bioRxiv preprint doi: https://doi.org/10.1101/133124; this version posted May 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Dense Bicoid Hubs Accentuate Binding along the Morphogen Gradient

Mustafa Mir[1], Armando Reimer[2], Jenna E. Haines[1], Xiao-Yong Li[3], Michael Stadler[1], Hernan Garcia[1,2,4], Michael B. Eisen[1,2,3,5], Xavier Darzacq[1,*]

[1]Department of Molecular and Cell Biology, University of California, Berkeley, CA, 94720, USA
[2]Biophysics Graduate Group, University of California, Berkeley, CA, 94720, USA
[3]Howard Hughes Medical Institute, University of California, Berkeley, CA, 94720, USA
[4]Department of Physics, University of California, Berkeley, CA, 94720, USA
[5]Department of Integrative Biology, University of California, Berkeley, CA, 94720, USA
*Correspondence to: darzacq@berkeley.edu

Abstract: Morphogen gradients direct the spatial patterning of developing embryos, however, the mechanisms by which these gradients are interpreted remain elusive. Here we perform *in vivo* single molecule imaging of the transcription factor Bicoid that forms a gradient along the anteroposterior axis of the early *Drosophila melanogaster* embryo. We observe that Bicoid binds to DNA with a rapid off-rate, such that its average occupancy at target loci becomes on-rate dependent, a property required for concentration-sensitive regulation. Surprisingly, we also observe abundant specific DNA binding in posterior nuclei, where Bicoid levels are vanishingly low. Live embryo imaging reveals "hubs" of local high Bicoid concentration that are dependent on the ubiquitous maternal factor Zelda. We propose that localized modulation of transcription factor on-rates via clustering provides a general mechanism to facilitate binding to low-affinity targets under limiting factor conditions, and that this may be a prevalent feature directing other developmental transcription networks.

Spatial patterning during embryonic development is widely accepted to be orchestrated through concentration gradients of regulatory molecules known as morphogens (*1*, *2*). The maternally deposited transcription factor (TF) Bicoid (BCD) in *Drosophila* was the first identified morphogen (*3*), and remains an iconic and widely studied developmental regulator. Bicoid is distributed in an exponentially decaying concentration gradient along the anteroposterior (A-P) axis of embryos and regulates the activity of ~100 genes associated with the formation of distinct spatial expression domains ranging from the anterior tip to the middle of the embryo (*4-7*). The ability of BCD and other morphogens to activate target genes in different locations across gradients is classically thought to arise from modulating the number and strength of cognate DNA binding sites within target enhancers (*8-10*), with sharp expression domain boundaries set through cooperative binding of BCD and/or the combinatorial actions of other transcription factors (*11*, *12*). The validity of these models and other essential mechanistic questions about how BCD, and morphogens in general, differentially activate genes along a concentration gradient has been challenging to resolve and would benefit from direct measurements of BCD-DNA interactions *in vivo* by single molecule imaging.

bioRxiv preprint doi: https://doi.org/10.1101/133124; this version posted May 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

