Scaling dictates the decoder structure

Jingxiang Shen[1], Feng Liu[1,2], Chao Tang[1,2,3]*

[1]Center for Quantitative Biology, Peking University, Beijing 100871, China.

[2]School of Physics, Peking University, Beijing 100871, China.

[3]Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China.

*Correspondence to: tangc@pku.edu.cn

ABSTRACT

Despite variability in embryo size, the tissue, organ and body plan develop *in proportion* with embryo size, known as the scaling phenomenon[1-7]. Scale-invariant patterning of gene expression is a common feature in development and regeneration, and can be generated by mechanisms such as scaling morphogen gradients[8,9] and dynamic oscillation[10,11]. However, whether and how static non-scaling morphogens (input) can induce a scaling gene expression (output) across the entire embryo is not clear[2,12,13]. Here we show that scaling requirement sets severe constraints on the geometric structure of the input-output relation (the decoder), from which information about the regulation and mutants' behavior can be deduced without going into any molecular details. We applied and tested our theory on *Drosophila* gap gene system[14-17] using the three maternal gradients (Bcd, Nos and Tor) as input. Remarkably, the parameter-free decoder correctly and quantitatively accounted for the gap gene expression patterns in nearly all morphogen mutants. It also revealed the regulation logic and the coding/decoding strategy of the gap gene system. Our work provides a general theoretical framework on a large class of problems where scaling output is induced by non-scaling input, as well as a unified understanding of scaling, mutants' behavior and regulation in the *Drosophila* gap gene and related systems.

In development and regeneration, a cell has to know its position in space in order to determine its fate. Often this positional information is encoded in a space-dependent signal like morphogen gradient[2,14,18]. The decoder in the cell (usually gene networks) reads the local value of the signal to infer its spatial position by expressing the appropriate genes, as exemplified in the French-flag model[2,19,20]. A common feature in these processes is that the gene expression pattern generated by the decoder is often scale-invariant with respect to fluctuations in the overall length of the system, which ensures the organism develops in the correct proportion[6,7,21-24]. A scale invariant output can be achieved with a scale invariant input. Indeed, this strategy is adopted by systems such as *Xenopus* embryo[8,25], *Drosophila* wing[9,26,27] and probably *Drosophila* dorsoventral axis[28], where the morphogen itself scales with the system size. However, achieving scale-invariant patterning from non-scaling morphogens is still largely an open question. A prototypical example is the scaling of *Drosophila* segmentation gene patterning[12,23,29-35], in which the gradients of the maternal morphogens seem to have fixed length constants that do not scale with the embryo size[29,33,35-38]. In this paper, we first establish a general theoretical framework. We show that the necessary and sufficient condition for a decoder to generate a scale-invariant expression pattern from non-scaling morphogens contains rich information on the geometric structure of the decoder. We then apply and test our theory on the *Drosophila* gap gene system, providing a comprehensive understanding on this classic scaling problem.

Geometric structure of the decoder and the fate map

We first illustrate our basic idea using a bi-gradient model[2,12,13]. In this simplified model, there are two morphogens, M_1 and M_2, having anterior and posterior exponential gradients with a fixed length constant λ (Fig. 1a, length of a standard-sized embryo is defined as L=1). For larger (smaller) embryos, the effective length constant should shrink (expand) if measured in the relative coordinate ($y \equiv x/L$, $M_1 = e^{-yL/\lambda}$, $M_2 = e^{-(1-y)L/\lambda}$, Fig. 1b). Consider for example the position y=0.35 as the boundary separating two different cell fates in the embryo. In a smaller (larger) embryo, as the absolute distance to both termini are shorter (longer) at this position, local levels of both morphogens are higher (lower). When L varies, the y=0.35 point traces out a line on the M_1-M_2 plane (Fig. 1c). Therefore, in order to achieve perfect scaling, this y-constant line must be followed by the decision boundary of the decoder, which maps a (M_1, M_2) pair to a gene expression state ("fate"). In other words, the requirement imposed by scaling is enough to determine the effective input-output relation (coloring scheme) of the decoder on the M_1-M_2 plane, no matter how the decoder is implemented

biochemically. The cell fate can also be represented by its equivalent position in wild-type (WT), denoted as $\tilde{y}$. In this case:

$$\tilde{y}(M_1, M_2) = \frac{\ln M_1}{\ln M_1 M_2}\,. \tag{1}$$

Once determined by its scale-invariant performance in WT embryos, the same decoder can be applied to mutant embryos where the maternal morphogen profiles are perturbed but the decoding machines are intact. In the case considered here, all *y* points belonging to a WT embryo of size *L* correspond to the hyperbolic curve $M_1 \cdot M_2 = e^{-L/\lambda}$ (Fig. 1c). On the other hand, consider a mutant embryo of which *L*=1 but the M_1 copy number is doubled:

$$M_1{}' = 2e^{-y_{mut}/\lambda},\ M_2{}' = e^{-(1-y_{mut})/\lambda}. \tag{2}$$

This mutant is also represented by a hyperbolic curve $M_1{}' \cdot M_2{}' = 2e^{-1/\lambda}$ (Fig. 1d). This curve is exactly the one corresponding to a WT embryo of size $L = 1 - \lambda \ln 2$, though points of the same cell fate locate at different *y*'s in WT and the mutant. For example, the blue-green boundary at *y*=0.35 in WT is shifted to 0.46 in the mutant. In general, for any y_{mut} in this mutant embryo, there always exists a corresponding position y_{WT} in WT embryo with the same morphogen values, hence the same cell fate $\tilde{y}$:

$$y_{WT}(y_{mut}) \equiv \tilde{y}(y_{mut}) = \frac{y_{mut} - \lambda \ln 2}{1 - \lambda \ln 2}\,. \tag{3}$$

We call this mapping $\tilde{y}(y_{mut})$ the "fate-map" of this mutant, which can also be obtained by directly substituting Eq. 2 into Eq. 1. This fate map can be tested experimentally by plotting the positions y_{WT} against y_{mut} for the expression domain boundaries (Fig. 1e).

