Shadow enhancers enable Hunchback bifunctionality in the Drosophila embryo

Max V. Staller, Ben J. Vincent, Meghan D.J. Bragdon, Javier Estrada, Zeba Wunderlich, Angela H. DePace[1]

Department of Systems Biology, Harvard Medical School, Boston, Massachusetts

[1]Corresponding Author: 200 Longwood Ave, Boston MA 02115, 617-432-7410

angela_depace@hms.harvard.edu

Key Words:
enhancer, computational model, dual transcriptional regulators, Drosophila development

Abstract

When bifunctional transcription factors activate and repress target genes within the same cell, these opposing activities must be encoded in regulatory DNA. Here, we use cellular resolution gene expression data and computational modeling to investigate Hunchback (Hb) bifunctionality in *Drosophila* embryogenesis. Previous computational models predicted that Hb both activated and repressed the enhancer controlling *even-skipped (eve)* stripes 3 and 7 *(eve3+7)*. We tested this hypothesis by measuring and modeling *eve* expression under multiple genetic perturbations and found that the *eve3+7* enhancer could not explain endogenous stripe 7 behavior. To explain this discrepancy, we measured the response of an extended *eve* stripe 2 enhancer that drives expression of *eve* stripes 2 and 7 *(eve2+7)*. We found that the behavior of endogenous stripe 7 is explained by the combined behavior of both enhancers, *eve3+7* and *eve2+7*. Bifunctionality arises from Hb activating the *eve2+7* enhancer and repressing the *eve3+7* enhancer. This pair can thus be considered “shadow enhancers” that both direct *eve* stripe 7, but respond to Hb in opposite ways. This example may illustrate a general way of encoding bifunctional regulation in the genome.

Introduction

Transcription factors (TFs) are typically categorized as activators or repressors, but many TFs can act bifunctionally by both activating and repressing expression of their target genes (1-5). In cases where TFs activate and repress targets in the same cells, bifunctionality must be locally encoded in regulatory DNA sequence. Determining the regulatory DNA sequence features that control TF bifunctionality will advance two fundamental challenges of decoding transcriptional networks: predicting expression patterns from regulatory sequence, and deciphering how network topology dictates systems-level properties including gene expression precision and robustness to genetic and environmental perturbations.

Here, we investigate how TF bifunctionality is encoded in regulatory DNA using a classic example: the *Drosophila* segmentation gene, *hunchback (hb)* (1, 6-10). Hb both activates and represses the seven-striped *even-skipped (eve)* gene expression pattern by acting on multiple enhancers, genomic regions responsible for tissue specific gene expression (11). Hb activates *eve* stripes 1 and 2 and represses stripes 3,4,5,6 and 7 (6, 8, 12). With our collaborators, we recently developed models of *eve* stripe regulation that suggested, consistent with previous models, that Hb bifunctionally regulates expression of *eve* stripes 3 and 7 (Fig. 1) (13, 14). Here, we test this hypothesis further using quantitative expression data in genetically perturbed embryos. We focus on whether Hb acts as both an activator and a repressor in the annotated *eve3+7* enhancer. We measured both the expression driven by the *eve3+7* enhancer in a reporter construct and the endogenous *eve* expression pattern at high resolution in embryos with perturbed Hb expression levels. We then used these data to challenge our computational models.

We found that Hb bifunctionality is encoded by separate enhancers that both direct *eve* stripe 7 expression. The first is the annotated *eve3+7* enhancer, where Hb acts as a repressor. The second is an extended piece of regulatory DNA encompassing the minimal *eve* stripe 2 *(eve2)* enhancer that drives expression of *eve* stripes 2 and 7 *(eve2+7)* (6, 15, 16). In the *eve2+7* enhancer, Hb acts as an activator. Therefore, *eve* stripe 7 is controlled by a pair of shadow enhancers, separate sequences in a locus that

drive overlapping spatiotemporal patterns (17). Notably, these shadow enhancers respond to Hb in opposite ways and therefore use different regulatory logic. The separation of activation and repression into distinct enhancers may be a general mechanism of encoding TF bifunctionality in the genome.

