Transcription factor clusters as information transfer agents

Rahul Munshi,[1,2] Jia Ling,[3,2] Sergey Ryabichko,[2] Eric Wieschaus,[2,4] and Thomas Gregor[1,2,5,*]

[1]*Joseph Henry Laboratories of Physics, Princeton University, Princeton, NJ 08544, USA*

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

[3]*Department of Molecular Biology, Princeton University, Princeton, NJ 08544, USA*

[4]*Department of Molecular Biology and Howard Hughes Medical Institute,
Princeton University, Princeton, NJ 08544, USA*

[5]*Department of Stem Cell and Developmental Biology, CNRS UMR3738 Paris Cité,
Institut Pasteur, 25 rue du Docteur Roux, 75015 Paris, France*

(Dated: March 6, 2024)

Deciphering how genes interpret information from the concentration of transcription factors (TFs) within the cell nucleus remains a fundamental question in gene regulation. Recent advancements have unveiled the heterogeneous distribution of TF molecules in the nucleus, posing challenges to the precise decoding of concentration signals. To explore this phenomenon, we employ high-resolution single-cell imaging of a fluorescently tagged TF protein, Bicoid, in living fly embryos. We show that accumulation of Bicoid in submicron clusters preserves the spatial information of the maternal Bicoid gradient, and that cluster intensity, size, and frequency offer remarkably precise spatial cues. We further discover that various known gene targets of Bicoid activation colocalize with clusters and that for the target gene Hunchback, this colocalization is dependent on its enhancer binding affinity. Modeling information transfer through these clusters suggests that clustering offers a more rapid sensing mechanism for global nuclear concentrations than freely diffusing TF molecules detected by simple enhancers.

INTRODUCTION

Transcription factors (TFs) play a pivotal role in regulating gene expression by interacting with DNA regulatory elements known as enhancers [1–3]. These enhancers often exhibit concentration-dependent behavior, activating or repressing gene expression only within specific TF concentration thresholds [4, 5]. The remarkable sensitivity of enhancers to subtle variations in the nuclear concentration of TF molecules implies that genes and enhancers carry out precise measurements of TF concentration [6, 7].

However, the challenge arises from the fact that in the nucleus, TF levels are often quite low and TF molecules are not uniformly distributed [8]. Instead, they assemble into dynamic transcriptional microenvironments referred to as transcriptional hubs [9–11]. These TF molecule accumulations are believed to form through transient clustering mechanisms [12–15] or through liquid-liquid phase separations (LLPS) , [16–18]. Separation of LLPS clusters reflects saturation kinetics, such that increasing concentration of the minor component results in increased size of droplets rather than an increase in the concentration within droplets [19, 20]. Whether droplet size provides a useful proxy for global nuclear concentration is unclear.

In this study, we aim to investigate whether these TF assemblies accurately reflect the nuclear concentration. We leverage the unique characteristics of the *Drosophila* TF Bicoid (Bcd), known for its varying concentration along the anterior-posterior (AP) axis of the early embryo [21]. Despite low nuclear concentrations, Bcd exhibits an extraordinarily reproducible profile, revealing precision in positional information comparable to the size of a single cell [22–24].

Various imaging approaches have unveiled that, similar to many other TFs, Bcd is not homogeneously distributed in the nucleus [25–27]. Instead, it forms numerous cluster-like droplets enriched with chromatin accessibility factors like Zelda [9] and actively transcribed canonical Bcd target genes, such as Hunchback, [28]. The higher concentrations within Bcd accumulations are believed to enhance transcription by increasing the local concentration near target enhancers [15, 29]. However, for these clusters to be functionally relevant to Bcd's well-characterized role in patterning, some features of the observed clusters must convey positional information with a precision similar to the nuclear concentration profile.

To decipher which features of Bcd accumulations maintain information about concentration, we developed a rigorous quantitative imaging strategy. Contrary to simple LLPS models, we found that cluster size remains independent of concentration, while the cluster concentration varies linearly with nuclear Bcd concentration. These clusters localize at the locus of active target genes, conferring information about cellular position precisely. We use these data to explore quantitively the impact of clustering on information transfer and discuss the circumstances where clustering might be a preferred mechanism as opposed to the gene interacting with the TF molecules freely diffusing in the nucleus.

* Correspondence: tg2@princeton.edu
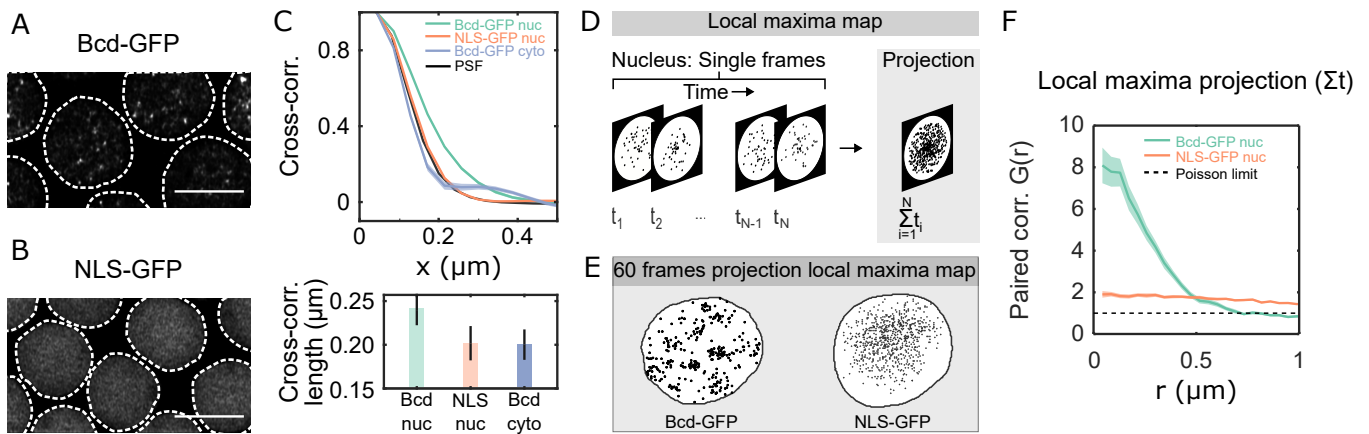
FIG. 1. **Quantitative characterization of nuclear Bcd heterogeneity.** (A-B) Confocal (*Zeiss-Airyscan*) images of cross-sections of Bcd-GFP (A), and NLS-GFP (B) expressing blastoderm nuclei in living *Drosophila* embryos (NC14). Scale bars are $5\,\mu m$. The broken lines represent a guide to the eye for the nuclear boundaries. (C) (TOP) Spatial cross-correlations computed on the nuclear pixels in 2D nuclear cross-section images expressing Bcd-GFP (green, 44 nuclei from 5 embryos) and NLS-GFP (orange, 27 nuclei from 3 embryos); and from pixels within the cytoplasm of Bcd-GFP expressing embryos (grey, 5 embryos). For comparison, the point-spread-function (PSF) of the objective is in black. Bottom panel shows mean and standard deviations of the computed correlation lengths: nucleoplasmic Bcd-GFP: $0.24 \pm 0.02\,\mu m$; nucleoplasmic NLS-GFP: $0.20 \pm 0.02\,\mu m$; and cytoplasmic Bcd-GFP: $0.20 \pm 0.02\,\mu$ m. (D) Schematic showing a map of local fluorescence intensity maxima inside a nucleus (left). The local maxima maps are extracted from individual frames of $\sim 30\,\text{s}$ long videos (60 frames) of nuclear cross sections ($1\mu m$ thick). All maps from a given video are projected onto a single frame to form the local maxima map (right). See also METHODS and FIG. S3. (E) Representative local maxima maps for a Bcd-GFP nucleus (left) and an NLS-GFP nucleus (right). (F) Radial distribution function (or pair-correlation function, $G(r)$) for the local maxima distribution expressed as a function of distance r from the center. $G(r)$ was calculated on time-projected (60 frames each) local intensity maxima centroid maps, averaged over multiple nuclei (same nuclei and embryo counts as in C). A distinct peak in G(r) indicates temporally persistent confinement of the local maxima, as seen for Bcd-GFP expressing nuclei. For NLS-GFP, the gradual reduction in the radial function indicates a gradual decline in intensity near the nuclear edges without the existence of any sub-micron accumulations. ($G(r) = 1$ corresponds to a perfectly uniform distribution).

RESULTS

Heterogeneity of nuclear TF distribution. We revisit the heterogeneous distribution of Bicoid (Bcd) within nuclei to establish quantitative insights. The distribution of Bcd within the nucleus comprises both freely diffusing molecules in intranuclear spaces and those engaging with chromatin [21, 30]. Cross-sectional images of Bcd-GFP nuclei ($1\,\mu m$ thick z-section) revealed multiple focal accumulations per cross-section (FIG. 1A, FIG. S1A-D and FIG. S5). Conversely, embryos expressing an NLS-GFP fusion construct, where molecules diffuse freely without chromatin interaction, showed no such heterogeneity (FIG. 1B).

Quantitative analysis of the focal accumulations' average sizes in these cross-sectional images, determined by pixel cross-correlation functions, revealed an average correlation length of 241 ± 17 nm for nuclear Bcd-GFP. In contrast, NLS-GFP expressing nuclei exhibited a smaller correlation length of 204 ± 16 nm, comparable to cytoplasmic Bcd-GFP (200 ± 17 nm) (FIG. 1C). Both nuclear NLS-GFP and cytoplasmic Bcd-GFP molecules are freely diffusing, and hence their correlation functions coincide with the microscope objective's point spread function (PSF, FIG. S1F). Nuclear Bcd-GFP however, forms focal accumulations larger than the diffraction limit.

To investigate if the Bcd focal accumulations are spatiotemporally persistent, we took short videos of nuclear cross-sections. Local GFP fluorescence intensity maxima were identified in each video frame, and all frames were combined to form projection maps (FIG. 1D). The projection maps revealed that the Bcd-GFP maxima tend to crowd inside confinement areas within the nucleus, contrasting with the dispersed maxima in NLS-GFP nuclei (FIG. 1E). Pair-correlation analysis [31] indicated an effective radius (ξ_{pair}) of the confinement area for Bcd-GFP nuclei as 369 ± 49 nm (FIG. 1F). No such correlation was detected in NLS-GFP nuclei. Within a confinement area, the density of local maxima is approximately eight times higher than outside (FIG. S4); 60% of frames recorded over a 30-second interval contain at least one maxima within the area. This implies that local maxima persist over the recording interval providing a minimal estimate of their lifetime and positional stability.

These results indicate that Bcd accumulations form persistent sub-micron clusters within the nucleus. To assess the potential of these clusters in transferring information to target genes, we proceed to characterize their biophysical properties.
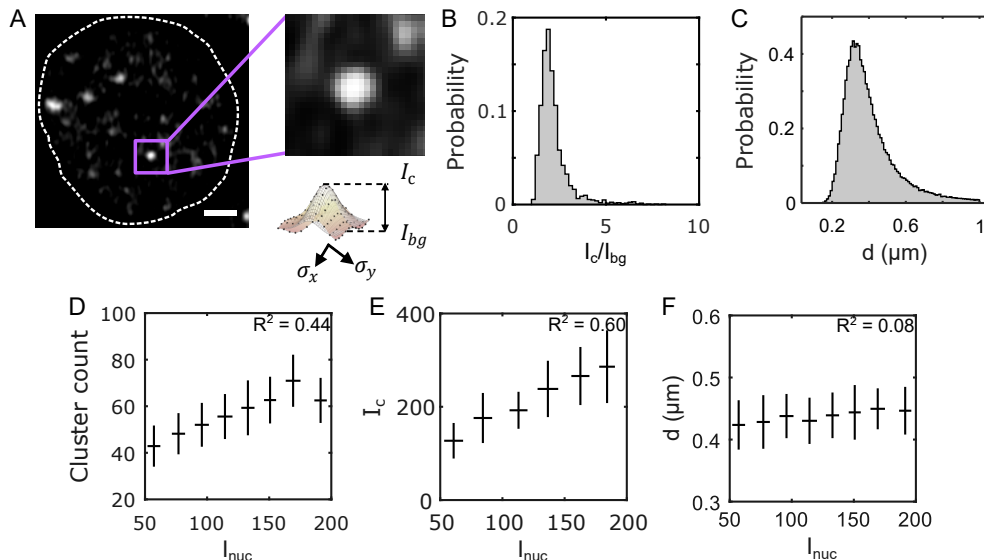
FIG. 2. **Biophysical properties of Bcd clusters** (A) A single nucleus showing Bcd-GFP heterogeneities. Close-up image (right) shows a single Bcd-GFP cluster. Cluster intensity fit with a 2D Gaussian (see profile below). The cluster intensity (I_c), the cluster background intensity (I_{bg}), and the linear cluster size d are extracted from this fit (Methods). (B) A histogram of the signal-to-noise ratio (I_c/I_{bg}) and (C) a histogram of the linear size (d) for 99671 clusters from 2027 nuclei in 14 embryos expressing Bcd-GFP. (D) Number of clusters per nucleus as a function of average nuclear Bcd-GFP intensity (mean $\pm$ std). (E) Average cluster intensity per nucleus as a function of average nuclear Bcd-GFP intensity (mean $\pm$ std). (F) Effective cluster diameter (d) as a function of average nuclear intensity.

Bcd cluster properties To characterize the biophysical properties of Bcd clusters in 3D, we took an approach different from the maxima detection approach used in the previous section. Since any cluster should be at least the size of the PSF along the x-y plane (> 4 pixels), an x-y cutoff of 3 pixels should eliminate any spurious spots. Also, from the persistence data (FIG. S4) and FIG. S1)), we argued that for an imaging frame time of ~ 500 ms, a cluster should span across at least two consecutive z frames, given the frame thickness is less than the z PSF (Methods). Using this approach, we identify between 40 and 70 analyzable clusters per nucleus.

Individual cluster parameters are determined from 2D Gaussian fits of the GFP intensity profile on the z-plane of the cluster centroid (see Methods and Supplement). These fits yield estimates for the effective linear size of the cluster (d) of the clusters (Methods). The average linear size per nucleus is $d = 400 \pm 140$ nm for all clusters (FIG. 2C). Notably, the histogram of the cluster size distribution reveals a vanishing left tail around the PSF limit, despite choosing a considerably smaller size cutoff than that limit. This implies that the detectable clusters are not diffraction-limited under our imaging conditions. It could be that the sub-diffraction clusters have very low intensities that are not fit for detection or are highly transient, rendering them undetectable. Either way, all such potential clusters would not be included in this study.

To gauge the cluster concentration, we introduce the parameter I_c, representing the cluster peak intensity, and I_{bg}), denoting the concentration of Bcd molecules in the

nuclear space surrounding the cluster (FIG. 2A and FIG. S6). The signal-to-background ratio I_c/I_{bg} offers insights into local Bcd concentration amplification within a cluster, with an average value of 2.2 ± 0.8 for close to 10^5 clusters (FIG. 2B).

Given the variation of the nuclear Bcd concentration along the AP axis of the embryo (given by the average nuclear Bcd-GFP intensity, I_{nuc}), we investigate how cluster properties correlate with I_{nuc}. Analysis reveals a strong dependence of the cluster count on nuclear concentration, exhibiting an almost two-fold drop between the anterior and posterior poles (FIG. 2D, also see FIG. S7). This could result from a drop in I_c and/or a reduction in d with decreasing I_{nuc}. From FIG. 2E and F, it is evident that while I_c shows a strong dependence on I_{nuc}, with an almost two-fold change between the poles, d varies only insignificantly. This is further elucidated by the fact that the distribution of d effectively remains the same at various ranges of I_{nuc} (FIG. S8B).