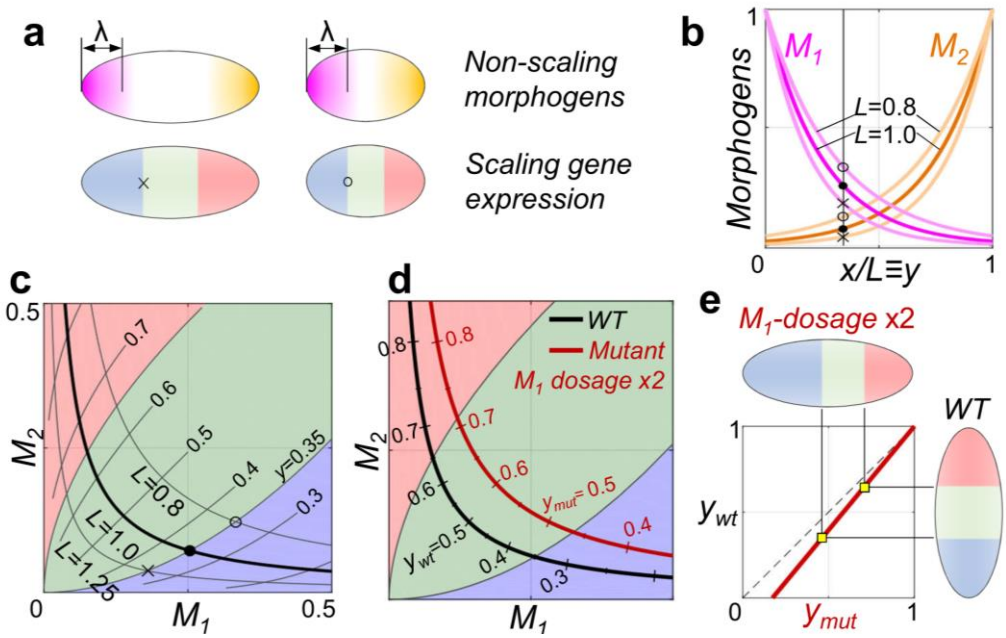

Fig. 1. Generating scaling output by reading the local values of two non-scaling gradients. (**a**) A schematic sketch with two embryos of different sizes. The two morphogen gradients are shown in the upper panel and the desired scaling output patterns are shown below. (**b**) The same relative position $y \equiv x/L$=0.35 has higher level of both morphogens in a smaller embryo (marked by "o") than a larger one ("x"). (**c**) In the M_1-M_2 space, an embryo of size L is represented by the curve $M_1 M_2 = e^{-L/\lambda}$. As L varies, the y=0.35 point traces out a line. (**d**) Maternal morphogen profiles in a mutant are also represented by a curve in the M_1-M_2 space. The red curve stands for an L=1 embryo where the M_1 dosage is doubled. The corresponding cell fates along this red curve, hence along the A-P axis of this mutant, can be directly read out. (**e**) The predicted "fate map" of the mutant in (c). All gene expression boundaries, if their positions in WT were plotted against their shifted positions in the mutant, should lie on the fate map. □

Drosophila gap gene system

Following the arguments presented in the last section, we next construct a decoder for the *Drosophila* gap gene system, using the three major morphogen gradients Bcd, Nos/mHb and Tor. Based on experimental data[33,35,38-43], the profiles of the morphogens are given by Eq. 4 below (see SI-1 and SI-4 for details) (Fig. 2a, Extended Data Fig. 1).

$$\begin{cases} Bcd(y) = L^{\beta} e^{-yL/\lambda_B} \\ mHb(y) = mHb_0 \left(1 + e^{\alpha L(y-1+(1-\lambda_H)/L)}\right)^{-1} \\ Tor(y) = e^{-yL/\lambda_T} + e^{-(1-y)L/\lambda_B} \end{cases} \tag{4}$$

The gap genes (*hb*, *Kr*, *kni* and *gt*) are the first set of zygotic segmentation genes downstream of the maternal morphogens, and the first to display a scaling pattern[29,34,35]. The *Drosophila* embryo has around 100 rows of cells along the Anterior-Posterior (A-P) axis before gastrulation. Nearly any two of them can be distinguished by their gap gene expression[44,45], so there are effectively around 100 different cell fates along the A-P axis and the fate map $\tilde{y}$ should be discussed at the resolution of 0.01. To visualize these cell fates by colors, we group them into 8 classes according to the dominantly expressing gap gene (Fig. 2b). A standard-sized WT embryo is then represented by a 1-d curve in the space of (Bcd, mHb, Tor). When L varies, this curve sweeps out a 2-d WT manifold. Following our arguments in Fig. 1, the ideal output of a scaling decoder on this WT manifold can be immediately determined – lines of constant y values should have the same cell fate (Fig. 2c).

In reality, *Drosophila melanogaster*[23] embryo length varies approximately between 450 and 570 μm (~±10%) across different fly stocks. To have a more realistic picture, we hereafter sample L from a normal distribution $L \sim N(1, 0.1)$ to simulate the effects of length fluctuation across *Drosophila* stocks. Another important aspect is noise in morphogen profiles. When morphogen level approaches zero it should have huge relative noise thus cannot carry useful information. We introduced explicitly a Poisson noise term to the morphogen values to reflect this effect (Methods). As a result, the "colored lines" in Fig. 2c now transform into the WT-point-cloud in Fig. 2d.

Interestingly, for the *Drosophila* case the boundaries between regions of different colors are effectively linear – any decoder that achieves scaling should effectively behave like a set of *linear* classifiers. Therefore, we fit the phenomenological decision boundaries with planes (Methods). Points of different colors seem to be separated satisfactorily in this way (Fig. 2d, e, Extended Data Fig. 2d).

The three maternal gradients can not only determine boundaries of gap gene expression domains, but also ~100 distinct A-P cell fates as mentioned above. Thus, we can construct a (approximately) continuous version of the decoding function using more linear classifiers of this kind. Along the A-P axis, we fit 100 such classification planes at 100 equally spaced y positions (Methods). The linear classifiers work sufficiently well with the decoding task. Fig. 2f shows the decoding result of the WT ensemble of Fig. 2d using the linear classifiers. Despite the presence of Poisson noise, positions ($\tilde{y}$) decoded by reading the maternal morphogen values are always close to the ground truth positions (y).

