Current Biology 24, 1283–1288, June 2, 2014 ©2014 Elsevier Ltd All rights reserved http://dx.doi.org/10.1016/j.cub.2014.04.028

Report

Maternal Origins of Developmental Reproducibility

Mariela D. Petkova,[1] Shawn C. Little,[2] Feng Liu,[1] and Thomas Gregor[1,3,*]
[1]Joseph Henry Laboratories of Physics, Princeton University, Princeton, NJ 08544, USA
[2]Department of Molecular Biology, Howard Hughes Medical Institute, Princeton University, Princeton, NJ 08544, USA
[3]Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544, USA

Summary

Cell fate decisions during multicellular development are precisely coordinated, leading to highly reproducible macroscopic structural outcomes [1–3]. The origins of this reproducibility are found at the molecular level during the earliest stages of development when patterns of morphogen molecules emerge reproducibly [4, 5]. However, although the initial conditions for these early stages are determined by the female during oogenesis, it is unknown whether reproducibility is perpetuated from oogenesis or reacquired by the zygote. To address this issue in the early *Drosophila* embryo, we sought to count individual maternally deposited *bicoid* mRNA molecules and compare variability between embryos with previously observed fluctuations in the Bicoid protein gradient [6, 7]. Here, we develop independent methods to quantify total amounts of mRNA in individual embryos and show that mRNA counts are highly reproducible between embryos to within ~9%, matching the reproducibility of the protein gradient. Reproducibility emerges from perfectly linear feedforward processes: changing the genetic dosage in the female leads to proportional changes in both mRNA and protein numbers in the embryo. Our results indicate that the reproducibility of the morphological structures of embryos originates during oogenesis, which is when the expression of maternally provided patterning factors is precisely controlled.

Results

Cells along the anterior-posterior (AP) axis of the developing *Drosophila* embryo determine their location by interpreting concentrations of morphogen molecules that correlate with AP position. One process leading to these molecular patterns (reviewed in [8]) originates in the female during oogenesis when maternal mRNA of the anterior determinant *bicoid* (*bcd*) is localized at the anterior pole of the egg. Upon egg activation, these mRNA molecules serve as sources for a protein gradient, which triggers a network of interacting genes that generate a cascade of increasingly diversified molecular spatial patterns, eventually specifying unique fates for each of the ~80 rows of cells along the AP axis [9]. The diffusion-driven, exponentially decaying concentration profile of Bcd protein is reproducible to within 10% from embryo to embryo [7], which is sufficient to encode position with 1.6% embryo length (EL) precision. Spatial precision is observed in all features of the subsequent molecular and morphologic patterns [10–13], including the first macroscopic structure in the embryo, the cephalic furrow [14–16]. It is unknown, however, whether such high reproducibility is also realized at the level of maternal mRNA. Is reproducibility of the protein gradient determined during oogenesis, or is it acquired in the embryo via specialized error-correcting mechanisms [5, 12, 17]? Here, we test whether the expression of *bcd* during oogenesis is controlled with 10% or better precision and determine the quantitative, mechanistic constraints on the amount of mRNA deposited into the oocyte.

To address whether the female confers reproducibility to the zygote through control of mRNA, we devised two strategies to quantify *bcd* mRNA molecules in individual embryos. Measuring reproducibility in intact embryos requires a measurement error that is low compared to the actual embryo-to-embryo fluctuations in mRNA numbers; we therefore sought to count individual molecules, which can only be achieved by an optical method. In wild-type embryos, optically resolving individual *bcd* mRNA molecules is hindered by the packaging of *bcd* mRNA into ribonuclear protein complexes containing variable multiples of mRNAs [18]. The formation of these particles requires the protein Staufen (Stau) [19]. Therefore, we optically measured *bcd* mRNA in embryos from *stau* mutant females (referred to hereafter as *stau*$^-$ embryos) and developed an alternative counting method based on bulk quantitative PCR (qPCR) measurements to confirm that our results also apply to wild-type.

Total mRNA Counts in Individual Intact Embryos

To optically identify individual mRNA molecules in whole-mount embryos and to assess embryo-to-embryo reproducibility, we extended a recently developed mRNA labeling method of fluorescence in situ hybridization (FISH) [18, 20]. We labeled *bcd* mRNAs with synthetic probes and then counted individual molecules and measured their fluorescence intensity by confocal microscopy (Figure 1; Figure S1 available online). In wild-type embryos, this technique revealed a bimodal intensity distribution of *bcd* mRNA particles (Figures 1A and S1B) held together by Stau [19]. We resolved these *bcd* complexes into individual mRNA molecules in *stau*$^-$ embryos (Figure 1B) [20], with a unimodal particle intensity distribution (Figure S1B). We counted $M_{bcd}^{\mathrm{FISH}} = (8.9 \pm 0.3) \times 10^5$ *bcd* mRNA molecules in individual *stau*$^-$ embryos; the error bar is the SE of the mean across n = 7 embryos (Supplemental Experimental Procedures).

To confirm that the number of *bcd* mRNA molecules in *stau*$^-$ embryos was comparable to that of wild-type, we modified a widely used PCR technique [21] to count molecules in wild-type and *stau*$^-$ embryos. This technique also allowed us to verify that the fluorescent particles in *stau*$^-$ embryos corresponded to individual mRNA molecules. In quantitative RT-PCR (qRT-PCR), mRNA molecules are chemically extracted from the sample, converted to DNA by reverse transcription, and subsequently quantified by real-time PCR amplification using a SYBR Green fluorescence reporter. Usually, qRT-PCR cannot measure absolute mRNA in biological samples, mainly due to challenges in quantifying the process of

*Correspondence: tg2@princeton.edu

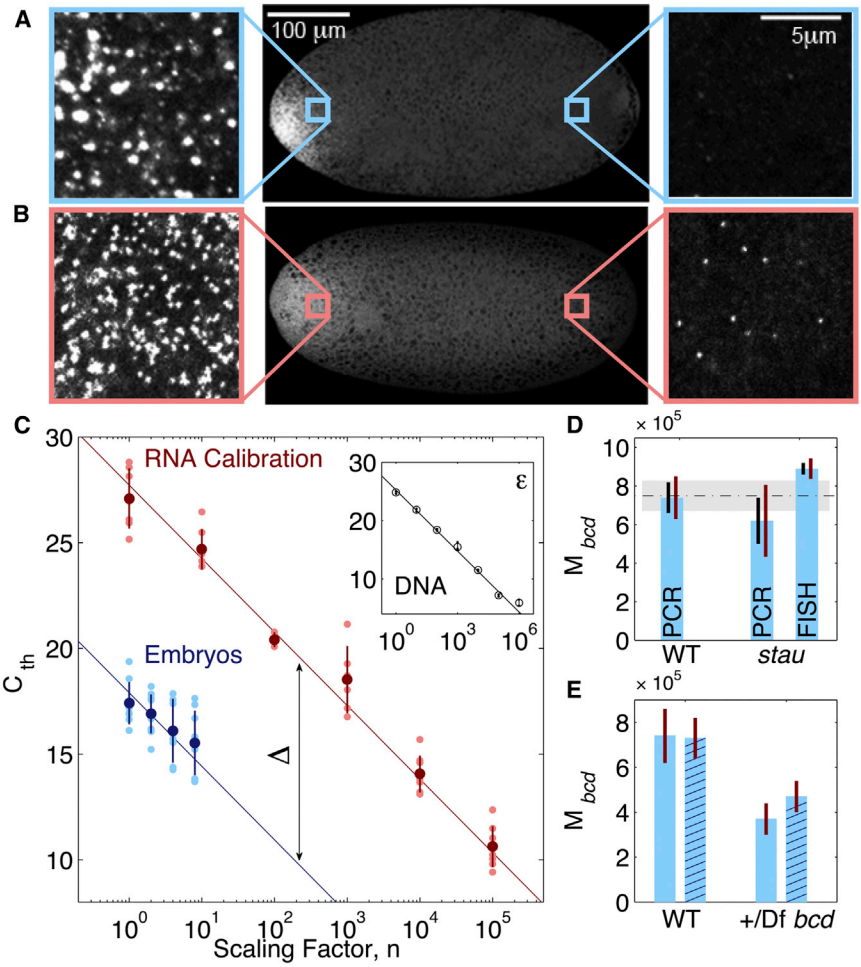

Figure 1. Two Independent Methods Measure ~750,000 *bcd* mRNA Molecules in Individual Embryos

(A and B) In situ total mRNA measurements using single-molecule counting. *bcd* mRNA molecules in fixed embryos are tagged with 90 fluorescently labeled oligonucleotide probes (20-mer) for wild-type (A) and *stau*$^{-}$ mutant (B) embryos. Particles containing multiple *bcd* mRNA molecules dissociate in *stau*$^{-}$ mutants (side panels), as apparent from unimodal particle intensity distribution (Figure S1).

(C) Bulk mRNA is extracted from individual samples (colored data points), converted to cDNA, and quantified via qPCR. Threshold cycle C_{th} refers to the PCR amplification cycle for which the sample fluorescence reaches an arbitrary predefined threshold at which all samples are compared. DNA dilution series (inset) measures the PCR amplification efficiency $\varepsilon = 10^{-1/\text{slope}}$ ($\varepsilon = 1.98$; error bars are smaller than data markers). An RNA calibration (red) is constructed from samples with $n \times M_{ref}$ synthetically generated mRNA molecules. An embryo series (blue) is constructed from samples containing n = (1, 2, 4, 8) embryos (10–30 min old). The slope of the blue and red lines is set by ε, and the offset (Δ) between the lines determines the number (M_{bcd}) of total mRNA molecules per embryo (Supplemental Experimental Procedures).

(D) qRT-PCR measurements of total *bcd* mRNA in wild-type and *stau*$^{-}$ embryos yield a mean of $7.4 \times 10^5 \pm (0.8, 1.1) \times 10^5$ (SEM, measurement error) and $6.2 \times 10^5 \pm (1.2, 1.9) \times 10^5$ molecules, respectively. SEM is shown as black bars. Measurement errors are estimated from the single-parameter fitting (red bars). With FISH in seven *stau*$^{-}$ embryos, $(8.9 \pm 0.3) \times 10^5$ molecules are detected (error is SEM; we estimate the overall FISH measurement error to be ~6%). Dashed horizontal line and gray shading correspond to the mean of the three measurements and the SEM, i.e., $M_{bcd} = (7.5 \pm 0.8) \times 10^5$.

(E) qRT-PCR measurements of total *bcd* mRNA counts from two independent large-scale experiments (empty and hashed bars) are $(7.4 \pm 1.2) \times 10^5$ and $(7.3 \pm 0.9) \times 10^5$ in wild-type embryos (two *bcd* DNA copies) and $(4.7 \pm 0.7) \times 10^5$ and $(3.7 \pm 0.7) \times 10^5$ in +/Df *bcd* embryos (one *bcd* DNA copy). The number of samples in each experiment is found in Table S1.

mRNA isolation [22, 23]. By quantifying all systematic errors along the different processing steps, we developed a scheme to accurately estimate *bcd* mRNA molecules in individual embryos.