Thus, one might speculate that droplet growth by coalescence at higher concentrations, a characteristic of LLPS condensates, might be absent in Bcd clusters [32]. This speculation is complemented by the observation that the dependence of the cluster concentration (I_c) on the nuclear concentration (I_{nuc}) is linear ($R^2 = 0.6$) (FIG. 2E), which contrasts with the switch-like dependence observed in liquid-liquid phase-separated condensates [20]. Notably, we observe clustering even in nuclei with very low Bcd concentrations, indicating the absence of a discernible threshold concentration triggering clus-
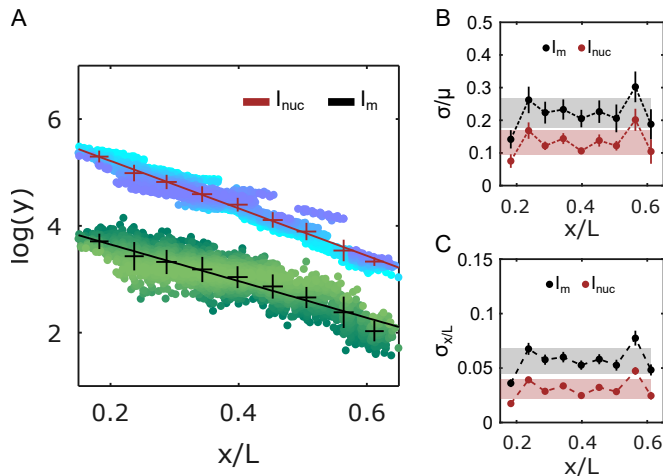
FIG. 3. **Precision of cluster positional information potential.** (A) Natural logarithms of Bcd-GFP nuclear intensity (I_{nuc}, crimson), and total Bcd-GFP cluster intensity (I_m, black) are binned and plotted against binned nuclear position x/L. The mean and standard deviation shown in the error bars are obtained by bootstrapping (14 embryos, 2027 nuclei). Exponential decay constants extracted from linear fits (solid lines) were $\lambda_{I_{nuc}} = 0.23 \pm 0.01$ EL, and $\lambda_{I_m} = 0.21 \pm 0.01$ EL. Each point in the data cloud (shades of blue and green) represents a single nucleus. (B) Errors (σ) in I_{nuc}, and I_m, normalized to the respective means (μ) are plotted against nuclear positions. Gray and red shades indicate the overall error in position estimate across the entire axis. (C) The effective error in the estimation of the nuclear positions using I_{nuc}, and I_m are plotted against nuclear position. Gray and red shades as in (B).

ter formation [33]. However, further investigation is warranted to ascertain whether the clusters analyzed represent a matured state where conventional LLPS rules no longer apply, or if detailed imaging, capturing cluster formation dynamics is needed to distinguish between these possibilities.

Do clusters contain enough positional information? Previously, we have established that the position of anterior nuclei in the early *Drosophila* embryo can be determined with a spatial precision of better than 1 % from nuclear Bcd concentration alone [34]. This precision stems from the collective contribution of all nuclear Bcd molecules reproducible to within 10 %. Given that clusters comprise only a small fraction of molecules within nuclei (FIG. S11B), we sought to investigate whether they could offer an accurate estimation of nuclear concentration and, consequently, the position of the nucleus along the anterior-posterior axis of the embryo.

To this end, we consider a representative molecular count of Bcd within a cluster ($I_m = 2I_c\sigma_x\sigma_y$), where σ_x and σ_y are characteristic cluster fit parameters (see FIG. 2A). From this quantity, an absolute count for the total Bcd molecules within an average cluster can be computed using previous estimates for absolute molec-

ular count conversions [21] (FIG. S11C).

When computing the nuclear average from all I_m in a given nucleus, this cluster-specific concentration decays exponentially with cell position as is typical for the Bcd gradient. The decay rate is comparable to the one of the overall nuclear Bcd concentration I_{nuc}, with a decay constant $\lambda_{I_m} = 0.26 \pm 0.02\,L$, which is statistically very similar to $\lambda_{I_{nuc}} = 0.23 \pm 0.03\,L$ (FIG. 3A). Thus, the molecular count of an average cluster mirrors the Bcd nuclear concentration gradient.

We calculated the corresponding variabilities (std/mean) across multiple nuclei located at identical spatial locations along the AP axis in multiple embryos for both I_{nuc} and I_m derived Bcd gradients (FIG. 3B). The overall variability in I_{nuc} and I_m were 14 ± 4 % and 22 ± 4 %, respectively. Despite the clusters representing only a small fraction of nuclear Bcd molecules ($5 - 10$ %, see FIG. S11), the I_m-derived Bcd gradient displayed remarkably low variability. This finding hints at the existence of tightly controlled mechanisms that regulate cluster formation.

As a morphogen, Bcd nuclear concentration imparts positional identity to a nucleus. To estimate the level of positional information contained in the cluster-derived Bcd gradient, we determine the positional precision $\sigma(x)$ from the gradient's concentration fluctuations $\delta c(x)$ using error propagation [34]. Then the positional precision is given by $\sigma(x) = \delta c(x)|\frac{c(x)}{dx}|^{-1}$, where $c(x)$ is the Bcd concentration.

Using I_m as the estimator for nuclear Bcd concentration, the positional error is $c(x) = 5.5 \pm 0.7\% L$ (FIG. 3E), which corresponds to a positional precision of roughly three cell diameters. This is significantly worse than the previously shown single-cell precision obtained when using the full nuclear Bcd concentration I_{nuc} as the estimator for the nuclear position. However, the estimation obtained from the cluster average I_m reflects the property of an average cluster. Individual clusters might confer positional information with varying accuracy, with the highest potentially being equivalent to I_{nuc}. However, for genes to access this information, the clusters must be in physical proximity to specific gene loci.

Cluster association with target genes. To elucidate the behavior of individual clusters around target gene transcription sites, we conducted three-dimensional imaging of labeled nascent mRNAs of putative target genes while imaging Bcd-GFP within the nuclei (FIG. 4A, B and FIG. S12). For each of the target genes (*hunchback*, *even-skipped*, *Krüppel*, *knirps*) [35, 36], Bcd accumulation was observed, with Bcd-GFP intensity peaks at the center of the nascent mRNA hotspot (FIG. 4C). The radii of Bcd-GFP accumulation around the four target genes were determined to be $485 \pm 39\,nm$, $546 \pm 23\,nm$, $394 \pm 44\,nm$, and $333 \pm 32\,nm$ (FIG. S14A). These radii were comparable to the radius of the average enrichment area shown in FIG. 1F (see FIG. S13A for a
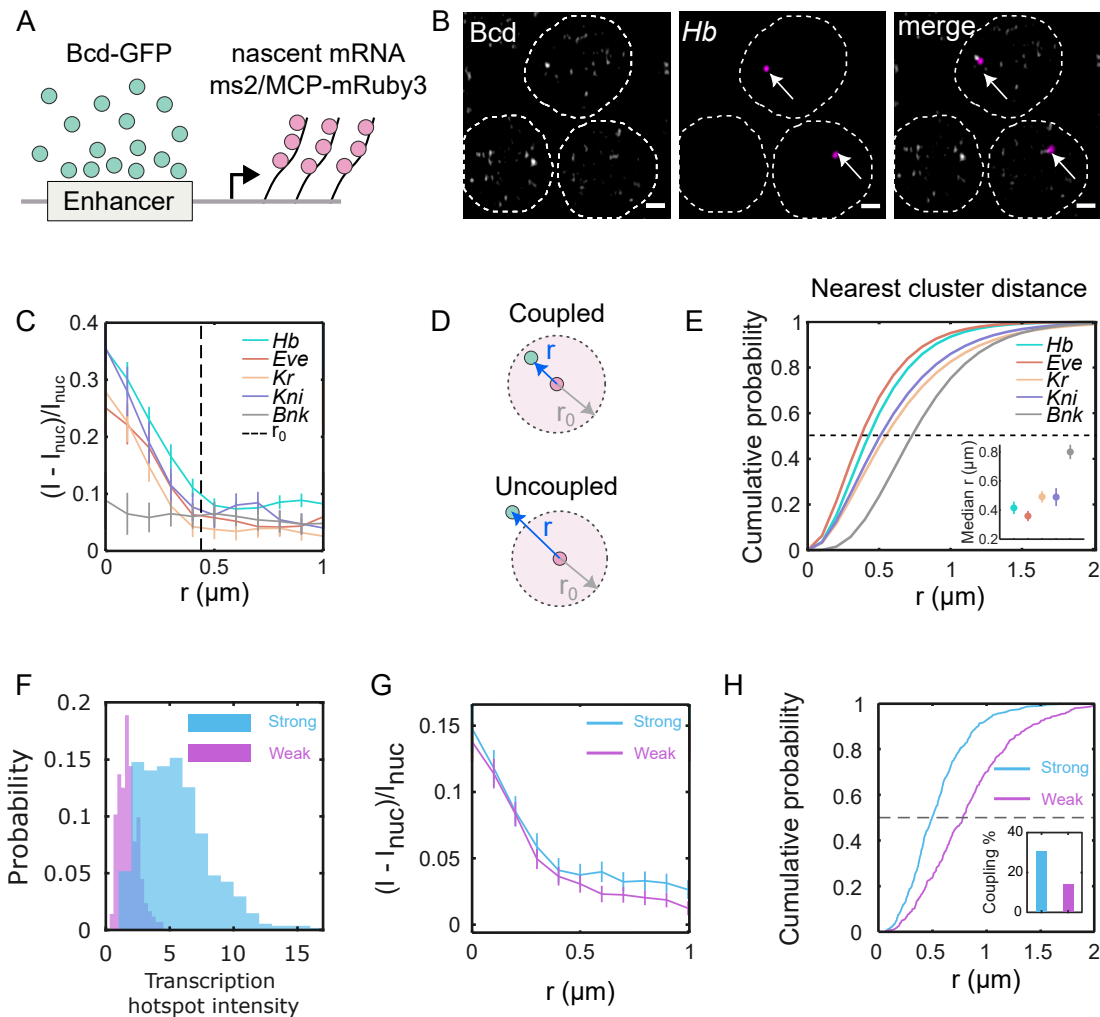
FIG. 4. **Bcd clusters co-localization with target genes is enhancer dependent.** (A) Cartoon showing scheme for dual color imaging with Bcd-GFP (green) and nascent transcription site labeled via the MS2/MCP system (red). (B) Images from embryos in NC14 showing nuclei expressing Bcd-GFP and *hb*-MS2/MCP-mRuby on sites of active transcription (arrows); scale bar is $1\,\mu m$. The dashed lines are guide to the eye for the nuclear boundaries. (C) Radial distribution of Bcd-GFP intensity with the centroid of the fluorescently labeled gene locus (i.e. hotspot) at the origin (see Methods). Data shown for canonical Bcd target genes, *hb*, *eve*, *kr*, *kni*, and the non-target gene *bnk*. Dashed line represents the average radius of accumulations from all genes combined ($r_0 = 0.44 \pm 0.05\,\mu m$). (D) Schematic showing the mRNA hotspot (red) and its nearest Bcd cluster (green). When the distance r between the nearest cluster and the hotspot is less than the Bcd accumulation radius r_0, the cluster is defined as being *coupled* to the gene; when it is greater than r_0 the cluster is assumed to be *uncoupled*. (E) Cumulative probability distributions of distances r for various genes; color code as in C. Dashed line determines median at EC50 (i.e. cumulative probability of 0.5). Inset: median distances for all genes. Errors are calculated from bootstrapping. (F) Histograms showing the transcription hotspot intensity from a strong and a weak enhancer construct driving an MS2-fusion reporter. The strong construct generates a 3.2-fold higher intensity, on average. (G) The radial distributions of relative Bcd-GFP intensity with the centroid of the transcription hotspot as the origin. The accumulation radii $0.36 \pm 0.05\,\mu m$ and $0.39 \pm 0.06\,\mu m$ for the strong and weak enhancer constructs respectively are statistically identical. (H) Cumulative probability distributions of distances r between the transcription hotspot and its nearest Bcd cluster. Black dashed line is at EC50. The median distances are $0.49 \pm 0.03\,\mu m$ and $0.78 \pm 0.05\,\mu m$ for the strong and weak constructs, respectively. Inset shows the fraction of coupled clusters for each construct (31 % and 13 %, respectively).

simulation-based representation).

In contrast, Bcd accumulation was not detected around a non-target gene, *bottleneck* [37] (FIG. 4C). Nor does Bcd accumulate around the geometric nuclear centers, which are considered random sites that are unassociated with a particular gene locus (FIG. S13B). Confirmed specificity was attained by imaging NLS-GFP in place of Bcd-GFP, where no accumulation was observed around the *hunchback* locus (FIG. S13B).

The presence of Bcd accumulation near target gene loci indicates that Bcd clusters tend to have a high probability of colocalizing with the gene loci. However, a TF

cluster may not be directly associated with the gene locus throughout the entire duration of active transcription of the gene locus. In such cases, the nearest TF cluster would be uncoupled from the gene (FIG. 4E), leading to a greater physical distance from the gene transcription site than a coupled cluster. The TF accumulation radius (FIG. S14A) gives a confinement radius within which a coupled cluster can be located. Utilizing this accumulation radius, a distance limit for cluster-gene coupling can be established, where any TF cluster located within that distance limit can be considered coupled to the respective gene.

The median 3D distances of the nearest Bcd clusters from the center of the genes (mRNA hotspots) were determined to be 418.6nm, 355.5nm, 500.5nm, and 485.8nm for *hunchback*, *even-skipped*, *Krüppel*, and *knirps*, respectively; and it was 804.3 nm for the non-target gene *bottleneck* (FIG. S14A). Applying the respective distance limits (FIG. 4D) to the cumulative probability plots of the nearest cluster distance distributions (FIG. 4E), we calculated the fraction of clusters coupled to the respective genes (an alternate technique yielding similar results is shown in FIG. S13C). The fractions of coupled clusters were 0.57, 0.73, 0.41, and 0.30, respectively, for *hunchback*, *even-skipped*, *Krüppel*, and *knirps* (FIG. S14B). Since an accumulation radius is not well-defined for *bottleneck*, no localization fraction could be determined for this gene.

These findings suggest that Bcd clusters tend to localize with target genes with a high probability. Additionally, we observed that localization is enhancer-dependent. When comparing a strong and a weak enhancer for the *hunchback* gene, the strong enhancer produces a much higher transcriptional output (FIG. 4F), even though the Bcd concentration at the site of transcription are similar for both (FIG. 4G). However, the Bcd clusters are closer to the transcription site (FIG. 4H) for the strong enhancer, which also has a higher fraction of clusters bound to the active target compared to the weak one (FIG. 4H, inset).

Given that Bcd clusters carry information about the nuclear position, genes can thus access this information by directly interacting with the clusters. However, genes can also interpret the same information by directly interacting with molecules diffusing in the nucleus. Thus, the question arises: why is clustering favored?

Clusters are fast information sensors. Nuclear Bcd concentration has to be interpreted by Bcd's target gene loci to accurately extract positional information from the morphogen gradient and trigger a transcriptional response accordingly. The ability to sense diffusing TF molecules naturally depends on the effective size of the sensor. Whether the size is of the order of a binding site (3nm), an enhancer (50nm), or the entire locus is unclear. We know from previous estimations that if a binding site is a relevant size metric, then the measured readout precision needs to invoke spatial averaging across multiple
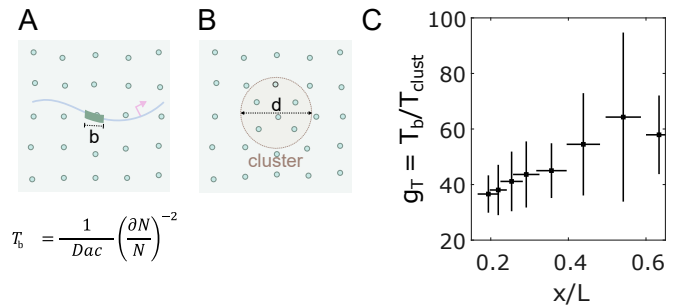

FIG. 5. **Clustering reduces time to precise concentration interpretation.** (A) Two cartoons showing Bcd molecules in the nucleus (green circles) and an enhancer with a binding site of length b (left) or a cluster of diameter d (right) embedded in the nuclear environment. Below is the equation for the time taken by a sensor of size a for nuclear concentration c with an accuracy of $\left(\frac{dN}{N}\right)$, where N is the number of molecules counted. (C) Reduction of time (g_T) to make an accurate ($\sim 10\%$) nuclear concentration estimation as a function of the nuclear position with the cluster as nuclear concentration sensor versus an enhancer binding site being the concentration sensor [34].

independent sensors, i.e. neighboring nuclei [34].