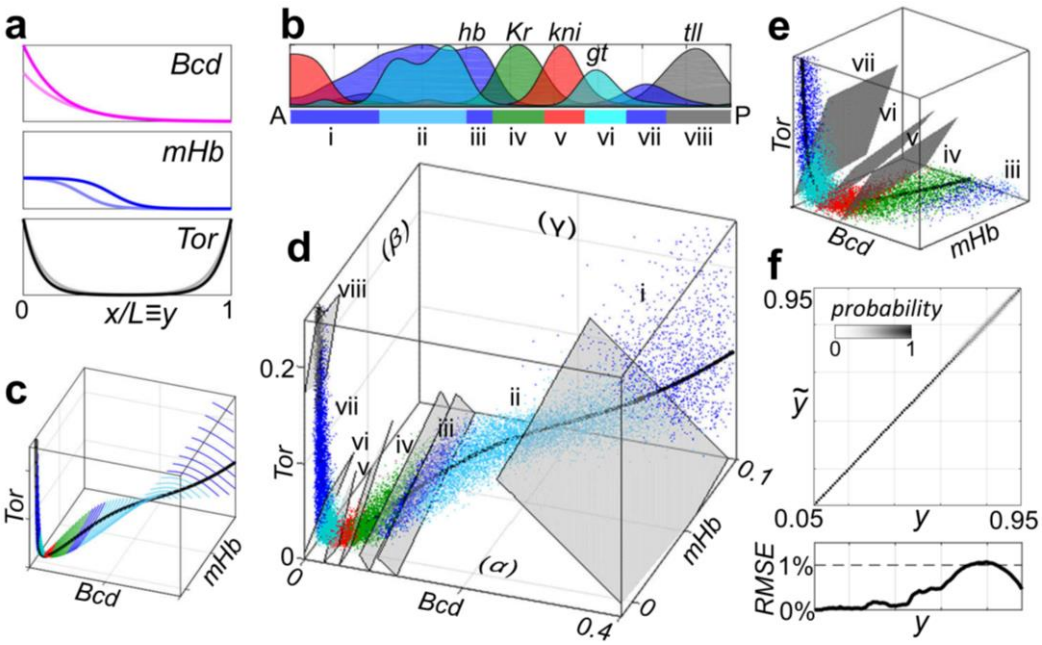

Fig. 2. The structure of a phenomenological decoder of *Drosophila*. (**a**) The three non-scaling maternal gradients in the coordinate normalized by embryo length $y \equiv x/L$ (Eq. 4). The morphogen profiles for a standard size WT embryo (L=1, darker lines) and a smaller one (L=0.8, lighter lines) are shown. (**b**) About 100 different cell fates along the A-P axis are grouped into 8 domains according to the gap gene expression. The normalized gap gene profiles are adopted from Ref.[49] and a Gaussian smoothing is applied. (**c**) In the (Bcd, mHb, Tor) space, the standard-sized WT embryo is represented by the black curve along which y varies from 0 to 1. When L changes, each point on this curve traces out a line representing the morphogen values at this y position in WT embryos of different sizes. Only the region corresponding to L=0.8~1.2 are plotted here. The y lines shown there have spacing 0.01 along the A-P axis. (**d**) A Poisson noise is added to the morphogen levels and the embryo length is sampled from a normal distribution, turning the 2-d WT manifold in (c) to a point cloud. The decision boundaries can be well approximated by a set of linear planes. (**e**) A magnification of (d). The iv-v, v-vi and vi-vii boundary planes are shown. (**f**) Positions decoded from the morphogen values ($\tilde{y}$) *vs*. the ground truth (y) for the WT point cloud in (d). Since gap gene expression is affected by the dorsoventral system when being very close to embryo termini, we only discuss y between 0.05 and 0.95 hereafter. The root-mean-square-error is hardly larger than 1%. □

The decoder quantitatively predicts phenotypes of morphogen mutants

We next demonstrate that such a decoder based on scaling has a remarkable power to predict nearly all phenotypes of morphogen mutants in *Drosophila*. Note that the values of (Bcd, mHb, Tor) in mutants may lie outside of the WT region (color point cloud in Fig. 2d). In this case, we linearly extrapolate the classification planes (SI-2).

Fig. 3a-c show the intersections of the linearly extrapolated classification planes with the bottom/left/back faces of the cube in Fig. 2d (marked by α/β/γ, on which Tor=0/Bcd=0/mHb=mHb_0), with the same color scheme as in Fig. 2. Maternal morphogen null mutants lie on these faces. For example, in the *nos*$^-$ mutant mHb is equal to mHb_0 throughout the entire embryo, corresponding to projecting the WT curve in the (Bcd, mHb, Tor) space onto the γ plane. It is clear in Fig. 3c that along this projected *nos*$^-$ curve, domain iv is followed immediately by domain vii, indicating the loss of abdominal *kni* (v) and *gt* (vi) domains. This is exactly the case observed in experiments[45-47]. As another example, consider the *bcd*$^-$*tor*$^-$. mHb is now the only morphogen gradient, decreasing from its maximum value to zero from the anterior to posterior pole. It is obvious in Fig. 3a-b that points on the mHb axis fall into the iv, v and vi domains successively, corresponding to three gap gene domains *K*r, *kni* and *gt*

appearing successively in this mutant embryo. As the Bcd=Tor=mHb=0 point is classified into domain vi, the *gt* domain should extend all the way to the posterior pole. This is also the pattern observed experimentally[45,48].

More than the predictions on presence or absence of certain gap gene domains, a continuous valued fate map $\tilde{y}(y_{mut})$ can be constructed by incorporating all the 100 classification planes. In Fig. 3d, the predicted fate map for WT and 11 maternal morphogen mutants are shown as black curves. To test our predictions against experiments, we identify the peak and boundary positions for all gap gene domains (Table S1) from the published quantitative measurements of gap gene profiles in WT and various mutants[40,45,49-51]. For each of them, its position in mutant is plotted against its WT positions (yellow squares in Fig. 3d). Remarkably, predictions and experiments are consistent in all cases.