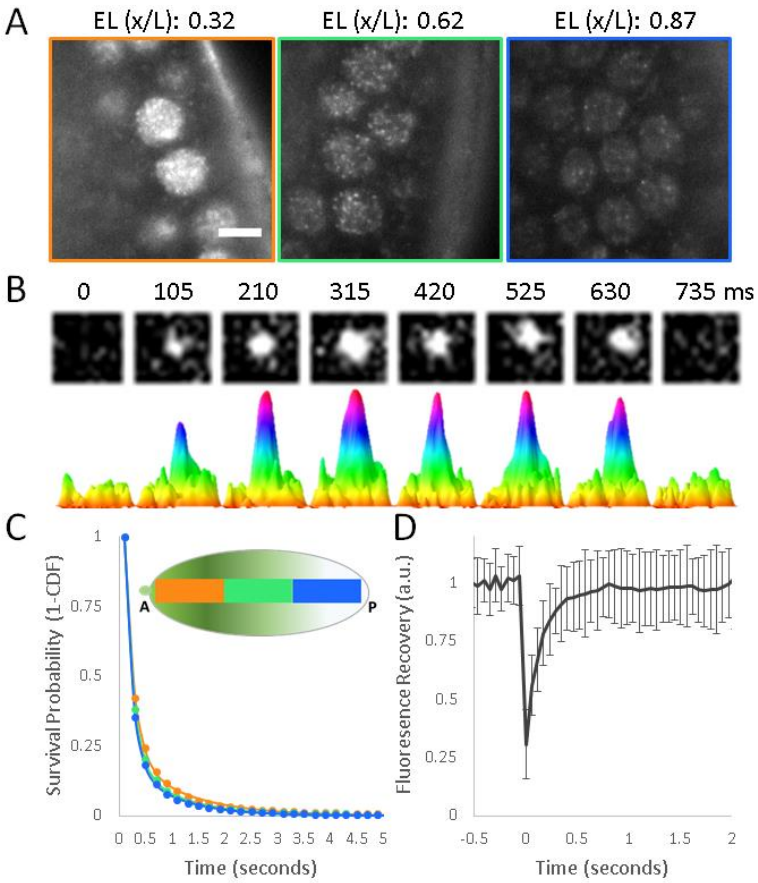

Figure 1. Single molecule kinetics of Bicoid in living *Drosophila* Embryos. (A) Raw images of Bicoid-GFP molecules in a living *Drosophila* embryo acquired with a 100 millisecond exposure time. Scale bar is 5 um. Positions along the A-P axis are shown as a fraction of the Embryonic Length (EL (x/L). (B) Example of a single molecule binding event. Top row shows raw images from a 1.2 x 1.2 µm area, bottom row shows corresponding surface plot representations to illustrate the signal-to-noise. (C) Survival probability (1-cummalative distribution function of trajectory lengths) for Bicoid binding (markers) in the Anterior (34 nuclei), Middle (70 nuclei) and Posterior (83 nuclei) segments of the embryo and corresponding fits to a two-exponent model (solid lines) show no significant differences. (D) FRAP curve for Bicoid shows a recovery time on the order of hundreds of milliseconds, error bars show standard deviation over 21 nuclei.

Single molecule imaging in living cells has been increasingly used in recent years to measure the dynamics of TF-DNA interactions (*13*). However, the techniques commonly used for single cells in culture are limited to regions close to the microscope coverslip, and are not suitable for whole embryos and thick tissues because of technical limitations that result in low signal-to-noise ratios. Here we utilized Lattice Light-Sheet Microscopy(*14*) (See Supplementary Materials and Fig. S1) to overcome these barriers and characterize the single molecule DNA-binding kinetics of BCD in developing *Drosophila melanogaster* embryos (Fig. 1A-B).

The classical morphogen models predict a difference in the average dissociation rates of Bicoid along the A-P axis. For example, genes that need to be activated at lower concentrations should have higher affinity sites resulting in lower off-rates to enable a higher time-average occupancy (i.e. residence time, RT). We therefore first performed single molecule imaging and tracking at long (100 millisecond) exposure times, effectively blurring out the fast moving (unbound) population, to estimate the residence times (See Supplementary Materials) of BCD binding in nuclei at all positions along the A-P axis(*15*).

Previous single molecule studies of transcription factors have consistently found two populations in residence time distributions, a short-lived population with RTs on the order of hundreds of milliseconds, and a longer-lived population with RTs on the order of 10's of seconds to minutes (*15-17*). These two populations have often been shown to be the non-specific and specific binding populations, respectively. We thus fit the survival probability (residence time) distributions of BCD binding events (Fig. 1C, S2) to a two-exponent model using data from the anterior, middle and posterior thirds of the embryo (See Supplementary Materials) . This analysis identified a short-lived population with RTs on the order of 100s of milliseconds and a longer lived population with RTs on the order of seconds (Fig. S2-S3), and with no significant dependence on position along the A-P axis for either population (Table S1), a finding not expected from the classical model.

bioRxiv preprint doi: https://doi.org/10.1101/133124; this version posted May 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

