc, the nuclear concentration of DI increases in the

beginning of interphase as Dl enters the nucleus after mitosis, then drops abruptly as the nuclear envelope breaks down before the next mitosis (14, 20). The long duration of nc 14 compared to the other ncs allows the nuclear concentration to reach a steady state value of ~ 120 nM.

According to our RICS measurements, the total nuclear concentration of Dl can be divided into two pools: freely diffusible Dl and slowly-moving Dl (26, 29, 35) (Fig. 1A; see Methods). The concentration of the freely diffusible pool of Dl was found to be maintained between roughly 20 and 30 nM during the blastoderm stage, reaching a steady state of  $\sim$ 30 nM during mid nc 14 (Fig. 1C).

The "slowly-moving" pool of Dl, which we defined as having a zero or nearly-zero diffusivity in our RICS analysis, is likely to be composed of Dl monomers or clusters bound to DNA. Both of these species would have a nearly zero diffusivity, and thus are indistinguishable in this method, being lumped together into this pool (indicated by a dashed circle around them in Fig. 1A). This pool could also, in principle, contain slowly-moving Dl with diffusivities significantly less than 1  $\mu$ m<sup>2</sup>/s, as diffusivities that low would be difficult to distinguish from zero. Therefore, to dissect the slowly-moving pool further, we performed cross-correlation RICS (ccRICS) by crosscorrelating the fluctuations between Dl-mNG and His2Av-RFP, (26, 29, 35) (see Methods). We found the population of Dl in the slowly-moving pool can be subdivided into the pool of Dl that cross-correlates with His2Av-RFP (likely representing a DNA-bound population (36, 37); referred to as the "correlated" pool henceforth), and the pool that does not (referred to as the "uncorrelated" pool) (35). The concentration of Dl in the correlated pool peaks in nc 12 and 13, then decreases in nc 14 to a steady state of roughly 30 nM (blue data in Fig. 1D; see also Fig. S1A). From these measurements alone, it is unclear what constitutes the remainder of the slowlymoving pool (not correlating with His2Av). Given the recent discovery of slowly-moving clusters of Bcd and Cic (33), we hypothesize the uncorrelated population we observe possibly includes slowly-diffusing oligomeric clusters of Dl (the presence of which is tested in the following sections). The concentration of this uncorrelated pool, determined from the difference between the concentrations of immobile Dl and correlated Dl, increases over time (blue data in Fig. 1E; see also Fig. S1B).

Both sub-populations of the slowly-moving pool are likely due to Dl binding interactions (either to DNA or to some unidentified structures). Therefore, we asked whether these binding interactions follow typical models of protein binding, such as TF/DNA interactions. These models are often loosely based on equilibrium binding affinity and are typically saturable functions of the free concentration, such as Hill functions (35, 38-42). In reality, the relation is complicated by the influence of several factors, including binding affinity, chromatin accessibility, and protein-protein interactions (1-4). Our measurements of the different subpopulations of Dl afford us the opportunity to test these models of DNA binding by constructing dose/response relationships between the correlated or uncorrelated concentrations of Dl and the free concentration. Surprisingly, we found that the binding of Dl to DNA (correlated concentration) is independent of the free concentration (blue data in Fig. 1F), rather than an increasing function of free concentration (as would be expected). This independence of the free Dl concentration suggests that a simple binding affinity model fails to predict DNA binding, and thus, transcriptional response. In contrast, the uncorrelated pool (blue data in Fig. 1G) could possibly be explained with a dose/response relationship between the uncorrelated concentration and the free concentration. To illustrate this, we fit a 4th-order Hill function to the data (see black dashed curve in Fig. 1G; see also Methods). We noted that, specifically within nc 14, the uncorrelated concentration does not change with variations in the free concentration. This apparent lack of relationship within nc 14 could result from saturation of the uncorrelated pool.

#### Reduced levels of DI result in an altered dose/response relationship

To further test the apparent dose/response relationship, approximated by a Hill function, we asked whether a reduction in Dl levels would result data points that fall on the same dose/response curve. To address this question, we used a Dl-mNG line that has the same protein sequence but has a 50% reduction in Dl levels (referred to as "reduced"; data not shown). Using this reduced fly line is advantageous because we are able to reduce the nuclear Dl levels across the whole time course (magenta in Fig. 1B) while holding the DV location, and hence the Toll-receptor associated signaling (43-45), constant.

As the only difference between the wildtype embryos and the reduced embryos is the total Dl concentration (and not Dl protein sequence or DV location), the base expectation would be a

reduction in the concentrations of all pools of Dl. While the two slowly-moving populations follow expectations and are generally decreased in the reduced line (magenta data in Fig. 1D, E), the free concentration in the reduced Dl line is, surprisingly, almost identical to that of the wt line (compare magenta and blue curves in Fig. 1C).

If the uncorrelated concentration is truly a simple function of the free concentration, we would expect the data points for both fly lines to fall on the same dose/response curve (uncorrelated concentration vs free concentration). However, we found the data points of the "reduced" embryos were shifted with respect to those from the wt embryos (magenta data in Fig. 1G) rather than falling on the same dose/response curve. To illustrate this, we fit another 4th-order Hill function to the data from the reduced line (see magenta dashed curve in Fig. 1G). As such, a simple dose/response relationship, such as a Hill function dependent only on free Dl concentration, fails to explain the observed behavior, which implies that both sub-populations of slowly-moving Dl are not primarily functions of the available (free) Dl concentration.

Given that the free concentration is the same between the two lines, even though the total concentration differs, our data suggest that there might exist a critical threshold of total nuclear Dl concentration (~30 nM). Below this threshold, all Dl would be freely diffusing, while any Dl concentration above this threshold becomes allocated to the slowly-moving pools (DNA-bound and/or mobile clusters), keeping the free Dl pool unchanged. Such a threshold could be the result of a highly ultrasensitive response to the total Dl concentration (not the free Dl concentration), calling into question whether models of affinity-based binding are appropriate for Dl. Another explanation of this threshold-based behavior would be an LLPS-driven process.

# Free concentration remains unchanged while other populations are dictated by bulk concentration along the DI gradient

To test whether the relationships we observed also hold across intermediate levels of Dl, and are not simply artifacts of measuring in two different fly lines, we leveraged the fact that the nuclear concentration of Dl varies with DV location. The natural gradient of Dl concentration, which is most intense at the ventral midline (DV coordinate of 0) and decays to near background levels roughly halfway around the embryo (DV coordinate of 0.5; Fig. 2A) (*14*, *19*), allows us to

measure the effect of continuous variations in total nuclear Dl concentration within the same fly line. Therefore, we imaged different locations along the DV axis during mid-to-late nc 14, when the concentrations of the pools of Dl are each roughly at steady state (Fig. 1B-E; see Methods). We grouped embryos by their DV coordinate, with at least two embryos per bin (width of bin = 0.05).

As expected, we found that the total nuclear concentration of Dl decreases as the DV location moves away from the ventral midline (Fig. 2B). The total Dl concentration peaks at the ventral midline (~120 nM) and drops to below 50 nM on the lateral side (DV = 0.45). Interestingly, we noted that the free concentration of Dl has nearly the same value of ~ 30 nM in all the DV locations (Fig. 2C), and that this concentration roughly agrees with the free concentrations we observed for the "wt" and "reduced" lines in the RICS time course (Fig. 1C).

The fraction of both pools of Dl composing the immobile fraction: the correlated fraction representing the DNA-bound pool determined by ccRICS (Fig. S2A) and the uncorrelated fraction representing the pool immobilized likely due to the formation of clusters (Fig. S2B) showed slight reduction with distance form ventral midline.

We found that the concentrations of both slowly-moving pools of DI decrease with increasing distance from the ventral midline, due to the decrease in the total nuclear concentration of DI (Fig. 2D and E). This is expected, as a higher concentration of bound DI is needed on the ventral side for driving the expression of target genes. Taken together with the nearly-constant concentration of free DI, we found no clear relationship between the concentrations of the slowly-moving pools of DI and the concentration of free DI (Fig. S2C,D). In contrast, the relationship between the concentrations of these pools of DI and the concentration of total DI is almost linear (Fig. 2F, G). Furthermore, a small extrapolation of these linear relationships reveals an x-intercept of ~30 nM of the total DI concentration. Thus, our *in vivo* measurements along the DV axis further support that there exists a threshold of total DI of ~30 nM that must be reached to facilitate the formation of the slowly-moving populations of DI (either by DNA binding or cluster formation), independent of the DV location. However, we cannot rule out that our

observed relationship between total Dl and the slowly-moving pools may be in the linear regime of a highly sigmoidal function.