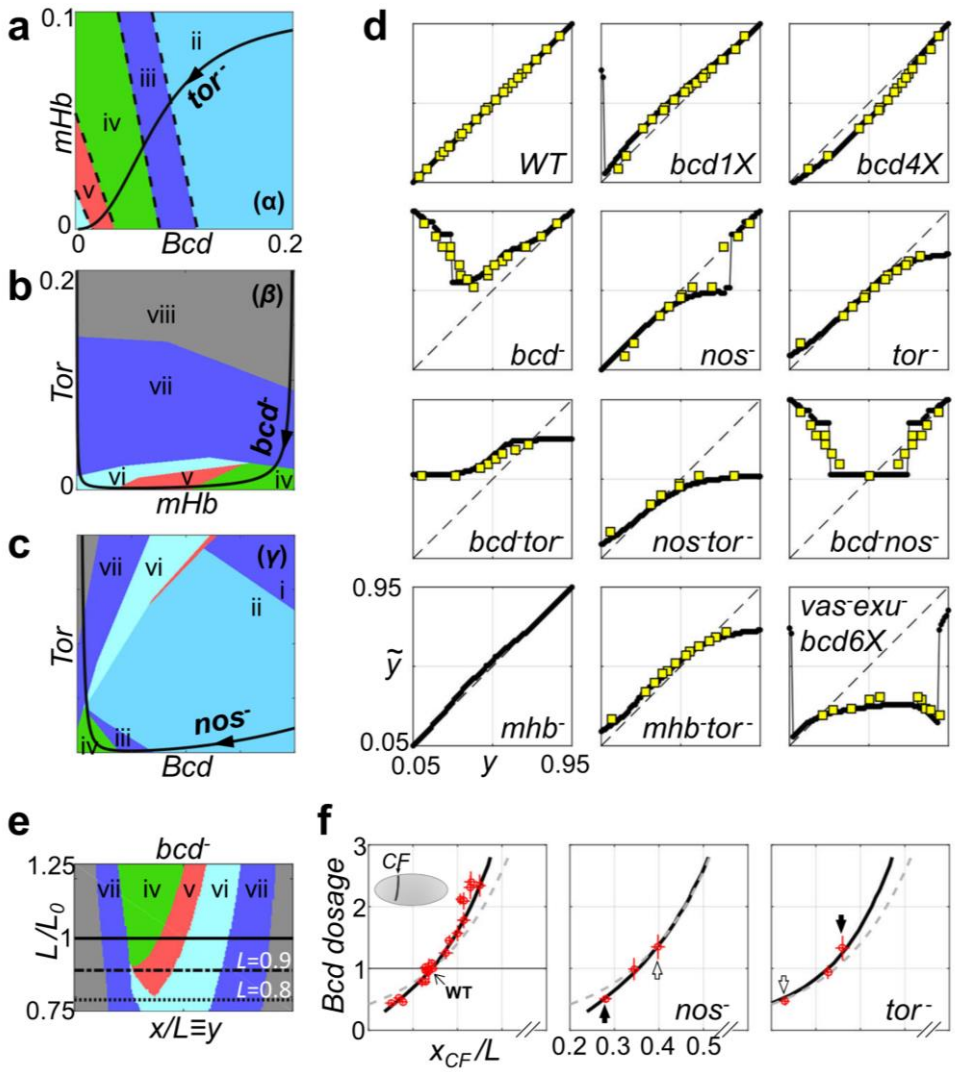

Fig. 3. Quantitative prediction on mutant fate map by the decoder. (**a-c**) The decoding results on the bottom/left/back faces (marked by α/β/γ) of the cube in Fig. 2d. Solid

black curves are projections of the L=1 WT curve onto these planes, which also represents the standard-sized mutant embryos *tor*$^-$/*bcd*$^-$/*nos*$^-$, respectively. Arrowheads on them are pointing from head to tail. The predicted gap gene expression in these mutants can be read out along these lines. (**d**) Fate map $\tilde{y}(y)$ predicted for WT and another 11 different mutants (black lines) and the comparison with experimental measurements (yellow squares, cited from published works, see main text for references). See SI-6 for discussion on the panels of *mhb*$^-$ and *vas*$^-$*exu*$^-$*bcd6X*. (**e**) The predicted domain boundaries in *bcd*$^-$ as L changes. As L shrinks from 1.0 to 0.9 to 0.8 (solid, dashed and dotted lines), the *Kr* (iv) and *kni* (v) domains disappear successively. (**f**) Shift of cephalic furrow (CF) under Bcd dosage change. Solid line is our prediction and red dots with error bars are from experimental data[40]. Dashed grey line shows the position of the same Bcd concentration as CF in WT. □

Our model can also account for some recent and more quantitative experiments, in which multiple factors such as Bcd dosage, morphogen mutation and change of the embryo length were at play. For example, the model predicts that missing Bcd destroys scaling completely and the gap gene pattern changes greatly with L. With decreasing embryo length, the *bcd*$^-$ embryo losses domains iv (*Kr*) and then v (*kni*) (Fig. 3e). A recent experiment[35] provides an excellent support for this prediction (Extended Data Fig. 5c). Another example is the shift of cephalic furrow (CF) ($\tilde{y} = 0.344$) under mutation of the posterior or terminal morphogens plus Bcd dosage change[40]. Our model quantitatively captured the shifts observed in experiment (Fig. 3f).

Decoder structure reveals underlying gene regulation

Predicted fate map can also be converted to a predicted gap gene pattern. For y_{mut} in a mutant embryo, the gap gene expression levels here can be approximated by the composite function $G(\tilde{y}(y_{mut}))$, where G stands for the WT gap gene pattern, and $\tilde{y}(y_{mut})$ is the fate map. Figs. 4a and c show two examples (solid lines). Measured profiles[45] are shown as dotted lines in lighter colors for comparison.