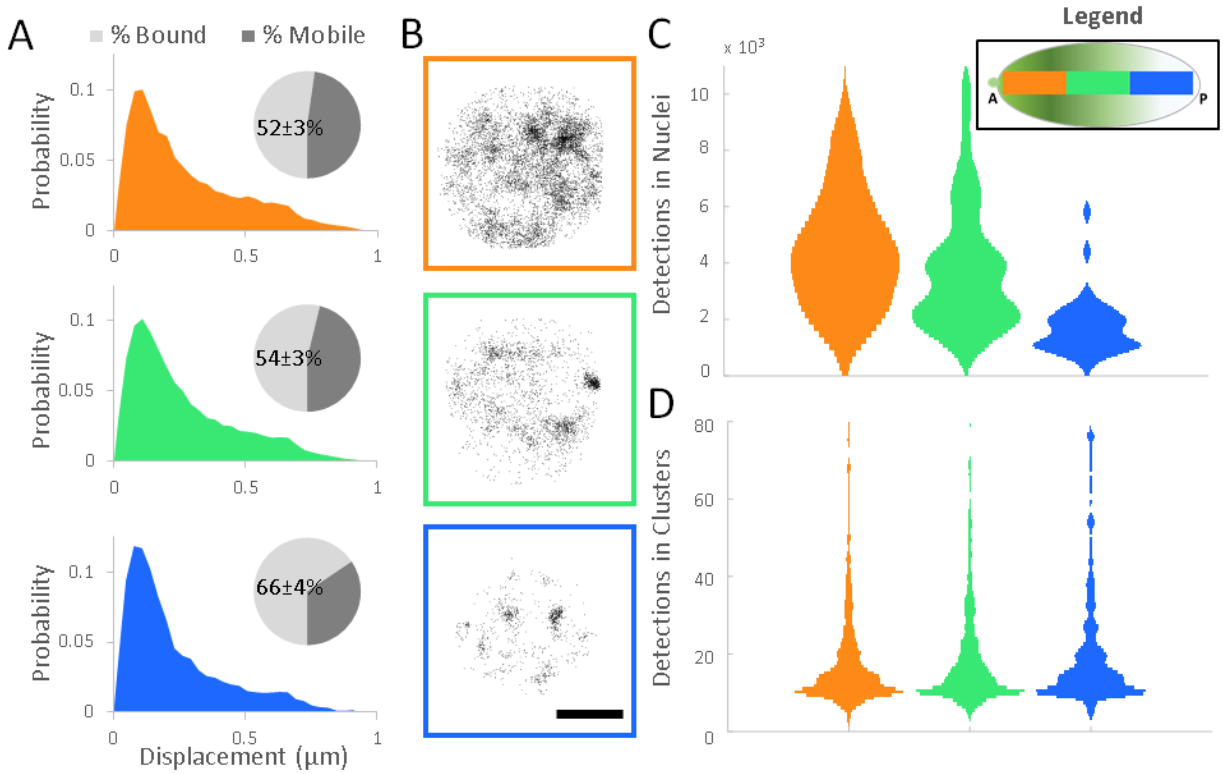

Figure 2. Local modulation of Bicoid Concentration (A) Normalized probability distributions of measured displacements in the Anterior (30 nuclei), Middle (67 nuclei), and Posterior (66 nuclei) positions of the embryos, pie charts show the estimated mobile and bound fractions from fits to a two population distribution with the bound population percent labelled with the standard error of the fit parameter (Fig. S5B). (B) Examples of the spatial distribution of all detections in nuclei along the A-P axis, scale bar is 2.5 µm. (C) Distribution of the number of detections in all nuclei. (D) Distributions of the number of detections in all clusters (see. Fig. S6).

Remarkably, the RT of the long-lived population is much shorter than lifetimes (10-60 sec) typically observed for other sequence specific DNA binding TFs using single molecule tracking. These measurements were cross-validated (See Supplementary Materials) with Fluorescence Recovery After Photobleaching (FRAP) experiments (Fig. 1D) confirming a remarkably fast halftime of recovery on the order of hundreds of milliseconds (Fig. S4). The dominance of the short-lived interactions highlights the preponderance of low affinity BCD binding sites and non-specific interactions, as previously suggested by genomic studies that captured this promiscuous binding behavior (*18, 19*). In addition to the expected large number of binding events in the anterior and middle segments of the embryo, we also observed a surprisingly large number of potentially specific binding events in the posterior-most nuclei where BCD has been reported to be at vanishingly low (<5nM posterior vs ~50 nM anterior) concentrations (*20*).