# Threshold concentrations in DI possibly indicate thermodynamic phase separation

The presence of a threshold concentration mentioned above is a phenomenon observed in LLPS (46). While this is a controversial topic, we asked whether any other hallmarks of LLPS exist in the Dl system. First, we noted that, in a pure-species system that undergoes phase separation, once a threshold concentration is reached, any increase in the total concentration changes relative volumes occupied by the phases, but not their concentrations (46). While this is only strictly true of a pure species system, it may be a reasonable approximation for the Dl system if other components do not strongly affect putative phase separation. If so, then the free population (representing the dispersed phase) and the slowly-moving ones (both representing condensed phases) should each have constant concentrations. We observed a constant concentration of the dispersed phase, consistent with LLPS (Figs. 1C, 2C). On the other hand, our measurements of the condensed phase concentrations have been averaged over the whole nuclear volume. Under the assumption of LLPS, these averaged concentrations would be equal to the product of the actual (fixed) concentrations and the volume fraction of the respective condensed phases (see Methods). Furthermore, the volumes occupied by these phases, and thus our measured concentrations, would be directly proportional to the total nuclear concentration of Dl above the threshold of 30 nM, consistent with our observations (Fig. 2F, G; see Methods).

Second, weak multivalent protein-protein interaction mediated by intrinsically disordered regions (IDRs) or low-complexity domains (LCDs) lead to LLPS (5, 46-49). We thereby used predictive algorithms such as PSPredictor (50), PhaSePred (51), FuzDrop (52), and PSPHunter (53) to estimate the propensity of LLPS in case of Dl. All the algorithms returned a high score indicating that Dl is likely to phase separate (see Table S1), as opposed to mNG, the fluorescent protein tagged to it, which shows very low propensity to phase separate. While the putative threshold and the IDRs intriguingly point to LLPS of Dl, showing that a system in fact phase separates is difficult, and further *in vivo* measurements to infer the material properties of the clusters and the mechanism of formation (46, 49) are needed before a definitive case can be made.

#### Single particle tracking reveals slowly-moving clusters of Dl

The results so far point to the presence of two pools of Dl with different diffusivities: fast (possibly free DI) and slow (possibly clusters of Dl and/or DNA-bound Dl, corresponding to the uncorrelated and correlated populations, respectively) (Fig. 1A). Any clusters of Dl, either those known to be bound to DNA (13, 23) or the putative clusters in the uncorrelated population, would be brighter than the single molecules of freely disusing Dl. Therefore, we performed single particle tracking using super-resolution imaging to identify and characterize the mobility of the clusters of Dl. We used the Zeiss Airyscan detector that allowed the fast imaging with an improved signal-to-noise ratio (SNR) compared to traditional confocal acquisitions, which is needed for particle tracking experiments (54). We imaged the ventral-most nuclei in the embryos at mid nc 14 with a rapid frame rate (90 ms per frame). In the average image over the whole time series, bright puncta were clearly visible for Dl-mNG, whereas no such puncta were observed in control embryos expressing NLS-eGFP (Fig. 3A). The Fiji Mosaic plugin was used to detect and track the particles (55) (see Methods), and we found clusters of Dl that lasted for several frames (Fig. 3A and Movie S2). We found that the total number of particles detected during the acquisition time decreased with increasing distance from the ventral nuclei (Fig. 3B). This suggests that more clusters are formed at high levels of Dl. A trend similar to that observed in case of the correlated and uncorrelated populations using RICS (Fig. 2D, E).

To analyze the movement of these particles, we used a step size distribution (SSD) approach, in which the step sizes of particles between two consecutive frames are recorded. The shape of the distribution of these step sizes can be analyzed to determine not only the diffusivities of the particles, but also whether multiple sub-populations with different diffusivities are present (*56*). Following this approach, we fitted a two-component Rayleigh distribution (see Methods) to the distribution of the step sizes (*33*) and found two populations with diffusion coefficients of 0.1  $\mu$ m<sup>2</sup>/s and 0.01  $\mu$ m<sup>2</sup>/s (27% and 73% of the total clusters respectively; Fig. 3C). These diffusivities are very similar to those recently reported for DNA-bound hubs and for mobile (slowly-moving) clusters, respectively, for the early *Drosophila* transcription factors Bcd and Cic (*33*). As such, we interpret the slower population with 0.01  $\mu$ m<sup>2</sup>/s diffusivity to be a DNA-bound population of DI and the relatively faster population with 0.1  $\mu$ m<sup>2</sup>/s diffusivity to be slowly-moving clusters of DI. To determine whether the two-component model is the best fit, we also

fitted one- and three-component distribution models to our data. The one-component distribution model could not explain the data (see Fig. S3A), and, in the three-component model, the first two components match those of the two-component model, while the third component is optimized to be zero percent of the total population (Fig. S3A). Notably, the population of free Dl (D ~ 3-4  $\mu$ m<sup>2</sup>/s (29)) is not detected in this experiment because free Dl would not form bright clusters and because the frame time is too long to detect a free diffuser. This suggests that, at a long enough frame time, slowly-moving clusters will be also blurred into the background and only the DNA-bound population will be detected (*10*, *57*). Indeed, when we used a longer frame time of 500 ms, we could only detect the population of particles bound to DNA (Fig. 3C).

Mean-squared displacement (MSD) (58, 59) analysis, although only applicable for the determination of diffusivity when a single population is expected, can still provide qualitative information about the type of motion of the particles. The slope in the MSD vs.  $\Delta t$  plot and slope of the MSD/ $\Delta t$  vs.  $\Delta t$  plot, where  $\Delta t$  is the lag time, provide insights into the diffusivity regime of the particles. For the 500 ms frame time, we found the slope to be less than 1 in the MSD vs.  $\Delta t$  plot and negative in the MSD/ $\Delta t$  vs.  $\Delta t$  plot (Fig. S3B), indicating the presence of the population in the sub-diffusion regime, which results from confined diffusion, consistent with movement of particles attached to the DNA (10, 55, 60, 61). Taken together, these results suggest that there are at least two separate populations of clusters of DI: DNA-bound and the slowly diffusing (Fig. 3D).

#### Modified RICS imaging parameters simultaneously detect three populations

Our RICS experiments described in a previous section showed there is a free population of DlmNG and a slowly-moving population (29). Using SPT, we found that the slowly-moving population is composed of at least two sub-populations: one bound to the DNA and another of slowly-moving Dl clusters with a diffusivity of roughly  $D = 0.1 \ \mu m^2/s$ . As such, the SPT and RICS analyses together have detected a total of at least three different populations on the basis of diffusivity: freely diffusing Dl, slowly-moving clusters, and DNA-bound Dl. However, the individual experiments could not detect all three populations together, as the free population detected in RICS moves too quickly to be detectable with SPT, while the slowly-moving subpopulations cannot be readily distinguished in our previous RICS experiments. To simultaneously detect all three populations (Fig. 4A), and, in particular, to test the hypothesis that there is a slowly-moving population of Dl, distinguishable from truly immobile Dl, we used a longer line time for the RICS acquisitions. The longer line time would give the slowly-moving clusters more time to diffuse to a new location during the raster scan, thus allowing us to detect them, while still allowing us to detect the free Dl. Therefore, to alter the line time while holding as many other imaging parameters (such as pixel dwell time, image size, and frame time) as constant as possible, we imaged 4096 x 1024 pixel areas (133 µm x 33 µm) along the ventral midline of nine nc 14 embryos multiple times under two different Cases (see Methods and Movie S3). In Case 1, we oriented the image acquisition vertically, such that the line time was  $\tau_{\ell} \sim 5$  ms (Fig. 4B), consistent with the line time in our RICS acquisitions described in a previous section. In Case 2, we flipped the orientation of the image acquisition, resulting in a line time of  $\tau_{\ell} \sim 17$  ms (Fig. 4C), which should be able to detect particles with a diffusivity as low as roughly 0.1 µm<sup>2</sup>/s.

To determine the expected behavior of the system if slowly-moving clusters are present, we simulated a two-component system (with a freely diffusing population and a truly immobile population, but no clusters), and a three-component system in which slowly-moving clusters are present (see Fig. S4A,B). Our simulations showed that, if slowly-moving clusters are not present, there would be very little difference in the slow direction cut of the ACF between Cases 1 and 2 (Fig. 4D, "Two component system"; see also Methods), only a small deviation at low pixel shifts (see red arrow in Fig. 4D). In addition, the large-pixel-shift tails of the ACF should not extend past the PSF (see gray arrow in Fig. 4D). On the other hand, if there is also a third population of slowly diffusing clusters of Dl, there should be a clear difference in the slow direction cut of the ACF between Cases 1 and 2 (compare blue curve to the red curve in Fig. 4E, "Three component system"). Furthermore, the ACFs of both Cases should extend beyond the tail of the PSF at large pixel shifts (black arrow in Fig. 4E). Our experimental results show that the slow direction of the ACF is clearly different between Cases 1 and 2, and both extend past the PSF, suggesting there are indeed three populations with different diffusivities present (Fig. 4F).