Further analysis reveals the connection between the decoder structure and the underlying gene regulation. Consider the *tor*$^-$ mutant as an example. As dictated by scaling, the decision boundaries *Kr-kni*, *kni-gt* and *gt-hb* should be inclined in 3-d, not perpendicular to the Bcd-mHb plane (Fig. 2e). This means that Tor should participate in positioning these boundaries to allow for scaling. Geometrically, when extrapolated to the Tor=0 plane following these inclined classification planes, a cell fate in WT should appear at a more posterior position than if being orthogonally projected

downward. Thereby in the *tor*$^-$ mutant, besides that posterior *hb* domain disappears, the rest of the abdominal domains should also shift posteriorly. This prediction is fully consistent with experiments (Fig. 4a). This example may help us to understand the ubiquitous "redundant" regulations in the gap gene network. According to the simplest interpretation, mHb gradient defines the anterior boundary of abdominal *gt* (vi) domain through inhibition; Gt inhibits *kni*, thereby setting the posterior boundary of *kni* (v) domain[48]. There is no "need", in principle, for Tor to be involved. But the observed shift of *kni* boundary in *tor*$^-$ clearly shows that in reality Tor contributes (probably through *tll*) to the repression on *kni*[47,52], setting its posterior boundary together with mHb. We propose that such seemingly redundant regulations actually tune the slope of *kni-gt* classification plane, so that it could align with the angle required by scaling (Fig. 4b).

Another example deals with the *bcd*$^-$ mutant (Fig. 4c). Bcd is well known to function in the anterior part. However, the region where Bcd plays a role seems to be much wider than naively expected – *bcd*$^-$ mutant affects even domain vi and vii near the posterior pole. This aspect of Bcd should also contribute to tuning the decision boundary orientations to allow for scaling, as it has clearly been captured by our scaling-based phenomenological model. To be more precise, the decision plane representing the *gt-hb* (vi-vii) boundary at $\tilde{y}$=0.75 is:

$$-0.85Bcd - 0.12mHb + 0.45Tor = Const.,$$

indicating that Tor should be effectively an activator for *hb* here while Bcd should play a repressive role (Fig. 4d). This speculation is consistent with existing biological knowledges. Tor is long known to activate the posterior *hb* domain through *tailless*(*tll*)[52]. And the effective repression by Bcd is most probably meditated by *Kr* and *kni*, upon whose mutation the posterior *hb* domain expands anteriorly[53]. Similar line of arguments can be applied to other cases of regulations (SI-10).

The decoder can be implemented with gene regulation

Scaling requirement can implicate regulation. To see if the converse is true, i.e., if gene regulation can produce scale-invariant decision boundary, consider a toy model with a static anterior morphogen *B* (like Bcd) of fixed length constant and a single gap gene *H* whose initial condition H_{init} serves as a posterior gradient (like mHb) (Fig. 4e). The desired output is to have *H* being activated in the anterior half with the boundary locating at *y*=0.5 regardless of the embryo length *L*. This requires that the morphogen levels marked by the black dots and empty triangles (Fig. 4e) should both lie on the *H* boundary. This can be achieved with the regulation network shown in Fig. 4f (inset).

From a dynamical perspective, this network is well known to have bi-stability and Fig. 4f shows a typical bifurcation diagram. Within the bistable range, the *H-high/H-low* states are separated by a critical line (red squares), which means exactly the decision boundary between cell fates. If the activation strengths are properly tuned, this critical line can be adjusted to have identical slope as the desired decision boundary in Fig. 4g.

The gap gene network has more degrees of freedom and more scaling boundaries, and does not necessarily have to reach a dynamical attracting point. But the general idea is the same. Based on the known gap gene network shown in Fig. 4h[17], we construct a differential equation model (see SI-11 for details). If fitted to the WT data[49] plus scaling requirement, the model can produce very similar decision boundaries as the phenomenological decoder (Fig. 4, i, j).