The lack of a detectable difference in the survival probability distributions along the A-P axis and the observation of significant binding in posterior nuclei prompted us to examine how much of the small BCD population remaining in the posterior embryo is actually bound. Since longer exposure times only allow detection of molecules bound for at least the span of the exposure, we performed single molecule tracking measurements at a decreased exposure time of 10 milliseconds. Through analyses of displacement distributions (Fig. S5) from single molecule trajectories (See Supplementary Materials) we estimated the fraction of BCD that is bound along the A-P axis (Fig. 2A) and found, counter-intuitively, that a greater fraction of

bioRxiv preprint doi: https://doi.org/10.1101/133124; this version posted May 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

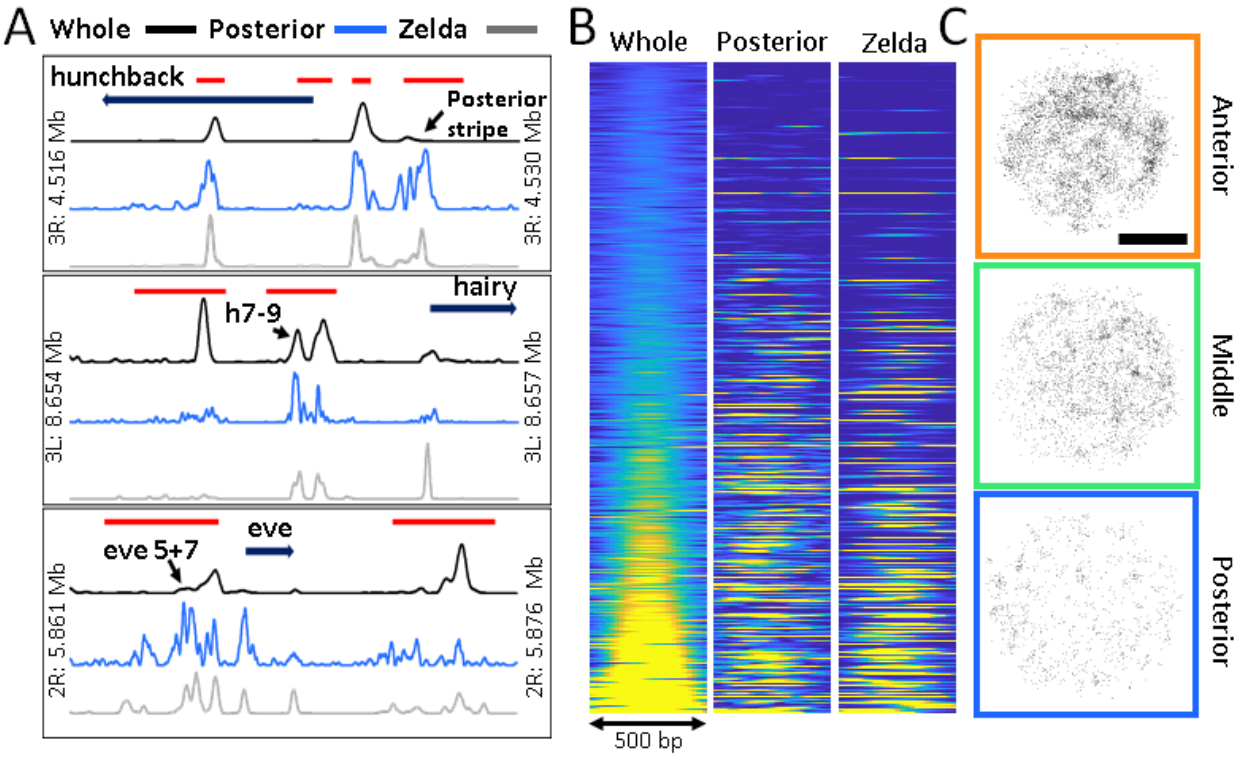

Figure 3. Zelda mediated Bicoid binding in the posterior embryo (A) Whole embryo and posterior third BCD, and whole embryo Zelda ChIP-seq signal enrichment at the hunchback, hairy, and eve gene loci. Blue arrows indicate gene bodies and red bars show known regulatory regions. Black arrows indicate regions known to regulate posterior expression. (B) ChIP-seq signal enrichment over a 500 bp window centered on BCD peaks called in the whole embryo data and sorted according to increasing signal of the whole embryo data, a total of 2145 peaks are shown, colors indicate enrichment over the background (blue) with all plots displayed on the same scale. (C) Examples of the spatial distribution of all detected bound molecules in nuclei along the A-P axis in ZLD- embryos, scale bar is 2.5 µm