In our strategy, the largest quantitative effect was achieved through controlling for losses associated with RNA isolation; mRNA molecules from homogenized embryos were compared to an mRNA reference calibration from a dilution series of synthetically generated *bcd* mRNA molecules undergoing the same procedure in parallel (Supplemental Experimental Procedures and Figure S2). To measure the number of *bcd* mRNAs by qRT-PCR, the mRNA reference calibration was compared to an embryo series with n = (1, 2, 4, 8) individuals. The comparison in Figure 1C shows two lines, the slope of which is determined by the PCR efficiency (ε), whereas their offsets (Δ) depend on the combined efficiency of mRNA isolation and reverse transcription (η). These quantities were measured with independent calibrations, which minimize our experimental error (Supplemental Experimental Procedures). Specifically, we first used a dilution series of *bcd* DNA molecules to precisely measure the slope ($S = -1/\log(\varepsilon)$), with an accuracy of better than 1%. We used this slope in order to perform one-parameter fits for the mRNA calibration and embryo series and thus determine the offset (Δ). The number of mRNAs per embryo is then given by $M_{bcd}^{PCR} = M_{ref}\varepsilon^{-\Delta}$, where M_{ref} is the number of synthetic mRNAs of the lowest member of the mRNA dilution series.

Using this technique, we found the number of *bcd* mRNA molecules in embryos from wild-type females to be $M_{bcd}^{PCR} = (7.4 \pm 1.1) \times 10^5$, where the error represents the measurement error on the mean number of mRNAs per embryo. *stau*$^{-}$ embryos contain a similar number of *bcd* mRNA molecules ($M_{bcd,stau^-}^{PCR} = (6.2 \pm 1.9) \times 10^5$; Figure 1D). These results confirm that, within measurement error, (1) single-molecule FISH counts indeed correspond to individual mRNAs, and (2) *bcd* mRNA counts in *stau*$^{-}$ mutant embryos are equivalent to counts in wild-type embryos.

These experiments produced three independent measures for the total *bcd* mRNA count in individual embryos: bulk qPCR measurements on wild-type and *stau*$^{-}$ embryos and FISH measurements on *stau*$^{-}$ embryos (Figure 1D). To assign a value for our overall estimate, we averaged the three independent measurements, yielding $M_{bcd} = (7.5 \pm 0.8) \times 10^5$ (SEM). The consistency among the three measures validates the FISH-based counting method in assessing embryo-to-embryo *bcd* mRNA count reproducibility.

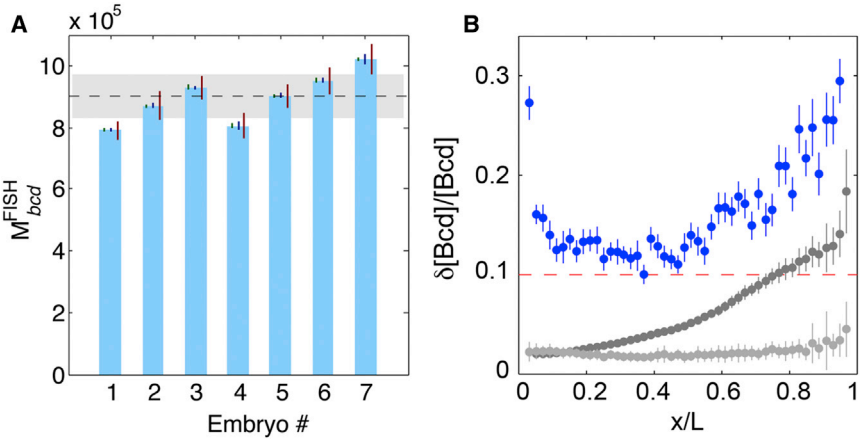

Figure 2. The *bcd* mRNA Source and the Bcd Protein Gradient Are ~10% Reproducible

(A) Total counts of *bcd* mRNA using FISH after density correction in seven individual *stau*$^{-}$ embryos. Mean (dashed black line) and SD (gray area) are (8.9 ± 0.8) × 10^5, which leads to a reproducibility of 0.8/8.9 ~ 9%.

(B) Reproducibility of Bcd protein profiles of 22 embryos expressing Bcd-GFP at wild-type Bcd levels (reference fly line Bcd2X$_A$ [14]). Average nuclear Bcd-GFP intensities from the midsagittal plane of all embryos (imaged live during nuclear cycle 14; on average 85 nuclei per embryo) are binned in 50 bins, over which the mean and SD were computed. For each bin, the SD divided by the mean (δ[*Bcd*]/[*Bcd*]) as a function of fractional egg length (*x*/*L*) is shown in blue (error bars are computed by bootstrapping with 15 embryos [7]). Gray and black lines show estimated contributions of measurement noise (see Supplemental Experimental Procedures). Dashed red line indicates 10% reproducibility mark, attained when subtracting gray from blue data.

Maternally Deposited mRNA Molecules Are as Reproducible as Zygotic Patterns

FISH allows counting of single molecules with measurement error low enough to assess embryo-to-embryo variability, which is impossible to assess with the PCR bulk measurements due to the large systematic error. We estimate our overall measurement error for FISH to be 6% of the total counts (Figure 2A and Supplemental Experimental Procedures). We measured the embryo-to-embryo reproducibility as the SD of *bcd* mRNA counts in seven *stau*$^{-}$ mutant embryos (Figure 2A). Strikingly, the result was 9% ± 2% (bootstrapping error). As can be seen in Figure 2B, this level of reproducibility is of the same order as that of the Bcd protein gradient in the anterior half of the embryo (~10% [7]). The similarity suggests that the low variability in *bcd* mRNA production is responsible for the reproducibility of the protein gradient. This is consistent with the idea that the large number of mRNA molecules minimizes the variability inherent to the processes of translation and protein transport, thus obviating requirements for error correction by a separate process.

mRNA and Protein Counts Scale Linearly with Maternal Gene Dosage

The consistency of the reproducibility levels in Figure 2 indicates that control over the mRNA source composition is sufficiently precise to generate a reproducible Bcd gradient. The question remains, however, how such reproducibility is achieved at the mRNA level, and, in particular, whether error correction and feedback mechanisms are required to ensure the correct degree of *bcd* gene expression during oogenesis. In *Drosophila*, 15 specialized germline cells, called nurse cells, remain associated with the oocyte by intercellular cytoplasmic bridges. Nurse cells synthesize maternal components necessary for early embryonic development at high rates and transport them to the growing oocyte over a >2.5-day period, enabling oogenesis to proceed rapidly (reviewed in [24]). High levels of gene expression are facilitated by extensive DNA replication. Such polyploidization generates multiple gene copies in nurse cells, though not necessarily covering the entire genome [25], rendering polyploidization a potentially error-prone process [26]. Thus, it is unclear what mechanisms generate the observed 10% reproducibility levels, or whether error correction is necessary.

To understand the mechanisms controlling the transitions from one molecular species to the next, we examined the link between the maternal *bcd* gene dosage and both *bcd* mRNA and Bcd protein numbers in the embryo. First, using qRT-PCR, we assessed how the number of deposited *bcd* mRNA molecules scales with *bcd* gene copy number. We compared our measurements in wild-type to a strain in which one copy of the *bcd* gene was deleted. The number of *bcd* mRNA molecules in embryos from these *bcd*-deficient (+/Df *bcd*) flies was 0.57 ± 0.14 times that of wild-type embryos with two *bcd* alleles (Figure 1E). The measurement is well within the measurement error of the expected factor of 2, consistent with previously observed differences in relative mRNA levels [27].

Although these measurements are consistent with a linear relationship between *bcd* dosage and actual *bcd* mRNA numbers in the embryo, to achieve a more precise measurement with more than two data points, we took advantage of a strategy exploiting the chromosome position effect [28] to generate small changes in the levels of maternal *bcd* gene product. Transgene constructs expressing Bcd-GFP fusion proteins are inserted at random locations in the fly genome, leading to distinct expression rates and quantitatively different spatial Bcd-GFP distributions [14]; i.e., the same integer copy number generates distinct transcript numbers compared to a calibrated transgenic reference fly line producing Bcd-GFP at wild-type levels (Figure 3, top inset; Figure S3). Using this strategy, we generated a set of six transgenic fly lines, with changes in gene expression levels at increments of 10%. Genetic combinations of individual *bcd-gfp* alleles yield expression levels in the resulting lines that are, within error bars, the sum of the expression levels of the original lines (gray data points in Figure 3), as observed in [14].

To test whether this linear relationship is also preserved at the level of *bcd* mRNA, we repeated our qPCR experiments to measure relative mRNA levels between wild-type and three fly lines expressing Bcd-GFP at 1.78-, 2.12-, or 2.4-fold-increased levels from a total of four or six transgenes (Figure S4). For the same fly lines, we measured the total amount of Bcd-GFP by optically calibrating GFP intensity to that of a purified GFP solution of known molarity [7]. As Figure 3 shows, a scatterplot of Bcd-GFP protein versus mRNA number demonstrates a linear relationship, perfectly matching previous

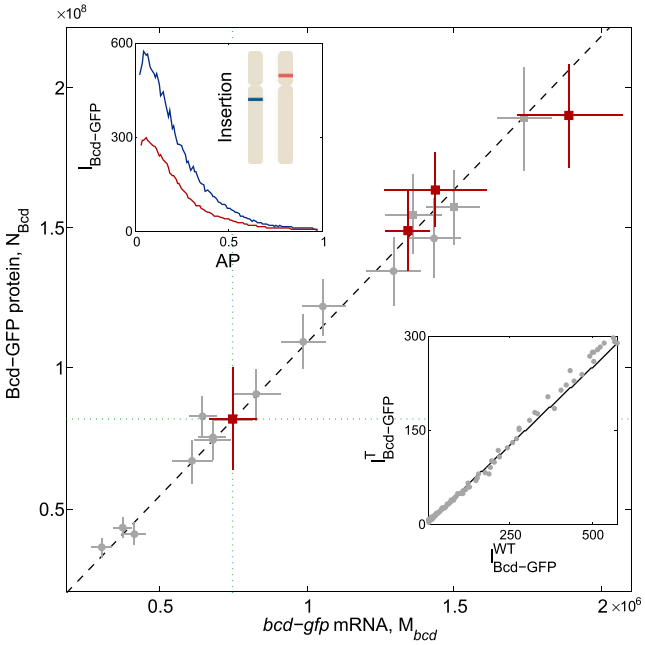

Figure 3. Total Bcd Protein in Embryos Scales Linearly with Maternal *bcd* Gene Expression

