Combining protein and mRNA quantification to decipher transcriptional regulation

Heng Xu[1–3], Leonardo A Sepúlveda[1,3], Lauren Figard[4], Anna Marie Sokac[1] & Ido Golding[1–3]

We combine immunofluorescence and single-molecule fluorescence *in situ* hybridization (smFISH), followed by automated image analysis, to quantify the concentration of nuclear transcription factors, number of transcription factors bound, and number of nascent mRNAs synthesized at individual gene loci. A theoretical model is used to decipher how transcription factor binding modulates the stochastic kinetics of mRNA production. We demonstrate this approach by examining the regulation of *hunchback* in the early *Drosophila* embryo.

© 2015 Nature America, Inc. All rights reserved.

Sequence-specific transcription factors bind to regulatory DNA elements and interact with the cell's transcriptional machinery to regulate mRNA production. Despite the central role of this process in cellular function, a quantitative mapping between transcription factor binding and the activity of the regulated gene is still lacking[1]. For this mapping to be obtained, transcriptional regulation needs to be characterized at the level of the individual gene copy, simultaneously capturing the stochastic events of transcription factor binding and mRNA production[2]. Here we present an approach to pursue this goal.

To demonstrate our method, we examined the regulation of the zygotic gene *hunchback* (*hb*) by the maternal transcription factor Bicoid (Bcd) in the early embryo of *Drosophila melanogaster*. As a morphogen, nuclear Bcd concentration naturally varies by >50-fold within a single embryo[3]; we could therefore examine its concentration-dependent activity without needing to externally perturb the system.

To simultaneously quantify *hb* mRNA production and Bcd protein concentration, we combined smFISH[4,5] with antibody staining (immunofluorescence). Combining smFISH and immunofluorescence is challenging, as reaction conditions for the two procedures conflict in multiple ways (Online Methods). We tested different protocol designs and found that smFISH labeling followed by immunofluorescence can preserve both the smFISH and immunofluorescence signals. We therefore developed a combined labeling protocol based on this design (Online Methods). Wild-type (OreR) embryos at nuclear cleavage cycles 11–14 were collected, labeled and imaged using laser scanning confocal microscopy (**Fig. 1a** and Online Methods). The images were analyzed as described below (**Supplementary Note**).

In the *hb* mRNA channel, we observed two distinct groups of foci (spots), distinguishable by their intensity and size (**Fig. 1b**). We identified the smaller, weaker spots as individual mature *hb* mRNAs, whereas the larger, brighter ones were mapped to the accumulation of multiple nascent mRNAs at the sites of active *hb* transcription[6]. The measured lifetime of both species following inhibition of transcription initiation supported this interpretation (**Supplementary Fig. 1** and Online Methods), and the measured numbers of active *hb* sites were consistent with previous reports[7–9] (**Supplementary Fig. 2**). We performed automated recognition of the individual fluorescent foci and extracted the intensity value corresponding to a single *hb* mRNA (**Supplementary Fig. 3** and **Supplementary Note**). We then used this value to convert the intensity of the smFISH signal at each transcription site to the total amount of *hb* mRNA (in units of mRNA molecules) at that gene locus[4]. The maximal measured level of *hb* transcription was in excellent agreement with a previous report[6] (**Supplementary Fig. 3**). Using a number of methods, we estimated our mRNA detection efficiency to be >80% and the resulting error in measuring nascent mRNA to be <10% (**Supplementary Fig. 4**).

In the Bcd protein channel, we verified that the antibody signal was proportional to Bcd concentration by measuring both the autofluorescence and immunofluorescence signal in a transgenic fly strain in which Bcd is fused to EGFP[10] (**Supplementary Fig. 5**). To convert the immunofluorescence signal to absolute Bcd concentration, we used two methods (**Supplementary Note**). In the first method, we automatically detected individual, diffraction-limited immunofluorescent spots in the cytoplasmic regions of the embryo (**Fig. 1c**). After discarding false positive spots (likely the result of nonspecific labeling by primary antibodies), we used the extracted single-protein intensity to convert nuclear fluorescence to Bcd concentration in each nucleus of a given embryo. In the second method, we made use of the spatial fluctuations of the fluorescence signal inside each nucleus (**Fig. 1c**). We used the slope of the variance-versus-mean curve in each embryo to obtain the single-Bcd fluorescence value and convert nuclear fluorescence to Bcd concentration. The two methods showed good agreement in estimating the maximal Bcd concentration in a given embryo (ratio of 1.01 with an s.d. of ~10%, 31 embryos; **Fig. 1d**). Measuring the maximal concentration of EGFP-Bcd in a transgenic fly strain yielded values that lie between the estimates of two previous studies[3,11] (**Supplementary Fig. 5**).