the BCD population is bound in more posterior positions of the embryo where BCD is present at the lowest concentrations (See Supplementary Materials). Furthermore, analyses of the spatial distribution of detections from the 100 millisecond dataset revealed a distinct clustering of binding events, which becomes more pronounced toward posterior positions (Figs. 2B, S6-S7). Remarkably, although the number of detections per nucleus follows the trend dictated by the global concentration gradient across the embryo (Fig. 2C), the distribution of BCD molecules detected per cluster or "hub" is maintained even in the posterior (Figs. 2D). These data suggest that the formation of locally dense BCD hubs at specific sites in nuclei across the A-P axis is conserved independent of the global gradient (See Supplementary Materials). This surprising finding raises the question: what mechanism exists to selectively enrich local concentrations to promote BCD binding at specific targets under conditions of very low BCD levels such as in posterior nuclei.

To test whether BCD is binding with specificity in the posterior embryo we analyzed its binding profiles in a spatially segregated manner (*21*) by comparing ChIP-seq profiles derived from individually dissected posterior thirds of embryos to previously published data from whole embryos (*22*). Our analysis revealed that BCD indeed binds to known targets in the posterior but with increased relative enrichment at specific enhancer elements (Figs. 3A). For example, in the *hunchback* locus, binding at the posterior stripe enhancer is highly enriched in nuclei from the dissected posterior third relative to the whole embryo. Intriguingly, genomic regions that exhibit a relative increase in BCD occupancy in the posterior are correlated with an enrichment of Zelda (ZLD) (Fig. 3A, S8), a ubiquitous

bioRxiv preprint doi: https://doi.org/10.1101/133124; this version posted May 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

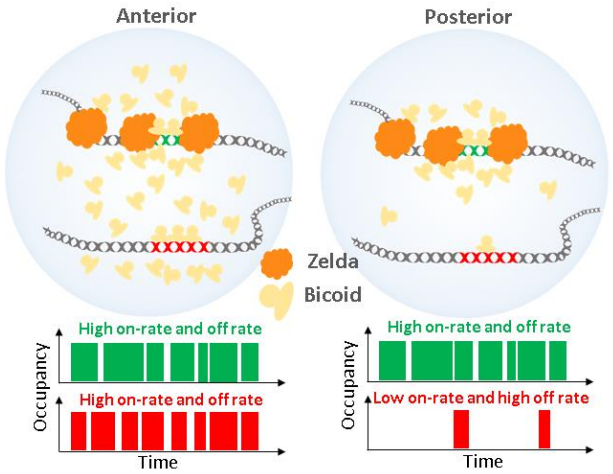

Figure 4. Model of Zelda mediation of the on-rate of Bicoid binding at specific loci in the posterior embryo. At high concentrations both specific and non-specific sites are highly occupied with transcriptional activity then regulated through the combinatorial action of opposing factors. At low concentrations loci with Zelda occupancy have an increased time averaged occupancy through the formation of hubs that enrich local concentrations.

activator often described as a pioneer factor active during early embryonic development (*23-26*). Genome wide analyses centered on BCD peaks detected in whole embryos reveal that binding of BCD to these same loci in posterior nuclei is highly correlated with ZLD co-binding (Figs. 3B, S8).

The posterior genomic data and the published evidence for Zelda's role in the regulation of chromatin accessibility (*23-29*) and its apparently differential effect on TF binding at low concentrations (*25*, *28*) naturally led us to hypothesize that the observed clustering of BCD binding events may be, in part, mediated by ZLD. We thus generated *zelda* null embryos and measured BCD binding at 100 millisecond exposures, and found an abolishment of BCD hubs in the posterior embryo but no significant change in RTs (Fig. 3C, Table S1). Although the exact mechanism by which ZLD mediates BCD hub formation and binding remains unclear, we speculate that a combination of protein-protein interactions facilitated by intrinsically disordered low-complexity domains of ZLD and its reported role in promoting chromatin accessibility may contribute to BCD clustering (Fig. 4).