Results

eve enhancer reporter patterns do not match the endogenous *eve* pattern

To determine if Hb bifunctionality is encoded in the annotated *eve3+7* enhancer, we compared the pattern of the endogenous *eve* stripes to the pattern driven by a *lacZ* reporter construct in two genetic backgrounds. We refer to the resulting patterns throughout the manuscript as "the *eve3+7* reporter pattern" and "the endogenous pattern" (Fig. 2). We examined both WT embryos and embryos where expression of Hb had been perturbed by removing *bcd*, one of its key regulators, using RNAi *(bcd* RNAi embryos) (18). We quantitatively measured expression patterns at cellular resolution using *in situ* hybridization, 2-photon microscopy, and an image processing toolkit developed specifically for *Drosophila* embryos (methods) (19, 20). We then averaged these data together into gene expression atlases (21). Importantly, the reporter construct isolates the activity of the annotated *eve3+7* enhancer while the endogenous pattern integrates the activity of the whole locus.

Our high resolution measurements revealed discrepancies between the endogenous pattern and the *eve3+7* reporter pattern (Fig. 2). In WT embryos, the *eve3+7* reporter pattern overlaps the endogenous *eve* stripes, but these stripes are broader, have uneven levels, and the peaks lie posterior to the endogenous peaks (Fig. 2). These discrepancies were more pronounced in *bcd* RNAi embryos than in WT embryos, especially for the anterior stripe (Fig. 2). When we tested reporters for other *eve* enhancers, we also found that they did not fully recapitulate the endogenous pattern (Fig. S1, S2).

To test if the discrepancies between the *eve3+7* reporter pattern and the endogenous pattern resulted from differences in *eve* and *lacZ* mRNA half-lives, we measured the expression driven by an *eve* locus BAC reporter where the coding sequence had been replaced with *lacZ* (a generous gift from Miki Fujioka). In both WT and *bcd* RNAi embryos, the peak positions and widths of the BAC reporter pattern were more faithful to the endogenous *eve* pattern, but still did not match exactly (Fig. 2, S1, S2). Differences between the endogenous and BAC reporter patterns must arise from differences in the transcripts. Differences between the BAC reporter and the isolated

enhancer reporter patterns must arise from regulatory DNA not included in the reporter constructs. These data indicate that the *eve3+7* reporter construct is missing relevant regulatory DNA. Additional regulatory DNA in the endogenous locus may respond to other TFs or respond to the same TFs differently, leading to differences in how the endogenous pattern is computed from the concentrations of its regulators. We therefore hypothesized that the endogenous locus and the *eve3+7* enhancer perform different computations to produce *eve* stripes 3 and 7. We next tested this hypothesis using computational models of *eve* regulation.

Computational models suggest that Hb activates and represses endogenous *eve* stripes 3 and 7, but only represses the *eve3+7* enhancer

With our collaborators, we previously identified two empirical computational models for the expression of *eve* stripes 3 and 7. (Fig. 1) (13). These models used logistic regression to directly relate the concentrations of input regulators to output expression in single cells by fitting a parameter for each regulator that reflects both TF strength and TF-DNA interactions (13). In the linear model, Hb has one parameter and only represses. In the quadratic model, Hb has two parameters, and activates at low concentrations while repressing at high concentrations (13). Both models perform equally well in WT embryos, but they make different predictions under genetic perturbation. Specifically, the quadratic model predicts published data from *hb* ventral misexpression experiments (see below).

To investigate patterns driven by individual enhancers, these models were fit on data parsed from the entire endogenous *eve* pattern because data for enhancer reporters was not available (13). To relate the results to individual enhancers, we employed a standard assumption: the endogenous expression of *eve* stripes 3 and 7 could be attributed to the activity of the annotated *eve3+7* enhancer. Here, we test this assumption explicitly by measuring and modeling the *eve3+7* reporter and endogenous patterns separately.

We compared the performance of the linear and quadratic models in WT and *bcd* RNAi embryos. As input regulators we used Hb protein and *gt, tll* and *kni* mRNA, and we used thresholded endogenous or *eve3+7* reporter mRNA data as our target output

patterns (Fig 1, methods). All of these regulators, especially Hb, are perturbed in *bcd* RNAi embryos (Fig. S3) (18). We report our modeling of the third time point, which is representative of results for other time points (Fig. S6), and evaluated model performance by computing the area under the receiver operating characteristic curve (AUC) (22).

We first addressed the endogenous *eve* pattern: we fit our models in WT embryos and used the resulting parameters to predict expression in *bcd* RNAi embryos. In this case, we found that the quadratic model more precisely predicted the perturbed endogenous *eve* pattern. Both models correctly predicted the positional shifts of stripe 7 and a wide anterior stripe, but the quadratic model was more accurate than the linear model (AUC_{linear} = 0.93, AUC_{quad} = 0.98) (Fig. 3F, Fig. S4). These analyses indicated that the quadratic model captured the activity of the whole locus by allowing Hb to both activate and repress *eve* stripes 3 and 7.