Bcd-GFP transgenes are inserted at different genome locations to generate alleles with varying *bcd* gene expression [14]. Insertions in two such fly lines (red and blue bands in top inset) are expressed at different rates, leading to quantitatively different Bcd protein intensity profiles $I_{Bcd-GFP}$ (Figure S3). Differences in protein concentration are determined by comparing average nuclear Bcd-GFP intensities across AP positions between a given transgenic test line (I^T) and a reference line (I^{WT}). The latter expresses *egfp-bcd* mRNA at the same level as endogenous *bcd* mRNA (Figure S4B) and produces Bcd-GFP at wild-type levels that restore normal patterning to *bcd* mutants. The slope of a line fitted to a scatterplot of intensities defines the Bcd-GFP expression level of the test line relative to the reference line (bottom inset; in this example, 21 and 108 embryos of the test and reference lines, respectively). The main panel shows a scatterplot of Bcd-GFP protein expression versus *bcd-gfp* mRNA expression measured by PCR (Figure S4A) in embryos from four fly lines (red squares). The relationship between protein expression and number of deposited mRNAs in the embryo is linear. For 18 fly lines, the gray data points correspond to Bcd-GFP expression of genetically combined alleles (wild-type calibrated to GFP standard) as a function of the sum of the separately measured *bcd-gfp* mRNA expression levels of the individual alleles (calibrated to wild-type mRNA counts). Wild-type is marked by vertical and horizontal dotted green lines. To calculate *bcd-gfp* mRNA expression, we measured Bcd-GFP protein expression [14] for each allele and converted it to mRNA expression (Figure S4A). Squares identify respective fly lines in four of the gray data sets and in the four red data sets. The linear relationship demonstrates that expression levels of individual alleles add independently in a linear feedforward manner. Gray error bars are as quantified previously [14]. Red error bars are SDs in expression level measurements; red error bars of data point of reference fly line ($Bcd2X_A$; crossing of dotted lines) represent calibration error to total molecular counts for protein and mRNA, respectively (Supplemental Experimental Procedures).

data from protein measurements [14]. Together, our observations indicate a scenario in which the transition from one molecular species to another is achieved in a perfectly linear feedforward manner. Specifically, small changes in the strength of the maternal *bcd* gene lead to proportional changes in both mRNA and protein counts in the embryo. All molecular processes from maternal gene expression to protein synthesis in the embryo must operate with very high precision, for which the protein measurements give an estimate: 10% changes in *bcd* strength at the level of maternal DNA are reflected as 10% changes in protein copy numbers.

Finally, measurements of both mRNA and protein numbers provide unique quantitative access to *bcd* translation. Bcd-GFP protein levels are linearly proportional to *bcd* and *gfp* mRNA amounts (Figure S4A), and the transgenic reference line produces the same amount of mRNA as the endogenous *bcd* locus, as determined by measuring *bcd* and *gfp* mRNA levels in the same embryos (Figure S4B). Previous intensity measurements of fluorescently tagged *bcd* mRNAs have shown that they are not degraded and remain stable throughout the first 2 hr of development (i.e., early nuclear cycle 14) [20]. During this period, Bcd protein is degraded uniformly with a lifetime of $\tau = 50$ min [29]. In the simplest model, Bcd translation is uniform, and the number of expressed proteins at time t is given by $N_{Bcd}(t) = k\tau M_{bcd}(1 - \exp(-t/\tau))$, where k is the translation rate and M_{bcd} is the number of *bcd* mRNAs. Using GFP calibration [7], we measured the total amount of Bcd-GFP in fixed embryos in which all Bcd-GFP proteins mature and become optically detectable [18]. We found a total of $N_{Bcd} = (8.2 \pm 1.8) \times 10^7$ (SD) Bcd-GFP molecules in approximately 2-hr-old embryos (i.e., $t_{14} = 146$ min). The value is larger than earlier estimates in live embryos [7] due to the maturation correction [18] and lower than semiquantitative biochemical measurements [29] (see Supplemental Experimental Procedures). Thus, with $M_{bcd} = (7.5 \pm 0.8) \times 10^5$ mRNA molecules, we calculated a translation rate of $k \approx 2$ proteins per mRNA per minute, which matches previously reported translation rates during the development of sea urchin embryos [30].

Discussion

We developed protocols to count the number of *bcd* mRNA molecules in individual *Drosophila* embryos. The total number is quite large, approaching one million molecules, and it is reproducible between embryos, with fluctuations of less than 10%. Such low-level fluctuations were previously observed for the Bcd protein gradient. Our present results support the idea that reproducibility is determined during oogenesis by females expressing a tightly controlled number of *bcd* mRNAs. These data show that the processes that establish the protein gradient, from maternal *bcd* DNA to Bcd protein in the zygote, are governed by linear feedforward mechanisms, which minimize requirements for error correction during gradient formation.

Our results are consistent with work showing that the segmentation gene network of early embryos acts as a relay to propagate molecular reproducibility [7, 11, 12, 14, 20, 31]. After 3 hr, cellular identities are generated with single-cell precision, as observed in gene expression patterns that arise downstream of maternal inputs and that result in morphogenetic events (e.g., the formation of the cephalic furrow). Here, we have added another element upstream of that cascade, demonstrating that reproducibility in the embryo is established during oogenesis through precise control over *bcd* mRNA expression. Hence, morphological features emerge reproducibly between embryos as a result of the precision with which the female controls initial patterning signals. The early segmentation cascade in the *Drosophila* embryo is therefore a striking example of a molecular network integrating precise initial conditions to achieve reproducible macroscopic outcomes. Thus, at least for wild-type embryos in the laboratory, early patterning does not require independent system(s) to monitor or correct fluctuations in molecular activities during

the transitions between mRNA and protein or between maternal inputs and zygotic outputs.

Given high levels of reproducibility, our finding that molecular signals are transmitted in a linear feedforward manner seems surprising. Feedforward processes have a propensity for escalating noise, yet we found that changes as low as 10% in *bcd* mRNA result in 10% changes in protein. Thus, the network establishing the Bcd gradient maintains precision at each transition between different molecular species. Particularly, reproducible *bcd* mRNA counts and the proportionality to maternal copy number indicate that the synthesis and transport of mRNA are precisely controlled and proceed essentially identically in different females. Here, control over molecular fluctuations is achieved by temporal and spatial averaging and by large molecule numbers. Nearly one million *bcd* mRNA molecules are generated, completely overriding Poisson fluctuations, matching other systems in which high mRNA levels reduce variability [32, 33]. Moreover, *bcd* is generated from multiple (>500) *bcd* DNA sources in polyploid nurse cells [34, 35] over a period of several days [24, 36]. This is conducive to spatial and temporal averaging, particularly when comparing this time course to the timescale of microscopic events, such as the synthesis of a single mRNA molecule.

Finally, the linear response in Bcd amounts to changes in the maternal gene copies indicates identical *bcd* translation kinetics across embryos and hence tight control over translation rates [37]. After fertilization, mRNA translation and establishment of the Bcd gradient require only ~1 hr. Thus, we expect that the translational apparatus operates at near maximal rate in the early embryo. However, the translation of *bcd* mRNA is an order of magnitude lower than the typical translation rate for eukaryotic organisms under optimal conditions [38, 39] (Supplemental Experimental Procedures). We suggest that to achieve reproducibility, embryos employ a lower-than-maximal translation rate with large numbers of source mRNA molecules, facilitating spatial and temporal averaging [7, 20, 40]. How these interactions are matched to produce reliable patterning remains unknown. Overall, the Bcd gradient as a paradigm for morphogen-mediated patterning now presents an ideal system to analyze how molecular networks are coordinated to maintain precision and reproducibility.

Experimental Procedures

FISH, confocal microscopy, and image analysis were performed as described in [18, 20], where capability of capturing all mRNA molecules in whole embryos was demonstrated. qRT-PCR quantification of transcripts in embryos was carried out with SYBR Green on an Applied Biosystems 7900HT Fast Real-Time PCR system using standard temperature protocol and automatic threshold detection. The absolute amount of *bcd* mRNA in extracts and of Bcd-GFP in live embryos is measured by calibration fluorescence standards (see Supplemental Experimental Procedures). All animal usage is under the approval of Princeton University's Institutional Animal Care and Use Committee.

Supplemental Information

Supplemental Information includes Supplemental Experimental Procedures, four figures, and two tables and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2014.04.028.

Acknowledgments

We thank E. Gavis, K. Sinsimer, S. Rutherford, and M. Tikhonov for advice and discussion; B. Bassler for providing the qPCR machine; and S. Blythe, E. Cox, K. Mody, A. Sgro, and J. Swan for comments. This work was supported by NIH Grants P50 GM071508 and R01 GM097275 and by Searle Scholar Award 10-SSP-274 to T.G.

Received: October 15, 2013
Revised: March 2, 2014
Accepted: April 11, 2014
Published: May 22, 2014

References

1. Thompson, D.W. (1942). On Growth and Form (Cambridge: Cambridge University Press).
2. Maynard-Smith, J. (1960). Continuous, quantized and modal variation. Proc. R. Soc. Lond. B. *152*, 397–409.
3. Lawrence, P.A. (1973). In Developmental Systems: Insects, *Volume 2*, Doane, W.W., Counce, S.J., and Waddington, C.H., eds. (New York: Academic Press), p. 157.
4. Kerszberg, M., and Wolpert, L. (2007). Specifying positional information in the embryo: looking beyond morphogens. Cell *130*, 205–209.
5. Arias, A.M., and Hayward, P. (2006). Filtering transcriptional noise during development: concepts and mechanisms. Nat. Rev. Genet. *7*, 34–44.
6. Driever, W., and Nüsslein-Volhard, C. (1988). A gradient of bicoid protein in Drosophila embryos. Cell *54*, 83–93.
7. Gregor, T., Tank, D.W., Wieschaus, E.F., and Bialek, W. (2007). Probing the limits to positional information. Cell *130*, 153–164.
8. Jaeger, J. (2011). The gap gene network. Cell. Mol. Life Sci. *68*, 243–274.
9. Gergen, J.P., Coulter, D., and Wieschaus, E.F. (1986). Segmental pattern and blastoderm cell identities. In Gametogenesis and the Early Embryo, J.G. Gall, ed. (New York: Alan R. Liss, Inc.), pp. 195–220.
10. Crauk, O., and Dostatni, N. (2005). Bicoid determines sharp and precise target gene expression in the Drosophila embryo. Curr. Biol. *15*, 1888–1898.
11. Dubuis, J.O., Samanta, R., and Gregor, T. (2013). Accurate measurements of dynamics and reproducibility in small genetic networks. Mol. Syst. Biol. *9*, 639.
12. Houchmandzadeh, B., Wieschaus, E., and Leibler, S. (2002). Establishment of developmental precision and proportions in the early Drosophila embryo. Nature *415*, 798–802.
13. Surkova, S., Kosman, D., Kozlov, K., Manu, Myasnikova, E., Samsonova, A.A., Spirov, A., Vanario-Alonso, C.E., Samsonova, M., and Reinitz, J. (2008). Characterization of the Drosophila segment determination morphome. Dev. Biol. *313*, 844–862.
14. Liu, F., Morrison, A.H., and Gregor, T. (2013). Dynamic interpretation of maternal inputs by the Drosophila segmentation gene network. Proc. Natl. Acad. Sci. USA *110*, 6724–6729.
15. Namba, R., Pazdera, T.M., Cerrone, R.L., and Minden, J.S. (1997). Drosophila embryonic pattern repair: how embryos respond to bicoid dosage alteration. Development *124*, 1393–1403.
16. Vincent, A., Blankenship, J.T., and Wieschaus, E. (1997). Integration of the head and trunk segmentation systems controls cephalic furrow formation in Drosophila. Development *124*, 3747–3754.
17. Gierer, A. (1991). Regulation and reproducibility of morphogenesis. Semin. Dev. Biol. *2*, 83–93.
18. Little, S.C., Tkačik, G., Kneeland, T.B., Wieschaus, E.F., and Gregor, T. (2011). The formation of the Bicoid morphogen gradient requires protein movement from anteriorly localized mRNA. PLoS Biol. *9*, e1000596.
19. Ferrandon, D., Elphick, L., Nüsslein-Volhard, C., and St Johnston, D. (1994). Staufen protein associates with the 3'UTR of bicoid mRNA to form particles that move in a microtubule-dependent manner. Cell *79*, 1221–1232.
20. Little, S.C., Tikhonov, M., and Gregor, T. (2013). Precise developmental gene expression arises from globally stochastic transcriptional activity. Cell *154*, 789–800.
21. Bustin, S.A. (2004). A–Z of Quantitative PCR (La Jolla: International University Line).
22. Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., et al. (2009). The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin. Chem. *55*, 611–622.
23. Nolan, T., Hands, R.E., and Bustin, S.A. (2006). Quantification of mRNA using real-time RT-PCR. Nat. Protoc. *1*, 1559–1582.