Our initial observation of the low affinity nature of BCD binding fits the conventional view of BCD as a concentration-dependent morphogen. We envision that high TF concentrations (in the anterior embryo) potentiate rapid on-rates along with a high chromatin occupancy. Our observed high off-rate also suggests that BCD is free to frequently sample non-specific sites. Thus, as BCD concentrations decrease posteriorly along the gradient, there should be progressively lower on-rates and the binding at specific sites should be correspondingly diminished. However, this simplistic model essentially consigns BCD to no posterior function and contradicts a wealth of evidence in the literature pointing to a specific role for BCD in the regulation of posterior gene expression (*8*, *11*, *12*, *30-32*). Although this paradox is inexplicable under the classical morphogen model, our evidence for local high-concentrations of BCD "hubs" capable of modulating specific binding with help from ZLD in the posterior provides a novel mechanism for TFs to carry out their regulatory role even when present at very low concentrations (Fig. 4). The formation of such clusters or hubs of TFs mediated by co-factors at specific genomic loci has recently been hypothesized to promote binding at low affinity enhancers in an orthogonal system (*33*), and may therefore, represent a general mechanism of transcriptional regulation evolved in eukaryotes to allow exquisite spatial and segmental modulation during development.

References

1. A. M. Turing, The Chemical Basis of Morphogenesis. *Philos T Roy Soc B* **237**, 37 (1952).
2. L. Wolpert, Positional Information and Pattern of Cellular Differentiation. *Biophysical Journal* **9**, A8 (1969).
3. W. Driever, C. Nusslein-Volhard, A gradient of bicoid protein in Drosophila embryos. *Cell* **54**, 83 (Jul 1, 1988).
4. W. Driever, C. Nusslein-Volhard, The bicoid protein determines position in the Drosophila embryo in a concentration-dependent manner. *Cell* **54**, 95 (Jul 1, 1988).
5. W. Driever, C. Nussleinvolhard, The Bicoid Protein Is a Positive Regulator of Hunchback Transcription in the Early Drosophila Embryo. *Nature* **337**, 138 (Jan 12, 1989).

bioRxiv preprint doi: https://doi.org/10.1101/133124; this version posted May 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