To quantitatively test our hypothesis that there is a slowly-moving population of Dl, we fit twoand three-component models to our data. If we fit these models to the simulations of a twocomponent system, the fitted parameter of the two-component model (the fraction slowlymoving) would not be statistically distinct between Cases 1 and 2. However, the fitted parameter in the three-component model (the fraction of mobile clusters) would be statistically distinct between Cases 1 and 2 (Fig. 4G), because the three-component model would not match the underlying two-component system. In contrast, if we fit these models to the simulations of a three-component system, the opposite would be true (Fig. 4H). In other words, if the model matches the system, fitted parameters should match between Cases 1 and 2, while if the model is a mismatch to the system, the fitted parameters should also be mismatched between Cases 1 and 2. When we fit the two models to our experimental data, we found a clear statistical difference between Cases 1 and 2 for the two-component model fit, but no statistical difference between Cases 1 and 2 for the three-component model fit (Fig. 4I). Therefore, the combination of RICS analyses for Cases 1 and 2 clearly allow for simultaneous detection of all three populations of Dl, including the slowly-moving clusters, which might be formed due to LLPS (Fig. 4A). Furthermore, our three-component model fits suggest that these slowly-moving clusters have a weighting of  $\phi_1 = 0.6$ . These results further verify our observations for the spatiotemporal variation using RICS and the SPT results.

#### Discussion

Despite decades of research, the molecular mechanisms behind the precise gene regulation by sequence-specific transcription factors, particularly mechanisms linking nuclear concentration of the transcription factor to its DNA binding (a crucial step in gene regulation), remain unresolved (22, 62). Transcriptional regulation is dependent on TF dynamics such as the total concentration, fraction bound, and hub formation (5). Therefore, in this study, we investigated how the bulk concentration of Dl relates to DNA binding and hub formation. Using experiments that span multiple length and time scales, we detected at least three different populations of Dl having distinct diffusivities present in the nuclei: freely diffusing (D ~  $3 \mu m^2$ ), previously unobserved slowly-moving clusters (D ~  $0.1 \mu m^2$ ), and a DNA-bound population (D ~  $0.01 \mu m^2$ ; see Fig. 5).

Recent work has suggested that clusters, or hubs, of transcription factors are significant in DNA binding. For example, DNA-bound hubs of the morphogen Bcd have been proposed recently to facilitate the transcriptional activity by locally increasing the concentration (10, 11) or by acting as fast sensor of concentration (63). Similarly, enrichment of puncta of Dl near individual transcription sites has been observed (13, 23). A recent study has shown that Bcd and Capicua (Cic), two TFs active in the early *Drosophila* embryo, when Dl is also active, form slowly-moving aggregates (D ~  $0.1 \mu m^2/s$ ), which might provide an efficient search mechanism for the target sites (33). In this work, we found a previously-unobserved population of slowly-moving clusters of Dl, suggesting that this might be a common mechanism of transcriptional regulation by morphogen gradients. Thus, gene regulation models in eukaryotes must take the formation of DNA-bound as well as slowly-moving hubs of TFs into account.

In addition to identifying three sub-populations of Dl, we also leveraged the variation of Dl concentration across nc 11-14, along the DV axis, and in two different fly lines to map the relationships among them. Strikingly, we found the free concentration of Dl remained roughly the same ( $\sim$ 30 nM), regardless of the bulk concentration, indicating the possibility of a threshold concentration of total Dl, above which further increases in Dl became exclusively allocated to the slowly-moving populations. Further supporting the presence of a threshold in bulk concentration, the dose/response maps connecting the two slowly-moving populations to the total Dl concentration were linear, extrapolating back to this same 30 nM threshold in total Dl concentration. Thus, our nucleus-level measurements, while not as specific as those that focus on individual enhancers (13), have provided crucial details regarding the mechanisms of DNA binding and cluster formation.

The presence of a threshold in the total nuclear concentration of Dl, above which no further free Dl is allocated, is a hallmark LLPS of biomolecular condensates (64). While the clusters of Dl forming from LLPS is an intriguing possibility, other work suggests the opposite, and overall, LLPS of transcription factors has been a hotly-debated topic in the past decade (48, 65–67). LLPS has been proposed to play a role in buffering biological fluctuations, aiding in robustness of certain pathways, or repressing pathways whose functioning requires noise (48, 68). However, quantitative methods to provide definitive evidence of LLPS are still lacking in the field

(reviewed in (49)). Therefore, regardless of the mechanistic status of Dl cluster formation, our *in vivo* quantitative measurements and methodologies provide a general, quantitative framework for testing properties of clusters for LLPS-like behavior by the detecting relationships among the constituent populations *in vivo*.

While may subpopulations detected in one experiment are similar to those detected in other experiments, we acknowledge these similar subpopulations may not necessarily be identical across experiments. For example, the correlated and uncorrelated populations from RICS/ccRICS experiments may not exactly map one-to-one onto the DNA-bound clusters (0.01  $\mu$ m<sup>2</sup>/s) and slowly-moving clusters (0.1  $\mu$ m<sup>2</sup>/s) detected in the SPT experiments, as some of the correlated RICS population could be DNA-bound monomers or dimers of Dl. Additionally, we acknowledge the possibility that there may be populations having intermediate composition and diffusivity not identified in our experiments. One further caveat is that clusters of fluorescently-labeled molecules, by definition, have a higher molecular brightness than the monomers and dimers, which affect the quantitative values of the RICS analyses(27, 35, 69). In particular, corrections for differing brightness values would result in a slight increase to the total concentration of Dl, and the concentration of the other subpopulations would also scale accordingly. Even so, our general findings, including the discovery of mobile clusters of Dl and the linear/threshold relationship of slowly-moving populations to the total Dl concentration, are not affected by these caveats.

The existence of clusters, both DNA-bound and mobile, have profound implications for our understanding of gene regulation. Put within the broader context, these clusters must be accounted for along with other crucial parameters that explain binding of a TF, such as the chromatin accessibility and cooperativity with pioneer-like factors. All together, our results support a model in which, during the early ncs, owing to its lower concentration, Dl binds only the accessible sites created by pioneer factors. During nc 14, when the nuclear concentration of Dl is high, binding is driven by the concentration of Dl, perhaps by clusters of Dl. This is in agreement with the proposed cooperativity between the pioneer factor Zld and Dl and the sporadic/delayed expression of Dl target genes in Zld mutant embryos (40, 70–72). Our results also suggest that, during nc 14, binding occurs when a threshold concentration of free Dl is

exceeded. As such, different populations of DI at varying concentrations within and across nuclei must be taken into account to understand its concentration-dependent transcriptional interpretation. More broadly, we anticipate that similar detailed measurements of nuclear subpopulations will be needed for other transcription factor systems, potentially pointing to generalizable rules of transcriptional regulation (*10*, *11*, *13*, *33*, *35*, *63*, *73*).

## **Materials and Methods**

## Drosophila strains

The fly strain used for all the imaging for Dl is yw/w; Dorsal-mNeonGreen /cyo; +. For the reduced level of Dl we used, yw/w; Dorsal-mNeonGreen-dsRed/cyo; +. The other fly strains used are: w[\*];  $P\{w[+mC]=His2Av-mRFP1\}II.2$  (Bloomington Drosophila Stock Center (BDSC) #23651) for nuclear segmentation and ccRICS using His2Av-RFP, w[\*];  $P\{w[+mC]=His2Av-EGFP.C\}2/SM6a$  (BDSC #24163) for using His2Av-eGFP as photobleaching control in the particle tracking experiments.

# Drosophila husbandry and sample preparation

The fly stocks were grown on standard cornmeal-molasses-yeast medium at 25°C. Fly cages were set with males and females of desired fly lines two days before imaging and kept at room temperature. Grape juice plates were streaked with yeast paste and placed on the bottoms of the cages for egg laying. The plates were changed once every day. On the day of imaging, the grape juice plates were placed for oviposition for an hour after which it was removed to collect the embryos. The embryos were washed from the plate into a mesh basket using deionized water. They were then dechorionated in bleach for 30 s and washed again with deionized water to remove residual bleach. The embryos were mounted in 1% solution of low melting point agarose (IBI Scientific, IB70051) in deionized water on a glass bottom Petri dish (MatTek, P35G-1.5-20-C). The mounting was done in such a way that the ventral surface of the embryo touched the coverslip in the Petri dish for imaging the ventral nuclei. For imaging other DV positions, the angle at which the embryo touched the coverslip was changed by rotating it slightly with a hair loop. The agarose was let to solidify after which deionized water was poured in the dish to avoid drying of the agarose covering the samples.