24. Spradling, A. (1993). Developmental genetics of oogenesis. In The Development of *Drosophila melanogaster*, *Volume 1*, Bate, M., and Arias, A.M., eds. (Cold Spring Harbor: Cold Spring Harbor Laboratory Press), pp. 1–70.
25. Claycomb, J.M., and Orr-Weaver, T.L. (2005). Developmental gene amplification: insights into DNA replication and gene expression. Trends Genet. *21*, 149–162.
26. Fox, D.T., Gall, J.G., and Spradling, A.C. (2010). Error-prone polyploid mitosis during normal Drosophila development. Genes Dev. *24*, 2294–2302.
27. Cheung, D., Miles, C., Kreitman, M., and Ma, J. (2011). Scaling of the Bicoid morphogen gradient by a volume-dependent production rate. Development *138*, 2741–2749.
28. Markstein, M., Pitsouli, C., Villalta, C., Celniker, S.E., and Perrimon, N. (2008). Exploiting position effects and the gypsy retrovirus insulator to engineer precisely expressed transgenes. Nat. Genet. *40*, 476–483.
29. Drocco, J.A., Grimm, O., Tank, D.W., and Wieschaus, E. (2011). Measurement and perturbation of morphogen lifetime: effects on gradient shape. Biophys. J. *101*, 1807–1815.
30. Bolouri, H., and Davidson, E.H. (2003). Transcriptional regulatory cascades in development: initial rates, not steady state, determine network kinetics. Proc. Natl. Acad. Sci. USA *100*, 9371–9376.
31. Dubuis, J.O., Tkacik, G., Wieschaus, E.F., Gregor, T., and Bialek, W. (2013). Positional information, in bits. Proc. Natl. Acad. Sci. USA *110*, 16301–16308.
32. Thattai, M., and van Oudenaarden, A. (2001). Intrinsic noise in gene regulatory networks. Proc. Natl. Acad. Sci. USA *98*, 8614–8619.
33. Ozbudak, E.M., Thattai, M., Kurtser, I., Grossman, A.D., and van Oudenaarden, A. (2002). Regulation of noise in the expression of a single gene. Nat. Genet. *31*, 69–73.
34. Dej, K.J., and Spradling, A.C. (1999). The endocycle controls nurse cell polytene chromosome structure during Drosophila oogenesis. Development *126*, 293–303.
35. King, R.C. (1970). Ovarian Development in *Drosophila melanogaster* (New York: Academic Press).
36. Becalska, A.N., and Gavis, E.R. (2009). Lighting up mRNA localization in Drosophila oogenesis. Development *136*, 2493–2503.
37. Bialek, W. (2012). Biophysics: Searching for Principles (Princeton: Princeton University Press).
38. Bonven, B., and Gulløv, K. (1979). Peptide chain elongation rate and ribosomal activity in Saccharomyces cerevisiae as a function of the growth rate. Mol. Gen. Genet. *170*, 225–230.
39. Qin, X., Ahn, S., Speed, T.P., and Rubin, G.M. (2007). Global analyses of mRNA translational control during early Drosophila embryogenesis. Genome Biol. *8*, R63.
40. Erdmann, T., Howard, M., and ten Wolde, P.R. (2009). Role of spatial averaging in the precision of gene expression patterns. Phys. Rev. Lett. *103*, 258101.

Current Biology, Volume *24*

Supplemental Information

Maternal Origins of Developmental Reproducibility

Mariela D. Petkova, Shawn C. Little, Feng Liu, and Thomas Gregor

SUPPLEMENTAL FIGURES

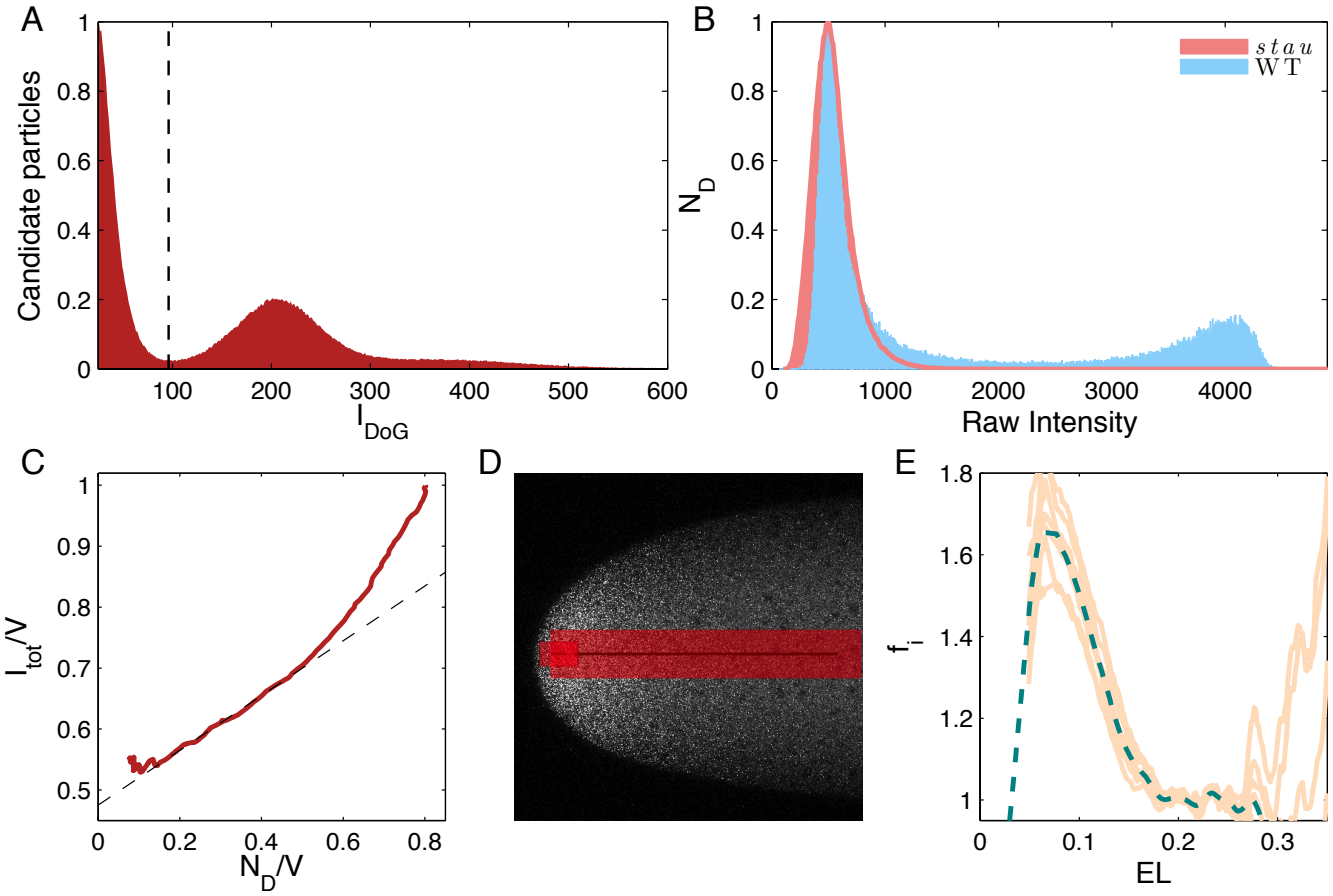

Figure S1. Measuring total number of *bcd* mRNA in whole mount *stau*⁻ embryos using smFISH, related to Figure 1. (**A**) Direct particle detection: rigid threshold (vertical dashed line) to the DoG intensities I_{DoG} of candidate particles separates false positives from directly detected particles N_{D}. The distribution is normalized to the maximum number of candidate particles. **(B)** Distribution of raw fluorescence (a.u.) per directly detected particle for wild-type (cyan) and *stau*⁻ (magenta) embryos. The low intensity peaks correspond to individual particles and the distributions have been aligned at these single molecule peaks (as fluorescence values between experiments differ by constant) for comparison. The distributions have been normalized by the respective maximum number of detected particles N_{D}. The unimodal *stau*⁻ distribution is consistent with the presence of single particles, while the bimodal distribution for wild-type embryos results from the presence of packaged mRNAs. **(C)** Total fluorescence detection of mRNAs in *stau*⁻ embryos: for moderate densities of mRNAs the total raw fluorescence I_{tot} per voxel V inside the embryo is linearly proportional to the density of directly detected particles N_{D}/V inside that voxel; at high mRNA densities multiple mRNAs can co-localize and be detected as a single particle leading to a nonlinear increase of I_{tot} with N_{D}/V. The slope α and offset β of the linear fit (dashed line) measure the intensity per mRNA particle and the background per voxel respectively. The number of true mRNA per voxel based on total fluorescence is $N_F = (I_{\mathrm{tot}} - \beta V)/\alpha$. (**D**) We sample an anterior sub-volume (shaded in red) by sliding a voxel of size 400 pxl x 400 pxl x 6 z-slices along the AP axis of the embryo to obtain α and β for the whole sub-volume. We then find the number of true mRNA per voxel based on total fluorescence $N_F(x)$ construct density correction function $f(x) = N_F(x)/N_D(x)$. For the anterior-most part <5%EL we sample with 200 pxl x 200 pxl x 6 z-slice voxels to avoid the edges of the embryo. (**E**) Individual density correction curves f_i (yellow) at subsequent z-depths within a sub-stack of the embryo are averaged and linearly interpolated to provide a global AP-dependent density correction function (blue dashed line) for the embryo (see Supplemental Experimental Procedures).