6. G. Struhl, K. Struhl, P. M. Macdonald, The Gradient Morphogen Bicoid Is a Concentration-Dependent Transcriptional Activator. *Cell* **57**, 1259 (Jun 30, 1989).
7. W. Driever, V. Siegel, C. Nusslein-Volhard, Autonomous determination of anterior structures in the early Drosophila embryo by the bicoid morphogen. *Development* **109**, 811 (Aug, 1990).
8. D. S. Burz, R. Rivera-Pomar, H. Jackle, 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 (Oct 15, 1998).
9. D. Lebrecht *et al.*, Bicoid cooperative DNA binding is critical for embryonic patterning in Drosophila. *P Natl Acad Sci USA* **102**, 13176 (Sep 13, 2005).
10. H. Xu, L. A. Sepulveda, L. Figard, A. M. Sokac, I. Golding, Combining protein and mRNA quantification to decipher transcriptional regulation. *Nat Methods* **12**, 739 (Aug, 2015).
11. A. Ochoa-Espinosa, D. Y. Yu, A. Tsirigos, P. Struffi, S. Small, Anterior-posterior positional information in the absence of a strong Bicoid gradient. *P Natl Acad Sci USA* **106**, 3823 (Mar 10, 2009).
12. H. Chen, Z. Xu, C. Mei, D. Yu, S. Small, A system of repressor gradients spatially organizes the boundaries of Bicoid-dependent target genes. *Cell* **149**, 618 (Apr 27, 2012).
13. Z. Liu, L. D. Lavis, E. Betzig, Imaging Live-Cell Dynamics and Structure at the Single-Molecule Level. *Molecular cell* **58**, 644 (May 21, 2015).
14. B. C. Chen *et al.*, Lattice light-sheet microscopy: Imaging molecules to embryos at high spatiotemporal resolution. *Science* **346**, 439 (Oct 24, 2014).
15. J. J. Chen *et al.*, Single-Molecule Dynamics of Enhanceosome Assembly in Embryonic Stem Cells. *Cell* **156**, 1274 (Mar 13, 2014).
16. D. Normanno *et al.*, Probing the target search of DNA-binding proteins in mammalian cells using TetR as model searcher. *Nat Commun* **6**, 7357 (2015).
17. A. S. Hansen, I. Pustova, C. Cattoglio, R. Tjian, X. Darzacq, CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *eLife* **6**, e25776 (2017/05/03, 2017).
18. X. Y. Li *et al.*, Transcription factors bind thousands of active and inactive regions in the Drosophila blastoderm. *Plos Biol* **6**, 365 (Feb, 2008).
19. A. Ochoa-Espinosa *et al.*, The role of binding site cluster strength in Bicoid-dependent patterning in Drosophila. *P Natl Acad Sci USA* **102**, 4960 (Apr 5, 2005).
20. A. H. Morrison, M. Scheeler, J. Dubuis, T. Gregor, Quantifying the Bicoid morphogen gradient in living fly embryos. *Cold Spring Harb Protoc* **2012**, 398 (Apr, 2012).
21. P. A. Combs, M. B. Eisen, Sequencing mRNA from Cryo-Sliced Drosophila Embryos to Determine Genome-Wide Spatial Patterns of Gene Expression. *Plos One* **8**, (Aug 12, 2013).
22. R. K. Bradley *et al.*, Binding Site Turnover Produces Pervasive Quantitative Changes in Transcription Factor Binding between Closely Related Drosophila Species. *Plos Biol* **8**, (Mar, 2010).
23. H. L. Liang *et al.*, The zinc-finger protein Zelda is a key activator of the early zygotic genome in Drosophila. *Nature* **456**, 400 (Nov 20, 2008).
24. M. M. Harrison, X. Y. Li, T. Kaplan, M. R. Botchan, M. B. Eisen, Zelda binding in the early Drosophila melanogaster embryo marks regions subsequently activated at the maternal-to-zygotic transition. *PLoS genetics* **7**, e1002266 (Oct, 2011).
25. Z. Xu *et al.*, Impacts of the ubiquitous factor Zelda on Bicoid-dependent DNA binding and transcription in Drosophila. *Genes Dev* **28**, 608 (Mar 15, 2014).
26. S. M. Foo *et al.*, Zelda potentiates morphogen activity by increasing chromatin accessibility. *Curr Biol* **24**, 1341 (Jun 16, 2014).
27. Y. J. Sun *et al.*, Zelda overcomes the high intrinsic nucleosome barrier at enhancers during Drosophila zygotic genome activation. *Genome Res* **25**, 1703 (Nov, 2015).
28. K. N. Schulz *et al.*, Zelda is differentially required for chromatin accessibility, transcription factor binding, and gene expression in the early Drosophila embryo. *Genome Res* **25**, 1715 (Nov, 2015).
29. X. Y. Li, M. M. Harrison, J. E. Villata, T. Kaplan, M. B. Eisen, Establishment of regions of genomic activity during the Drosophila maternal to zygotic transition. *eLife* **3**, (Oct 14, 2014).
30. R. Rivera-Pomar, X. Lu, N. Perrimon, H. Taubert, H. Jackle, Activation of posterior gap gene expression in the Drosophila blastoderm. *Nature* **376**, 253 (Jul 20, 1995).
31. S. Small, A. Blair, M. Levine, Regulation of two pair-rule stripes by a single enhancer in the Drosophila embryo. *Dev Biol* **175**, 314 (May 1, 1996).
32. A. La Rosee, T. Hader, H. Taubert, R. Rivera-Pomar, H. Jackle, Mechanism and Bicoid-dependent control of hairy stripe 7 expression in the posterior region of the Drosophila embryo. *The EMBO journal* **16**, 4403 (Jul 16, 1997).
33. J. Crocker *et al.*, Nuclear Microenvironments Modulate Transcription From Low-Affinity Enhancers. *bioRxiv*, (2017).

Acknowledgments: We thank Astou Tangara and the Betzig group at HHMI Janelia Research campus for help in constructing the Lattice Light Sheet Microscope. We thank Robert Tjian for extensive discussions and advice along the course of this work, and for his help in writing this manuscript. We thank Elizabeth Eck for her help in generating the *zelda*- germ line clones. We thank members of the Darzacq, Tjian, Garcia, and Eisen labs for suggestions and discussion. This work was supported by the California Institute of Regenerative Medicine (CIRM) LA1-08013 and the National Institutes of Health (NIH) UO1-EB021236 & U54-DK107980 to X.D., by the Burroughs Wellcome Fund Career Award at the Scientific Interface, the Sloan Research Foundation, the Human Frontiers Science Program, the Searle Scholars Program, and the Shurl & Kay Curci Foundation to H.G., a Howard Hughes Medical Institute investigator award to M.E., J.H. and A.R are supported by NSF Graduate Research Fellowships.