Here we consider the possibility that the TF cluster functions as a sensor and that the transcription output of a gene (its sensing of the gradient) is a reflection of the Bcd content of the cluster rather than interaction between single Bcd molecules at individual target gene enhancers. In previous analyses focused on enhancers, the time required for interpreting nuclear concentration was estimated using a molecular sensing argument that goes back to the work of Berg and Purcell [34, 38]. To apply this approach to Bcd's heterogeneous distribution in nuclei, we treated a cluster as a sphere with an effective diameter d, the concentration of Bcd molecules inside the cluster as c_{clust}, and the diffusion constant of Bcd as D (FIG. 5B). The time T_{clust} required for the cluster to precisely mirror the global nuclear concentration (the cluster sensing event that interprets nuclear concentration with an accuracy of $\left(\frac{\partial N}{N}\right)$) can be compared to the time T_{b} required for a simple binding site in an enhancer of linear size b, to measure nuclear concentration (FIG. 5A).

The ratio $T_{\text{b}}/T_{\text{clust}}$ yields insight into the comparative sensing times. If clusters function as concentration sensors, the average cluster in an anterior nucleus in the embryo could sense nuclear concentration approximately 37.5 ± 5.1 times faster than a single binding site (FIG. 5C), owing to the larger sensor size and ~ 2 fold concentration amplification within a cluster (FIG. 2B). Hence, an average cluster can interpret nuclear concentration in ~ 3 minutes, which is the timescale relevant to the activation of the target genes [34, 39]. In this scenario, no spatial averaging across neighboring nuclei or cells would be necessary. How Bcd content in the cluster interacts mechanistically with the target genes to govern transcription is unclear but points to an important area of future

research.

DISCUSSION

In this study, we employed quantitative imaging techniques in live embryos to elucidate the role of subnuclear compartmentalization, particularly clustering, in preserving the information carried by signaling molecules within the cell nucleus. Our findings highlight the potential significance of tight correlations between regulatory proteins and transcriptional outcomes, which may be mediated by transcription factor (TF) clustering within the transcriptional microenvironment. Notably, clustering facilitates the interpretation of concentration within transcriptionally relevant timescales, eliminating the need for invoking spatial averaging theories to explain the rapid regulation of transcriptional outputs [34, 40].

In such transcriptional microenvironments, various components such as mediators, chromatin-modifying agents, and PolII coexist and interact with each other. For instance, Bcd interacts with DNA via its DNA binding domain [41] and cooperatively with other proteins via its activation domain [42]. Our results suggest that the size of Bcd clusters is not solely dependent on the nuclear concentration of Bcd and that clusters can occur even at extremely low concentrations, contrary to classic liquid-liquid phase separation (LLPS) assemblies [19, 20, 27].

A more precise understanding of how Bcd clusters achieve nuclear concentration dependence should focus on the rate of approach of molecules at the cluster boundary, and how that rate is influenced by parameters relevant to diffusion [38]. Of these, parameters we assume that the diffusion constant and the surface area of clusters remain constant along the AP axis, leaving the nuclear concentration as the only variable. Once captured, the Bcd molecules reside within the cluster for a duration governed by their residence time before escaping [43, 44]. Cluster intensities will stabilize when capture rates are balanced by escape rates, thus allowing the final intensities of clusters to reflect overall nuclear concentration. Further theoretical exploration is warranted to elucidate the constraints on rate constants governing these processes.

Enhancers interact with TF molecules through binding events at DNA binding elements [1, 45]. Clustering amplifies local concentration, thereby increasing the probability of binding events at these sites and allowing enhancers to assay cluster concentration more rapidly. If Bcd's diffusion within the cluster is similar to its general diffusion within the nucleus, the 2 fold increase in concentration in clusters will halve the measurement time at enhancers. To achieve the 37 fold increase calculated in FIG. 5, the cluster size must provide the important advantage and somehow the sensor function of the cluster must therefore operate at levels higher than simple enhancer binding. Enhancers dispersed along the genome communicate with gene promoters through DNA compaction mechanisms [46, 47], allowing them to share the same gene microenvironment as TF clusters [48]. This involvement of multiple enhancers in gene regulation adds complexity to the interpretation of nuclear TF concentration. Understanding how multiple binding sites within the same enhancer and the existence of multiple enhancers for the same target gene affect the measurement of nuclear TF concentration, and how those measurements are converted into transcriptional outputs, such as transcriptional bursts remains an open challenge.

Our analysis tests the feasibility of obtaining quantitative data on cluster properties from fluorescently labeled living embryos. Further studies should exploit this possibility to measure the lifetime of individual clusters and the dynamics of their association with target genes in both active and inactive states. Using systematically designed synthetic loci, one can also test the influence of locus complexity on concentration sensing, as well as cluster lifetimes. Understanding the information flow across length scales, from the comparatively larger scale of a cell nucleus to a gene locus from an information-theoretic perspective could yield valuable insights. In conclusion, While our work focuses on a particular TF in a model organism, the principles governing such clusters are likely to be applicable across diverse biological systems, including mammalian systems.

[1] F. Spitz and E. E. Furlong, Transcription factors: from enhancer binding to developmental control, Nature reviews genetics **13**, 613 (2012).

[2] A. Panigrahi and B. W. O'Malley, Mechanisms of enhancer action: the known and the unknown, Genome Biology **22**, 108 (2021).

[3] J. Trojanowski and K. Rippe, Transcription factor binding and activity on chromatin, Current Opinion in Systems Biology **31**, 100438 (2022).

[4] N. Rosenfeld, J. W. Young, U. Alon, P. S. Swain, and M. B. Elowitz, Gene Regulation at the Single-Cell Level, Science **307**, 1962 (2005).

[5] J. Park, J. Estrada, G. Johnson, B. J. Vincent, C. Ricci-Tam, M. D. Bragdon, Y. Shulgina, A. Cha, Z. Wunderlich, J. Gunawardena, and A. H. DePace, Dissecting the sharp response of a canonical developmental enhancer reveals multiple sources of cooperativity, eLife **8**, e41266 (2019).

[6] J. Jaeger, The gap gene network, Cellular and Molecular Life Sciences **68**, 243 (2011).

[7] N. Haroush, M. Levo, E. F. Wieschaus, and T. Gregor, Functional analysis of the Drosophila eve locus in response to non-canonical combinations of gap gene expression levels, Developmental Cell 10.1016/j.devcel.2023.10.001 (2023).

[8] M. A. Grande, I. v. d. Kraan, L. d. Jong, and R. v. Driel, Nuclear distribution of transcription factors in relation to sites of transcription and RNA polymerase II, Journal of Cell Science **110**, 1781 (1997).

[9] 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, and M. Lagha, Temporal control of gene expression by the pioneer factor Zelda through transient interactions in hubs, Nature Communications **9**, 5194 (2018).

[10] A. Tsai, M. R. Alves, and J. Crocker, Multi-enhancer transcriptional hubs confer phenotypic robustness, eLife **8**, e45325 (2019).

[11] C.-Y. Cho and P. H. O'Farrell, Stepwise modifications of transcriptional hubs link pioneer factor activity to a burst of transcription, Nature Communications **14**, 4848 (2023).

[12] A. J. Wollman, S. Shashkova, E. G. Hedlund, R. Friemann, S. Hohmann, and M. C. Leake, Transcription factor clusters regulate genes in eukaryotic cells, eLife **6**, e27451 (2017).

[13] B. T. Donovan, A. Huynh, D. A. Ball, H. P. Patel, M. G. Poirier, D. R. Larson, M. L. Ferguson, and T. L. Lenstra, Live-cell imaging reveals the interplay between transcription factors, nucleosomes, and bursting, The EMBO Journal **38**, 10.15252/embj.2018100809 (2019).

[14] J. V. W. Meeussen, W. Pomp, I. Brouwer, W. J. de Jonge, H. P. Patel, and T. L. Lenstra, Transcription factor clusters enable target search but do not contribute to target gene activation, Nucleic Acids Research 10.1093/nar/gkad227 (2023).

[15] K. Kawasaki and T. Fukaya, Functional coordination between transcription factor clustering and gene activity, Molecular Cell **83**, 1605 (2023).

[16] 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, and R. A. Young, Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains, Cell **175**, 1842 (2018).

[17] D. Hnisz, K. Shrinivas, R. A. Young, A. K. Chakraborty, and P. A. Sharp, A Phase Separation Model for Transcriptional Control, Cell **169**, 13 (2017).

[18] M.-T. Wei, Y.-C. Chang, S. F. Shimobayashi, Y. Shin, A. R. Strom, and C. P. Brangwynne, Nucleated transcriptional condensates amplify gene expression, Nature Cell Biology **22**, 1187 (2020).

[19] K. Shrinivas, B. R. Sabari, E. L. Coffey, I. A. Klein, A. Boija, A. V. Zamudio, J. Schuijers, N. M. Hannett, P. A. Sharp, R. A. Young, and A. K. Chakraborty, Enhancer Features that Drive Formation of Transcriptional Condensates, Molecular Cell **75**, 549 (2019).

[20] D. M. Mitrea, M. Mittasch, B. F. Gomes, I. A. Klein, and M. A. Murcko, Modulating biomolecular condensates: a novel approach to drug discovery, Nature Reviews Drug Discovery **21**, 841 (2022).

[21] C. E. Hannon, S. A. Blythe, and E. F. Wieschaus, Concentration dependent chromatin states induced by the bicoid morphogen gradient, eLife **6**, e28275 (2017).

[22] M. Petkova, S. Little, F. Liu, and T. Gregor, Maternal Origins of Developmental Reproducibility, Current Biology **24**, 1283 (2014).

[23] G. Tkačik, J. O. Dubuis, M. D. Petkova, and T. Gregor, Positional Information, Positional Error, and Read-Out Precision in Morphogenesis: A Mathematical Framework, Genetics **199**, genetics.114.171850 (2014).

[24] B. Zoller, S. C. Little, and T. Gregor, Diverse Spatial Expression Patterns Emerge from Unified Kinetics of Transcriptional Bursting, Cell **175**, 835 (2018).

[25] A. Abu-Arish, A. Porcher, A. Czerwonka, N. Dostatni, and C. Fradin, High mobility of bicoid captured by fluorescence correlation spectroscopy: implication for the rapid establishment of its gradient, Biophysical Journal **99**, L33 (2010).

[26] T. Athilingam, A. V. Nelanuthala, C. Breen, T. Wohland, and T. E. Saunders, Long-ranged formation of the Bicoid gradient requires multiple dynamic modes that spatially vary across the embryo, bioRxiv , 2022.09.28.509874 (2022).

[27] M. Mir, A. Reimer, J. E. Haines, X.-Y. Li, M. Stadler, H. Garcia, M. B. Eisen, and X. Darzacq, Dense Bicoid hubs accentuate binding along the morphogen gradient, Genes & Development **31**, 1784 (2017).

[28] M. Mir, M. R. Stadler, S. A. Ortiz, C. E. Hannon, M. M. Harrison, X. Darzacq, and M. B. Eisen, Dynamic multifactor hubs interact transiently with sites of active transcription in drosophila embryos, Elife **7**, e40497 (2018).

[29] T.-Y. Chen, A. G. Santiago, W. Jung, L. Krzemiński, F. Yang, D. J. Martell, J. D. Helmann, and P. Chen, Concentration- and chromosome-organization-dependent regulator unbinding from DNA for transcription regulation in living cells, Nature Communications **6**, 7445 (2015).

[30] Z. Xu, H. Chen, J. Ling, D. Yu, P. Struffi, and S. Small, Impacts of the ubiquitous factor Zelda on Bicoid-dependent DNA binding and transcription in Drosophila, Genes & Development **28**, 608 (2014).

[31] S. L. Veatch, B. B. Machta, S. A. Shelby, E. N. Chiang, D. A. Holowka, and B. A. Baird, Correlation Functions Quantify Super-Resolution Images and Estimate Apparent Clustering Due to Over-Counting, PLoS ONE **7**, e31457 (2012), 1106.6068.

[32] J. Söding, D. Zwicker, S. Sohrabi-Jahromi, M. Boehning, and J. Kirschbaum, Mechanisms for Active Regulation of Biomolecular Condensates, Trends in Cell Biology **30**, 4 (2020).

[33] S. F. Shimobayashi, P. Ronceray, D. W. Sanders, M. P. Haataja, and C. P. Brangwynne, Nucleation landscape of biomolecular condensates, Nature **599**, 503 (2021).

[34] T. Gregor, D. W. Tank, E. F. Wieschaus, and W. Bialek, Probing the Limits to Positional Information, Cell

10.1016/j.cell.2007.05.025 (2007).

[35] W. Driever, G. Thoma, and C. Nüsslein-Volhard, Determination of spatial domains of zygotic gene expression in the Drosophila embryo by the affinity of binding sites for the bicoid morphogen, Nature **340**, 363 (1989).

[36] C. Nüsslein-Volhard and E. Wieschaus, Mutations affecting segment number and polarity in Drosophila, Nature **287**, 795 (1980).

[37] E. D. Schejter and E. Wieschaus, bottleneck acts as a regulator of the microfilament network governing cellularization of the Drosophila embryo, Cell **75**, 373 (1993).

[38] H. Berg and E. Purcell, Physics of chemoreception, Biophysical Journal **20**, 193 (1977).

[39] T. Lucas, H. Tran, C. A. P. Romero, A. Guillou, C. Fradin, M. Coppey, A. M. Walczak, and N. Dostatni, 3 minutes to precisely measure morphogen concentration, PLOS Genetics **14**, e1007676 (2018).

[40] T. Erdmann, M. Howard, and P. R. t. Wolde, Role of Spatial Averaging in the Precision of Gene Expression Patterns, Physical Review Letters **103**, 258101 (2009), 0907.4289.

[41] S. D. Hanes and R. Brent, DNA specificity of the bicoid activator protein is determined by homeodomain recognition helix residue 9, Cell **57**, 1275 (1989).

[42] D. S. Burz, R. Rivera-Pomar, H. Jäckle, and S. D. Hanes, Cooperative dna-binding by bicoid provides a mechanism for threshold-dependent gene activation in the drosophila embryo, The EMBO journal **17**, 5998 (1998).

[43] M. L. Heltberg, J. Miné-Hattab, A. Taddei, A. M. Walczak, and T. Mora, Physical observables to determine the nature of membrane-less cellular sub-compartments, eLife **10**, e69181 (2021).

[44] A. S. Lyon, W. B. Peeples, and M. K. Rosen, A framework for understanding the functions of biomolecular condensates across scales, Nature Reviews Molecular Cell Biology **22**, 215 (2021).

[45] K. Wagh, D. A. Stavreva, A. Upadhyaya, and G. L. Hager, Transcription Factor Dynamics: One Molecule at a Time, Annual Review of Cell and Developmental Biology **39**, 277 (2023).

[46] M. Levo, J. Raimundo, X. Y. Bing, Z. Sisco, P. J. Batut, S. Ryabichko, T. Gregor, and M. S. Levine, Transcriptional coupling of distant regulatory genes in living embryos, Nature **605**, 754 (2022).