[1]Verna & Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, Texas, USA. [2]Center for the Physics of Living Cells, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA. [3]Center for Theoretical Biological Physics, Rice University, Houston, Texas, USA. [4]Integrative Molecular and Biomedical Sciences Graduate Program, Baylor College of Medicine, Houston, Texas, USA. Correspondence should be addressed to I.G. (golding@bcm.edu).

RECEIVED 18 DECEMBER 2014; ACCEPTED 5 MAY 2015; PUBLISHED ONLINE 22 JUNE 2015; DOI:10.1038/NMETH.3446

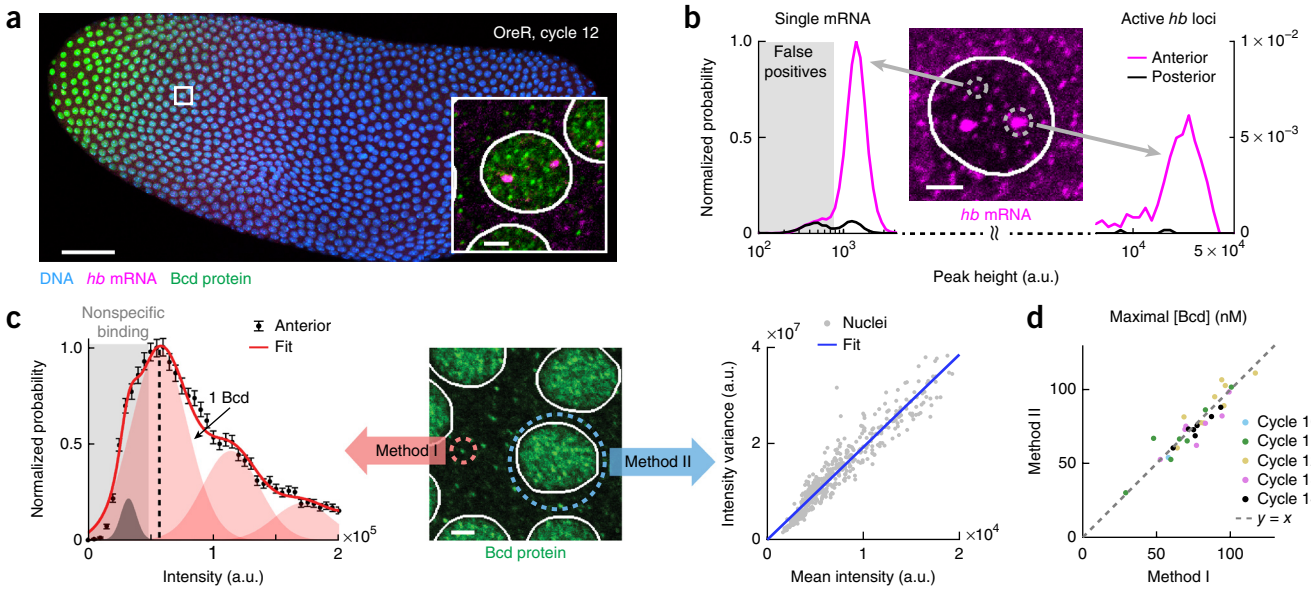

Figure 1 | Simultaneous quantification of Bcd protein and *hb* mRNA in a single embryo. (**a**) Confocal image of a wild-type *Drosophila* embryo labeled for Bcd protein, *hb* mRNA and DNA at nuclear cleavage cycle 12. Scale bar, 50 µm. Inset, magnified view of a single anterior nucleus. Scale bar, 2 µm. (**b**) Intensity histograms of smFISH spots in the anterior (0.13–0.38 embryo length (EL)) and posterior (0.63–0.88 EL) parts of a single embryo (>40,000 spots). The histograms were used to discard false positive spots and to discriminate individual mRNAs from sites of active transcription. Center, smFISH image of a single nucleus in the anterior region. Scale bar, 2 µm. a.u., arbitrary units. (**c**) Two methods for converting the immunofluorescence signal to absolute Bcd concentration. Center, immunofluorescence signal in the anterior part of the embryo. Scale bar, 2 µm. Left, intensity histogram of cytoplasmic Bcd spots at the anterior (>10,000 spots). Error bars, s.e.m. By fitting the histogram to the sum of Gaussian functions, we identified the intensity corresponding to a single Bcd protein. Right, relation between the mean and variance of Bcd fluorescence intensity of individual nuclei (>500 nuclei). The data were fitted to a straight line, and the slope was used to calculate the intensity of a single labeled Bcd protein. (**d**) Average Bcd concentration in the brightest 1% of nuclei of individual embryos (maximal [Bcd]), calculated using the two methods (see **c**). Data from 31 embryos, nuclear cycles 10–14.