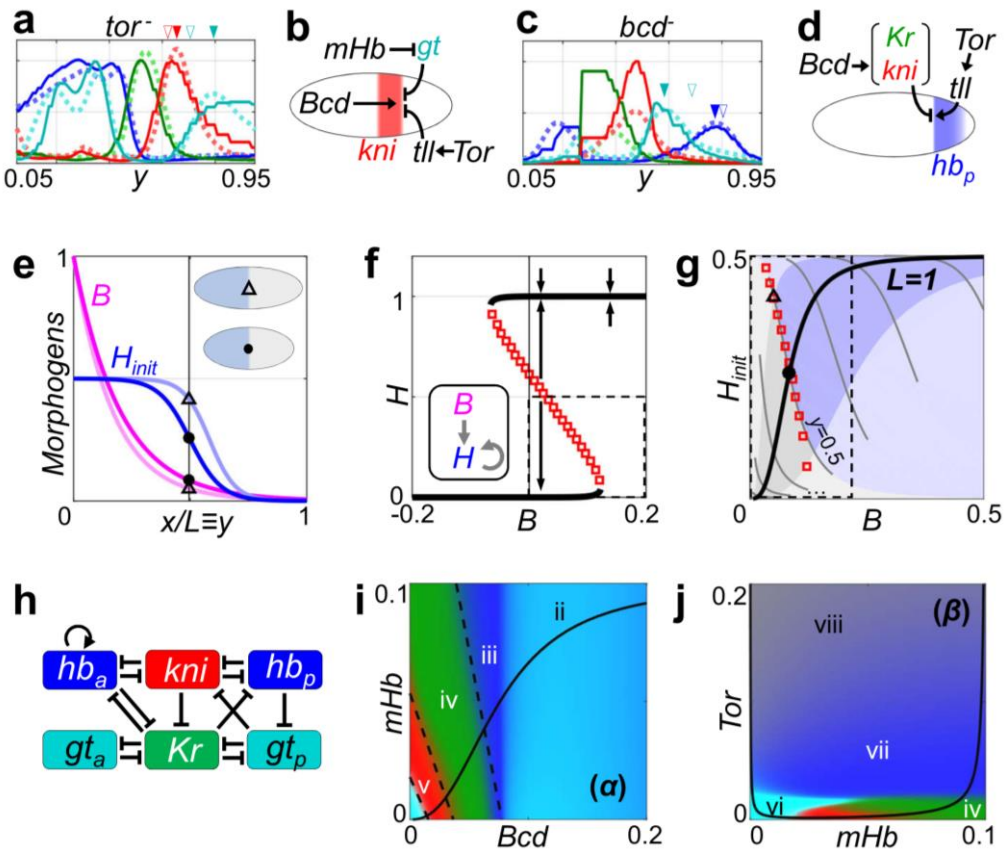

Fig. 4. Connecting the decoder with gene regulation. (**a**, **c**) The predicted gap gene pattern (solid lines) in *tor*$^-$ (a) and *bcd*$^-$ (c) are compared to measurements[45] (dashed lines). The peak positions of the posterior *kni*, *hb* and *gt* domains are marked by filled arrowheads, and their WT positions by empty arrowheads. (**b**) Redundant regulation on the posterior boundary of *kni* domain as required by scaling. (**d**) The vi-vii boundary ($\tilde{y}$=0.75) is set by the balance between the opposing gradients Bcd and Tor. (**e**) The profiles of two morphogens in a toy model, shown for two embryos of standard size (darker lines) and larger size (lighter lines). (**f**) The bifurcation diagram for the network

in the inset. A cell can reach one of the two bi-stable states depending on the values of the two morphogens, separated by the unstable manifold (red squares). This unstable manifold is the decision boundary separating the Hb-high and Hb-low fates. Its slope can be tuned by the activation strengths. (**g**) Within the realistic L range (highlighted region), The y=0.5 line is well approximated by the decision boundary in (f) (red squares). (**h**) The known gap gene network[17]. (**i**, **j**) The same plots as Fig. 3a and b, but using an ODE model of (h). □

Concluding remarks

We developed a theoretical framework on scale-invariant gene patterning from non-scaling signals and illustrated its application and utility in the *Drosophila* gap gene system, providing a unified understanding to scaling, regulation and the many observed phenomena in the *Drosophila* system. Since our framework is rather general and does not depend on specific molecular details, it can be readily extended to other systems (see an example in SI-13 on the long-germband insect *Megaselia abdita*).

References

1 Morgan, T. H. Two Embryos from One Egg. *The Scientific Monthly* **18**, 529-546 (1924).

2 Wolpert, L. Positional information and the spatial pattern of cellular differentiation. *J Theor Biol* **25**, 1-47, doi:10.1016/s0022-5193(69)80016-0 (1969).

3 Cooke, J. Scale of body pattern adjusts to available cell number in amphibian embryos. *Nature* **290**, 775-778, doi:10.1038/290775a0 (1981).

4 Umulis, D. M. & Othmer, H. G. Mechanisms of scaling in pattern formation. *Development* **140**, 4830-4843, doi:10.1242/dev.100511 (2013).

5 Ishimatsu, K. *et al.* Size-reduced embryos reveal a gradient scaling-based mechanism for zebrafish somite formation. *Development* **145**, doi:10.1242/dev.161257 (2018).

6 Almuedo-Castillo, M. *et al.* Scale-invariant patterning by size-dependent inhibition of Nodal signalling. *Nat Cell Biol* **20**, 1032-1042, doi:10.1038/s41556-018-0155-7 (2018).

7 Capek, D. & Muller, P. Positional information and tissue scaling during development and regeneration. *Development* **146**, doi:10.1242/dev.177709 (2019).

8 Ben-Zvi, D., Shilo, B. Z., Fainsod, A. & Barkai, N. Scaling of the BMP activation gradient in Xenopus embryos. *Nature* **453**, 1205-1211, doi:10.1038/nature07059 (2008).

9 Ben-Zvi, D., Pyrowolakis, G., Barkai, N. & Shilo, B. Z. Expansion-repression mechanism for scaling the Dpp activation gradient in Drosophila wing imaginal discs. *Curr Biol* **21**, 1391-1396, doi:10.1016/j.cub.2011.07.015 (2011).

10 Cooke, J. Control of somite number during morphogenesis of a vertebrate, Xenopus laevis. *Nature* **254**, 196-199, doi:10.1038/254196a0 (1975).

11 Lauschke, V. M., Tsiairis, C. D., Francois, P. & Aulehla, A. Scaling of embryonic patterning based on phase-gradient encoding. *Nature* **493**, 101-105, doi:10.1038/nature11804 (2013).

12 Houchmandzadeh, B., Wieschaus, E. & Leibler, S. Precise domain specification in the developing Drosophila embryo. *Phys Rev E Stat Nonlin Soft Matter Phys* **72**, 061920, doi:10.1103/PhysRevE.72.061920 (2005).

13 McHale, P., Rappel, W. J. & Levine, H. Embryonic pattern scaling achieved by oppositely directed morphogen gradients. *Phys Biol* **3**, 107-120, doi:10.1088/1478-3975/3/2/003 (2006).

14 St Johnston, D. & Nusslein-Volhard, C. The origin of pattern and polarity in the Drosophila embryo. *Cell* **68**, 201-219, doi:10.1016/0092-8674(92)90466-p (1992).

15 Monk, N. Development: dissecting the dynamics of segment determination. *Curr Biol* **14**, R705-707, doi:10.1016/j.cub.2004.08.043 (2004).

16 Jaeger, J. Modelling the Drosophila embryo. *Mol Biosyst* **5**, 1549-1568, doi:10.1039/b904722k (2009).

17 Jaeger, J. The gap gene network. *Cell Mol Life Sci* **68**, 243-274, doi:10.1007/s00018-010-0536-y (2011).

18 Gurdon, J. B. & Bourillot, P. Y. Morphogen gradient interpretation. *Nature* **413**, 797-803, doi:10.1038/35101500 (2001).

19 Crauk, O. & Dostatni, N. Bicoid determines sharp and precise target gene expression in

the Drosophila embryo. *Curr Biol* **15**, 1888-1898, doi:10.1016/j.cub.2005.09.046 (2005).