[47] H. Chen, M. Levo, L. Barinov, M. Fujioka, J. B. Jaynes, and T. Gregor, Dynamic interplay between enhancer–promoter topology and gene activity, Nature Genetics **50**, 1296 (2018).

[48] 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, and R. A. Young, Coactivator condensation at super-enhancers links phase separation and gene control, Science **361**, 10.1126/science.aar3958 (2018).

[49] F. Liu, A. H. Morrison, and T. Gregor, Dynamic interpretation of maternal inputs by the Drosophila segmentation gene network, Proceedings of the National Academy of Sciences **110**, 6724 (2013).

[50] H. Garcia, M. Tikhonov, A. Lin, and T. Gregor, Quantitative Imaging of Transcription in Living Drosophila Embryos Links Polymerase Activity to Patterning, Current Biology **23**, 2140 (2013).

[51] J. R. Bateman, A. M. Lee, and C.-t. Wu, Site-specific transformation of Drosophila via phiC31 integrase-mediated cassette exchange., Genetics **173**, 769 (2006).

[52] S. C. Little, G. Tkačik, T. B. Kneeland, E. F. Wieschaus, and T. Gregor, The Formation of the Bicoid Morphogen Gradient Requires Protein Movement from Anteriorly Localized mRNA, PLoS Biology **9**, e1000596 (2011).

[53] H. Xu, L. A. Sepúlveda, L. Figard, A. M. Sokac, and I. Golding, Combining protein and mRNA quantification to decipher transcriptional regulation, Nature Methods **12**, 739 (2015).

[54] J. P. Bothma, H. G. Garcia, S. Ng, M. W. Perry, T. Gregor, and M. Levine, Enhancer additivity and non-additivity are determined by enhancer strength in the Drosophila embryo, eLife **4**, e07956 (2015).

[55] E. El-Sherif and M. Levine, Shadow Enhancers Mediate Dynamic Shifts of Gap Gene Expression in the Drosophila Embryo, Current Biology **26**, 1164 (2016).

MATERIALS AND METHODS

Fly husbandry and genetics

Drosophila fly lines expressing bcd-eGFP from [49] were used as the starting point. In all such lines, the endogenous Bcd was replaced with a null phenotype Bcd^{E1}. Stable stocks expressing NLS-MCP-mRuby3 ; Bcd-eGFP-bcd ^{E1} were created. Virgins from these stocks were then crossed with males expressing reporter constructs with the gene regulatory regions, while the gene body was substituted with MS2 stem-loop cassettes and LacZ.

For the synthetic enhancers, the following scheme was used: A 472 base pair (bp) fragment spanning the modified *hb* proximal enhancer and the *hb* P2 basal promoter was synthesized by IDT and ligated into the piB-hbP2-P2P-MS2-24x-lacZ-αTub3'UTR construct [50] between the restriction sites HindIII and NcoI. In the resulting reporter construct the hb promoter drives the expression of 24 copies of the MS2 loops and is followed by the lacZ coding sequence. The number of MS2 loops in the reporter was verified by Sanger sequencing. In the strong enhancer reporter, 8 suboptimal Bcd binding sites were converted to the consensus sequence TAATCC, resulting in a total of 11 strong Bcd binding sites. In the weak enhancer reporter, all 3 consensus sequence TAATCC were converted to the suboptimal Bcd binding site TAAGCT, resulting in a total of 11 weak Bcd binding sites. Both constructs were integrated into the 38F1 landing site on chromosome II of the fly line FC31 (y+); 38F1 (w+) using FC31 integrase-mediated cassette exchange [51]. All fly lines from which males were crossed and their sources are tabulated in TABLE I.

Sample preparation

Embryos were harvested on apple juice plates, using protocols mentioned earlier [49]. Staged two-hour-old embryos were dechorioned by hand by rolling them over a tape band (Scotch). Dechorioned embryos were placed on the lateral side on a mounting membrane lined with glue. The glue was prepared by submerging 10 cm of Scotch tape in 4 ml Heptane for 48 hours in a shaker at 37^oC. A drop of glue was placed on the mounting membrane, gently smeared evenly, and was then allowed to air dry, before placing the dechorioned embryos. After the embryos were placed on the membrane, they were submerged in a mixture of halocarbon oil (60% Halocarbon27, 40% Halocarbon700, Sigma), and then covered with a $25 \times 25\,\mathrm{mm}^2$ glass coverslip (Corning).

Imaging

Three-dimensional stacks of fluorescence images were acquired using the fast airyscan mode of Zeiss LSM 880 microscope, run by Zen Black 2.3, SP1 software. A Plan-Apochromat 63x/1.4 oil immersion objective (Zeiss) was used for all measurements. For GFP excitation, 488 nm line of Argon laser was used (140 μW), while mRuby3 excitation was done by a 561 nm diode pumped solid state laser (36 μW). Laser power at the back aperture of the objective was measured with a Thorlabs power meter at the beginning of each measurement session. A MBS 488/561 beamsplitter was used to combine the beams. Two sets of emission filters were used: BP 420-480 / BP 495-550 was used for GFP emission, while BP 495-550 / LP 570 was used for mRuby3. The effective emission wavelengths were 515 nm for GFP and 578 nm for mRuby3. Detector gain of 740 was used for all imaging cases. The voxel size was fixed at $43 \times 43 \times 200\,\mathrm{nm}^3$ for all 3D measurements. For 2D single-plane videos, however, a z section of 1000 nm was used. The frame times were 497 ms for each frame for both color channels, with a pixel dwell time of 0.744 μs. Each image frame was 1044×1044 pixels, or $45 \times 45\,\mu$m for the 3D acquisitions. No averaging was done. Imaging was done using the "Fast Airyscan" mode, and the final images were obtained after applying the "Airyscan Processing" within the Zen software.

Imaging was done on embryos in nuclear cycle #14, mostly from the 20th to the 35th minute after mitosis. The nuclei at the surface of the embryo facing the glass coverslip were imaged. A total z depth of $\sim 14\,\mu$m was imaged, split into 70 z-frames. This thickness was enough to scan entire nuclei, leaving some space at the top as well as the bottom. The z-stacks were imaged in four patches of $45 \times 45\,\mu$m for each embryo, at different positions along the AP axis. Each patch spanned across ~ 40 nuclei.

Embryo fixation

The fixation procedure aims to retain the clusters while immobilizing them in space. Since we are dealing with near diffraction limit puncta with a low signal-to-noise ratio, it is best to avoid artifacts typically introduced by stained antibodies, which are usually used to visualize fixed samples. Hence, we aimed to use the fluorescence of the monomeric eGFP in Bcd-GFP expressing fly embryos to visualize clusters. This presents us with two challenges: 1) preserving the fluorescence of GFP after fixation, 2) preserving the clusters themselves.

We solved both issues using exclusively freshly dissolved methanol-free formaldehyde (to a final concentration of 4%) for embryo fixation and by minimal exposure to organic solvents like Heptane, Methanol, or Ethanol. With these modifications to a standard protocol [52], fixation and visualization of Bcd clusters in the embryos were achieved.

Pixel cross-correlation

For this, the pixels along the x and y axes were separately autocorrelated with themselves. The `crosscorr` function of `MATLAB` was used for this purpose. This program uses the cross-correlation function, which is usually used to find the similarity between a time series and a lagged version of another. Here, we used either the same pixel row (x) or the same pixel column (y) in place of both the time series functions. Using pixel rows, (x), we get the correlation function (c) to be:

$$c_{x,x} = \begin{cases} \frac{1}{T} \sum_{t=1}^{T-k} (x_n - \bar{x})(x_{n+k} - \bar{x}) & \text{k=0,1,2, ...} \\ \frac{1}{T} \sum_{t=1}^{T-k} (x_n - \bar{x})(x_{n-k} - \bar{x}) & \text{k=0,-1,-2, ...} \end{cases} \tag{1}$$

This is repeated over all the rows and the average is then calculated. The columns are similarly treated and the average is calculated. The correlation lengths calculated along the rows or columns were the same for all images. The rows and the column data are then combined to obtain the overall image average. The average function was fitted with an exponential, $y(x) = a + b \cdot exp(-c \cdot x)$ and the "correlation length" was computed by $\lambda_{corr} = x_0 + \frac{log(2)}{c}$. Subsequently, the error in the correlation length is given by, $\sigma_\lambda = \lambda \cdot \frac{\sigma_c}{c}$. This operation is selectively done for either the pixels exclusively within or outside the nuclear masks in the images.

Local maxima detection

We locate local maxima in images of blastoderm nuclei indiscriminately, i.e. the resulting maxima may or may not represent a cluster. The process is broadly divided into two steps.

First, the nuclear pixels are selected and an Otsu thresholding is performed. Only the pixels above the threshold are retained and the rest are converted to "not a number", (NaN). This removes the background pixels, leaving the candidates for maxima. A new set of μ and σ are calculated from the resulting non-NaN pixels. The image obtained is then rescaled to an interval $[0,1]$. This process is repeated n times, iteratively to obtain a set of images $I^{i \in [1,n]}$.

In the second step of processing, local thresholding is done on the final image obtained from the previous step. For this, a filter window of 25×25 pixels is created. The image is then convolved with a normalized matrix, of the size corresponding to the filter window. This is done to apply a moving windowed mean (μ_k) filter on the image. Similarly, a moving standard deviation σ_k filter is applied to the same image. From this, we can create a matrix of local thresholds of $\mu_k + \sigma_k$. Pixels under values determined by this local threshold matrix are set to zero. Thus a new image is obtained. The image is rescaled between $[0,1]$ and the process is repeated iteratively m times, so that we have a set of images $I'^{\, i \in [1,n], \, j \in [1,m]}$.

To determine the optimal m and n iterations, we calculate the structural similarity index (ssim) values, by comparing each binarized image $I'^{\, i \in [1,n], \, j \in [1,m]}_{bin}$ to the binarized image $I'^{\, i \in [1,n], \, 1}_{bin}$. In-built `MATLAB` function `ssim` is used for this. The ssim value drops with each $j \in [1,m]$, while the total "spots" detected does not significantly change. This is because each local thresholding operation retains a slightly smaller area around the local maxima, than the previous iteration. This also enables the splitting of two neighboring local maxima peaks after sufficient iterations. This is reflected in the positive change in the number of spots detected as the total spot area decreases. This iteration i, j sets the optimal thresholding criterion for spot detection.

Localization of the maxima is trivial from the final images and can be done with very high accuracy, by binarizing and using `regionprops` to find the maxima centroids.

Pair auto-correlation

Consider randomly distributed points in space, described by Poisson process. In the same field, consider locations with significantly higher density of points which are also highly localized. This can be approximated by Gaussian point processes superimposed on a Poisson point process. To estimate the average density and effective size of these regions with Gaussian processes, we employ the pair correlation function [31].

The density function of the points expressed in polar coordinates is given by $\rho(\vec{r})$. We can calculate the pair correlation function for such a point distribution by [31]:

$$g(\vec{r}) = < \rho(\vec{R})\rho(\vec{R} - \vec{r})) > /\rho^2 \tag{2}$$

Here, ρ is the average density. In practice, this correlation function can be calculated using Fast Fourier Transforms applied to an image I containing the point distribution.

$$g(\vec{r}) = \frac{1}{\rho^2} \cdot \frac{FFT^{-1}(|FFT(I)|^2)}{FFT^{-1}(|FFT(W)|^2)} \tag{3}$$

The I is a sparse matrix with 1 s at the locations of the maxima and 0 s elsewhere. The quantity, W is a window matrix adjusted to fit within the area of a nuclear cross-section taken as a convex hull.

The time projection of Bcd-GFP local intensity maxima has both randomly dispersed points as well as focal accumulations of points in space. The randomly distributed points can be considered to be representative of a Poisson process. The accumulations, however, can be considered Gaussian functions, convolved with hypothetical singularities. If we consider the density of the Poisson process to be 1, and the Gaussian process peak density to be ρ' above 1, we get the expression:

$$g(r) = \rho' exp(-r/\sigma^2) + 1 \tag{4}$$

Here σ denotes the size of the focal accumulation of the maxima, representing the Gaussian processes. This expression can then be used to fit the autocorrelation function to derive the effective width of the function as well as infer the increase in density within these Gaussian accumulations.

Finding a nucleus' position

To find the position of a nucleus in the embryo, a coordinate system relevant to the embryo needs to be defined first. First two images are acquired, one of the anterior and the other of the posterior ends of the embryo. The imaging is done on a single z-plane passing through the midsagittal plane of the embryo. The imaging conditions are kept the same as that of the nuclear images. This gives us the location of the anterior (x_0, y_0) and posterior (x_L, y_L) tips of the embryo (L signifying the length of the embryo) in the microscope stage coordinates. The line joining these two points represent the anterior-posterior (AP) axis of the embryo, and is the x-axis in the embryo co-ordinate system (X'). The perpendicular to this line, passing through (x_0, y_0) is the y-axis (Y'). Hence, the anterior end is $(0, 0)$ and the posterior end is $(L, 0)$ in the embryo coordinates.

Now, if the centroid of a nucleus is (x_i, y_i) in the microscope stage coordinates, the distance of the nucleus from (x_0, y_0) is $r_i = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2}$ and the angle made by r_i with the AP axis is given by $\theta_i = arctan(y_i - y_0)/(x_i - x_0)) - arctan((y_L - y_0)/(x_L - x_0))$. Hence, the location of the nucleus in the X' coordinates is given by $x'_i = r_i \cdot cos(\theta_i)$, or:

$$x'_i = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2} \cdot cos \left(tan^{-1} \left(\frac{y_i - y_0}{x_i - x_0} \right) - tan^{-1} \left(\frac{y_L - y_0}{x_L - x_0} \right) \right) \tag{5}$$

This value can be computed for each nuclear controid. For pooling nucleus by position, nuclei with x'_i within the position bin edges are accumulated.

Segmentatation of a "filled" nucleus.

Automated segmentation of Bcd-GFP expressing nuclei was done using the following scheme: First, the raw images were contrast adjusted using `imadjustn`, then filtered with a median filter, `medfilt3`, followed by a Gaussian filter,

`imgaussfilt3`. A cuboidal structural element was then used for a series of morphological transformations to the resulting images. First, an erosion was applied (`imerode`), followed by a reconstruction, (`imreconstruct`), and dilation, (`imdilate`). The complement of the reconstruction transformation was then blurred with a Gaussian filter (`imgaussfilt3`). The resulting image was closed (`imclose`) and eroded. A binary mask for the nucleus pixels was subsequently obtained. Finally, watershed segmentation (`watershed`) was applied to separate any conjoined neighboring nuclei. Labels are then assigned to the nuclear masks.

Locating clusters (in 3D)

A detection technique 2D local maxima localization was introduced earlier. This technique indiscriminately detects all local intensity peaks, which include noise spikes along with protein clusters. The difference between a "real" puncta and a "noise related local maximum" is that the spot size for a real puncta is at least as large as the point spread function of the microscope, while a noise related maximum is likely to be smaller. To capture this difference, we chose a pixel size which oversamples a diffraction limited spot. With an x-y size of $43\,nm$, a diffraction limited spot is spanned by $4-5$ pixels in any direction. Hence, a cutoff limit of 3 pixels should effectively differentiate a real puncta from a noise related maximum. For z-slices, we chose a thickness of $200\,nm$. The PSF width along the z direction being larger than $500\,nm$, any maxima that is does not span atleast 2 z-slices is likely a noise related maxima. Hence we chose 2 pixels as the cutoff limit along the z axis.