To analyze *hb* regulation, we first plotted the number of nascent *hb* mRNAs versus nuclear Bcd concentration for all *hb* loci (within the range 0.25–0.7 embryo length (EL)) in each embryo (**Fig. 2a**). Binning the data along the Bcd axis yielded the gene regulation function[12] (**Fig. 2a**, **Supplementary Fig. 6** and **Supplementary Note**). We fitted the gene regulation function of each embryo to a Hill function[3] (**Fig. 2a** and **Supplementary Fig. 6**). The Hill coefficient h was approximately constant during cycles 11–13, $h = 6.6 \pm 0.4$ (designates mean ± s.e.m. throughout) (**Fig. 2b**). A possible interpretation is that *hb* transcription is activated through the cooperative binding of six or seven Bcd proteins at the *hb* regulatory region[3,13]. To test this idea, we performed the same analysis for a fly strain in which a *lacZ* reporter gene is driven by three high-affinity Bcd binding sites fused to a minimal promoter (*bcd3-lacZ*)[14] (Online Methods). The analysis yielded $h = 3.1 \pm 0.2$ (**Fig. 2b** and **Supplementary Fig. 7**), i.e., a Hill coefficient approximately equal to the known number of Bcd binding sites in this construct. These results suggest that the measured Hill coefficient in wild-type embryos estimates the number of Bcd proteins that bind cooperatively to regulate *hb* expression.

Our next goal was to use the measured variability in number of nascent *hb* mRNAs (**Supplementary Fig. 8** and **Supplementary Note**) to extract information about the way Bcd modulates the stochastic kinetics of promoter activity. This was done by comparing the observed mRNA copy-number statistics to the prediction of a theoretical model describing *hb* kinetics. In the model, the gene stochastically switches between an inactive (OFF) and an active (ON) state, and transcription initiation is possible only in the active state. ON and OFF switching and transcription initiation are characterized by Poissonian rates k_{ON}, k_{OFF} and k_{INI}, respectively[2]. Transcription initiation is followed by mRNA elongation (with speed V_{EL}) and release, both modeled as deterministic processes (**Fig. 2c**). We solved the model to obtain the steady-state distribution of nascent *hb* mRNAs per gene locus as a function of the kinetic parameters (**Supplementary Note**).

To compare model predictions with the observed *hb* statistics, we binned the entire single-locus data set from each embryo into multiple Bcd-concentration windows (**Fig. 2c**). We then applied maximum-likelihood parameter estimation to fit the experimental nascent mRNA distribution in each window (**Supplementary Fig. 9** and **Supplementary Note**). We found that good agreement between theory and experiment was achieved by assuming that only k_{ON} is affected by Bcd concentration, whereas k_{OFF} and k_{INI} remain constant (**Fig. 2d** and **Supplementary Figs. 9** and **10**). In support of the fitting procedure, the extracted value for the maximal transcription rate, $k_{TX} = k_{INI}\, k_{ON}^{max}/(k_{ON}^{max} + k_{OFF}) = 18.9 \pm 0.4\ min^{-1}$, was consistent with previous estimates[9,15]. As another test, we used our calculated parameters to simulate nascent mRNA accumulation at a single *hb* locus and found good agreement with the results of a recent live-embryo study[7] (**Supplementary Fig. 11** and **Supplementary Note**).

Our analysis indicated that the dependence of k_{ON} on Bcd concentration during cycles 11–13 can be approximated by a Hill function (**Fig. 2d** and **Supplementary Fig. 12**), with a Hill coefficient $h = 6.5 \pm 0.3$ (**Fig. 2e**), very close to what we found for the gene regulation function above (**Fig. 2b**). This observation suggests a direct relation between cooperative Bcd binding and the switching of *hb* to an active transcriptional state (**Supplementary Fig. 13**). Performing the same analysis for *bcd3-lacZ* yielded a Hill coefficient of 3.0 ± 0.1 (**Fig. 2e** and **Supplementary Fig. 12**), i.e., equal

© 2015 Nature America, Inc. All rights reserved.