## **Confocal imaging**

# **Spatial and temporal variation RICS**

All the imaging was performed using LSM 900 (Carl Zeiss, Germany) confocal laser scanning microscope. For the confocal imaging C-Apochromat 40x/1.2 water immersion Korr objective, 488 nm laser for the mNeonGreen, 561 nm laser for RFP, and GaAsP-PMT detector were used. Embryo stage was estimated based on the His2Av-RFP signal. Image acquisition was started when the embryos reached the desired stage: nc 10 for the temporal variation study and mid nc 14 for the spatial variation study. A frame size of  $1024 \times 1024$  pixels at 5x zoom was used. This corresponded to a pixel size of 31.95 nm and a pixel dwell time of 2.06 µs at a frame time of 5.06 s. The acquisition was continued until the embryo gastrulated as indicated by the nuclear morphology.

To minimize photobleaching, the ROI was moved to a different part of the embryo after every 12 frames ensuring that all the locations are aligned in the same DV coordinate. RICS analysis on the 12 frame acquisitions of time series enabled the precise determination of the temporal dynamics of the total nuclear concentration and binding of Dl.

# **RICS** analysis and fitting of ACF models

We performed RICS analysis according to previously-published protocols (29, 35). In brief, two dimensional (2D) RICS autocorrelation functions of the nuclear fraction of each frame were built using a fast Fourier transform protocol in Matlab, and these ACFs were averaged together for all frames in a given group. The result was a time series of 2D ACFs, each of which corresponded to a given grouping of 7-12 frames. Background subtraction was performed on the fly by examining the histogram of intensities and fitting a Gaussian to the lowest intensity pixels in the image.

The theoretical 2D RICS ACF,  $G(\Delta x, \Delta y)$ , is the following:

$$G(\Delta x, \Delta y) = A\left(\frac{1}{1 + \frac{4D(\tau_p \Delta x + \tau_\ell \Delta y)}{w_0^2}}\right) \sqrt{\frac{1}{1 + \frac{4D(\tau_p \Delta x + \tau_\ell \Delta y)}{w_z^2}}} \exp\left(-\frac{(\Delta x^2 + \Delta y^2)\Delta r^2}{w_0^2 + 4D(\tau_p \Delta x + \tau_\ell \Delta y)}\right), \quad (1)$$

where A is the amplitude; D is the diffusivity;  $w_0$  and  $w_z$  are the radii of the point spread function (PSF) in the xy plane and the axial (z) direction, respectively;  $\Delta r$  is the xy size of a pixel;  $\tau_p$  and  $\tau_\ell$  are the pixel dwell time (determined by the scan speed) and line time (determined by a combination of the scan speed and number of pixels in the width of the image), respectively; and  $\Delta x$  and  $\Delta y$  are the pixel shifts in the fast and slow directions, respectively.

In practice,  $A = \gamma/(V_{PSF}\bar{c})$ , where  $\gamma = \sqrt{2}/4$  is a factor that accounts for the uneven illumination airy unit,  $\bar{c}$  is the average concentration in the confocal volume, and  $V_{PSF} = \pi^{\frac{3}{2}} w_0^2 w_z$  is the effective volume of the point spread function.

The fast direction of the data-derived 2D ACF, at each time point, was used to fit the fast direction cut of the theoretical 2D ACF,  $G(\Delta x, 0)$ , which reduces to a Gaussian equation that approximates the microscope's PSF:

$$G(\Delta x, 0) \approx (A - B) \exp\left(-\frac{\Delta x^2 \Delta r^2}{w_0^2}\right) + B,$$
 (2)

To improve the robustness of the fit, a small, adjustable background constant, B, was added and  $w_0$  was allowed to vary slightly. The parameter B was constrained to have a magnitude less than  $10^{-3}$ . This first fitting step gave us an estimate of the ACF amplitude, A, and hence, the average total concentration.

After obtaining an estimate of *A*, we held it fixed and used the entire 2D ACF to fit  $G_{2c}(\Delta x, \Delta y)$ , which is a linear combination of two components, a freely diffusing population and a slowly-moving population (29):

$$G_{2c}(\Delta x, \Delta y) = (A - B)[\phi G_0(\Delta x, \Delta y) + (1 - \phi)G_1(\Delta x, \Delta y)] + B, \qquad (3)$$

where  $G_0(\Delta x, \Delta y)$  is given by Equation 1 with A = 1 and D = 0,  $G_1(\Delta x, \Delta y)$  is given by Equation 1 with A = 1 and D non-zero, and the linear combination weight,  $\phi$ , is the fraction slowly-moving. Note that the parameter B in Equation 3 may have a different value from the one found in Equation 2. To avoid overfitting, we assumed  $D = 3 \ \mu m^2/s$ .

#### Fitting ccRICS cross-correlation functions (CCFs) to estimate correlated binding

We performed analysis of ccRICS according to previously published work (29, 35). In brief, the 2D cross correlation function (CCF) between Dl-mNG and His2Av-RFP was computed through a fast Fourier transform protocol in Matlab. This 2D CCF was then used to fit a zero-diffusion 2D model of cross correlation (29):

$$G_{cc}(\Delta x, \Delta y) = (A_{cc} - B) \exp\left(-\frac{(\Delta x \Delta r - d_x)^2 + (\Delta y \Delta r - d_y)^2}{\widehat{w}_0^2} - \frac{d_z^2}{\widehat{w}_z^2}\right) + B, \qquad (4)$$

where  $d_x$ ,  $d_y$ ,  $d_z$  represent the displacements between the centers of the two PSFs in the *x*, *y*, and *z* directions, respectively;  $\hat{w}_0$  and  $\hat{w}_z$  are the average PSF sizes for the planar (xy) and axial (z) directions, respectively;  $A_{cc}$  is the amplitude of the CCF;  $\Delta r$  is the pixel size (see Equation 1); and *B* is a background parameter (see Equations 2 and 3). The axial displacement was previously estimated by imaging fluorescent beads (35) and the factor  $\exp(d_z^2/\hat{w}_z^2)$  was found to be roughly equal to 1.02, implying the factor could be safely ignored. The average PSF sizes are defined as  $\hat{w}_0^2 = 0.5(w_{0,g}^2 + w_{0,r}^2)$  and  $\hat{w}_z^2 = 0.5(w_{z,g}^2 + w_{z,r}^2)$ , and the subscripts *g* and *r* denote the green and red channels, respectively.

The fraction of Dl-mNG correlated to His2Av-RFP (the correlated fraction) is related to the ratio of  $A_{cc}$  and the ACF amplitude of the red channel,  $A_{red}$ :

correlated fraction = 
$$\frac{\widehat{w}_0^2 \widehat{w}_z}{w_{0,r}^2 w_{z,r}} \frac{A_{cc}}{A_{red}}$$
. (5)

# Relationship between total concentration and condensed phase

Under the assumption of LLPS, to derive the relationship between free Dl (putative dispersed phase) and the correlated and uncorrelated pools of the slowly-moving population (putative condensed phases 1 and 2, respectively), we used the pure species behavior as the foundation. In this case, all phases in equilibrium would have a fixed concentration with respect to varying total concentration, and variations in total concentration would change the volume partitioning among the two phases. We define  $\bar{C}_0$ ,  $\bar{C}_1$ , and  $\bar{C}_2$  as the fixed concentrations of the dispersed phase and the two condensed phases, respectively. We further define  $V_0$ ,  $V_1$ , and  $V_2$  as the volumes occupied by the dispersed phase and the two condensed phases, respectively. We further define  $V_0$ ,  $V_1$ , and  $V_2$  as the volumes occupied by the dispersed phase and the two condensed phases, respectively. We further define  $V_0$ ,  $V_1$ , and  $V_2$  as the volumes occupied by the dispersed phase and the two condensed phases, respectively. We further define  $V_0$ ,  $V_1$ , and  $V_2$  as the volumes occupied by the dispersed phase and the two condensed phases, respectively, and note that the total volume of the nucleoplasm available to Dl is  $V_{tot} = V_0 + V_1 + V_2$ .

As our measurements of the free, correlated, and uncorrelated concentrations ( $C_{meas,0}$ ,  $C_{meas,1}$ , and  $C_{meas,2}$ , respectively) are averages over the entire volume,  $V_{tot}$ , they relate to their respective fixed concentrations by:

$$C_{meas,0} = \bar{C}_0 V_0 / V_{tot}, \quad (6)$$
  

$$C_{meas,1} = \bar{C}_1 V_1 / V_{tot}, \quad (7)$$
  

$$C_{meas,2} = \bar{C}_2 V_2 / V_{tot}, \quad (8)$$

respectively. If  $V_1 + V_2 \ll V_0$ , then Equation (6) reduces to  $C_{meas,0} = \overline{C}_0$ , which explains how our measurements of free Dl remain roughly fixed (at 30 nM) with varying total Dl concentration.