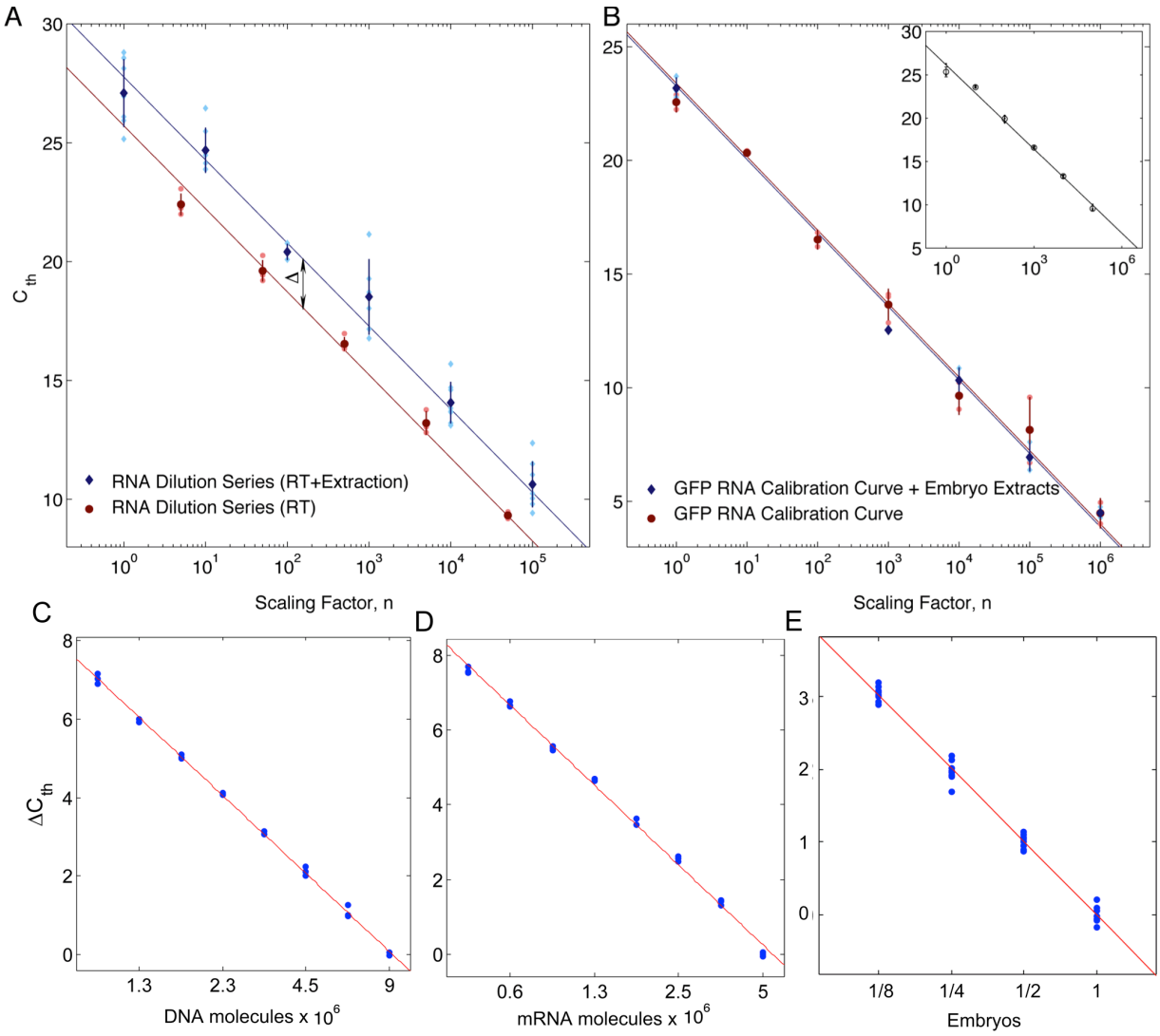

Figure S2. qRT-PCR calibration, related to Figure 1. (A) The efficiency η' of RNA isolation is measured by constructing RNA calibration series (blue diamonds) and standard RNA curve (red circles) using identical dilution series of synthetic *bcd* RNA. The efficiency is calculated from the slope and the offset between the two lines: $\eta' = \varepsilon^{-\Delta}$, with $\varepsilon = 1.98 \pm 0.05$ from *bcd* DNA series; here we find $\eta' = 0.3 \pm 0.1$. **(B)** The presence of other nucleic acids and ribonucleases in the embryo could affect downstream reactions, resulting in differences between the yield from synthetic mRNA used for the RNA calibration series and that from mRNA extracted from embryos. To check for this effect, twofold dilution series of synthetic *eGFP* mRNA over 7 orders of magnitude (starting at 10^6 molecules) are used to construct an eGFP RNA calibration series (red circles) and *eGFP* RNA calibration curve for which we add extracts from 8 embryos (ages 10–40 min) to each sample during the TRIzol step (blue diamonds). The efficiency of PCR for eGFP is calculated from series of eGFP DNA plasmid (inset), with $\varepsilon = 2.04 \pm 0.04$ corresponding to a slope of $-1/\log(\varepsilon) = -1.403 \pm 0.004$. The offset for the RNA calibration series is 23.39 ± 0.46 and for the RNA calibration series with embryo extracts is 23.26 ± 0.36. The two curves lie on top of each other well within error bars, confirming that the presence of other molecular species does not affect the mRNA measurements in our setup. In both (A) and (B) PCR is performed on all samples on the same plate and each RNA series is fitted with one-parameter linear fits to the data with slopes equal to the slope of the respective DNA series. **(C-E)** To determine *bcd* PCR efficiency at molecular numbers similar to the number of *bcd* mRNAs found in embryos, standard curves were constructed by twofold dilution of known numbers of input plasmid DNA molecules containing the *bcd* open reading frame (C), in vitro transcribed *bcd* mRNA molecules (D), or lysates of individual embryos (E). Change in threshold cycle (ΔC_{th}) is shown as a function of the amount of input material. Red lines possess slopes of -0.99 (-0.97,-1.01) (C), -1.04 (-1.01,-1.07) (D), and -1.01 (-0.97,-1.04) (E), with 95% confidence intervals reported in brackets. The slopes correspond to efficiencies of 1.97 (1.96, 2.01), 2.06, (2.01,2.1) and 2.01 (1.96, 2.06) respectively. The efficiency estimates for all three types of extracts are within confidence intervals, which is consistent with our expectation that PCR amplification is the same for each extract and the DNA, reference mRNA and embryo series are parallel.

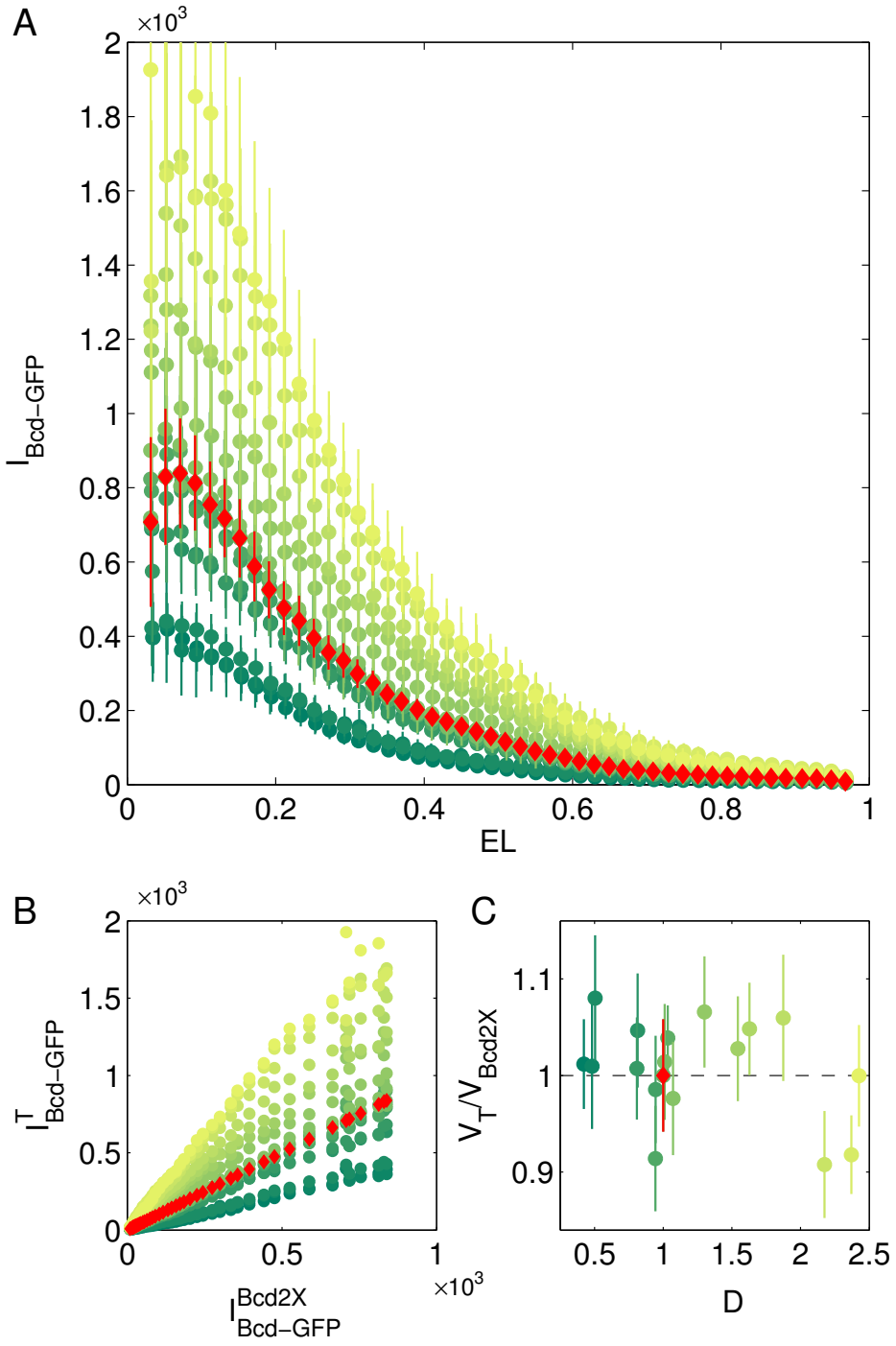

Figure S3. Bcd-GFP dosage and embryo volume measurements in transgenic fly lines, related to Figures 2 and 3. (**A**) Intensity of Bcd-GFP fluorescence along the anterior-posterior axis in all transgenic fly lines. The reference Bcd2X$_A$ fly line is shown in red. Error bars are across embryos. (**B**) The Bcd-GFP intensity of each transgenic fly line, $I^{T}_{\mathrm{Bcd-GFP}}$, is plotted against that of the reference fly line, $I^{\mathrm{Bcd2X}}_{\mathrm{Bcd-GFP}}$, to measure the ratio of their Bcd-GFP concentrations (i.e. their strength of *bcd* expression, *D*). (**C**) The embryo volume for each fly line V_T, normalized to the volume of the reference Bcd2X$_A$ line, V_{Bcd2X}. Error bars are across embryos.