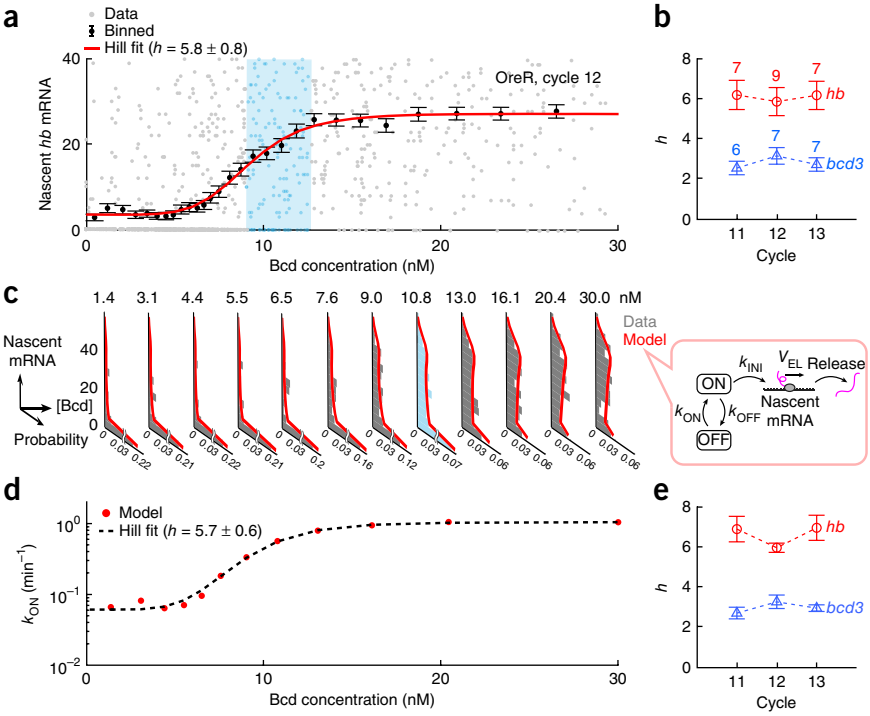

Figure 2 | The regulatory relationship between Bcd concentration and *hb* transcription. (**a**) The number of nascent mRNAs at individual *hb* loci plotted against nuclear Bcd concentration (single embryo, >600 nuclei, 0.25–0.7 EL). The single-locus data were binned along the Bcd axis (mean ± s.e.m.) and fitted to a Hill function. One window of Bcd concentration (corresponding to one histogram in **c**) is highlighted in cyan. (**b**) The Hill coefficient of the gene regulation function. Error bars represent s.e.m. from the indicated number of embryos. (**c**) Histograms of nascent *hb* mRNA at different Bcd concentration windows (200 data points per window; the mean Bcd concentration in each window is indicated above the histogram), covering the entire data set of **a**. The histogram corresponding to the highlighted window in **a** is marked in cyan. Each histogram was fitted to a two-state model of *hb* transcription (right). (**d**) The rate of stochastic gene activation, k_{ON}, estimated using the procedure above, is plotted against nuclear Bcd concentration and fitted to a Hill function. (**e**) The Hill coefficient of k_{ON} activation, calculated for the same embryos as in **b**. Error bars, s.e.m.

to the number of Bcd binding sites, thus directly supporting the connection between Bcd binding and k_{ON} modulation. We note that even at the highest Bcd concentration, the *hb* gene is active only $k_{ON}/(k_{ON} + k_{OFF}) = 53\% \pm 1\%$ of the time, which is consistent with the higher-than-Poisson noise measured at the anterior part of the embryo (**Supplementary Fig. 8** and ref. 6).

We next set out to directly measure Bcd binding at the *hb* gene locus. We used the presence of nascent mRNA (smFISH signal) to identify the spatial positions of *hb* gene copies in the nucleus (**Fig. 3a**). We then defined Bcd enrichment at the *hb* locus as the difference between the Bcd immunofluorescence signal in the vicinity of the nascent transcripts and the signal elsewhere in the nucleus[16]. In a single nucleus, the measured enrichment is predicted to have too high an uncertainty to be significant, owing to the Poissonian statistics of molecule positions in space (**Supplementary Note**). However, binning the data over ≳100 gene loci should allow the signal to emerge above the noise, and this is indeed what we observed (**Supplementary Fig. 14**). To convert the enrichment signal to the number of bound Bcd molecules, we normalized by the corresponding value measured for the *bcd3-lacZ* construct, for which the number of bound Bcd proteins at saturating concentration was assumed to be 3 (**Supplementary Fig. 14** and **Supplementary Note**).

Applying the procedure above, we measured an average binding of 5.36 ± 0.10 Bcd proteins at the *hb* locus (calculated from >21,000 loci in 31 embryos; **Supplementary Fig. 14**), which is close to the numbers we estimated above from the gene regulation function and the stochastic analysis of *hb* transcription. In contrast, measuring Bcd binding at a random position in the same nuclei yielded a value of 0.04 ± 0.09. We compared our estimation of Bcd binding at three gene loci (*hb*, *Act5C* and *nullo*) to published ChIP-chip data[17] (**Supplementary Note**) and found good agreement between the two data sets (**Fig. 3b**).