The total concentration relates to the phase volumes by:

$$C_{tot} = C_{meas,0} + C_{meas,1} + C_{meas,2}$$
$$= \frac{\bar{C}_0 V_0 + \bar{C}_1 V_1 + \bar{C}_2 V_2}{V_{tot}}, \qquad (9)$$

where, in the second line, we have used Equations 6-8. Under the continued assumption that  $V_1 + V_2 \ll V_0$ , Equation 9 can be reduced to:

$$C_{tot} \approx \bar{C}_0 + \frac{\bar{C}_1 V_1}{V_{tot}} + \frac{\bar{C}_2 V_2}{V_{tot}},$$
 (10)

As the three phase concentrations and  $V_{tot}$  are constant, Equation 10 implies that  $V_1$  and  $V_2$ , and thus also  $C_{meas,1}$  and  $C_{meas,2}$ , are in a direct linear relationship with the difference between the total concentration,  $C_{tot}$ , and the concentration threshold,  $\overline{C}_0$ .

#### Varying line time experiments

In the experiments with varied line times, the pixel size and pixel dwell time were kept the same as the previous experiments at 31.95 nm and 2.06  $\mu$ s respectively. The line time was varied by changing the frame size. For the shorter line time (Case 1), the frame size was 1024 × 4096 pixels, resulting in a line time of 5 ms. For the same embryo, the longer line time (Case 2) was obtained by rotating the ROI by 90° (exchanging the x- and y-directions of the image), resulting in a frame size of 4096 × 1024 pixels and a line time of 17 ms. This ensured that the imaging parameters will be conserved and each embryo will be imaged with two different line times, the

short and the long one, at the exact same DV position. Thus, with the same number of pixels, pixel size, and pixel dwell time the images with frame size of 4096 x 1024 pixels have a line time just under 4 times longer than that of the images with 1024 x 4096 for the same DV location. To minimize and account for photobleaching, the ROI was rotated after 12 frames and for some of the embryos the longer line time experiments were performed first and then the shorter line time ones whereas for the other embryos the order was reversed.

#### Simulations of varying line time experiments

We simulated the ACFs of both Cases (the short line time and long line time) for two different systems. The first system was a two component system with freely diffusing Dl-mNG and DNA-bound Dl-mNG (zero diffusion), simulated using Equation 3. We set the baseline diffusivity of free Dl to be 3  $\mu$ m<sup>2</sup>/s and the baseline slowly-moving fraction ( $\phi$ ) to be 0.75.

The second system was a three component system with the extra component being a slowlymoving cluster of Dl-mNG. In this system, we assumed the slowly-moving fraction, set at 0.75 for the two-component system, was split into a baseline of  $\phi_0 = 0.25$  for the fraction bound to DNA and a baseline of  $\phi_1 = 0.50$  for the fraction of slowly-moving clusters, consistent with our experimental measurements of the correlated and uncorrelated fractions (see Fig. S1):

 $G_{3c}(\Delta x, \Delta y) = (A - B)[\phi_0 G_0(\Delta x, \Delta y) + \phi_1 G_1(\Delta x, \Delta y) + (1 - \phi_0 - \phi_1) G_2(\Delta x, \Delta y)], \quad (11)$ where  $G_0, G_1, G_2$  are each given by Equation 1 with A = 1 and with baseline diffusivities of

zero, 0.1  $\mu$ m<sup>2</sup>/s, and 3  $\mu$ m<sup>2</sup>/s, respectively.

We ran 20 simulations of each system with 10% random noise added to the parameters. After the simulated ACFs were constructed according to Equation 3 (two-component system) or Equation 11 (three-component system), we added 2% Gaussian noise to the ACFs simulate experimental error. The smoothness of the resulting ACFs were comparable by-eye to the smoothness of our typical experimental ACFs. The means of the 20 simulations are found in Fig. 4C and D, with error bars being the standard deviation.

We fit each of these simulations to either a two-component model (Equation 3) or a threecomponent model (Equation 11). To avoid overfitting, we allowed only one fitted parameter in each case:  $\phi$  for the two-component model and  $\phi_1$  for the three-component model. In both cases, we fixed the diffusivity of free Dl-mNG to be 3  $\mu$ m<sup>2</sup>/s (29). For the three-component model, we additionally fixed  $\phi_0$  to be 0.25 and the diffusivity of slowly-moving clusters to be 0.1  $\mu$ m<sup>2</sup>/s.

#### RICS analysis of varying line time experiments and fitting of ACF models

The RICS analysis of our time course data for Cases 1 and 2 in the varying line time experiments was performed as described above. Once the data-derived ACFs were computed, they were used to fit a two-component model (Equation 3) and a three-component model (Equation 11). As described above, to avoid overfitting, we held fixed all parameters except for  $\phi$  for the two-component model and  $\phi_1$  for the three-component model, and we fixed the diffusivity of free Dl-mNG to be 3  $\mu$ m<sup>2</sup>/s and the diffusivity of slowly-moving clusters to be 0.1  $\mu$ m<sup>2</sup>/s. We fixed  $\phi_0$  to be equal to the value of correlated fraction for that time course.

#### Super resolution imaging

The particle tracking experiments require rapid frame rates and high spatial resolution, and as such, we used the fast Airyscan detector for improved spatial resolution and signal-to-noise ratio (SNR) (54). All the images were taken on the ventral side of the embryos at mid nc 14. Plan-Apochromat 63x/1.4 Oil DIC M27 objective and 488 nm excitation laser for the Dl-mNeonGreen were used. All the images have a pixel size of 42 nm. For the step size distribution experiments the rapid frame rate of 87 ms was obtained by having a frame size of  $167 \times 167$  pixels and the timeseries consisted of 35 frames. Whereas the longer duration of 500 ms was obtained by having a frame size of 50 frames to detect the DNA-bound clusters that last for a longer time.

## Particle detection and tracking

The Fiji Mosaic plugin was used for particle detection and tracking (55), using the following parameters: particle radius 6 pixels, cutoff score 0%, intensity percentile 0.22%. For particle linking, the maximum step length was set at 10 pixels and the link range 1 frame. Any trajectory lasting less than three steps was discarded.

#### Localization precision

To determine the localization precision, we imaged immobilized fluorescent beads (*33*, 74, 75). The beads (Fluoresbrite® YO Carboxylate Microspheres 0.20  $\mu$ m, Polysciences, Inc., Catalog No. 19391-10) were washed twice with deionized water and then mounted in low melting point agarose on a glass bottom Petri dish in the same way as the embryos to emulate the experimental conditions. The imaging was then performed using the same microscope parameters as those used for the embryos for the 87 ms frame rate. Since localization precision depends on SNR (76, 77) which in turn depends on laser power, we used 8 different laser power levels for the imaging: 0.025%, 0.05%, 0.075%, 0.1%, 0.2%, 0.3%, 0.5%, and 0.8%. For each laser power, 10 different time series having 35 frames each at 10 different locations were acquired. The mean diffusivity was calculated using the mean squared displacement (MSD) method (*33*, 74) instead of the step size distribution method since the beads are immobilized thus expected to have long-term trajectories and we do not expect multiple populations with different diffusivities. We calculated the localization precision for all the acquisitions from the diffusivity ( $\sigma^2 = 4D\Delta T$ ) obtained from the Fiji Mosaic plugin (*33*, *55*, *74*). The localization precision as a function of SNR was then empirically fit to the following equation (*55*, 77):

$$\sigma^2 = \frac{A}{SNR} + \frac{B}{SNR^2}, \qquad (12)$$

where the first term is due to background noise and the second term is due to shot noise. SNR is defined as the ratio of the mean intensity of the signal and standard deviation of the background. The constants A and B were determined by fitting Equation 12 to the data in the plot of  $\sigma^2$  vs SNR plot using lsqcurvefit solver in Matlab (Fig. S3C). The mean intensity of the signal in the embryos was determined at the coordinates of detected particles obtained from the Mosaic plugin and the standard deviation of the intensity outside the detected particles inside the nucleus was determined. The ratio gives the SNR for our samples. The localization precision was then calculated using Equation 12 at the determined SNR.