We next addressed the *eve3+7* reporter pattern: again, we fit our models in WT embryos and used the resulting parameters to predict expression in *bcd* RNAi embryos. To our surprise, the linear model was more accurate than the quadratic model in this case (AUC_{linear} = 0.90, AUC_{quad} = 0.87) (Fig. 3L). Although neither model captured the dorsal-ventral modulation of the pattern, the linear model accurately predicted the posterior boundary of the anterior stripe. We controlled for several factors that may confound model performance. We assessed sensitivity to changes in regulator concentrations, refit the models in *bcd* RNAi embryos alone, and refit the models on all of the data, none of which changed our conclusions (Fig. S5 and S6, Supplemental Note 1). Although these differences in model performance were subtle, the results supported our hypothesis that the *eve3+7* reporter pattern is regulated differently than the endogenous pattern. Specifically, they suggested that the endogenous pattern required bifunctional Hb while the *eve3+7* reporter pattern required only Hb repression.

Based on these results, we hypothesized that Hb activation is encoded in regulatory DNA outside the annotated *eve3+7* enhancer. The differences in model performance were not conclusive of their own accord, prompting us to seek another perturbation to validate this hypothesis. We therefore returned to the perturbation that

previously distinguished the linear and quadratic models, ventral mis-expression of *hb* (13, 14).

_hb_ mis-expression confirms that the endogenous _eve_ pattern and the _eve3+7_ reporter pattern respond to Hb differently

In Ilsley et al., we preferred the quadratic model because it qualitatively predicted the behavior of a classic mis-expression experiment. Mis-expressing *hb* along the ventral surface of the embryo (*sna::hb* embryos) causes *eve* stripe 3 to retreat and bend and stripe 7 to bend and expand (23). In simulations of this perturbation, the quadratic model predicted this behavior while the linear model did not (Fig. 4 E and F reproduced from (13)). We hypothesized that if Hb bifunctionality is encoded in the whole locus but not in the annotated *eve3+7* enhancer, the endogenous and *eve3+7* reporter patterns would respond differently to *hb* misexpression. To test this hypothesis, we repeated this perturbation and measured both the endogenous *eve* pattern and the *eve3+7* reporter pattern quantitatively at cellular resolution (Fig. 4).

In *sna::hb* embryos, the endogenous pattern behaved differently from the *eve3+7* reporter pattern. As previously observed, the endogenous *eve* stripe 3 retreated from the ventral Hb domain and bent posteriorly; the endogenous stripe 7 did not retreat at all, and instead expanded and bent anteriorly. Both of these behaviors were predicted by the quadratic model applied to simulated mis-expression data (Fig. 4 B and E). By contrast, in the *eve3+7* reporter pattern, both stripes 3 and 7 retreated from the ventral Hb domain, consistent with the predictions of the linear model applied to simulated mis-expression data (Fig. 4 C and F) (13). Thus, under this mis-expression perturbation, the linear model predicted the behavior of the *eve3+7* reporter pattern while the quadratic model predicted the behavior of the endogenous pattern. The subtle quantitative difference between the two models that we saw in *bcd* RNAi embryos was corroborated by a strong qualitative difference in *sna::hb* embryos. These data indicate that Hb represses the *eve3+7* reporter pattern and both activates and represses the endogenous pattern. Hb repression is encoded in the *eve3+7* enhancer; we next sought to identify the regulatory DNA in the *eve* locus that directs stripe 7 expression, but is activated by Hb.

Two shadow enhancers encode bifunctional Hb regulation of *eve* stripe 7

We hypothesized that Hb activation of *eve* stripe 7 was encoded in regulatory DNA near the *eve2* enhancer. We focused on the *eve2* enhancer for several reasons: Hb is known to activate the *eve2* enhancer (7); longer versions of *eve2* drive stripe 7 in some embryos (15, 16, 24, 25); orthologous *eve2* enhancers from other species sometimes drive stripe 7 expression (26, 27); finally, in *sna::hb* embryos the border of the expanded stripe 7 appeared to be set by *Krüppel (Kr)*, a known regulator of *eve2* (Fig. S7) (6, 28). We measured expression driven by an extended version of the minimal *eve2* enhancer construct that drove a robust stripe 7 pattern (Table S2); we call this enhancer reporter construct *eve2+7*. Since both the *eve3+7* and *eve2+7* enhancers drive stripe 7 expression, they can be considered a pair of shadow enhancers for stripe 7 (17).