20 Bier, E. & De Robertis, E. M. BMP gradients: A paradigm for morphogen-mediated developmental patterning. *Science* **348**, aaa5838, doi:10.1126/science.aaa5838 (2015).

21 Briscoe, J. & Small, S. Morphogen rules: design principles of gradient-mediated embryo patterning. *Development* **142**, 3996-4009, doi:10.1242/dev.129452 (2015).

22 Shilo, B. Z. & Barkai, N. Buffering Global Variability of Morphogen Gradients. *Dev Cell* **40**, 429-438, doi:10.1016/j.devcel.2016.12.012 (2017).

23 Lott, S. E., Kreitman, M., Palsson, A., Alekseeva, E. & Ludwig, M. Z. Canalization of segmentation and its evolution in Drosophila. *Proc Natl Acad Sci U S A* **104**, 10926-10931, doi:10.1073/pnas.0701359104 (2007).

24 Bollenbach, T. *et al.* Precision of the Dpp gradient. *Development* **135**, 1137-1146, doi:10.1242/dev.012062 (2008).

25 Ben-Zvi, D. & Barkai, N. Scaling of morphogen gradients by an expansion-repression integral feedback control. *Proc Natl Acad Sci U S A* **107**, 6924-6929, doi:10.1073/pnas.0912734107 (2010).

26 Restrepo, S. & Basler, K. Morphogen gradients: expand and repress. *Curr Biol* **21**, R815-817, doi:10.1016/j.cub.2011.08.041 (2011).

27 Hamaratoglu, F., de Lachapelle, A. M., Pyrowolakis, G., Bergmann, S. & Affolter, M. Dpp signaling activity requires Pentagone to scale with tissue size in the growing Drosophila wing imaginal disc. *PLoS Biol* **9**, e1001182, doi:10.1371/journal.pbio.1001182 (2011).

28 Rahimi, N. *et al.* A WntD-Dependent Integral Feedback Loop Attenuates Variability in Drosophila Toll Signaling. *Dev Cell* **36**, 401-414, doi:10.1016/j.devcel.2016.01.023 (2016).

29 Houchmandzadeh, B., Wieschaus, E. & Leibler, S. Establishment of developmental precision and proportions in the early Drosophila embryo. *Nature* **415**, 798-802, doi:10.1038/415798a (2002).

30 Manu *et al.* Canalization of gene expression in the Drosophila blastoderm by gap gene cross regulation. *PLoS Biol* **7**, e1000049, doi:10.1371/journal.pbio.1000049 (2009).

31 Vakulenko, S., Manu, Reinitz, J. & Radulescu, O. Size Regulation in the Segmentation of Drosophila: Interacting Interfaces between Localized Domains of Gene Expression Ensure Robust Spatial Patterning. *Phys Rev Lett* **103**, doi:10.1103/PhysRevLett.103.168102 (2009).

32 de Lachapelle, A. M. & Bergmann, S. Precision and scaling in morphogen gradient read-out. *Mol Syst Biol* **6**, 351, doi:10.1038/msb.2010.7 (2010).

33 Cheung, D., Miles, C., Kreitman, M. & Ma, J. Scaling of the Bicoid morphogen gradient by a volume-dependent production rate. *Development* **138**, 2741-2749, doi:10.1242/dev.064402 (2011).

34 Wu, H., Manu, Jiao, R. & Ma, J. Temporal and spatial dynamics of scaling-specific features of a gene regulatory network in Drosophila. *Nat Commun* **6**, 10031, doi:10.1038/ncomms10031 (2015).

35 Huang, A., Rupprecht, J. F. & Saunders, T. E. Embryonic geometry underlies phenotypic variation in decanalized conditions. *Elife* **9**, doi:10.7554/eLife.47380 (2020).

36 Holloway, D. M., Harrison, L. G., Kosman, D., Vanario-Alonso, C. E. & Spirov, A. V. Analysis of pattern precision shows that *Drosophila* segmentation develops substantial

independence from gradients of maternal gene products. *Dev Dynam* **235**, 2949-2960, doi:10.1002/dvdy.20940 (2006).

37 Sprenger, F. & Nusslein-Volhard, C. Torso Receptor Activity Is Regulated by a Diffusible Ligand Produced at the Extracellular Terminal Regions of the Drosophila Egg. *Cell* **71**, 987-1001, doi:10.1016/0092-8674(92)90394-R (1992).

38 Coppey, M., Boettiger, A. N., Berezhkovskii, A. M. & Shvartsman, S. Y. Nuclear trapping shapes the terminal gradient in the Drosophila embryo. *Curr Biol* **18**, 915-919, doi:10.1016/j.cub.2008.05.034 (2008).

39 Gregor, T., Tank, D. W., Wieschaus, E. F. & Bialek, W. Probing the limits to positional information. *Cell* **130**, 153-164, doi:10.1016/j.cell.2007.05.025 (2007).

40 Liu, F., Morrison, A. H. & Gregor, T. Dynamic interpretation of maternal inputs by the Drosophila segmentation gene network. *Proc Natl Acad Sci U S A* **110**, 6724-6729, doi:10.1073/pnas.1220912110 (2013).

41 Poustelnikova, E., Pisarev, A., Blagov, M., Samsonova, M. & Reinitz, J. A database for management of gene expression data in situ. *Bioinformatics* **20**, 2212-2221, doi:10.1093/bioinformatics/bth222 (2004).

42 Pisarev, A., Poustelnikova, E., Samsonova, M. & Reinitz, J. FlyEx, the quantitative atlas on segmentation gene expression at cellular resolution. *Nucleic Acids Res* **37**, D560-D566, doi:10.1093/nar/gkn717 (2009).

43 Kim, Y. *et al.* MAPK substrate competition integrates patterning signals in the Drosophila embryo. *Curr Biol* **20**, 446-451, doi:10.1016/j.cub.2010.01.019 (2010).

44 Dubuis, J. O., Tkacik, G., Wieschaus, E. F., Gregor, T. & Bialek, W. Positional information, in bits. *Proc Natl Acad Sci U S A* **110**, 16301-16308, doi:10.1073/pnas.1315642110 (2013).