# **Step size distribution**

The possibility that multiple populations with different diffusivities might be present and the presence of trajectories exhibiting short-term motion led us to use the step size distribution for the estimation of diffusivities. The length of the steps taken between consecutive frames were calculated and the step size distribution of the particles was obtained. The probability density

function of the step sizes was fit with a two-component Rayleigh mixture model (10, 33, 58, 78):

$$P(r,\Delta t) = \sum_{i=1}^{k} f_i \frac{r}{2(D_i \Delta t + \sigma^2)} \exp\left(-\frac{r^2}{4(D_i \Delta t + \sigma^2)}\right), \quad (13)$$

Where  $f_i$  is the fraction of the population i (and thereby  $\sum_{i=1}^{k} f_i = 1$ ),  $D_i$  is the diffusivity of the population i,  $\Delta t$  is the time between the steps, the localization precision (from the immobilized fluorescent beads experiments) is  $\sigma^2 = 4.5 \times 10^{-4} \,\mu m^2$ , and r is the step size. The number of components was changed to 1 and 3 (Fig. S3A) to justify the use of two-component model. The values of diffusivities and fraction of the population in Equation 13 were estimated using the Matlab optimization function fmincon.

## References

- 1. F. Spitz, E. E. M. Furlong, Transcription factors: from enhancer binding to developmental control. *Nature reviews genetics* **13**, 613–626 (2012).
- 2. A. L. Todeschini, A. Georges, R. A. Veitia, Transcription factors: Specific DNA binding and specific gene regulation. *Trends in Genetics* **30**, 211–219 (2014).
- 3. K. J. Brennan, M. Weilert, S. Krueger, A. Pampari, H. yun 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. *Developmental Cell* **58**, 1898-1916.e9 (2023).
- 4. S. Inukai, K. H. Kock, M. L. Bulyk, Transcription factor–DNA binding: beyond binding site motifs. *Current Opinion in Genetics and Development* **43**, 110–119 (2017).
- 5. F. Lu, T. Lionnet, Transcription Factor Dynamics. Cold Spring Harbor Perspectives in Biology 13 (2021).
- 6. M. Stortz, D. M. Presman, V. Levi, Transcriptional condensates: a blessing or a curse for gene regulation? *Communications Biology* **7** (2024).
- 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).
- 8. 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).
- 9. K. Kawasaki, T. Fukaya, Functional coordination between transcription factor clustering and gene activity. *Molecular Cell* **83**, 1605-1622.e9 (2023).
- 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, e40497–e40497 (2018).
- 11. 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 & development* **31**, 1784–1794 (2017).
- 12. 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, e28975–e28975 (2017).

- 13. S. Fallacaro, A. Mukherjee, M. A. Turner, H. G. Garcia, M. Mir, Transcription factor hubs exhibit gene-specific properties that tune expression. Biophysics [Preprint] (2025). https://doi.org/10.1101/2025.04.07.647578.
- G. T. Reeves, N. Trisnadi, T. V. Truong, M. Nahmad, S. Katz, A. Stathopoulos, Dorsal-Ventral Gene Expression in the Drosophila Embryo Reflects the Dynamics and Precision of the Dorsal Nuclear Gradient. *Developmental Cell* 22, 544–557 (2012).
- 15. A. M. Huang, J. Rusch, M. Levine, An anteroposterior Dorsal gradient in theDrosophila embryo. *Genes & development* **11**, 1963–1973 (1997).
- G. T. Reeves, A. Stathopoulos, Graded dorsal and differential gene regulation in the Drosophila embryo. *Cold Spring Harbor perspectives in biology* 1, a000836–a000836 (2009).
- 17. S. Roth, D. Stein, C. Nüsslein-Volhard, A gradient of nuclear localization of the dorsal protein determines dorsoventral pattern in the Drosophila embryo. *Cell* **59**, 1189–1202 (1989).
- 18. M. Markstein, P. Markstein, V. Markstein, M. S. Levine, Genome-wide analysis of clustered Dorsal binding sites identifies putative target genes in the Drosophila embryo. *Proceedings of the National Academy of Sciences* **99**, 763–768 (2002).
- 19. L. M. Liberman, G. T. Reeves, A. Stathopoulos, Quantitative imaging of the Dorsal nuclear gradient reveals limitations to threshold-dependent patterning in Drosophila. *Proceedings of the National Academy of Sciences* **106**, 22317–22322 (2009).
- 20. R. DeLotto, Y. DeLotto, R. Steward, J. Lippincott-Schwartz, Nucleocytoplasmic shuttling mediates the dynamic maintenance of nuclear Dorsal levels during Drosophila embryogenesis. (2007).
- 21. A. E. Schloop, P. U. Bandodkar, G. T. Reeves, Formation, interpretation, and regulation of the Drosophila Dorsal/NF-κB gradient. *Current Topics in Developmental Biology* **137**, 143–191 (2020).
- 22. C. A. Rushlow, S. Y. Shvartsman, Temporal dynamics, spatial range, and transcriptional interpretation of the Dorsal morphogen gradient. *Current Opinion in Genetics and Development* 22, 542–546 (2012).
- 23. 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. *Current Biology* 29, 1387-1393.e5 (2019).
- 24. M. A. Digman, C. M. Brown, P. Sengupta, P. W. Wiseman, A. R. Horwitz, E. Gratton, Measuring fast dynamics in solutions and cells with a laser scanning microscope. *Biophysical Journal* **89**, 1317–1327 (2005).

- 25. M. A. Digman, P. Sengupta, P. W. Wiseman, C. M. Brown, A. R. Horwitz, E. Gratton, Fluctuation correlation spectroscopy with a laser-scanning microscope: Exploiting the hidden time structure. *Biophysical Journal* **88**, L33–L36 (2005).
- 26. M. A. Digman, E. Gratton, Analysis of diffusion and binding in cells using the RICS approach. *Microscopy research and technique* **72**, 323–332 (2009).
- 27. A. E. Schloop, S. V. Hiremath, R. Shaikh, C. M. Williams, G. T. Reeves, Spatiotemporal dynamics of NF-κB/Dorsal inhibitor IκBα/Cactus in Drosophila blastoderm embryos. *bioRxiv*, 2022–2024 (2024).
- 28. C. M. Brown, R. B. Dalal, B. Hebert, M. A. Digman, A. R. Horwitz, E. Gratton, Raster image correlation spectroscopy (RICS) for measuring fast protein dynamics and concentrations with a commercial laser scanning confocal microscope. *Journal of Microscopy* 229, 78–91 (2008).
- H. Al Asafen, N. M. Clark, E. Goyal, T. Jacobsen, S. Siddika Dima, H.-Y. Chen, R. Sozzani, G. T. Reeves, Dorsal/NF-κB exhibits a dorsal-to-ventral mobility gradient in the Drosophila embryo. doi: 10.1101/320754.
- 30. E. L. Elson, D. Magde, Fluorescence correlation spectroscopy. I. Conceptual basis and theory. *Biopolymers* 13, 1–27 (1974).
- 31. D. Magde, W. W. Webb, E. L. Elson, Fluorescence correlation spectroscopy. III. Uniform translation and laminar flow. *Biopolymers* **17**, 361–376 (1978).
- 32. O. Krichevsky, G. Bonnet, Fluorescence correlation spectroscopy: the technique and its applications. *Rep. Prog. Phys.* 65, 251–297 (2002).
- 33. L. Zhang, L. Hodgins, S. Sakib, A. Verbeem, A. Mahmood, C. Perez-Romero, R. A. Marmion, N. Dostatni, C. Fradin, Both the transcriptional activator, Bcd, and repressor, Cic, form small mobile oligomeric clusters. *Biophysical Journal* (2024).
- 34. S. N. Carrell, M. D. O'Connell, T. Jacobsen, A. E. Pomeroy, S. M. Hayes, G. T. Reeves, A facilitated diffusion mechanism establishes the Drosophila dorsal gradient. *Development* (*Cambridge*) **144**, 4450–4461 (2017).
- 35. S. S. Dima, G. T. Reeves, Bulk-level maps of pioneer factor binding dynamics during Drosophila maternal-to-zygotic transition. *bioRxiv*, 2024.09.06.611749-2024.09.06.611749 (2024).
- 36. K. Bacia, P. Schwille, Practical guidelines for dual-color fluorescence cross-correlation spectroscopy. *Nature protocols* **2**, 2842–2856 (2007).
- T. Weidemann, M. Wachsmuth, M. Tewes, K. Rippe, J. Langowski, Analysis of ligand binding by two-colour fluorescence cross-correlation spectroscopy. *Single Molecules* 3, 49– 61 (2002).