We found that Hb activates stripe 7 expression in the *eve2+7* enhancer. In *sna::hb* embryos, the stripe 7 region of the *eve2+7* reporter pattern expanded, recapitulating the bulging behavior observed in the endogenous *eve* pattern (Fig. 4B and D). We conclude that Hb activates endogenous *eve* stripe 7 through the *eve2+7* enhancer. Taken together, our results indicate that Hb controls *eve* stripe 7 expression by activating and repressing distinct enhancers.

Discussion

We uncovered a pair of shadow enhancers in the *eve* locus that both direct expression of stripe 7 using distinct regulatory logic: one uses Hb as an activator, the other uses Hb as a repressor. We measured expression of the endogenous *eve* locus and transgenic reporter constructs at cellular resolution under two genetic perturbations. We contextualized our results by comparing two computational models with different roles for Hb*:* a linear model where Hb is a dedicated repressor and a quadratic model where Hb is a bifunctional regulator (13). Guided by the modeling, we found that stripe 7 is encoded by two enhancers: Hb represses the *eve3+7* enhancer and activates the *eve2+7* enhancer. These two shadow enhancers therefore use Hb in different ways to position overlapping patterns. This form of regulatory redundancy may be a general way to encode TF bifunctionality.

Expression patterns driven by reporter constructs do not precisely match the endogenous pattern

"Veteran enhancer-bashers, and those who carefully read the papers, know that 'minimal' enhancer fragments do not always perfectly replicate the precise spatial boundaries of expression of the native gene…" (17). Our data clearly support this often neglected aspect of enhancer reporter constructs. One explanation offered for such discrepancies is different mRNA half-lives. We controlled for this possibility with a BAC reporter where the *eve* coding sequence was replaced with *lacZ* and found better, but not perfect, agreement with the endogenous pattern. We conclude that transcript sequence features contribute to the differences between reporter and endogenous patterns, but that additional regulatory DNA in the locus also plays a role. This result highlights the limitations of enhancer reporter constructs for recapitulating behavior of endogenous loci and the importance of using BAC reporters or genomic editing to study loci with multiple enhancers (29-31).

eve stripe 7 is encoded by two shadow enhancers

Early efforts to dissect the regulatory architecture of the *eve* locus suggested that stripe 7 activity was distributed over DNA encompassing both the *eve3+7* and *eve2* enhancers (15, 16). We find that there are at least two regions of regulatory DNA that position stripe 7 using different regulatory logic. The classically annotated *eve3+7* enhancer is repressed by Hb (8, 23, 32), while the *eve2+7* enhancer, which encompasses the minimal *eve2* enhancer, is activated by Hb. We think that activation of *eve2+7* is direct because it is clear from *in vivo* DNA binding data and binding site mutagenesis that Hb directly activates the minimal *eve2* enhancer (7, 33). The redundancy in *eve* stripe 7 regulation may confer robustness to genetic or environmental stresses (34-36), may increase synchrony or precision (37), may facilitate temporal refinement of patterns (38), or may arise from genetic drift (39).

It is likely that these two enhancers are differentially sensitive to additional TFs. Previous studies have revealed the *eve3+7* enhancer is activated by the spatially uniform TFs Stat92E and Zelda, the anterior boundary of stripe 7 is set by Kni repression, and the posterior boundary is set by Hb repression (8, 23, 32, 40). The minimal *eve2* enhancer is activated by Bcd and Hb, its anterior boundary is set by Slp1 and Gt, and its posterior boundary set by Kr (7, 24, 28, 41). Kr appears to set the boundary of both the expanded endogenous *eve* stripe 7 and the *eve2+7* reporter pattern in *sna::hb* embryos (Fig. S7), but *eve3+7* is not sensitive to Kr, as *eve* stripe 3 sits directly beneath the Kr pattern and this enhancer has no predicted Kr binding sites (Fig. S7). In agreement with others, we speculate that the anterior boundary of *eve* stripe 7 in *eve2+7* may be set by Gt (25). However, we cannot rule out the possibility that this boundary is set by limiting levels of Hb activation or by Kni repression.

The molecular mechanism by which Hb represses *eve3+7* and activates *eve2+7* remains unclear. One hypothesis is that other TFs convert Hb from a repressor into an activator. For example, there is experimental evidence for activator synergy between *bcd* and *hb* (42) and Hb/Cad activator synergy has been proposed based on computational work (43). Another proposed mechanism is that monomeric Hb is an activator, but Hb dimers are repressors (14, 44). High and low Hb concentrations may also be correlated with some other spatially varying factor in the embryo, such as

phosphorylation by the MAPK pathway in the poles (45), but data from *sna::hb* embryos reduces the likelihood of this possibility. Testing these hypotheses will require quantitative data in additional genetic backgrounds and mutagenesis of individual binding sites in the two enhancers.