45 Petkova, M. D., Tkacik, G., Bialek, W., Wieschaus, E. F. & Gregor, T. Optimal Decoding of Cellular Identities in a Genetic Network. *Cell* **176**, 844-855 e815, doi:10.1016/j.cell.2019.01.007 (2019).

46 Eldon, E. D. & Pirrotta, V. Interactions of the Drosophila gap gene *giant* with maternal and zygotic pattern-forming genes. *Development* **111**, 367-378 (1991).

47 Rothet, M., Wimmer, E. A., Pankratz, M. J., Gonzalez-Gaitan, M. & Jackle, H. Identical transacting factor requirement for knirps and knirps-related gene expression in the anterior but not in the posterior region of the *Drosophila* embryo. *Mechanism of Development* **46**, 169-181 (1994).

48 Struhl, G., Johnston, P. & Lawrence, P. A. Control of Drosophila body pattern by the hunchback morphogen gradient. *Cell* **69**, 237-249, doi:10.1016/0092-8674(92)90405-2 (1992).

49 Dubuis, J. O., Samanta, R. & Gregor, T. Accurate measurements of dynamics and reproducibility in small genetic networks. *Mol Syst Biol* **9**, 639, doi:10.1038/msb.2012.72 (2013).

50 Wimmer, E. A., Cohen, S. M., Jäckle, H. & Desplan, C. *buttonhead* does not contribute to a combinatorial code proposed for *Drosophila* head development. *Development* **24**, 1509-1517 (1997).

51 Ochoa-Espinosa, A., Yu, D., Tsirigos, A., Struffi, P. & Small, S. Anterior-posterior positional

information in the absence of a strong Bicoid gradient. *Proc Natl Acad Sci U S A* **106**, 3823-3828, doi:10.1073/pnas.0807878105 (2009).

52 Bronner, G. & Jackle, H. Control and function of terminal gap gene activity in the posterior pole region of the Drosophila embryo. *Mech Dev* **35**, 205-211, doi:10.1016/0925-4773(91)90019-3 (1991).

53 Surkova, S. *et al.* Quantitative dynamics and increased variability of segmentation gene expression in the *Drosophila Kruppel* and *knirps* mutants. *Dev Biol* **376**, 99-112, doi:10.1016/j.ydbio.2013.01.008 (2013).

Methods

Fitting the linear classification planes.

The WT point cloud (Fig. 2d) used for fitting the linear classification planes are generated as follows. First, the A-P axis is discretized into 101 points y=0% to 100%. For each of the y position, we sample 400 embryo length values from the normal distribution L~N(1, 0.1) and calculate the corresponding (Bcd, mHb, Tor) levels using equation (4). The three morphogen values are noted by $\boldsymbol{m}=(m_1, m_2, m_3)$ for convenience. Obviously, $0<m_{1,3}<1$ and $0<m_2<\mathrm{mHb}_0$ (mHb_0=0.1 here).

Then, a Poisson noise is added to each m_i by assuming the actual number of molecules is a Poisson variable n_i with $\langle n_i\rangle=N*m_i$, and the final morphogen value with noise is n_i/N. The maximum molecule number N controls the noise magnitude, and set to N=1000 throughout the main text. Note that we do not claim that the *actual* morphogen noise follows independently identically Poisson distribution. And N does *not* correspond directly to the number of molecules per nucleus. Instead, N=1000 is chosen to make the positional error of Bcd gradient close to that measured by Gregor, et. al.[39] (about 1~2% embryo length in the anterior half of embryo). Our model is not sensitive to the exact value of N, see Extended Data Fig. 4, g-h for the results for N=500 or 2000.

The entire point cloud in Fig. 2d thus consists of 101 subsets $\boldsymbol{m}|_y$, each of them have 400 points. The 100 classification planes are located at y=0.5%, 1.5%, …, 99.5%, denoted as classifiers #1, …, #100. Since gap gene expression is affected by the dorsoventral system when being very close to embryo termini, only classifiers #6 to #95 are considered in the main text. Each of the planes should perform the *local* classification task of distinguishing adjacent $\boldsymbol{m}$ point sets. For example, the plane locates at y=2.5% should go through the *noise-free* y=2.5% point for *standard-sized* embryo $\boldsymbol{m}|_{L=1,y=2.5\%}$ by itself. And its orientation is such that it can best distinguishing the point sets $\bigcup_L \bigcup_{y=\{0\%,1\%,2\%\}} \boldsymbol{m}|_y$ against $\bigcup_L \bigcup_{y=\{3\%,4\%,5\%\}} \boldsymbol{m}|_y$. To find the best plane orientation numerically, we simply enumerate the Euler angles θ and φ of its normal vector at the resolution of 1° and find the one with the highest classification accuracy.

The noise due to finite sampling and discretizing θ and φ can be eliminated by averaging the classification plane orientations for 25 repeats of the above sampling and fitting steps.

Predicting the fate-map function with the ensemble of linear classifiers.

In principle, 100 well-separated planes should divide the morphogen space (the cube with $0<m_1<1$, $0<m_2<0.1$, $0<m_3<1$) into 101 slices, corresponding to $\tilde{y}$=0% to 100%. If a query $\boldsymbol{m}$ point falls into the slice $\tilde{y}$=n%, it locates on the posterior side of classification planes #1 through #n, and on the anterior side of planes #n+1 to #100. Therefore, the corresponding cell fate $\tilde{y}$ can be read out from the classification results of all the linear classification planes (SI-2).

In some situations, however, the 100 linear classifiers could have contradictory outputs. Say, a point may be classified to the posterior side by classifier #70 but to the anterior side by #30. We introduce a "posterior dominance rule" to tackle this difficulty. Anytime when this happens, output of the anterior classifier (#30 here) is always ignored. The reason for us to introduce the posterior dominance rule is simple – some anterior classification planes may intersect with the much more posterior region of the point cloud, vary far from where they were fitted (Extended Data Fig. 3b). This posterior dominance rule works well, and finally yields all the results in Figs. 2-3.