- 38. R. P. Zinzen, K. Senger, M. Levine, D. Papatsenko, Computational models for neurogenic gene expression in the Drosophila embryo. *Current Biology* **16**, 1358–1365 (2006).
- 39. D. Papatsenko, M. Levine, The Drosophila gap gene network is composed of two parallel toggle switches. *PLoS One* **6**, e21145–e21145 (2011).
- 40. J. S. Kanodia, H. L. Liang, Y. Kim, B. Lim, M. Zhan, H. Lu, C. A. Rushlow, S. Y. Shvartsman, Pattern formation by graded and uniform signals in the early Drosophila embryo. *Biophysical Journal* **102**, 427–433 (2012).
- 41. P. U. Bandodkar, H. Al Asafen, G. T. Reeves, Spatiotemporal control of gene expression boundaries using a feedforward loop. *Developmental Dynamics* **249**, 369–382 (2020).
- 42. A. A. Jermusyk, N. P. Murphy, G. T. Reeves, Analyzing negative feedback using a synthetic gene network expressed in the Drosophila melanogaster embryo. *BMC Syst Biol* **10**, 85 (2016).
- 43. A. M. Whalen, R. Steward, Dissociation of the dorsal-cactus complex and phosphorylation of the dorsal protein correlate with the nuclear localization of dorsal. *The Journal of cell biology* **123**, 523–534 (1993).
- 44. D. S. Schneider, K. L. Hudson, T. Y. Lin, K. V. Anderson, Dominant and recessive mutations define functional domains of Toll, a transmembrane protein required for dorsal-ventral polarity in the Drosophila embryo. *Genes Dev.* 5, 797–807 (1991).
- 45. D. Stein, S. Roth, E. Vogelsang, C. Nu<sup>°</sup>sslein-Volhard, The polarity of the dorsoventral axis in the drosophila embryo is defined by an extracellular signal. *Cell* **65**, 725–735 (1991).
- 46. S. Alberti, A. Gladfelter, T. Mittag, Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. *Cell* **176**, 419–434 (2019).
- 47. W. J. de Jonge, H. P. Patel, J. V. W. Meeussen, T. L. Lenstra, Following the tracks: How transcription factor binding dynamics control transcription. *Biophysical Journal* **121**, 1583–1592 (2022).
- 48. S. F. Banani, H. O. Lee, A. A. Hyman, M. K. Rosen, Biomolecular condensates: Organizers of cellular biochemistry. *Nature Reviews Molecular Cell Biology* **18**, 285–298 (2017).
- 49. D. T. McSwiggen, M. Mir, X. Darzacq, R. Tjian, Evaluating phase separation in live cells: diagnosis, caveats, and functional consequences. *Genes and Development* **33**, 1619–1634 (2019).
- 50. X. Chu, T. Sun, Q. Li, Y. Xu, Z. Zhang, L. Lai, J. Pei, Prediction of liquid–liquid phase separating proteins using machine learning. *BMC bioinformatics* **23**, 72–72 (2022).
- 51. Z. Chen, C. Hou, L. Wang, C. Yu, T. Chen, B. Shen, Y. Hou, P. Li, T. Li, Screening membraneless organelle participants with machine-learning models that integrate

multimodal features. *Proceedings of the National Academy of Sciences* **119**, e2115369119–e2115369119 (2022).

- 52. A. Hatos, S. C. E. Tosatto, M. Vendruscolo, M. Fuxreiter, FuzDrop on AlphaFold: visualizing the sequence-dependent propensity of liquid–liquid phase separation and aggregation of proteins. *Nucleic acids research* **50**, W337–W344 (2022).
- J. Sun, J. Qu, C. Zhao, X. Zhang, X. Liu, J. Wang, C. Wei, X. Liu, M. Wang, P. Zeng, Precise prediction of phase-separation key residues by machine learning. *Nature Communications* 15, 2662–2662 (2024).
- 54. J. Huff, The Airyscan detector from ZEISS: confocal imaging with improved signal-to-noise ratio and super-resolution. (2015).
- 55. I. F. Sbalzarini, P. Koumoutsakos, Feature point tracking and trajectory analysis for video imaging in cell biology. *Journal of Structural Biology* **151**, 182–195 (2005).
- 56. F. Simon, L. E. Weiss, S. Van Teeffelen, A guide to single-particle tracking. *Nat Rev Methods Primers* **4**, 66 (2024).
- 57. A. S. Hansen, I. Pustova, C. Cattoglio, R. Tjian, X. Darzacq, CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *eLife* **6** (2017).
- 58. A. S. Hansen, M. Woringer, J. B. Grimm, L. D. Lavis, R. Tjian, X. Darzacq, Robust modelbased analysis of single-particle tracking experiments with Spot-On. *Elife* 7, e33125– e33125 (2018).
- I. Izeddin, V. Récamier, L. Bosanac, I. I. Cissé, L. Boudarene, C. Dugast-Darzacq, F. Proux, O. Bénichou, R. Voituriez, O. Bensaude, M. Dahan, X. Darzacq, Single-molecule tracking in live cells reveals distinct target-search strategies of transcription factors in the nucleus. *eLife* 2014 (2014).
- I. Izeddin, V. Récamier, L. Bosanac, I. I. Cissé, L. Boudarene, C. Dugast-Darzacq, F. Proux, O. Bénichou, R. Voituriez, O. Bensaude, M. Dahan, X. Darzacq, Single-molecule tracking in live cells reveals distinct target-search strategies of transcription factors in the nucleus. *eLife* 3, e02230 (2014).
- 61. M. J. Saxton, Wanted: A Positive Control for Anomalous Subdiffusion. *Biophysical Journal* **103**, 2411–2422 (2012).
- 62. G. Tkačik, T. Gregor, The many bits of positional information. *Development* **148**, dev176065 (2021).
- 63. R. Munshi, J. Ling, S. Ryabichko, E. F. Wieschaus, T. Gregor, Transcription factor clusters as information transfer agents. *Sci. Adv.* **11**, eadp3251 (2025).

- 64. D. M. Mitrea, M. Mittasch, B. F. Gomes, I. A. Klein, M. A. Murcko, Modulating biomolecular condensates: a novel approach to drug discovery. *Nature Reviews Drug Discovery* **21**, 841–862 (2022).
- 65. A. A. Hyman, C. A. Weber, F. Jülicher, Liquid-liquid phase separation in biology. *Annual* review of cell and developmental biology **30**, 39–58 (2014).
- 66. D. Hnisz, K. Shrinivas, R. A. Young, A. K. Chakraborty, P. A. Sharp, A Phase Separation Model for Transcriptional Control. *Cell* **169**, 13–23 (2017).
- 67. L. Peng, E. M. Li, L. Y. Xu, From start to end: Phase separation and transcriptional regulation. *Biochimica et Biophysica Acta Gene Regulatory Mechanisms* **1863** (2020).
- 68. A. Eldar, M. B. Elowitz, Functional roles for noise in genetic circuits. *Nature* **467**, 167–173 (2010).
- J. R. Lakowicz, Ed., "Fluorescence Correlation Spectroscopy" in *Principles of Fluorescence Spectroscopy* (Springer US, Boston, MA, 2006; http://link.springer.com/10.1007/978-0-387-46312-4\_24), pp. 797–840.
- H.-L. Liang, C.-Y. Nien, H.-Y. Liu, M. M. Metzstein, N. Kirov, C. Rushlow, The zinc-finger protein Zelda is a key activator of the early zygotic genome in Drosophila. *Nature* 456, 400–403 (2008).
- 71. C. Y. Nien, H. L. Liang, S. Butcher, Y. Sun, S. Fu, T. Gocha, N. Kirov, J. R. Manak, C. Rushlow, Temporal coordination of gene networks by Zelda in the early Drosophila embryo. *PLoS Genetics* **7** (2011).
- 72. S. M. Foo, Y. Sun, B. Lim, R. Ziukaite, K. O'Brien, C. Y. Nien, N. Kirov, S. Y. Shvartsman, C. A. Rushlow, Zelda potentiates morphogen activity by increasing chromatin accessibility. *Current Biology* 24, 1341–1346 (2014).
- 73. 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. *Developmental Cell* 58, 1610-1624.e8 (2023).
- 74. C. Jiang, S. X. Dou, P. Y. Wang, H. Li, Quantifying transport dynamics with threedimensional single-particle tracking in adherent cells. *STAR Protocols* **3** (2022).
- H. Shen, L. J. Tauzin, R. Baiyasi, W. Wang, N. Moringo, B. Shuang, C. F. Landes, Single Particle Tracking: From Theory to Biophysical Applications. *Chemical Reviews* 117, 7331– 7376 (2017).
- 76. U. Endesfelder, S. Malkusch, F. Fricke, M. Heilemann, A simple method to estimate the average localization precision of a single-molecule localization microscopy experiment. *Histochemistry and Cell Biology* 141, 629–638 (2014).