At this point, one might point towards the fact that for live imaging of mobile structures (like subdiffusing clusters in our case), the likelihood of detecting a cluster in two consecutive frames is subject to the frame rate. To account for this, we refer to FIG. S4F, where we see that the detection probability of a cluster in frames imaged $\sim 500\,ms$ apart is $> 70\,\%$. Hence, for z-slices imaged $\sim 500\,ms$ apart with thickness less than the PSF width, we should be able to detect the cluster with a high degree of reliability.

To identify only relevant puncta-like entities in the nucleus, we employed the following technique on the raw images of Bcd-GFP nuclei. First, morphological top-hat filtering was applied using a "disk" as the structural element to the 3D raw images of the nuclei using `imtophat`. The transformed image thus obtained was used to detect local intensity maxima peaks. For this, the top 1 percentile pixels within a nucleus were selected from the transformed images. Joined neighboring spots were then separated by applying a watershed algorithm (`watershed`).

It must be noted that the "spots" thus detected are not the real spots, but rather peaks detected in the morphologically transformed image. However, it can be argued that the centroid of the local maxima peaks from the raw images are preserved through this transformation. Hence the detected "spots" can be used to locate the real cluster peaks. This information can be extracted by using the `regionprops3` function. For this, the spot mask is obtained from the spot segmentation in the morphologically transformed image, and the mask is then applied to the raw image. Intensity weighted centroids in 3D of the voxels within each mask are then calculated using `WeightedCentroid` on the raw image. This gives the peak position of each cluster.

However, not all clusters thus detected are retained for further analysis. A size-based thresholding (as mentioned above) is then performed such that if the x-y cross section of a detected spot is less than 3×3 pixels wide and the z depth is not at least 2 pixels wide, the spot is discarded. That brings the threshold volume to $3 \times 3 \times 2 = 18$ pixels. A corresponding effective spot diameter d can be calculated from the threshold volume, such that $d = (6/\pi \times vol)^{1/3}$. The threshold diameter turns out to be 3.25 pixels wide, which converted to absolute units gives, $138.2\,nm$, which is significantly less than the 3D PSF of the microscope. Thus, using this technique we identify the locations (only) of the "real" puncta in the Bcd-GFP nuclei.

Fitting clusters

Cluster fitting is done on the raw image pixels to extract cluster relevant parameters like size and peak intensity. First, the plane corresponding to z coordinate of the cluster centroid is chosen and pixels within a square window with the cluster centroid as the center are extracted. The window length set at $2w + 1$ pixels, where w is ~ 12 pixels for a typical window size of $\sim 1 \times 1\,\mu m^2$. A 2-dimensional Gaussian function is then used as the fitting function to fit the intensity profile within the window [53]. Fitting is done using `lsqcurvefit`, employing the `levenberg-marquardt` algorithm. The fitting equation is:

$$f(x,y) = I_p exp(-(a(x-x_0)^2 + 2b(x-x_0)(y-y_0) + c(y-y_0)^2)) + I_{bg} \tag{6}$$

where,

$$a = \frac{cos^2\theta}{2\sigma_1^2} + \frac{sin^2\theta}{2\sigma_2^2}$$
$$b = \frac{sin2\theta}{4\sigma_1^2} + \frac{sin2\theta}{4\sigma_2^2} \qquad (7)$$
$$c = \frac{sin^2\theta}{2\sigma_1^2} + \frac{cos^2\theta}{2\sigma_2^2}$$

The pixel value at the center pixel of the window is used as the original guess for I_p, and a guess of $I_{bg} \approx I_{nuc}$ are used. The initial guesses for both σ_1 and σ_2 are $w/2$, while $\theta = 0$ is used for the rotational angle. The bounds for the σs are set to w^2 and the bound to x_0, y_0 are set to $-w : w$. The bounds for θ is $0 < \theta < \pi/4$. The cluster candidates whose fitting parameters do not satisfy the criteria are automatically discarded.

Cluster parameters

We consider three cluster parameters here: cluster size, cluster peak intensity and cluster total intensity. An effective cluster size can be estimated from the σ obtained from the 2D Gaussian fit. While a cluster is a three-dimensional entity spanning over multiple imaging sections, we only do a two-dimensional fitting of the intensity profile of the plane passing through the intensity-weighted center along the z-axis. The imaging conditions are such that the resolution along z is significantly worse than the resolution along the x-y axes (worse by ~ 5 times). Therefore, any fitting along the z-axis will yield significantly higher error.

To calculate the effective size of clusters we consider the Gaussian spread along x and y directions (σ_x, σ_y). An expression for an effective radius (r_{eff}) can be obtained from these by:

$$r_{eff} = \sqrt{\sigma_1^2 + \sigma_2^2} \qquad (8)$$

as the effective size of the three-dimensional cluster. In the above calculation, the Gaussian width is assumed to be a projection on a section of the obloid representing the real spot.

The peak intensity, I_c is given by the peak intensity, I_p. The total intensity, or the area under the 2D Gaussian gives a measure for the total number of molecules in the clusters. The total intensity is given as:

$$I_m = 2\pi I_c \sigma_1 \sigma_2 \qquad (9)$$

Hence, the expression for total intensity is independent of the expression for cluster size, even though it contains the expression for σs. Another thing to be noted is that I_c and I_{bg} are separately obtained from the fits, and hence the I_c is automatically background corrected. Bounds of the fitting parameters are supplied and failure to return parameters within the bounds render the fits unusable, however, it is a rare occurence $(< 2\,\%)$.

Slope calculation

To calculate the properties of the average cluster in a nucleus, the nuclear average of the cluster properties are first calculated. The average data is pooled across all embryos and then plotted against the respective nuclear positions (x/L) or average nuclear concentration (I_{nuc}). This data can be then divided into position or nuclear concentration bins. The concentration or the position bins were evaluated by fixing the total number of bins and accumulating data points within the bins by using `histcounts`. The mean and standard deviations of the nuclear average data per bin are found by bootstrapping. It needs to be emphasized that this does not reflect the standard deviation of the within a nucleus, but rather the standard deviation within a bin.

The average cluster properties like I_c, d and I_m (or their respective natural logarithms) are plotted against the nuclear average concentration (or the nuclear positions). The data is then fitted with a linear regression model and the coefficient of determination, R^2, the slope of the linear fit as well as the error in slope are obtained from the fit parameters.

The cluster properties are assumed to linearly depend on the nuclear concentration, while the dependence with the nuclear position is assumed to exponential. Hence, the slope of the linear fit of the log of cluster properties plotted against the nuclear positions can be used to calculate the exponential decay constant: $\lambda = -(1/slope) \pm \sigma_{slope}/slope^2$.

Error in nuclear property estimation using cluster property

To calculate error in nuclear concentration estimation σ_c using cluster properties like I_c, d and I_m, we use the slope obtained from the linear fits (s) and the error in the cluster property estimation in each concentration bin σ_i. We have $\sigma_c = \sigma_i/s$, while the error associated with each σ_c is $\sigma_{\sigma_c} = \sigma_c \cdot \sqrt{(\sigma_{\sigma_i}/\sigma_i)^2 + (\sigma_s/s)^2}$. Here, σ_{σ_i} is the error associated with the cluster property determination and σ_s is the error in slope. Similarly, the error in nuclear position estimation σ_p using cluster properties can be computed by using the λ as follows: $\sigma_p = \lambda \cdot \sigma_i/i$, while the error in σ_p is given by $\sigma_{\sigma_p} = \sigma_p \cdot \sqrt{(\sigma_{\sigma_i}/\sigma_i)^2 + (\sigma_i/i)^2 + (\sigma_\lambda/\lambda)^2}$. All these expressions are derived by using laws of error propagation.

Empirical fitting function for cluster count

We can empirically arrive at a model to depict the concentration dependence of the number of clusters detected per nucleus. We base our model on the probability of a cluster being bound to the seeding site. This is given by $p_{on} = t_{on}/(t_{on} + t_{off})$, where t_{on} is the time for which the cluster is "bound", and hence detectable, and t_{off} is the time for which the cluster is "unbound", and hence not detectable. To calculate these, we make three simplistic assumptions. 1. The total number of seeding sites in a nucleus is fixed, N. 2. The "bound" time, t_{on} for a cluster i is a property of the seeding site only, $n(i)$. 3. The "unbound" time, t_{off} is a property of the diffusion parameters and is inversely proportional to the nuclear concentration, c, and hence to I_{nuc}. Therefore $\sum_{i=1}^{N} p_{on}(i,c)$ gives the total number of clusters detected per nucleus.

For the fitting function in FIG. S7E, $N = 80$ has been used, and $n(i)$ has been randomly generated.

Segmentation of a "hollow" nucleus

Segmentation was also done to the MCP-mRuby3 expressing nuclei. Contrary to the Bcd-GFP expressing nuclei, the MCP-mRuby3 expressing nuclei are devoid of fluorophores and hence have a lower signal than the internuclear space. To segment such nuclei, the brightness of the images was first adjusted using `imadjustn`, and a three-dimensional median filter was applied. This was followed by a Gaussian filter and an extended maxima transformation was done (`imextedndedmax`). The resulting image was then binarized and the inverse of the resulting binary image was created. subsequently, a three-dimensional kernel was used to convolve the binary image. A threshold was then applied to the resulting image to which watershed segmentation was finally applied. the resulting image was opened with a cuboidal structural element, and any holes within the bright structures were filled to provide the final mask for the nuclei.

A "hollow" nucleus can be mapped to a "filled" nucleus if the corresponding centroids lie within a distance of less than half a nuclear length from each other. If the "hollow" and the "filled" masks cannot be mapped onto one another, the nucleus is discarded.

Transcription hotspot detection

The transcription hotspots appear as distinctly bright spots in the nucleus against a significantly dark background. First, the nuclear bounds are found using the hollow nuclear segmentation scheme. To identify the transcription hotspots within the "hollow" nucleus, a Difference of Gaussian (DoG) algorithm was first applied to the raw images. The image obtained was then convolved with the raw image, and rescaled. A threshold based on the intensity of the nuclear pixels, (I_{nuc}) was applied such that any pixels with an intensity $< (4\sigma_{I_{nuc}} + \mu_{I_{nuc}})$ were discarded. Following this, the masks of potential transcription hotspots were obtained to which a size cutoff of 18 pixels was applied.

Radial intensity profile and coupling fraction calculation

The intensity-weighted centroids of the transcription hotspots were obtained using `regionprops` after applying the transcription hotspot masks to the nuclear pixels in the MCP channel (mRuby3). The intensity of Bcd-GFP is calculated as a function of distance from the transcription hotspot centroid, I_r. For this, only the intensity along the

x-y plane passing through the hotspot centroid is considered. I_r is then obtained by averaging the values of pixels lying within the ring r and $r + 0.1\ \mu m$. Data for each I_r is pooled from multiple nuclei across many embryos to calculate the average I_r profile.

The accumulation radius for a gene is calculated from its average I_r profile. For this, the profile data is fitted with a double Gaussian function:

$$f(x) = k_0 e^{\left(-\frac{(x-x_0)}{a_0}\right)^2} + k_1 e^{\left(-\frac{(x-x_1)}{a_1}\right)^2} \tag{10}$$

The accumulation radius r_0 is simply the FWHM width of the first of the two Gaussians and the error in radius is calculated from the fitting error. The distance of the intensity-weighted centroid of the closest Bcd-GFP cluster is measured from the centroid of the transcription hotspot. A histogram of all such distances is plotted and the cumulative probability function is calculated directly from the histogram or a spline fit to the histogram. The cumulative probability value at $x = r_0$ gives the coupling fraction for a gene.

Alternatively, the histogram of the nearest neighbor TF cluster distances from the mRNA hotspot can be fitted with double Gaussians. The first peak corresponds to the nearest neighbor cluster which is coupled to the gene, and the second, weaker peak is that of the cluster which is the nearest uncoupled cluster. The intersection of these two Gaussian fits marks the boundary $r_0^{'}$, such that only clusters inside this boundary are coupled.

The radius r_0 comes partially from broadening, as the chromatin is not stationary but rather subdiffusive in the nuclear space. This incorporates motion blurring of point sources during video capture. another source of broadening is the fact that the MS2 hotspot is not a point source itself. A hotspot consists of multiple MS2 stem loops, which are spread across the gene body and might span a few kilobases at any moment. Also, along with the nascent mRNA the stem loops project out of the gene body and have their degrees of freedom. Also, since transcription is a kinetic process, the stem-loops linearly traverse along the gene body with the speed of transcription (about $2kB/min$). These factors combined are the biological source of broadening of the MS2 hotspot signal. Hence, the hotspot can be approximated as a point source convolved with a Gaussian that incorporates all forms of broadening.

Estimation of molecules per cluster and the total cluster fraction

Estimation of the total Bcd molecules in the nucleus has been done in a previous study using Western blots [21]. However, the construct was such that the Bcd concentration was uniform throughout the embryo instead of decaying exponentially along the axis. However, since the total Bcd molecules in the case of the flat expressing line remain the same as the total expressing molecules in the wild-type gradient, we can arrive at a relation equating the two quantities as:

$$N_{0_{flat}} \int_0^L dx = N_0 \int_0^L e^{-x/\lambda} dx \tag{11}$$

Here, $N_{0_{flat}}$ is the average concentration of Bcd in a nucleus in the flat expression lines used in the study, while N_0 is the concentration at $x = 0$ for the wild type line, expressing an exponential gradient with a length constant of λ.

Solving this, we get $N_0 = N_{0_{flat}} L/\lambda$. Using $N_{0_{flat}} \sim 8000$, and $\lambda \sim 0.2$, we get $N_0 \equiv 40000$ molecules. Approximating an average nuclear diameter to $5\ \mu m$, the average density of Bcd molecules in the nucleus turns out to be 600 molecules/μm^3. Considering that an average cluster is 2.2 times higher Bcd concentration than the nucleoplasm, the Bcd concentration inside a cluster at the anterior of the embryo turns out to be 1320 molecules/μm^3. The average diameter of a cluster is $\sim 0.4\ \mu m$, making the average volume, $0.03\ \mu m^3$. Knowing the molecular density within a cluster and the cluster volume, we can calculate the number of molecules per cluster as 37 in the anterior of the embryo. We show an estimate of the molecules per embryo along the embryo axis in FIG. S11C.

Derivation of the concentration sensing limit

We consider the cluster to be a sphere with an effective diameter d, the concentration of Bcd molecules inside the cluster is c_{clust}, and the diffusion constant of Bcd is represented by D. The fractional error δN in counting N molecules by the cluster is given by

$$\frac{\partial N}{N} = \left(\frac{6}{D \cdot d \cdot c_{clust} \cdot T_{clust}}\right)^{1/2} \tag{12}$$

Rearranging EQ. 12 we get

$$T_{clust} = \frac{6}{D \cdot d \cdot c_{clust}} \left(\frac{\partial N}{N} \right)^{-2} \tag{13}$$

The time, T_{clust} in EQs. 12 and 13 represent the duration for which a cluster needs to "make the measurement" to interpret the nuclear concentration with an accuracy of $\left(\frac{\partial N}{N} \right)$.

The equivalent time a binding site of length a would take to measure the nuclear concentration is given by:

$$T_{en} = \frac{1}{D \cdot a \cdot c_{nuc}} \left(\frac{\partial N}{N} \right)^{-2} \tag{14}$$

Here, c_{nuc} is the nuclear concentration. Using EQs. 13 and 14 we obtain the ratio:

$$\frac{T_{en}}{T_{clust}} = \frac{d}{6a} \left(\frac{c_{clust}}{c_{nuc}} \right) \tag{15}$$

The length of a typical binding site can be considered to be $a = 3.4nm$ [34]. However, d, c_{nuc}, and c_{clust} are dependent on the nuclear position, making T_{en}/T_{clust} a position dependent quantity.

The point spread function

To compute the Point Spread Function of the microscope (PSF), 100 nm mounted fluorescent polystyrene beads (Thermo Fisher Catalog# T14792) were imaged with 1000 nm z-slice thickness with the Airyscan mode on Zeiss LSM 880 microscope. Pixel cross-correlation was done on the images. The correlation length gave the point spread function of the microscope.