Next, we examined how Bcd binding at *hb* varied with nuclear Bcd concentration. The measured curves (**Fig. 3c** and **Supplementary Fig. 15**) exhibited plateaus suggestive of distinct binding states, with sharp transitions between these plateaus, consistent with the highly cooperative binding by Bcd at *hb*[18]. When comparing

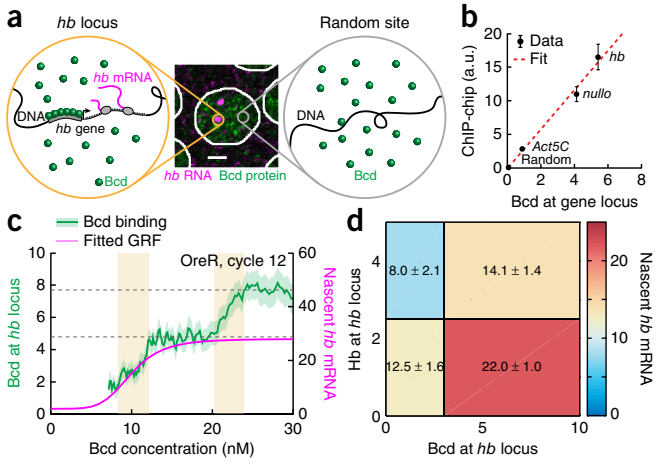

Figure 3 | Transcription factor binding at the *hb* locus. (**a**) To detect Bcd binding at the *hb* locus, we subtracted the Bcd signal at a random part of the nucleus (right, gray circle) from the signal in the vicinity of the gene (left, orange circle). (**b**) Comparison of our measured Bcd binding values with published chromatin immunoprecipitation (ChIP)-chip data (Berkeley *Drosophila* Transcription Network Project: http://bdtnp.lbl.gov/Fly-Net/). Error bars, s.e.m. (**c**) Bcd binding at the *hb* locus (shading indicates s.e.m.) and nascent *hb* mRNA (Hill fit of the gene regulation function, GRF) as a function of nuclear Bcd concentration (data from 9 embryos). Dashed gray lines highlight discrete binding plateaus. The binding curve exhibits two sharp increases (yellow shading). The left one is accompanied by the activation of *hb*, whereas the right one is not. (**d**) *hb* transcription as a function of Bcd and Hb binding. The numbers of nascent *hb* mRNAs and bound Bcd and Hb were measured for >7,600 loci in 18 embryos (nuclear cycles 11–14, 0.25–0.7 EL). The data were binned by nuclear position and then by Bcd and Hb binding levels. The mean number of nascent *hb* mRNAs in each bin is color coded as well as written (mean ± s.e.m.).

© 2015 Nature America, Inc. All rights reserved.

the Bcd binding curve to the *hb* gene regulation function (**Fig. 3c** and **Supplementary Fig. 15**), we observed that *hb* activation coincides with a sharp increase in bound Bcd, to approximately five molecules. This value is again roughly consistent with our previous conclusion that Bcd activates *hb* transcription through the cooperative binding of six or seven molecules. At higher Bcd concentrations, the transition to other binding states (**Fig. 3c** and **Supplementary Fig. 15**) is not accompanied by a noticeable change in *hb* transcription. We speculate that these binding states correspond to the occupation of additional Bcd binding sites, possibly at other *hb* enhancers[19].

Our analysis above of *hb* activity and Bcd binding showed distinct differences between nuclear cleavage cycles 11–13 and cycle 14 (**Supplementary Figs. 2**,**6**,**7** and **15**), consistent with previous reports that *hb* expression in cycle 14 is subject to additional regulation[19]. To further probe *hb* regulation during cycle 14, we focused on the role of its own gene product, the Hb transcription factor. Hb is believed to regulate *hb* expression, but it is debated whether Hb serves as an activator[20] or a repressor[19] and how it interacts with other transcription factors such as Bcd[13]. We simultaneously measured the binding of Bcd and Hb at each active *hb* locus (**Supplementary Figs. 16** and **17** and **Supplementary Note**). Analysis of the full data set (Bcd binding, Hb binding and nascent *hb* mRNA) revealed that higher Bcd binding is accompanied by higher *hb* expression, whereas higher Hb binding is accompanied by lower *hb* expression (**Fig. 3d** and **Supplementary Fig. 18**). Thus, the simultaneous binding data confirmed the positive regulatory effect of Bcd while demonstrating a repressive effect of Hb binding.