- 77. R. E. Thompson, D. R. Larson, W. W. Webb, "Precise Nanometer Localization Analysis for Individual Fluorescent Probes" (00063495/02/05/2, 2002); http://rlab.sourceforge.net.
- 78. D. Mazza, A. Abernathy, N. Golob, T. Morisaki, J. G. McNally, A benchmark for chromatin binding measurements in live cells. *Nucleic Acids Research* **40** (2012).

# Acknowledgments

S.S.D. was supported by NIH R01 Award R01GM151409 and by NSF Award MCB-2105619.

M.A.T. was supported by the National Science Foundation's Graduate Research Fellowship Program (NSF GRFP).

H.G.G. was supported by NIH R01 Awards R01GM139913 and R01GM152815, by the Koret-UC Berkeley-Tel Aviv University Initiative in Computational Biology and Bioinformatics, by a Winkler Scholar Faculty Award, and by the Chan Zuckerberg Initiative Grant CZIF2024-010479. H.G.G. is also a Chan Zuckerberg Biohub Investigator (Biohub – San Francisco). G.T.R. was supported by NIH R01 Award R01GM151409 and by NSF Award MCB-2105619.

# **Author contributions**

Conceptualization: S.S.D and G.T.R Imaging and analysis: S.S.D and G.T.R Generation of fly lines: All authors Investigation: S.S.D and G.T.R Visualization: S.S.D and G.T.R Supervision: G.T.R Writing—original draft: S.S.D and G.T.R Writing—review & editing: All authors

# **Competing interests**

The authors declare no competing interests.

# Data and materials availability

Image analysis pipeline: https://github.com/gtreeves/RICS\_timecourse\_pipeline All the image files (.czi) will be deposited in the "Texas Data Repository".

# Figures

**Figure 1:** Temporal variation of biophysical parameters of Dl in a fly line having wildtype level of Dl (wt) and in a line where the level is reduced (reduced) in the ventral-most nuclei.

Figure 2: Spatial variation of biophysical parameters of Dl mid nc 14.

Figure 3: Single particle tracking to characterize Dl clusters.

**Figure 4:** Variation of the line time of the RICS experiments to capture the slowly-moving clusters of Dl.

Figure 5: Model of Dl-DNA binding as indicated by multiple experimental techniques.

# **Supplementary Materials**

# Figures

**Figure S1:** Dynamics of fraction of different pools of Dl from nc 10 until gastrulation for wt and reduced fly lines.

Figure S2: Variation of fraction of different pools of Dl mid nc 14 at different DV locations.

Figure S3: Model fits to the particle tracking parameters.

Figure S4: Simulations of individual ACFs.

# Tables

**Table S1:** Phase separation propensity scores of Dl and mNG, the fluorescent protein tagged to it, using multiple algorithms.

# **Movie Files**

**Movie S1:** Image acquisition on the ventral side of the embryo from nc 10 until late nc 14 for RICS analysis. Mitosis not shown, as only interphase was used for the analysis.

Movie S2: Detection of clusters that stay for multiple frames and the trajectories followed by them.

**Movie S3:** Image acquisitions showing Case 1 (short line time) and Case 2 (long line time) for the simultaneous detection of all three populations using RICS.

bioRxiv preprint doi: https://doi.org/10.1101/2025.04.24.650477; this version posted April 27, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

bioRxiv preprint doi: https://doi.org/10.1101/2025.04.24.650477; this version posted April 27, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



**Figure 1:** Temporal variation of biophysical parameters of Dl in a fly line having wildtype level of Dl (wt) and in a line where the level is reduced (reduced) in the ventral-most nuclei.

(A) Representative image showing the Dl gradient (wt) and RICS acquisition in the ventral-most nuclei. We expect to see three populations: free, DNA-bound, and clusters of Dl as depicted in the figure. Populations inside the dashed circle are identified as similar (correlated in this case) using RICS.

(B-E) Dynamics of different pools of Dl from nc 10 until gastrulation including total nuclear concentration (B), free concentration (C), correlated concentration (D), and uncorrelated concentration (E). The dynamics of the free population of Dl from nc 10 until gastrulation show only minor differences between the two fly lines compared to the other pools. Curves: mean values (n = 8 for wt embryos, n = 8 for reduced embryos). Error bars: S.E.M.

(F-G) Dose/response map between free and the low-diffusivity populations including correlated population (F) and uncorrelated population (G).



Figure 2: Spatial variation of biophysical parameters of Dl mid nc 14.

(A) Snap shots of Dl-mNG in nc 14 embryos, illustrating the relationship between Dl nuclear intensity and DV coordinate in the embryo cross-section (left panel) and surface (right panel). (B-E) Variation of different pools of Dl in ventral to lateral regions of the embryo including total nuclear concentration (B), free concentration (C), correlated concentration (D), and uncorrelated concentration (E). The concentration of free Dl is only a weak function of DV location. Curves: mean values. Error bars: S.E.M. Numbers on curve in (B) indicate sample size at that DV location. (F-G) Dose/response map between total nuclear concentration of Dl and the low-diffusivity populations: correlated population (F) and uncorrelated population (G)

bioRxiv preprint doi: https://doi.org/10.1101/2025.04.24.650477; this version posted April 27, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



Figure 3: Single particle tracking to characterize Dl clusters.

- (A)Representative average image of a time series for Dl-mNG and NLS-eGFP and representative trajectories of the particles.
- (B) Number of clusters detected during acquisition on different DV locations as showed in the left panel. Numbers in the right panel indicate sample size.
- (C) Step size distribution and 2-component model fitting for the determination of diffusivities and compositions of different populations for trajectories detected in 90 ms frame rate (left panel, n = 140 nuclei) and 500 ms frame rate (right panel, n = 118 nuclei).
- (D) Bound and slowly-moving clusters are identified using SPT.



**Figure 4:** Variation of the line time of the RICS experiments to capture the slowly-moving clusters of Dl.

- (A)Populations present as indicated by the variation of the line time of the RICS experiments. Populations inside the dashed circle are identified as similar (bound in this case) in this experiment.
- (B) Case 1: Acquisition of a 133 μm x 33 μm area of the embryo along the ventral midline, oriented such that the line time is roughly 5 ms. Arrows denote exaggerated laser path: the horizontal arrow represents line scanning, and the dashed diagonal arrow represents line retracing. Together, the lengths of these arrows roughly correspond to the line time.
- (C) Case 2: The same embryo is imaged with the microscope axes reoriented by 90°, ensuring the line time is much longer (17 ms). The arrows are similar to those in (A), illustrating the longer line time in Case 2.
- (D) Simulations of a two-component system (free Dl + immobile Dl) showing the averaged, normalized slow-direction ACFs for Cases 1 and 2. The simulated PSF is shown in gray. The Case 2 ACF visually differs from Case 1 only slightly and only at pixel shifts 1 and 2. Neither ACF extends past the PSF (gray arrow). Number of simulations for each Case: 20. Error bars: SEM.
- (E) Simulations of a three-component system (free Dl + slowly-moving clusters + immobile Dl) showing the averaged, normalized slow-direction ACFs for Cases 1 and 2. The simulated PSF is shown in gray. The ACFs of the two Cases clearly differ at intermediate pixel shifts (η between 1 and 10, red arrow), and both ACFs extend past the PSF (black arrow). Number of simulations for each Case: 20. Error bars: SEM.
- (F) Experimental data with unknown number of components. The weighted average of normalized slow-direction ACFs for Cases 1 and 2 differ significantly at pixel shifts 1-9 (red arrow). Both ACFs have tails that extend past the PSF (black arrow). N = 9 embryos with at least two acquisitions of each Case. Error bars: weighted SEM.
- (G) Model fits to the simulated two-component system. Fitting a two-component model to a two-component system produced the same values of fitted parameter between Cases 1 and 2. However, mismatching a three-component model fitted to a two-component

system produced statistically different values of the fitted parameter between the two Cases.

- (H) Model fits to the simulated three-component system. Mismatching a two-component model to a two-component system produced statistically different values of the fitted parameter between Cases 1 and 2. However, a three-component model fitted to a threecomponent system produced statistically indistinguishable values of the fitted parameter between the two Cases.
- (I) Two- and three-component models fitted to the experimental data. The fits to the twocomponent model produced statistically different values of the fitted parameter between Cases 1 and 2. However, fitting the three-component model produced statistically indistinguishable values of the fitted parameter between the two Cases.

bioRxiv preprint doi: https://doi.org/10.1101/2025.04.24.650477; this version posted April 27, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.





The total population of Dl in the nuclei is composed of free and slow populations as indicated by RICS. The slow population is composed of DNA-bound population (indicated by ccRICS) and slowly-moving clusters. SPT can differentiate between the DNA-bound clusters and slowly-moving clusters. Finally, long line time RICS can detect the presence of all the constituent populations together. Both the experimental techniques and the different populations are color coded.