I. SUPPLEMENTAL FIGURES
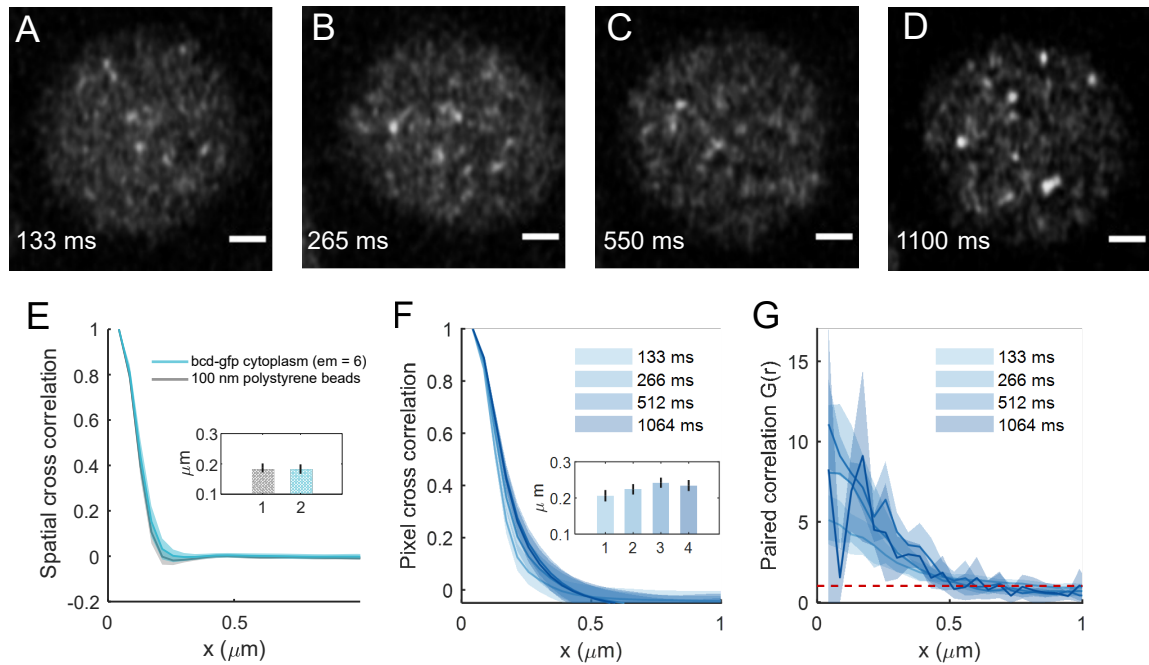
FIG. S1. **Imaging frame times and correlation lengths.** (A-D) Representative images of nuclei expressing Bcd-GFP, at different frame times (mentioned on panels). The frame time is increased by increasing the pixel dwell time of imaging. Increasing pixel dwell time blurs the more transient (or mobile) features into the background, leaving only the brighter, more stable ones in the image. The scale bars are $1\mu m$ long. (E) Pixel cross-correlation plotted for different frame times. (mean and standard error of mean are shown as error bars in the inset). The correlation lengths are statistically indifferent, indicating that motion induced blurring does not cause significant broadening of resolvable structures even at the frame time of $\sim 1\,s$ (F) Pixel correlation plot of cytoplasmic Bcd-GFP. The correlation length is $0.187 \pm 0.015\,\mu m$, a value comparable to the PSF width, indicating that Bcd in the cytoplasm is predominantly diffusive.

Bcd-GFP (confocal)   Bcd-GFP (airyscan)

Cic-Ven (airyscan)   Gro-GFP (airyscan)   His2B-RFP (airyscan)
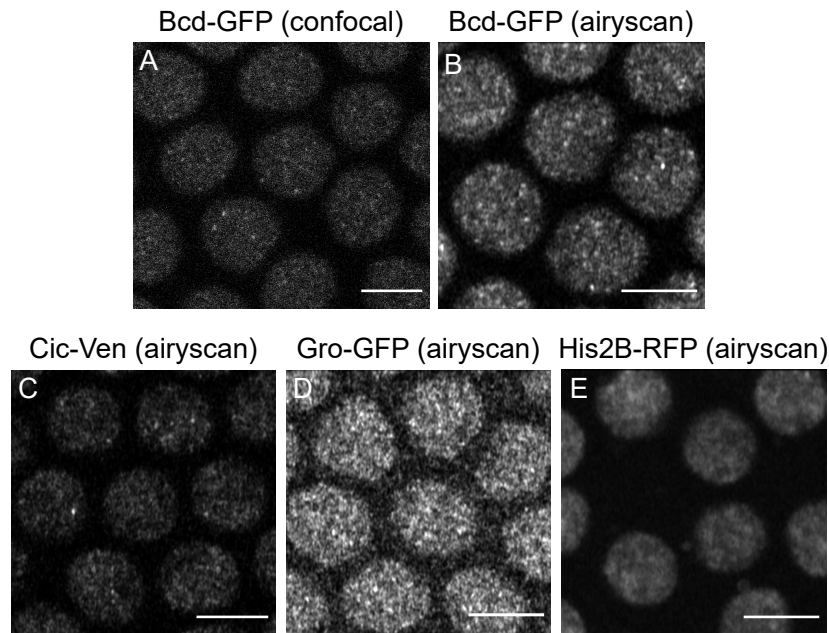
FIG. S2. **Micrographs of labeled proteins in the nuclei of _Drosophila_ embryo.** (A-B) Traditional confocal (A) as well as _Airyscan_ confocal images showing cross sections of Bcd-GFP expressing nuclei of live NC14 embryos. (C-E) _Airyscan_ confocal images showing nuclei expressing Capicua tagged with Venus, Groucho tagged with monomeric eGFP, and Histone2B tagged with RFP in live NC14 embryos.
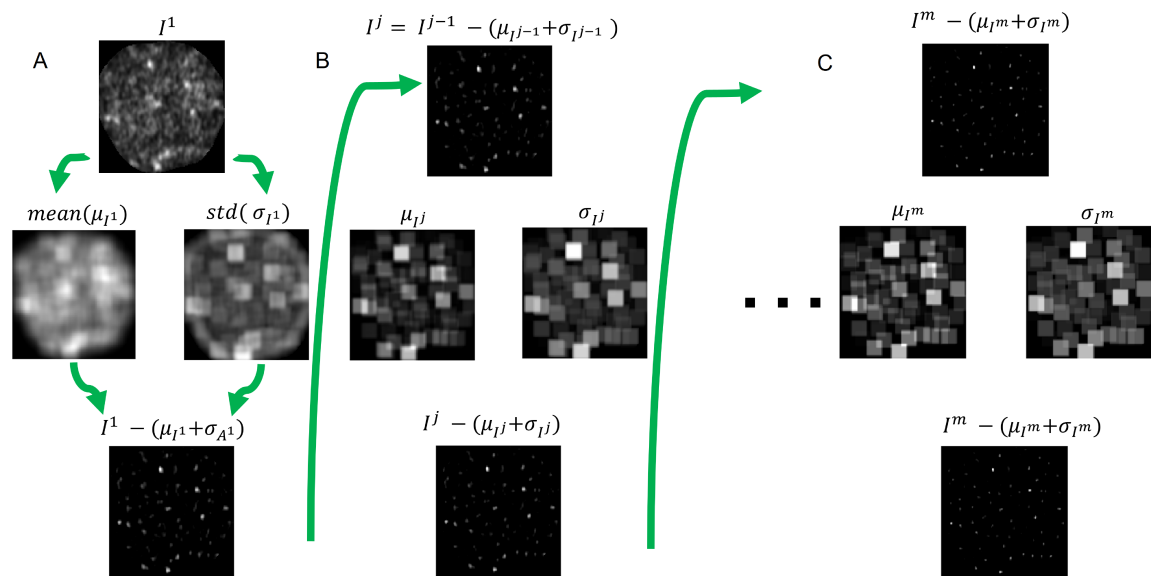
FIG. S3. **Local maxima detection technique.** (A) Image shows the raw image of a nucleus expressing Bcd-GFP (TOP). Windowed mean filtered image derived from (MIDDLE, LEFT) and windowed standard deviation filtered image derived from the top (MIDDLE, RIGHT). The image obtained after adding the images in the MIDDLE row and subtracting the result from the image at the TOP is shown in the BOTTOM. This image is the result for one layer of local filter application, and is used for the second layer of local filter application in B. (B) The image obtained in (A, BOTTOM) is recursively filtered, following the same scheme in A to obtain B (BOTTOM). This operation is repeated n times to obtain the final filtered image in C (BOTTOM).
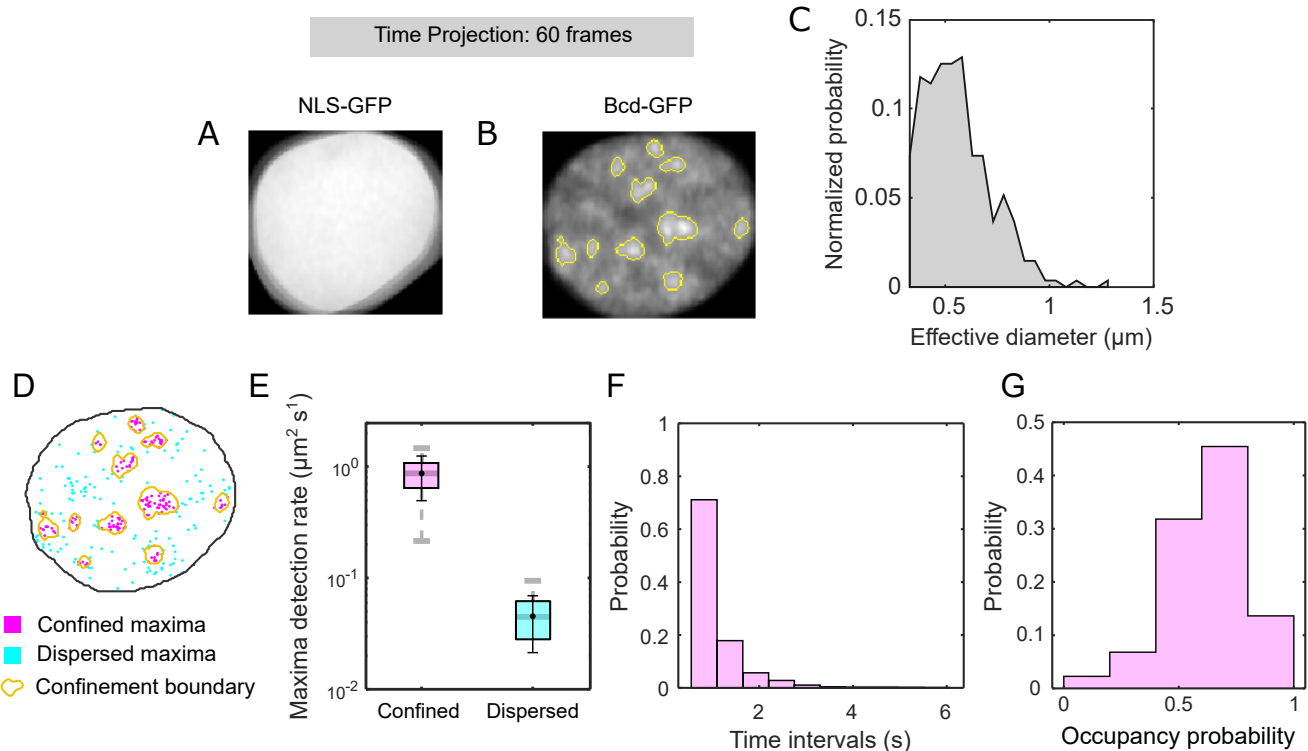
FIG. S4. **Characterization of local maxima accumulation over time.** (A, B) Each frames of movies of nuclei expressing (A) NLS-GFP and (B) Bcd-GFP (60 video frames, ~ 30 seconds) are rescaled to $[0, 1]$. The nuclei are then registered and projected onto a single frame. The projection images in A and B show clear visual differences. The image in A is highly uniform, suggesting that any heterogeneity in individual frames are random and average out over time. Whereas, the image in B shows distinct subnuclear accumulations. These accumulations are segmented and the accumulation boundaries are shown as yellow line overlays. These boundaries represent confinement areas within which local maxima tend to appear in high frequency in the movies. (C) If the confinement areas in B are approximated by circles, the histogram of the respective diameters is shown in C. The $mean \pm std$ of the diameters of the confinement area is $529 \pm 164\ nm$. (D) We can now revisit the local intensity maxima map projected from all the video frames (FIG. 1E, left). Utilizing the confinement area boundaries (in yellow), we can segregate the maxima into either confined (magenta) or dispersed (cyan) maxima. Visually, the magenta maxima occur with higher spatial density than the cyan ones. (E) This disparity in density can be quantified. Box plots show the rate of detection of confined (magenta) and dispersed (cyan) maxima in the projection maps (44 nuclei, 12 embryos). Boxes extend from the 25th to the 75th percentile, while the horizontal divider marks the median. Mean and standard deviations are overlaid in black (0.87 ± 0.37 and 0.04 ± 0.02 for confined and dispersed respectively). These represent the spatial density of maxima, although, the density is obtained by projecting all maxima over the time span of the video. Hence, the spatial density of maxima per unit time can be obtained by divided by the total time of the video (30 s). Thus we obtain the unit of $\mu m^2/s$. (F) Data in E accounts for the density of maxima accumulated over the length of a video, without presenting any estimate for the persistence of a single maximum within the confinement areas. To quantify persistence, we look for the time interval between the appearance of at least one maximum inside a confinement area. The histogram, shows that a maximum is observed inside a confinement area every consecutive frame ($\sim 500\ ms$) with a probability of ~ 0.7. This implies that a maximum can be continuously detected inside a confinement zone for a period of atleast 30 s. The time limit of 30 s arises out of chromatin motion which causes DNA bound spots to drift out of the plane in that time frame. (G) To summarize, we draw this histogram, which shows that the overall fraction of time for which a confinement area is occupied during the time span of 30 s. The mean obtained from this histogram is $\sim 65\%$.
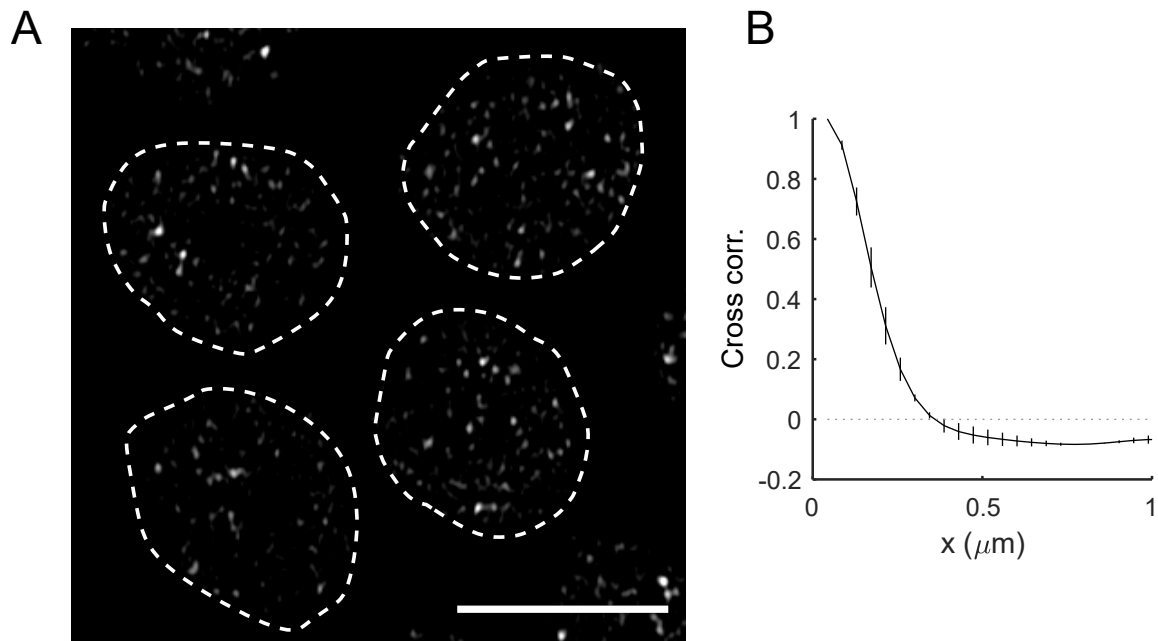
FIG. S5. **Bcd-GFP clusters in fixed embryos.** (A) Image shows cross section of nuclei expressing Bcd-GFP fixed using Formaldehyde. The dashed lines are guides to the eye for the nuclear boundaries. The scale bar is $5\mu m$ long. (B) Figure shows pixel cross correlation function (mean±std, n = 15 nuclei) on the nuclear pixels of a fixed Bcd-GFP embryo. Exponential fit to the plot gives a correlation length of $0.24 \pm 0.01 \, \mu m$, which is similar to the correlation length in live embryos.
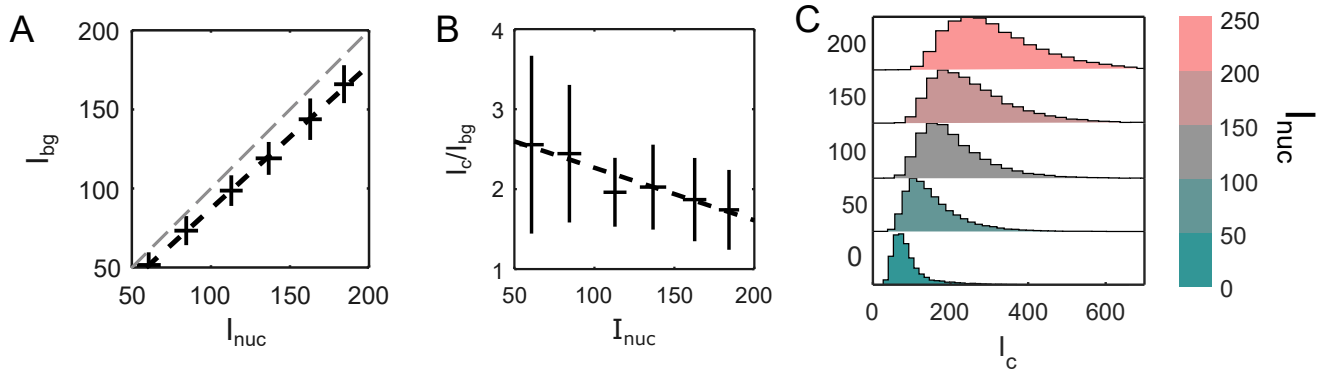