We have shown that the simultaneous quantification of transcription factors and nascent mRNA in individual nuclei and individual gene loci, followed by three layers of analysis—average transcriptional response, mRNA statistics and transcription factor binding—provides a powerful approach for investigating the stochastic process of gene regulation. In our analysis, we took advantage of the natural variation in Bcd concentration across nuclei in a single embryo. Yet even in the absence of a well-defined gradient (as found for morphogens and the transcription factors they regulate[21]), the natural cell-to-cell fluctuations in transcription factor levels are often significant[22] and span—*ipso facto*—the physiologically relevant range of regulatory concentrations, thereby extending the scenarios that can be investigated using this approach. When needed, a transcription factor can also be artificially expressed (or depleted) in order to scan a broader range of concentrations[23]. The ability to delineate the combinatorial activity of multiple transcription factors acting on the same gene is an especially promising application of our approach, as this remains a major challenge to our ability to decipher the genetic networks that drive cell behavior.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank J. Reinitz (University of Chicago) for the gift of anti-Hb antibody, and the following people for their generous advice: A. Boettiger, J. Elf, H. Garcia, D. Larson, S. Little, J. Ma, A. Raj, A. Sanchez, E. Segal, D. van Dijk and all members of the Golding and Sokac labs. Work in the Golding lab was supported by grants from the US National Institutes of Health (NIH R01 GM082837), US National Science Foundation (PHY 1147498, PHY 1430124 and PHY 1427654) and Welch Foundation (Q-1759). H.X. is supported by the Burroughs Wellcome Fund Career Award at the Scientific Interface. A.M.S. and L.F. are supported by a grant from the NIH (R01 GM115111), a Computational and Integrative Biomedical Research (CIBR) Center Seed Award and the Curtis Hankamer Basic Research Fund Award from Baylor College of Medicine. We gratefully acknowledge the computing resources provided by the CIBR Center of Baylor College of Medicine.

AUTHOR CONTRIBUTIONS

H.X., L.A.S. and I.G. conceived the experimental and analysis methods. H.X. and L.A.S. developed image and data analysis algorithms. H.X. performed the experiments, developed algorithms and theoretical models, and analyzed the data. L.F. performed fly-injection experiments. A.M.S. provided guidance on fly biology and microscopy. I.G. supervised the project. H.X., A.M.S. and I.G. wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Segal, E. & Widom, J. *Nat. Rev. Genet.* **10**, 443–456 (2009).
2. Sanchez, A. & Golding, I. *Science* **342**, 1188–1193 (2013).
3. Gregor, T., Tank, D.W., Wieschaus, E.F. & Bialek, W. *Cell* **130**, 153–164 (2007).
4. Femino, A.M., Fay, F.S., Fogarty, K. & Singer, R.H. *Science* **280**, 585–590 (1998).
5. Raj, A., van den Bogaard, P., Rifkin, S.A., van Oudenaarden, A. & Tyagi, S. *Nat. Methods* **5**, 877–879 (2008).
6. Little, S.C., Tikhonov, M. & Gregor, T. *Cell* **154**, 789–800 (2013).
7. Lucas, T. *et al. Curr. Biol.* **23**, 2135–2139 (2013).
8. Porcher, A. *et al. Development* **137**, 2795–2804 (2010).
9. Garcia, H.G., Tikhonov, M., Lin, A. & Gregor, T. *Curr. Biol.* **23**, 2140–2145 (2013).
10. Gregor, T., Wieschaus, E.F., McGregor, A.P., Bialek, W. & Tank, D.W. *Cell* **130**, 141–152 (2007).
11. Abu-Arish, A., Porcher, A., Czerwonka, A., Dostatni, N. & Fradin, C. *Biophys. J.* **99**, L33–L35 (2010).
12. Rosenfeld, N., Young, J.W., Alon, U., Swain, P.S. & Elowitz, M.B. *Science* **307**, 1962–1965 (2005).
13. Lopes, F.J., Spirov, A.V. & Bisch, P.M. *Dev. Biol.* **370**, 165–172 (2012).
14. Ronchi, E., Treisman, J., Dostatni, N., Struhl, G. & Desplan, C. *Cell* **74**, 347–355 (1993).
15. Boettiger, A.N. & Levine, M. *Cell Rep.* **3**, 8–15 (2013).
16. He, F., Ren, J., Wang, W. & Ma, J. *PLoS ONE* **6**, e19122 (2011).
17. Li, X.Y. *et al. PLoS Biol.* **6**, e27 (2008).
18. Ma, X., Yuan, D., Diepold, K., Scarborough, T. & Ma, J. *Development* **122**, 1195–1206 (1996).
19. Perry, M.W., Bothma, J.P., Luu, R.D. & Levine, M. *Curr. Biol.* **22**, 2247–2252 (2012).
20. Lopes, F.J., Vieira, F.M., Holloway, D.M., Bisch, P.M. & Spirov, A.V. *PLoS Comput. Biol.* **4**, e1000184 (2008).
21. Ashe, H.L. & Briscoe, J. *Development* **133**, 385–394 (2006).
22. Eldar, A. & Elowitz, M.B. *Nature* **467**, 167–173 (2010).
23. Niwa, H., Miyazaki, J. & Smith, A.G. *Nat. Genet.* **24**, 372–376 (2000).

© 2015 Nature America, Inc. All rights reserved.

ONLINE METHODS

Fly strains. Oregon-R (OreR) strain was used as the wild type. *egfp-bcd* strain[10] was obtained from the Bloomington *Drosophila* Stock Center (#29018).