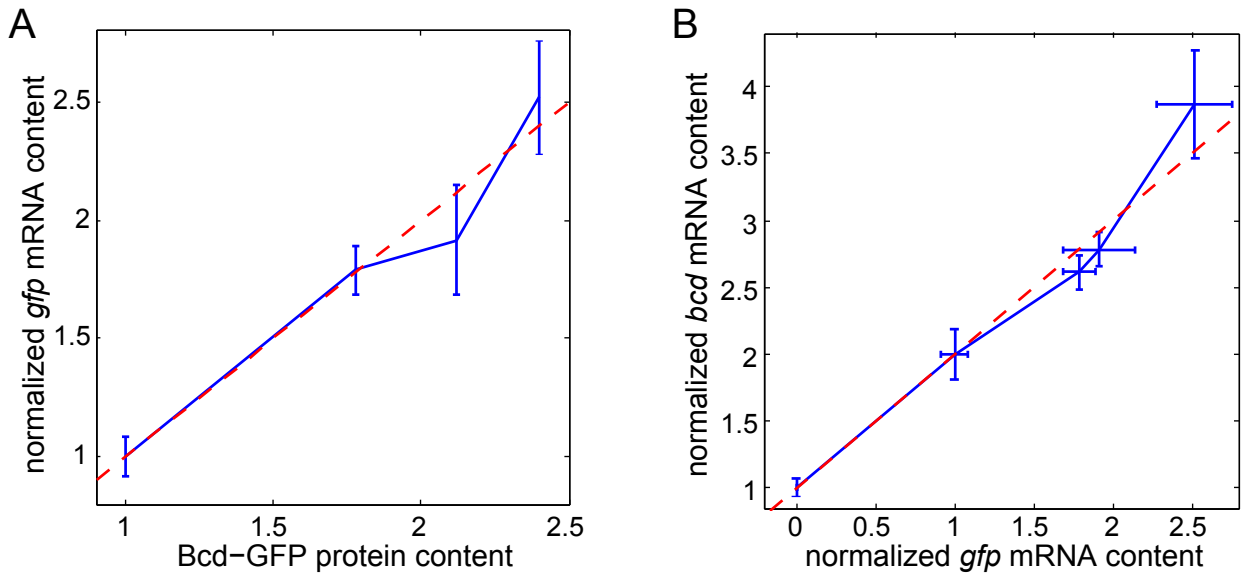

Figure S4. Linear relationship between Bcd-GFP protein, *gfp* mRNA, and *bcd* mRNA content, related to Figure 3. **(A)** Transgenic *gfp* mRNA and Bcd-GFP protein content are proportional. *gfp* mRNA content was measured by relative quantitative RT-PCR in a collection of 0–1h old embryos from four transgenic lines expressing Bcd-GFP in known quantities in a bcd^{E1} mutant background [S1]. Protein levels are shown relative to a reference line containing 2 *gfp-bcd* transgenes that restore normal pattern to bcd^{E1} mutant embryos. The reference line is indicated on the plot at a protein content of 1.0. Three additional lines are shown expressing Bcd-GFP at 1.78, 2.02, or 2.4-fold increased levels from a total of 4 or 6 transgenes. *gfp* mRNA was measured relative to *tubulin*56 mRNA as described [S2], and was then normalized to the content of the reference line. Error bars are standard error of the mean of 6 technical replicates. Red dashed line is line of slope 1, expected for perfect correlation of protein and mRNA. **(B)** *bcd* mRNA and *gfp* mRNA content is proportional. Total *bcd* mRNA content relative to *tub*56 was measured from the same embryos as in A and from a nontransgenic, wild-type line. After normalization to wild-type, *bcd* mRNA was plotted as a function of *gfp* mRNA content. Because all transgenic lines express endogenous *bcd* mRNA from the mutant bcd^{E1} locus, the four transgenic lines possess 1 additional complement of *bcd* mRNA compared to *gfp* mRNA. Red dashed line indicates a line of slope 1 with intercept at (0,1). Error bars are standard error of mean *gfp* (x-axis) and *bcd* (y-axis) mRNA content.

SUPPLEMENTAL TABLES

Genotype	$M_{bcd} \pm \sigma_{bcd}$	$\langle M_{bcd} \rangle \pm SEM$	$N_{\mathrm{samp}}^{\mathrm{emb}}$	Embryos	$N_{\mathrm{samp}}^{\mathrm{ref}}$
+/Df *bcd*	$(3.7 \pm 0.7)\times10^5$		92	449	36
+/Df *bcd*	$(4.7 \pm 0.7)\times10^5$	$(4.2 \pm 0.5)\times10^5$	38	156	39
WT	$(7.4 \pm 1.2)\times10^5$		35	129	39
WT	$(7.3 \pm 0.9)\times10^5$	$(7.4 \pm 0.8)\times10^5$	76	654	59
stau$^-$	$(7.2 \pm 2.1)\times10^5$		11	96	18
stau$^-$	$(5.1 \pm 1.3)\times10^5$	$(6.2 \pm 1.2)\times10^5$	13	45	15

Table S1. qRT-PCR measurements of *bcd* mRNA in +/Df *bcd*, wild-type (WT) and *stau*$^-$ embryos, related to Figure 1. Here $N_{\mathrm{samp}}^{\mathrm{emb}}$ and $N_{\mathrm{samp}}^{\mathrm{ref}}$ refer to the total number of samples with embryos and reference RNA molecules respectively. M_{bcd} represents *bcd* mRNA counts measured in each experiment, and σ_{bcd} is the fitting error for the experiment. The mean counts $\langle M_{bcd} \rangle$ for each genotype is obtained by averaging the individual experiments and the standard error on the mean is obtained by error propagation (see text). The relatively low number of *stau*$^-$ samples reflects the difficulty in obtaining large numbers of homozygous mutant embryos.

Embryo #	$M_{\mathrm{bcd}}^{\mathrm{smFISH}}$	$\sigma_D/M_{bcd}^{smFISH}$	$\sigma_f/M_{bcd}^{smFISH}$	$\sigma_V/M_{bcd}^{smFISH}$
1	1.02×10^6	0.6%	0.8%	3.8%
2	9.52×10^5	0.6%	1.0%	5.1%
3	9.02×10^5	0.9%	0.5%	4.1%
4	8.06×10^5	0.9%	2.0%	4.8%
5	9.30×10^5	0.8%	0.8%	4.2%
6	8.71×10^5	0.7%	0.8%	4.5%
7	7.91×10^5	0.6%	1.4%	4.8%

Table S2. Total counts of *bcd* mRNA using smFISH in seven individual *stau*$^-$ embryos, related to Figure 2. Three systematic sources of measurement error are reported: 1) the number of directly detected particles $\sigma_D/M_{bcd}^{smFISH}$, **2)** computing the total fluorescence function $\sigma_f/M_{bcd}^{smFISH}$, and 3) locating the midsagittal plane of the embryo $\sigma_V/M_{bcd}^{smFISH}$ (see text).

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

FISH Image Analysis

We use custom MATLAB software to directly detect mRNA particles as described in [S3]. Briefly, the z-slices in each stack are aligned to correct for microscope drift and each z-slice is then filtered by a circular difference-of-Gaussian (DoG) roughly matched to the size of a single *bcd* mRNA particle to identify candidate particles (local maxima that exceed a low threshold). The candidate particles are arranged in columns and particles with *n=3* shadows (resulting from the axially extended PSF of the microscope) on successive planes are considered as true mRNA particles. Since the DoG threshold is set low to capture all true mRNA, arranging candidate particles in columns yields false positives, which are removed by setting a rigid threshold (Figure S1A).

The distribution of particle intensities in wild-type embryos is consistent with a population containing both packaged and individual mRNA, while in *stau*$^{-}$ embryos the packaged mRNA are resolved (Figure S1B). The anterior region of the *stau*$^{-}$ embryos (i.e. $x/L < 20\%$ EL) is densely populated with mRNA and multiple mRNA can co-localize and be detected as a single particle. To control for this artifact, we complement the direct detection method with a total-fluorescence counting approach described in [S2]. Briefly, this approach relies on sampling the embryo with voxels of volume *V* to obtain the total fluorescence I_{tot} and the number of directly detected particles (true mRNA particles, see above) within the voxel, N_{D}. Over a certain range of mRNA densities the intensity density I_{tot}/V scales linearly with the density of individual particle spots N_{D}/V (Figure S1C). The slope α of this relationship is equal to the intensity per spot while the offset β estimates the fluorescence background per voxel. The resulting number of true mRNA molecules in a voxel estimated using total-fluorescence is given by $N_F = (I_{\mathrm{tot}} - \beta V)/\alpha$.

We employ the total-fluorescence method to construct a global density correction function $f(x)$ along the anterior-posterior (AP) axis of the embryo such that the number of true mRNA molecules is estimated as $N_{\mathrm{true}}(x) = f(x)N_D(x)$. To construct $f(x)$, we first slide a voxel along the AP axis as shown in Figure S1D, and estimate α and β for this sub-volume; $f(x) = N_F(x)/N_D(x)$ is calculated for each voxel. A global density correction function for the anterior is obtained by averaging density correction functions constructed at multiple z-depths (Figure S1E).

In all seven embryos, the density of mRNA decreases monotonically along the AP axis; therefore in the anterior-most regions ($< 15\%EL$) the total fluorescence method yields a density correction factor $f(x) > 1$. In regions of moderate densities ($15\text{–}20\%EL$) the direct particle detection and total fluorescence method are equally reliable, i.e. $f(x) = 1$. Finally, as the density of mRNAs decreases further ($> 25\%EL$), background noise fluctuations override the signal and the total fluorescence method becomes unreliable, i.e. $f(x)$ takes arbitrary values. Therefore, we compute the true number of mRNA as $N_{\mathrm{true}}(x) = f(x)N_D(x)$ in the anterior-most region of each embryo (i.e. $x < 15\%EL$ and $f(x) > 1$) and $N_{\mathrm{true}}(x) = N_D(x)$ in the remaining volume. To estimate the measurement error, we construct global density correction functions by bootstrapping with replacement the individual density correction functions, evaluating the number of fluorescence corrected particles in the anterior volume and computing their standard deviation.

Finally, the total number of mRNA per individual embryo is twice as high ($2\times N_{\mathrm{true}}$), since only half of the embryo is imaged. The number of mRNA for each of the seven embryos is shown in Table S1; the mean number per embryo is $M_{\mathrm{emb}} = 8.9\times10^5$, with standard deviation of 0.8×10^5 counts (or 9% reproducibility) across the seven embryos and standard error on the mean of 0.3×10^5 counts. The error on the reproducibility is $\pm2\%$, calculated by bootstrapping the standard deviation with replacement.