FIG. S6. **Molecular enrichment within clusters.** (A) Plot of average Bcd-GFP cluster background (I_{bg}) against average nuclear Bcd-GFP intensity ($R^2 = 0.95$). The dashed grey line represents $y = x$. (B) The nuclear average of the ratio of the peak cluster intensity to the respective background intensity (I_c/I_{bg}), plotted as a function of the average nuclear intensity I_{nuc}. The ratio is higher for lower nuclear concentrations. (C) Histograms showing the distribution of cluster peak intensities, I_c for different average nuclear intensity bin, I_{nuc}s. The color bar represents ranges of I_{nuc}
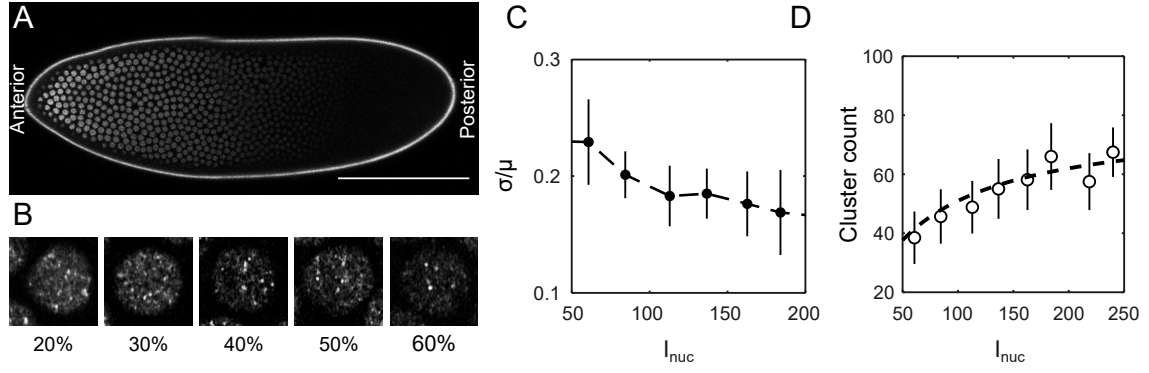
FIG. S7. **Cluster count.** (A) Cross-section of a *Drosophila* embryo in nuclear cycle #14 showing nuclei expressing Bcd-GFP. the imaging plane was near the surface of the embryo. The anterior and posterior ends of the embryo are marked. The higher GFP intensities in the anterior nuclei indicate a higher concentration of Bcd. The scale bar is $100\,\mu$m. (B) Left-to-right are zoomed images of cross-sections of individual nuclei expressing Bcd-GFP. The horizontal positions of the nuclei, measured from the anterior end are expressed as the percentage of embryo lengths below the respective panels. Images are $8 \times 8\,\mu$m in size. (C) The coefficient of variation (CV) of the cluster count per nucleus expressed as the standard deviation normalized to the mean values. Values are computed from pooled data (14 embryos, 2027 nuclei) discretized into nuclear Bcd-GFP intensity (I_{nuc}) bin. With an average CV of $< 20\%$, the cluster count displays remarkable reproduciblility across embryos. (E) To speculate about the nuclear concentration (I_{nuc}) dependence of the cluster counts, we employ an emperical relation (described in methods). The dashed line shows the fit of the cluster count data using that empirical relation.
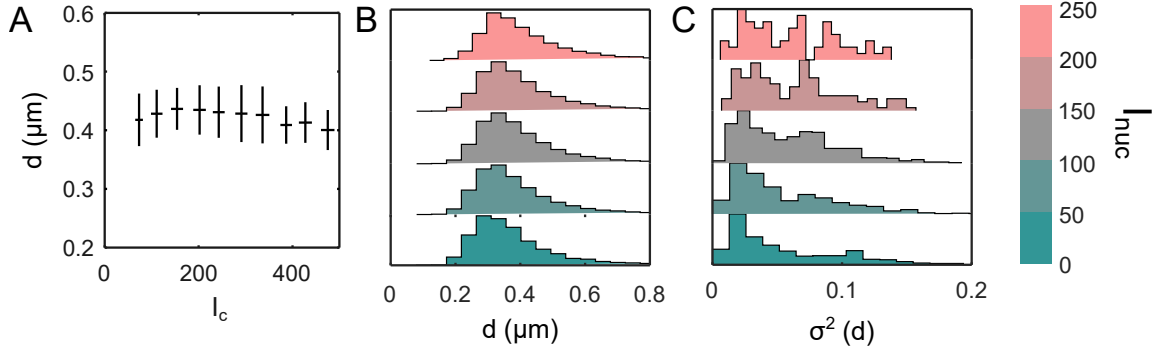

FIG. S8. **Cluster size.** (A) Is the cluster size dependent on the concentration of the cluster? To answer this, this figure shows the dependence of the cluster diameter, d on the cluster intensity, I_c. The cluster size, d is independent of I_{nuc}. (B) Clusters are sorted according to the I_{nuc} of their parent nucleus, and the histograms of cluster diameter, d are plotted for clusters belonging to various I_{nuc} range. The I_{nuc} range is indicated by the colorbar to the right of C. Interestingly, the histograms remain similar across a large range of I_{nuc}. This indicates that the cluster size is indeed independent of the nuclear concentration. (C) Here we visualize the variability of of the variance in cluster size within a nucleus. This figure shows the histograms of variances of cluster size per nucleus. The nuclei are sorted according to the range of I_{nuc}. (Colorbar to the right indicates the I_{nuc} range).
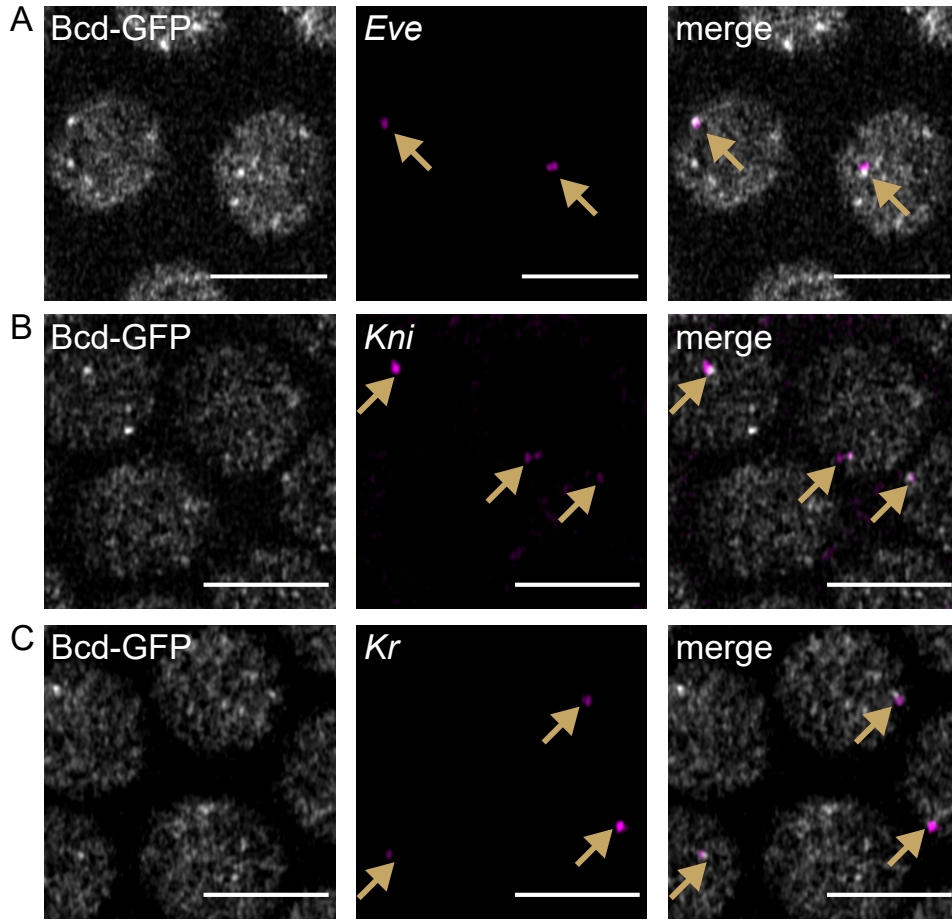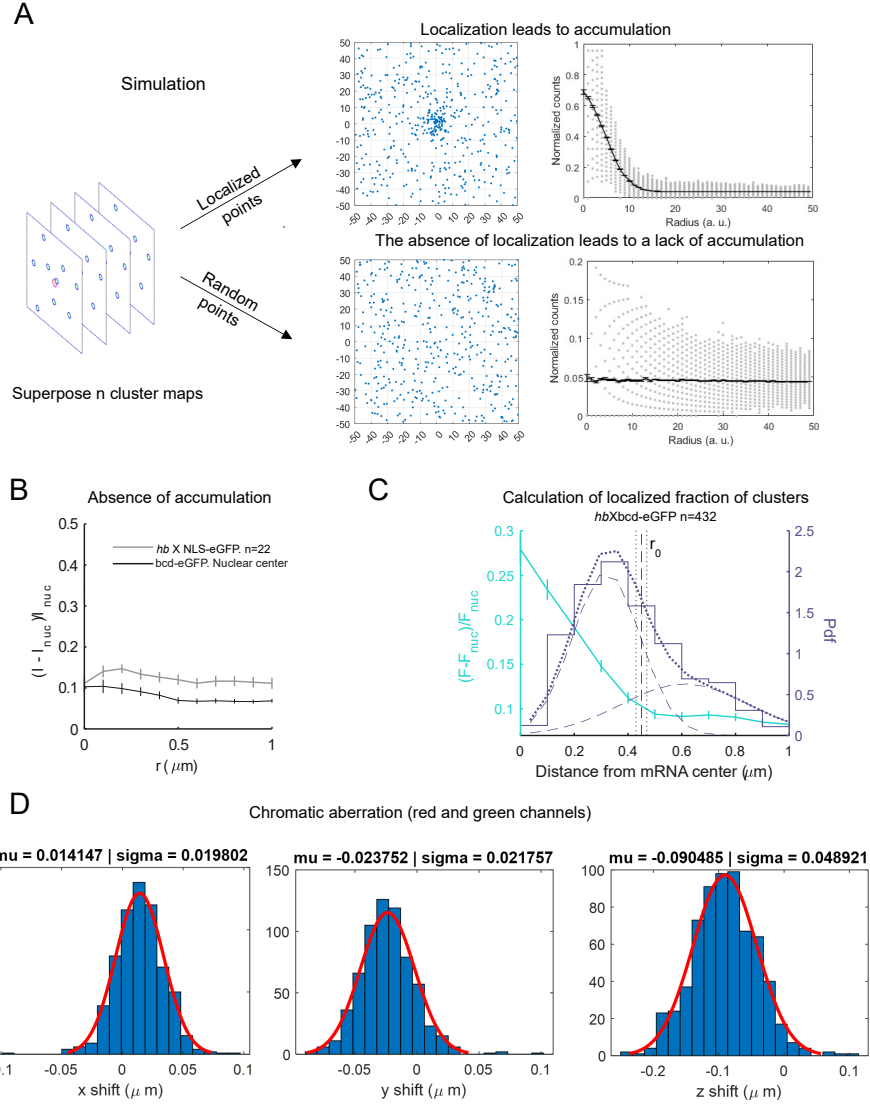
FIG. S9. **Accuracy of nuclear position determination using various cluster parameters.** (A) Natural logarithm of Bcd-GFP peak cluster intensity, I_c expressed as a function of fractional egg length x/L. The exponential decay constant extracted from linear fit is $\lambda = 0.32 \pm 0.03$ EL. (B) The Coefficient of Variation (CV) of I_c expressed as the coefficient of variation. The average CV is $\sim 30\%$. (C) The precision of nuclear position determination using the average I_c is $\sim 10\%$, which is equivalent to ~ 4 nuclear width. this shows that I_c is less precise in determination of the nuclear position that I_m. (D) Natural logarithm of Bcd-GFP peak cluster intensity, I_{bg} expressed as a function of fractional egg length x/L. The exponential decay constant extracted from linear fit is $\lambda = 0.25 \pm 0.02$ EL, which is statistically similar to the gradient of I_{nuc}. (E) The Coefficient of Variation (CV) of I_{bg} expressed as a function of x/L. The average CV is $< 20\%$. (F) The precision of nuclear position determination using the average I_{bg}. The precision is $\sim 4\%$, which is equivalent to ~ 1.5 nuclear width, which is equivalent to the precision in I_{nuc}. (G) Natural logarithm of Bcd-GFP peak cluster intensity, d expressed as a function of fractional egg length x/L. The exponential decay constant extracted from linear fit is $\lambda = 10.7 \pm 0.9$ EL. (H) The Coefficient of Variation (CV) of d expressed as a function of x/L. The average CV is $\sim 10\%$. (I) The precision of nuclear position determination using the average d. The precision is $\sim 100\%$, which is equivalent to ~ 1 embryo length. This shows that d can not sense nuclear position.