Construction of P[*bcd3-lacZ*] and *bcd3-lacZ* strain. The plasmid P[*bcd3-lacZ*] was constructed as follows. The oligonucleotide sequence containing three Bcd binding sites (AGGTTCTAATCC CGGTCTAATCCCTCGAGTCTAATCCCATGAGTCGACG)[14] was synthesized as an insertion in pIDTSMART-KAN, between the EcoRI and BamHI sites (performed by Integrated DNA Technologies), and subcloned into pCaSpeR-hs43-lacZ (ref. 24; plasmid obtained from the *Drosophila* Genomics Resource Center) upstream of the *hsp70* minimal promoter. P-element transformation (ref. 25; performed by BestGene) was used to generate *bcd3-lacZ* transgenic fly lines. A stock with the transgene inserted on chromosome 2 was used in this study.

α-Amanitin injection. For RNA lifetime measurement, α-amanitin injection was performed following the method of ref. 26. Briefly, embryos were collected at about 2 h after deposition, dechorionated in 50% bleach and staged by eye under the stereoscope. α-Amanitin was dissolved in nuclease-free water (Ambion AM9937) at a concentration of 1 mg/ml and injected into stage 4 or 5 embryos (~50 pl per embryo) using an Eppendorf FemtoJet Express microinjector. Embryos were then incubated at room temperature for a delay time period (0, 2, 4, 6, 8 or 30 min) before being fixed. *hb* mRNA in the embryo was labeled and imaged as described below.

Embryo collection, fixation and storage. All embryos were collected at 25 °C (following the protocol of ref. 27). EGFP-Bcd expressing embryos were fixed in 8% (v/v) paraformaldehyde/phosphate buffered saline (PBS):heptane (4 ml:4 ml) for 30 min, hand devitellinized and stored in 1 ml 1× PBS (with 0.1% (w/v) BSA and 0.1% (v/v) Triton X-100) at 4 °C. All other embryos were fixed in 4% (v/v) paraformaldehyde/PBS:heptane (4 ml:4 ml) for 15 min, vortexed in methanol: heptane (4 ml: 4 ml) for 30 s for devitellinization and stored in 1 ml methanol at −20 °C.

Simultaneous labeling using smFISH and immunofluorescence. The reaction conditions for smFISH and immunofluorescence are conflicting in the following ways. (i) The typical smFISH hybridization solution is based on saline-sodium citrate (SSC) buffer[5], whereas immunofluorescence hybridization solution is usually based on PBS[28]. (ii) smFISH hybridization requires higher reaction temperature and harsher chemicals (≥30 °C, usually with formamide[5]) than immunofluorescence (4–24 °C, no harsh chemicals[28]). (iii) smFISH is RNase sensitive, whereas many immunofluorescence reagents, including blocking reagents and antibodies, are usually not RNase free.

To combine smFISH and immunofluorescence labeling, we tested three different protocol designs: (1) smFISH and immunofluorescence labeling in the same reaction[5], (2) immunofluorescence labeling followed by smFISH[29–31] and (3) smFISH labeling followed by immunofluorescence[6,16]. For each protocol, we characterized the following quantities: (i) the average smFISH intensity of the *hb* gene at the anterior side of the embryo (cycles 11–13, 0.25–0.4 EL) and (ii) the exponential slope (decay length) of the Bcd immunofluorescence signal. We found that, compared to smFISH- and immunofluorescence-only labeling, both design 1 and design 2 resulted in a decrease of the *hb* smFISH signal as well as a decrease of the Bcd immunofluorescence decay length, indicating that both smFISH and immunofluorescence signals have been distorted. In contrast, design 3 preserved both the smFISH intensity and the immunofluorescence slope. We therefore developed a combined labeling protocol based on design 3.

In this protocol, the smFISH labeling procedure was adapted from previous protocols[5,32]. Sets of DNA oligonucleotides complementary to the target gene (48 probes for *hb*, 72 probes for *lacZ*, 48 probes for *Act5C* and 20 probes for *nullo*; **Supplementary Table 1**) were designed and ordered (Biosearch Technologies). *hb* probes were directly synthesized with 3′ TAMRA labeling; all other probes were synthesized with 3′ amine modification and were conjugated to various fluorophores in the lab, as described in ref. 32. Before hybridization, embryos were rehydrated four times (10 min each) in 1 ml PBTx (1× PBS, 0.1% (v/v) Triton X-100) and washed twice (5 min each) in 1 ml hybridization wash buffer (2× SSC, 20% (w/v) formamide, 0.1% (v/v) Triton X-100). Hybridization was then performed at 30 °C for 16 h by incubating embryos with 100 nM of probes (or roughly 2 nM per probe) in 50 µl hybridization buffer (2× SSC, 20% (w/v) formamide, 0.2 mg/ml BSA(RNase free, Ambion AM2616), 2 mM ribonucleoside vanadyl complex (NEB S1402S), 1 mg/ml *Escherichia coli* tRNA (Sigma 83854-100MG), 0.1 g/ml dextran sulfate). A shaker was used to facilitate mixing. Hybridized embryos were washed three times (30 min each) in hybridization wash buffer at 30 °C and twice (10 min each) in 2× SSC (with 0.1% (v/v) Triton X-100) at room temperature. All washing steps were performed with nutation.