FISH measurement error

We estimate three types of measurement errors for mRNA counts using FISH: (1) direct detection, (2) total fluorescence detection, and (3) identification of the embryo's midsagittal plane. To directly detect particles, we set a rigid threshold to separate noise from true particles (Figure S1A) and discard the left-hand tail of the particle intensity distribution, which leads to underestimation of the direct particle counts N_D. To estimate the magnitude of this effect for each embryo, we fit a Gaussian to the distribution of directly detected particles and compute the counts with intensities below the threshold (Table S2, $\boldsymbol{\sigma_D/M_{bcd}^{smFISH}}$). Next we control for co-localization of mRNA particles in the anterior pole of the embryo by using a total-fluorescence detection method to construct a global density correction function $f(x)$ (Figure S1D). The density correction function is computed as the average of the individual density correction functions f_i at subsequent z-depths of a sub-stack within the embryo. We find the variability in the total fluorescent counts by calculating global density functions by bootstrapping with replacement the individual density correction functions, evaluating the total fluorescent counts N_f for each bootstrapped global density function and calculating their standard deviation (Table S2, $\boldsymbol{\sigma_f/M_{bcd}^{smFISH}}$). Finally, because we image only half of the embryo, we estimate the uncertainty in locating the midsagittal plane of the embryo. All seven embryos are imaged with 420nm z-spacing and pixel resolution of 75.7nm and ~ 65 z-depths per half embryo. If the embryo has length L, the width W, and height H then the volume of the embryo can be assumed to be that of an ellipsoid and calculated using the formula $V = \frac{1}{6}\pi LWH$. We estimate 2 z-depths of error in locating the midsagittal plane, or uncertainty of $\Delta H = 820\ \mathrm{nm}$ and uncertainty in the volume of $\Delta V = \frac{1}{2}\pi LW\Delta H$. (Table S2, $\boldsymbol{\sigma_V/M_{bcd}^{smFISH}}$).

qRT-PCR calibration.

To provide reference to *bcd* mRNA copy numbers in embryos, we use calibration with dilutions series of synthetic *bcd* mRNA. The aliquots in the series are processed identically to the embryo samples, including TRIzol extraction followed by reverse transcription. The *bcd* gene has 5 transcripts (splice variants): bcd-RE (2547bp), bcd-RF (2532bp), bcd-RG (2525bp), bcd-RD (2510bp), and bcd-RA (1490bp) [S4]. We synthesized reference *bcd* mRNA from the Bcd-RD transcript cloned in a pCRII-TOPO plasmid using an mMESSAGE mMACHINE SP6 kit (Invitrogen) following manufacturer's protocol. The concentration in grams per volume of the synthesized reference RNA was determined by absorbance spectroscopy measurements of RNA dilution series and converted to copy numbers per volume using the molecular weight of single-stranded bcd-RD RNA (MW=484167 g/mole). For copy number quantification the embryo samples and the RNA calibration series are amplified via qPCR using SYBR Green reporter on the same plate with a dilution series of Bcd-RD plasmid. The primer pair is designed to amplify all *bcd* transcripts with the exception of the bcd-RA transcript which does not generate a functional Bcd protein and whose existence is weakly supported [S4].

We focus on optimizing each step toward maximal reproducibility and minimizing the effects of other molecular species endogenous to samples extracted from embryos. We remove non-nucleic species by the use of TRIzol and PLG gel (5 Prime), preserve only nucleic species with glycogen (Ambion) and remove contaminating traces of the first extraction step by drying out the nucleic acid carrying glycogen pellet. The nucleic acid species are both DNA and RNA, but DNA contamination is not a concern since the samples with embryos of ages 10–30 min, which have very few nuclei (at most 8 nuclei per embryo) and the primer pair has been designed to span an exon-exon junction. The presence of mRNA species could affect the efficiency of reverse transcription (RT), and to prevent this potential problem we used *bcd*-specific primers, avoiding effects of non-specific reverse transcription of other mRNA species. Therefore, at qPCR both samples from embryo extracts and synthetic RNA calibration contain cDNA only from *bcd* molecules.

RNA calibration provides reference to *bcd* mRNA copy-numbers in embryos under the following assumptions (i) the efficiency of RNA isolation is identical for both embryonic and synthetic RNA

molecules, (ii) both RNA isolation and reverse transcription are linear with combined efficiency η. Consistent with the latter assumption, the relationship between input number of reference RNA molecules and qPCR is linear on a log-linear plot RNA calibration series over 5 orders of magnitudes. By comparing our calibration series with RNA series that has undergone just reverse transcription prior to PCR, we find that only (0.3±0.1) of the RNA are successfully isolated (Figure S2A). We also confirm that there is no difference in the chemical efficiencies between samples with embryos and RNA calibration series due to the presence of other molecular species endogenous to embryonic samples. We synthesized reference *eGFP* mRNA transcript and constructed *eGFP* RNA calibration series, which we compared to *eGFP* RNA calibration series where each sample is spiked with homogenates of embryos during TRIzol extraction. Specifically, we synthesized reference *eGFP* mRNA from *eGFP* transcript cloned in pCRII-TOPO plasmid using mMESSAGE mMACHINE SP6 kit (Invitrogen) following the manufacturer's protocol. To obtain embryo homogenates, embryos of ages 10-60 minutes were collected and homogenized in TRIzol such that the equivalent of 8 embryos was added to each *eGFP* RNA sample. We did not observe differences between the samples with and without embryo extracts (Figure S2B) which confirms that the presence of other molecular species in samples from embryos does not affect the mRNA measurements in our setup. Finally, we confirmed that the efficiency of PCR amplification $\boldsymbol{\varepsilon}$ is the same for cDNA from reference *bcd* mRNA molecules, *bcd* mRNA extracted from embryos and *bcd* DNA plasmid (Fig S2C-E).

After RNA isolation and reverse transcription, a sample with $n{\times}M_{ref}$ mRNA molecules yields $N_{cDNA} = n\eta M_{ref}$ cDNA molecules and a sample with n embryos with M_{bcd} mRNA per embryos yields $N_{cDNA} = n\eta M_{bcd}$. During qPCR, a sample with N_{cDNA} molecules is amplified with efficiency ε. The number of DNA molecules at amplification cycle C is $N(\mathrm{C}) = N_{cDNA} \times \varepsilon^{C}$. By recording the cycle C_{th} at which the sample crosses a fluorescence threshold, we obtain an equation for the RNA calibration series:

$$C_{th}^{ref} = -\frac{1}{\log(\boldsymbol{\varepsilon})}\log\left(n\eta M_{ref}\right) + \frac{1}{\log(\boldsymbol{\varepsilon})}log(N_{th})$$

Here N_{th} is the number of DNA molecules corresponding to the fluorescence threshold. Analogous equation for a sample with n embryos each with M_{bcd} per embryo is given by:

$$C_{th}^{emb} = -\frac{1}{\log(\boldsymbol{\varepsilon})}\log(n\eta M_{bcd}) + \frac{1}{\log(\boldsymbol{\varepsilon})}log(N_{th})$$

The two lines have identical slopes, i.e. $S = -1/log(\boldsymbol{\varepsilon})$, which is measured with high precision by the DNA dilution series. Note that the lines are parallel and separated by an offset Δ regardless of the choice of logarithmic base. The number of mRNA per embryo is found by subtracting the two equations:

$$M_{bcd}^{pcr} = M_{ref}\varepsilon^{-\Delta}$$

The fitting error σ_{bcd}^{pcr} in estimating the mean number of mRNA per embryo M_{bcd}^{pcr} is calculated using standard error propagation methods:

$$\frac{\sigma_{bcd}^{pcr}}{M_{bcd}^{pcr}} = \sqrt{\frac{{\sigma_{ref}}^2}{{\mathrm{M}_{ref}}^2} + {\sigma_{\Delta}}^2\ln(\varepsilon)^2 + \frac{{\sigma_{\boldsymbol{\varepsilon}}}^2}{\boldsymbol{\varepsilon}^2}\Delta^2}$$

The terms under the root are the contributions from the measurement error in the number of reference RNA ($\sigma_{ref}{\sim}5\%$), the fitting error in measuring the offset between the linear fits to the calibration and embryo series (σ_{Δ}), and the PCR efficiency (σ_{ε}). To extract the fitting errors, first the DNA series is fit to a two-parameter linear equation $y = a + Sx$ using least squares to obtain an estimate for S and the fitting error σ_S. Second the RNA calibration and embryo series are both fit to linear equations: $y = a + Sx$, where S is found from the fit to the DNA series and only the intercept a is fit using least squares (one-parameter fit). The offset between the RNA calibration series and the embryo series is $\Delta =$

$a_{calib} - a_{emb}$, and the uncertainty is $\sigma_\Delta = \sqrt{(\sigma^2_{a,calib} + \sigma^2_{a,emb})}$. The PCR efficiency and its uncertainty are calculated from the slope S and from the fitting error to the slope σ_S for each DNA series. Specifically, $\varepsilon = exp(-1/S)$, and we choose to use natural log so the fractional error in the efficiency is $\sigma_\varepsilon/\varepsilon = \left|\frac{1}{S}\left(\frac{\sigma_S}{S}\right)\right| \approx 1\%$. (if base 10 is used instead, $\sigma_\varepsilon/\varepsilon = 2.303\left|\frac{1}{S}\left(\frac{\sigma_S}{S}\right)\right|$). The contribution of ~1% error in the efficiency ε to the overall uncertainty in the number of *bcd* mRNA is proportional to the offset Δ, i.e. for $\Delta \approx 25 - 30$ cycles the uncertainty in efficiency translates into $(\sigma_\varepsilon/\varepsilon)\Delta \approx 25 - 30\%$ uncertainty in the amplified amount.

bcd mRNA counts in +/Df *bcd*, Ore-R (wild-type) and *stau*$^-$ fly lines are summarized in Table S2. The number of *bcd* mRNA counts per genotype is the mean of the measurements from two independent experiments and the standard error on the mean (SEM) reported is calculated using error propagation as $\sqrt{\sum_{i=1}^{2}\frac{\sigma_i^2}{2^2}}$, where σ_i is the fitting error from each experiment whose values are shown in the second column of Table S2.

Protein quantification in embryos of transgenic fly lines.

We measure the total amount of Bcd-GFP molecules in individual embryos for all transgenic fly lines relative to our reference line Bcd2X$_A$, which expresses Bcd at wild-type level [S1]. Bcd-GFP intensity data and dosage measurements for all transgenic fly lines (Figure S3A) were obtained from [S1]. The total number of Bcd-GFP molecules in each of the fly lines N^T_{Bcd} is calibrated to that of the Bcd2X$_A$ line, N_{Bcd}, by multiplying the ratio of their Bcd-GFP concentrations (i.e. their dosage, D, see Ref. [S1], Figure S3B) by the ratio of embryo volumes, i.e. $N^T_{\mathrm{Bcd}} = D \times N_{\mathrm{Bcd}} \times V^T/V_{\mathrm{Bcd2X}}$. The shape of the embryos is assumed to be a prolate spheroid and the volume is calculated as $V = \frac{1}{6}\pi L W^2$, where *L*, and *W* are the lengths of the major and minor axes of the embryo in the midsagittal plane, extracted by image analysis. The embryo volume from each fly line is normalized to that of the Bcd2X$_A$ line (Figure S3C).