FIG. S10. **Accuracy of nuclear concentration determination using various cluster parameters.** (A) Total cluster intensity, I_m expressed as a function of nuclear Bcd-GFP intensity I_{nuc}. (B) The Coefficient of Variation (CV) of I_m expressed as a function of I_{nuc}. The average variability is $\sim 15\%$. (C) The precision of nuclear concentration (I_{nuc}) determination using the average I_c is 0.23 ± 0.07. However, the precision in concentration prediction is $\sim 15\%$ for the anterior nuclei. (D) Peak cluster intensity, I_c expressed as a function of I_{nuc}. The slope of the linear fit is $slope = 1.19 \pm 0.10$ EL. (E) The CV of I_c expressed as a function of I_{nuc}. The average CV is $\sim 25\%$. (F) The precision of nuclear concentration (I_{nuc}) determination using the average I_c is 0.43 ± 0.20, equivalent to $\sim 40\%$ error. Thus, I_c is less precise in determination of the I_{nuc} that I_m. (G) Background cluster intensity, I_{bg} expressed as a function of I_{nuc}. The slope of the linear fit is $slope = 0.99 \pm 0.02$ EL. (H) The CV of I_{bg} expressed as a function of I_{nuc}. The average CV is $\sim 10\%$, making it a highly reproducible quantity, similar to I_{nuc}. (I) The error in estimating of nuclear concentration (I_{nuc}) using the average I_{bg} is $< 10\%$ in the anterior. (J) The cluster diameter, d expressed as a function of nuclear Bcd-GFP intensity I_{nuc}. (K) The CV of d expressed as the coefficient of variation. The average variability is $< 10\%$. (L) The precision of nuclear concentration (I_{nuc}) determination using the average d is $> 100\%$, making cluster size an extremely imprecise metric for estimating I_{nuc}.
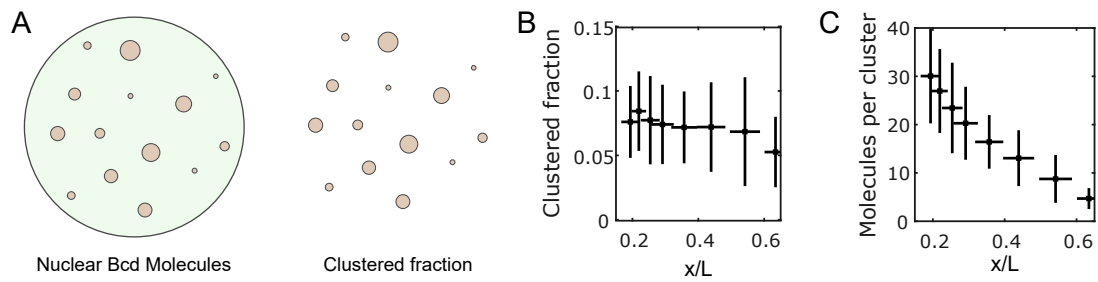
FIG. S11. **Estimation of the number of molecules per cluster.** (A) A cartoon depicting the broad classification of distribution of the Bcd molecules in the nucleus. The green depicts the freely diffusing Bcd molecules, while the brown (also separately shown on the right) shows the clustered fraction. (B) Figure shows the fraction of Bcd molecules in the clustered fraction per nucleus ($mean \pm std$) as a function of the nuclear position in the embryo. It must be noted that the clustered fraction is defined by the limits of detection in this study. (C) The average molecules of Bcd per cluster ($mean \pm std$) is plotted as a function of the nuclear position in the embryo (Methods). The average number of molecules drops from about 30 at the anterior to 5 in the posterior half of the embryo. This also denotes the lower limit of detection, setting the limit to ~ 5 molecules per cluster.
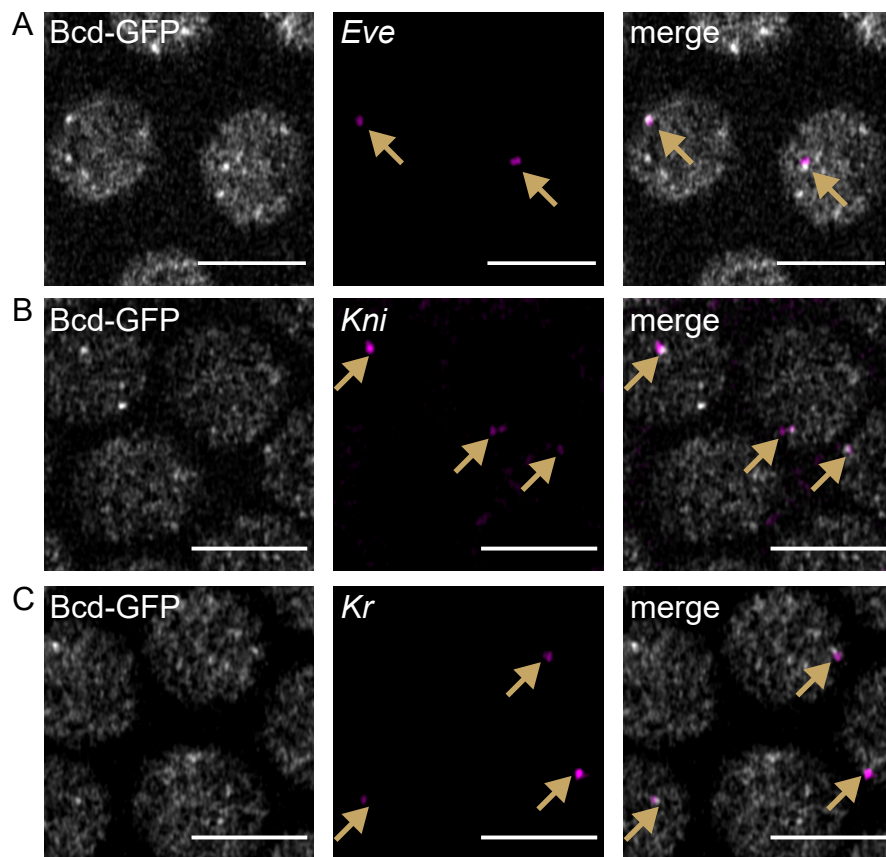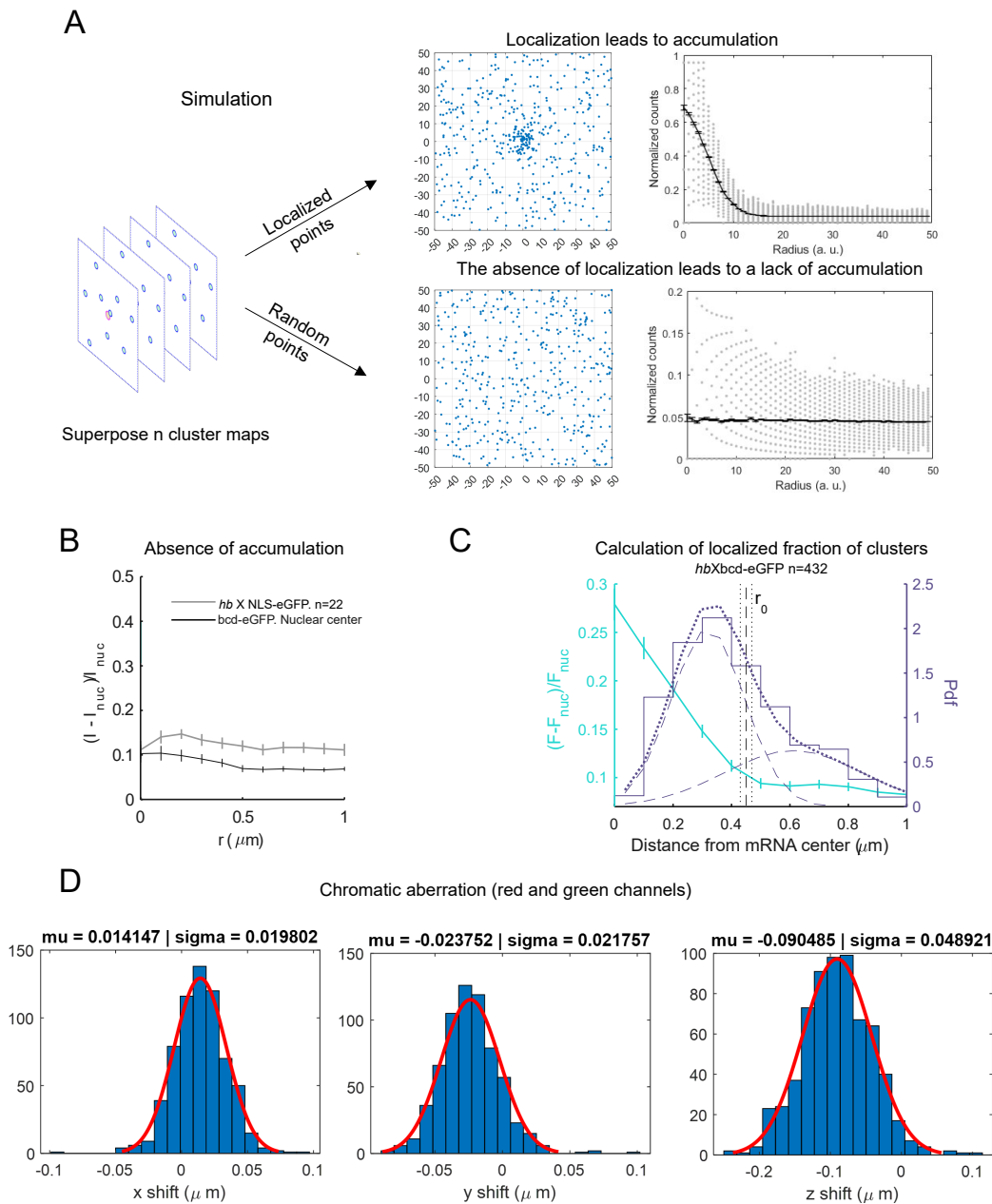

FIG. S12. **Micrographs of Bcd clusters colocalizing with nascent transcription hotspots.** The left panels show the raw images of nuclei expressing Bcd-GFP. The middle panels show the same nuclei expressing (A) *Eve*, (B) *Kni* and (C) *Kr*. The right panels shows the two images from the left and the middle overlaid. Arrows indicate the transcriptional hotspots.

FIG. S13. **Localization of Bcd cluster with the nascent transcription hotspot.** (A) (Left) Cartoons representing a nascent transcription hotspot (pink) and TF molecules (blue). (Right) With a nascent transcription hotspot at the center of the simulation window, TF molecules preferentially accumulate near the origin, (TOP) giving a Gaussian profile for the radial cluster density. The size of the Gaussian function (FWHM) gives the spread of the accumulation. Molecules appearing within this limit can be considered to be coupled to the hotspot at the origin in some capacity. (Bottom) Absence of a transcription hotspot at the origin leads to a random distribution of molecules, yielding a flat profile (RIGHT). Examples of this can be seen in B. (B) NLS-eGFP does not accumulate at the *hb* transcription site (gray). Also, Bcd-GFP does not accumulate at a random site in the nucleus (nuclear center). (C) Since the TF accumulation radius defines in some sense the "limit of influence" of a target gene, any TF cluster appearing within this limit should be considered coupled to the gene. To test this hypothesis, the histogram of the distance of the Bcd-GFP clusters nearest to *hb* transcription hotspot is plotted. The histogram is then fitted with a double Gaussian, and the position of intersection of the two Gaussian kernels is compared with the "limit" obtained earlier. The two quantities are within $\sim 50\,nm$ of each other. (D) The three plots show the histograms of shifts in the intensity-weighted centers of polystyrene beads measured in two color channels, red and green. A Gaussian fit to each histogram gives the corresponding σ which serves as a measure of chromatic aberration along each image axis. All three shifts are sub-pixel.
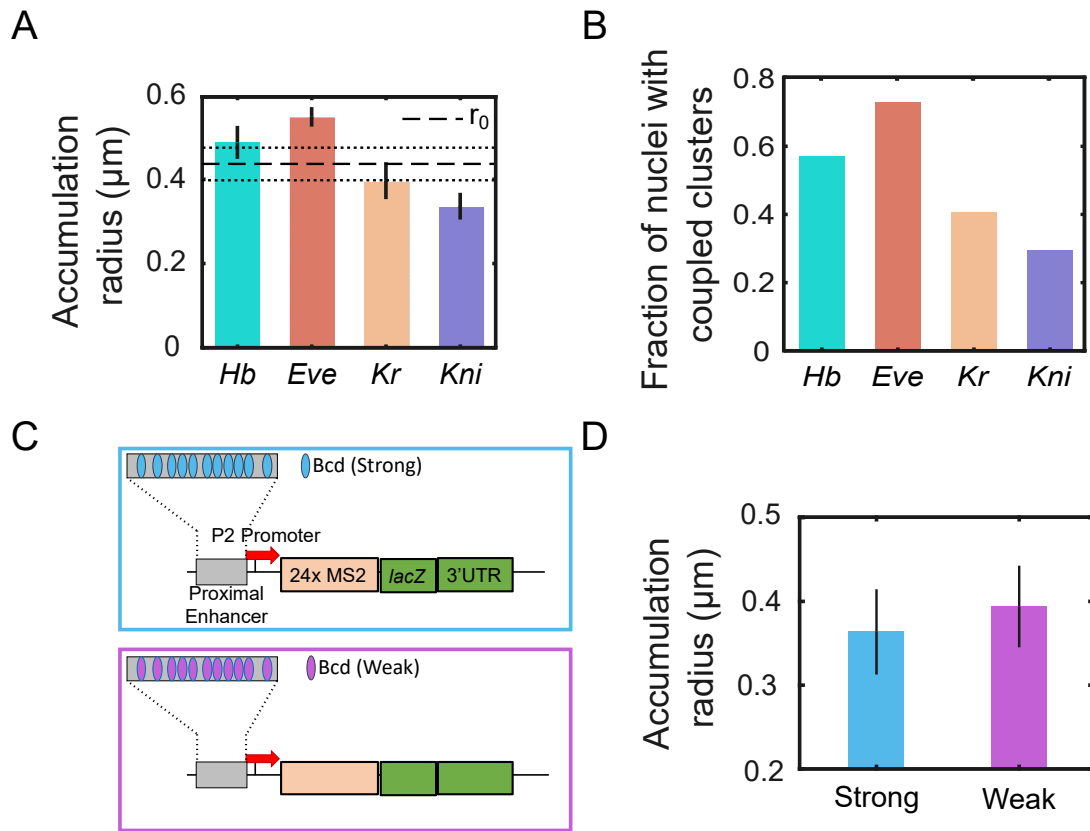
FIG. S14. **Cluster coupling characteristics with the target genes.** (A) Bars representing 2σ from the Half normal fits of Bcd-GFP accumulation (see Methods) for each target gene (labeled on the x axis). All bars indicate standard errors. The horizontal line shows the average (dashed) and standard error bounds(dotted) from the entire dataset, calculated via bootstrapping. The average radius is also shown in C by a vertical dashed line. (B) The fraction of nuclei in which expressing target genes are associated with a coupled cluster. The corresponding genes are labeled on the x axis. (C) Schematic shows the two constructs, one driven by an enhancer with strong Bcd binding sites, and the other with an enhancer composed of weak Bcd binding sites. (D) Bars representing 2σ from the Half normal fits of Bcd-GFP accumulation for the two constructs shown in C.

II.   SUPPLEMENTAL TABLES

Fly line	Source
hb BAC<MS2	Bothma *et al.*[54]
kni BAC<MS2	Bothma *et al.*[54]
eve MS2	Chen *et al.*[47]
kr MS2	El-Sherif *et al.*[55]
bnk MS2	This work
P2 (Strong) MS2	This work
P2 (Weakg) MS2	This work

TABLE I. **List of MS2 stem loop fly lines.**