Immunofluorescence was performed after smFISH, using a protocol adapted from refs. 28,33. Briefly, embryos were washed four times (10 min each) in 1 ml PBTx and blocked in 1 ml PBT-B (1× PBS, 20% (v/v) western blocking reagent (Roche 11921673001), 2 mM ribonucleoside vanadyl complex (NEB S1402S), 0.1% (v/v) Triton X-100) at room temperature for 1 h. 500 µl of rabbit anti-Bcd primary antibody (Santa Cruz Biotechnology SC-66818, 1:50 (v/v) dilution in PBT-B) or guinea pig anti-Hb primary antibody[28] (gift of J. Reinitz, 1:200 (v/v) dilution in PBT-B) were preabsorbed (incubated with stage 15 or 16 wild-type embryos at 4 °C for 20 h) and then hybridized with the embryos of interest at 4 °C for 20 h. This was followed by four additional washes (10 min each) in 1 ml PBTx and a 1-h blocking in 1 ml PBT-B. Next, embryos were incubated with 500 µl of either goat anti-rabbit or goat anti–guinea pig IgG secondary antibodies conjugated with either Alexa 488 or Alexa 647 (Invitrogen, A11008, A21244, A11073, 1:500 (v/v) dilution in PBT-B) at room temperature for 1 h and washed four times (10 min each) in 1 ml PBTx. All steps were performed with nutation.

DNA staining and embryo mounting. After immunofluorescence labeling, embryos were stained with 1 ml Hoechst 33342 (Invitrogen P36236, 1:10,000 (v/v) dilution in PBTx) for 10 min on a nutator and washed four times (10 min each, with nutation) in 1 ml PBTx before mounting in Aqua-Poly/Mount (Polysciences 18606). Imaging was performed at least 36 h after mounting to allow complete solidification of the sample.

© 2015 Nature America, Inc. All rights reserved.

Microscopy. 16-bit, high-resolution (*xy* pixel size, 83 nm; *z* spacing, 350 nm) 3D image stacks were acquired on a Zeiss LSM 710 laser scanning confocal microscope using a 63× oil-immersion objective (1.4 NA) and a pinhole size of 1 Airy unit. Embryos at the mitotic interphase of nuclear cleavage cycles 11–14 were selected on the basis of the number and shape of the nuclei (Hoechst signal) and the appearance of bright spots in the smFISH channel. Three adjacent image stacks, with typical size of ~2,300 × 2,700 × 23 voxels each, were acquired automatically and stitched together (using the "tiling" function of the LSM710 imaging software) to cover the cortex layer of each embryo. Before image acquisition, the microscope was warmed up for at least 3 h to avoid possible drift of the focal plane.

Data and code availability. Image processing and data analysis were performed using custom Matlab scripts (**Supplementary Software**) and are described in detail in the **Supplementary Note**. The raw images themselves are available online at http://1drv.ms/1ztXI1a.

24. Thummel, C. & Pirrotta, V. *Drosoph. Inf. Serv.* **71**, 150 (1992).
25. Spradling, A.C. & Rubin, G.M. *Science* **218**, 341–347 (1982).
26. Edgar, B.A., Weir, M.P., Schubiger, G. & Kornberg, T. *Cell* **47**, 747–754 (1986).
27. Figard, L. & Sokac, A.M. *J. Vis. Exp.* **49**, e2503 (2011).
28. Kosman, D., Small, S. & Reinitz, J. *Dev. Genes Evol.* **208**, 290–294 (1998).
29. Toledano, H., D'Alterio, C., Loza-Coll, M. & Jones, D.L. *Nat. Protoc.* **7**, 1808–1817 (2012).
30. Zimmerman, S.G., Peters, N.C., Altaras, A.E. & Berg, C.A. *Nat. Protoc.* **8**, 2158–2179 (2013).
31. Namekawa, S.H. & Lee, J.T. *Nat. Protoc.* **6**, 270–284 (2011).
32. Skinner, S.O., Sepúlveda, L.A., Xu, H. & Golding, I. *Nat. Protoc.* **8**, 1100–1113 (2013).
33. Jaeger, J. *et al. Nature* **430**, 368–371 (2004).

© 2015 Nature America, Inc. All rights reserved.

doi:10.1038/nmeth.3446