The total number of Bcd-GFP molecules in individual embryos of the reference line Bcd2X$_A$ is measured by optically calibrating its GFP fluorescence with a 54 nM eGFP solution. Purified eGFP, a gift of H.S. Rye (Texas A&M University), was overproduced in *E. coli* (BL21) from a *trc* promoter and purified as described in Ref. [S5]. The protein concentration was determined spectroscopically using serial dilutions and a dual-beam spectrophotometer with the absorption wavelength set to 488nm to avoid counts of potential non-fluorescent eGFP molecules. In particular, the protein concentration of eGFP can be spectroscopically determined using either 1) the absorption of eGFP fluorphore at the wavelength of 488 nm with the molar extinction coefficient of 55,000 $M^{-1}cm^{-1}$, or 2) the absorption of aromatic residues in eGFP at the wavelength of 280 nm with the molecular extinction coefficient of 21,890 $M^{-1}cm^{-1}$. The former measures only fluorescent eGFP molecules. The latter measures both fluorescent eGFP and potential non-fluorescent eGFP molecules. We chose to report the absorption at 488 nm. In addition, we also measured the absorption at 280 nm. The concentrations determined using both ways were consistent with each other, i.e., the percentage of non-fluorescent eGFP is nearly zero. Calibration measurement was performed on fixed and hand-peeled embryos (to allow for proper eGFP maturation, see details in [S3]) that were imbedded in the purified eGFP solution. Molecules get into hand-peeled embryos as long as a quick alcohol wash step (ethanol, not methanol, to preserve eGFP fuorescence) is part of the processing procedure. Hence the solutions of both fixed Bcd-GFP inside the embryo and of free-floating eGFP are identical, and therefore the pH should be the same for both.

Assuming an overall cylindrically symmetric Bcd-GFP distribution in the embryo [S6], we rotate the image plane around the major egg axis to reconstruct, in software, the whole embryo. Since the embryos shrink during fixation and are flattened during the mounting procedures, the total intensity

measured at the midsaggital plane is corrected to account for these deformations. The fixed embryos are assumed to have an ellipsoid shape with volume calculated as $V = \frac{1}{6}\pi LWH$ and measured average length $L = (365 \pm 19)\,\mu m$, width $W = (135 \pm 4)\,\mu m$ and height $H = (88 \pm 3)\,\mu m$. Assuming uniform compression, the intensity at the midsaggital plane is corrected by multiplying each pixel with a correction factor $k = H/W$. The embryo is then reconstructed by rotation about its major axis and the mean fluorescence intensity per whole flattened Bcd2X$_A$ embryo is thus found to be $I = (127 \pm 7)$ counts/pixel. After subtracting the auto-fluorescence background $B = (74 \pm 6)$ counts/pixel as measured from identically processed wild-type embryos (i.e. without Bcd-GFP expression), the mean intensity of the Bcd-GFP fluorescence per Bcd2X$_A$ embryo is estimated as $I_{\mathrm{Bcd-GFP}} = I - B = (53 \pm 9)$ counts/pixel. Under the same imaging conditions, the mean intensity of 54 nM eGFP solution is $I_{\mathrm{eGFP}} = (58 \pm 1)$ counts/pixel, thus the average Bcd-GFP concentration of the Bcd2X$_A$ line is $C = I_{\mathrm{Bcd-GFP}}/I_{\mathrm{eGFP}} \times 54 = (50 \pm 10)$ nM, which corresponds to $(29 \pm 6)\ \mathrm{molecules}/\mu m^3$.

The total number of Bcd-GFP molecules in a Bcd2X$_A$ embryo at about 16 min into nuclear cycle 14 is $N_{\mathrm{Bcd}} = C * V = (6.7 \pm 1.5) \times 10^7$. This number is a lower bound on the actual total Bcd-GFP molecule count as we have not included intensity attenuation due to optical path penetration in our samples. We estimate this effect to account for an additional ~21%, bringing the total amount to $(8.2 \pm 1.8) \times 10^7$ molecules. This number is about twofold higher than that obtained with similar optical measurements [S7], and twofold less than measurements based on semi-quantitative western blotting [S8]. Further quantitative control experiments are necessary to clarify discrepancies in systematic errors for the three independent measurements.

To test whether the concentration ratio of different fly lines in the dosage measurement is affected by the maturation effect in live embryos, we also checked the average Bcd-GFP concentration in fixed embryos of a fly line with twice the wild-type Bcd dosage. We find that the concentration ratio is consistent with the dosage ratio, suggesting that that maturation is independent of Bcd-GFP dosage.

bcd mRNA translation rate calculation.

The number of *bcd* mRNA M_{bcd} deposited by the female has been shown to remain constant during the time needed to establish the Bcd protein gradient [S3]. Bcd-GFP protein levels correlate with *bcd* and *gfp* mRNA amounts across transgenic lines with different numbers of transgenes (Figure S4A). In addition, the transgenic Bcd2X$_A$ reference line produces the same amount of mRNA as the endogenous *bcd* locus (Figure S4B), as determined by measuring *bcd* and *gfp* mRNA levels relative to the maternally supplied mRNA *tubulin*56 [S2]. The equivalency of endogenous and transgenic *bcd* production, combined with the observation that the transgenic Bcd2X$_A$ reference line is normally patterned, allows us to calculate the translation rate of endogenous *bcd* mRNA using the number of molecules measured in wild-type embryos and the amount of GFP protein measured in the reference line.

The Bcd protein is degraded uniformly during the first 13 nuclear cycles [S8] and the dynamics of Bcd synthesis is described by

$$\frac{dN_{\mathrm{Bcd}}(t)}{dt} = kM_{bcd} - \frac{N_{\mathrm{Bcd}}(t)}{\tau},$$

Here $N_{\mathrm{Bcd}}(t)$ is the number of Bcd protein molecules at time *t*, *k* is the rate of translation and τ is the protein lifetime. Solving the differential equation, we obtain the Bcd protein number as a function of time:

$$N_{\mathrm{Bcd}}(t) = k\tau M_{bcd}\left(1 - \exp\left(-\frac{t}{\tau}\right)\right).$$

The Bcd protein gradient was imaged at 16 ± 2 min after entry into mitotic division 13, specifically, at $t_{14} \approx 146$ min after fertilization [S1]. The lifetime of the protein during this period is $\tau \approx 50\,min$ [S8].

Therefore, with our measurements of $M_{bcd} = (7.5 \pm 0.8)\times 10^5$ *bcd* mRNA molecules and $N_{\mathrm{Bcd}} = (8.2 \pm 1.8)\times 10^7$ Bcd proteins at t_{14} we obtain a translation rate, in units of proteins per mRNA per second:

$$k = \frac{N_{\mathrm{Bcd}}(t_{14})}{\tau M_{bcd}\left(1 - \exp\left(-\frac{t_{14}}{\tau}\right)\right)} \approx 0.03.$$

The calculation of the *bcd* mRNA translation rate *k* assumes that the mRNA molecules are uniformly translated. However, there is evidence that temporally regulated poly-adenylation of *bcd* mRNA can lead to a 15–30 min delay in the activation of translation and to a time-dependent translation rate with a monotonic increase over the following 1.5h time window [S9]. However, even taking these effects into account, our calculation still yields a translation rate that is well below the maximally available one.

If we assume that the translational apparatus operates at near maximal rates in the early embryonic environment, where molecular patterns are generated with extreme speed (i.e. the first 9 cycles of nuclear divisions last just 8 minutes each [S10], and the Bcd protein gradient stabilizes within less than an hour [S6]), we can estimate an upper limit on the translation rate, $k_{\max}$, from known quantities. In particular, under optimal growth conditions typical translation rates for eukaryotic organisms are of the order of ~10 amino acids per second [S11]. Thus the ~500 amino acid-long Bcd protein should be translated at $k_{\max} = 0.26$ proteins per mRNA per second (with 13 ribosomes simultaneously working on a *bcd* mRNA molecule [S12]), which is an order of magnitude higher than the translation rate we obtained above. Therefore, the translational apparatus clearly performs at submaximal translation rates, and possibly compensates utilizing a source with a large number of mRNA molecules.

Supplemental References

S1. Liu, F., Morrison, A.H., and Gregor, T. (2013). Dynamic interpretation of maternal inputs by the Drosophila segmentation gene network. Proceedings of the National Academy of Sciences of the United States of America *110*, 6724-6729.

S2. Little, S.C., Tikhonov, M., and Gregor, T. (2013). Precise developmental gene expression arises from globally stochastic transcriptional activity. Cell *154*, 789-800.

S3. Little, S.C., Tkacik, G., Kneeland, T.B., Wieschaus, E.F., and Gregor, T. (2011). The formation of the Bicoid morphogen gradient requires protein movement from anteriorly localized mRNA. PLoS biology *9*, e1000596.

S4. Marygold, S.J., Leyland, P.C., Seal, R.L., Goodman, J.L., Thurmond, J., Strelets, V.B., and Wilson, R.J. (2013). FlyBase: improvements to the bibliography. Nucleic acids research *41*, D751-757.

S5. Rye, H.S., Burston, S.G., Fenton, W.A., Beechem, J.M., Xu, Z., Sigler, P.B., and Horwich, A.L. (1997). Distinct actions of cis and trans ATP within the double ring of the chaperonin GroEL. Nature *388*, 792-798.

S6. Gregor, T., Wieschaus, E.F., McGregor, A.P., Bialek, W., and Tank, D.W. (2007). Stability and nuclear dynamics of the bicoid morphogen gradient. Cell *130*, 141-152.

S7. Drocco, J.A., Wieschaus, E.F., and Tank, D.W. (2012). The synthesis-diffusion-degradation model explains Bicoid gradient formation in unfertilized eggs. Physical biology *9*, 055004.

S8. Drocco, J.A., Grimm, O., Tank, D.W., and Wieschaus, E. (2011). Measurement and perturbation of morphogen lifetime: effects on gradient shape. Biophysical journal *101*, 1807-1815.

S9. Salles, F.J., Lieberfarb, M.E., Wreden, C., Gergen, J.P., and Strickland, S. (1994). Coordinate initiation of Drosophila development by regulated polyadenylation of maternal messenger RNAs. Science (New York, N.Y.) *266*, 1996-1999.

S10. Foe, V.E., and Alberts, B.M. (1983). Studies of nuclear and cytoplasmic behaviour during the five mitotic cycles that precede gastrulation in Drosophila embryogenesis. Journal of cell science *61*, 31-70.

S11. Bonven, B., and Gullov, K. (1979). Peptide chain elongation rate and ribosomal activity in Saccharomyces cerevisiae as a function of the growth rate. Molecular & general genetics : MGG *170*, 225-230.

S12. Qin, X., Ahn, S., Speed, T.P., and Rubin, G.M. (2007). Global analyses of mRNA translational control during early Drosophila embryogenesis. Genome biology *8*, R63.