The quadratic model is a superposition of two computations

Models are not ends in and of themselves, but merely means to formalize assumptions and develop falsifiable hypotheses (46, 47). The quadratic model, which includes Hb as a bifunctional TF, accurately predicts the behavior of the locus in all cells of WT and perturbed embryos, but it does not predict the behavior of either individual enhancer. The interpretation in Ilsley et al. that Hb bifunctionality is a feature of the canonical *eve3+7* enhancer was based on a common assumption: that the endogenous expression pattern could be attributed to the annotated enhancer. Here we show that Hb bifunctionality is encoded in separate enhancers. The quadratic model works because it combines the critical features of the *eve3+7* enhancer and the *eve2+7* enhancer, effectively behaving as a superposition of the two activities. In the future, we plan to develop computational models of each enhancer and uncover how these two (or more) activities are combined.

Conclusion

We tested the hypothesis that Hb bifunctionality is encoded in the *eve3+7* enhancer and discovered that it is actually encoded in two separate enhancers that respond to Hb in opposite ways. We show that expression patterns driven by annotated enhancers differ from the endogenous pattern, especially under perturbation, and that these differences can be due to relevant, yet unannotated, regulatory DNA. The stripe 7 shadow enhancers reside in a classic example of a modular locus (48), implying that regulatory complexity may be pervasive. Since the enhancers are active in the same cells, Hb bifunctionality must be encoded in their DNA sequences. This example provides an opportunity to uncover sequence features governing Hb bifunctionality, which will improve our ability to interpret regulatory DNA and infer connections in gene regulatory networks.

Materials and Methods

Fly Work

The *bcd* RNAi gene expression atlas is described in Staller et al. 2014 (submitted) and available at depace.med.harvard.edu. Briefly, we combined short hairpin RNA knockdown of *bcd* with *in situ* hybridization and 2-photon imaging and automated image segmentation (21, 49-51). Hb protein stains used a guinea pig anti-*hb* from John Reinitz (University of Chicago, IL). Embryos were partitioned into six time points using the degree of membrane invagination (0-3%, 4-8%, 9-25%, 26-50%, 51-75%, and 76-100%) which evenly divide the ~60 min blastoderm stage (20). All enhancer reporters are in pBOY and integrated at attP2 (26, 52) (Table S1). The *eve* locus bacterial artificial chromosome (BAC) *lacZ* reporter was a gift from Miki Fujioka (Thomas Jefferson University, PA). It begins 6.4 kb upstream of the *eve* transcriptional start site (TSS) and ends 11.3 kb downstream of the *eve* TSS. The *eve* coding sequence has been replaced by *lacZ* and the adjacent gene, *TER94*, has been fused to GFP. It is effectively a reporter for the whole *eve* locus. *hb* ventral misexpression was performed as described in Clyde et al., 2003 using two copies of the transgene on chromosome 2.

Building the coarsely aligned *sna::hb* gene expression atlas.

We determined the genotype of the *sna::hb* embryos by examining the *eve* or *fushi-tarazu (ftz)* mRNA patterns. Embryos were aligned morphologically to create a coarsely registered gene expression atlas (21). Data is available at depace.med.harvard.edu.

Logistic modeling of enhancer gene regulatory functions

The logistic modeling framework was developed and described in detail previously (13). All modeling was performed in MatLab (MathWorks, Natick, MA) using the DIP image toolbox (diplib.org) and the PointCloudToolBox (bdtnp.lbl.gov). Ilsley et al. used protein data for Hb and Gt, whereas we used Hb protein and *gt* mRNA data. For genes where we used mRNA data, the mRNA and protein patterns are known to be correlated (21, 53). For the enhancer *lacZ* reporters, we thresholded cells to be ON or OFF by creating a histogram of the expression data (50 bins), identifying the bin with

the most counts and adding one standard deviation. Our ON set included all cells expressing the reporter, and our OFF set includes all other cells. All regulators are maintained as continuous values.

To threshold the endogenous WT *eve* pattern into ON and OFF cells we used 0.2 for all time points (13). To threshold the endogenous *eve* patterns in the *bcd* RNAi atlas, we used the lowest threshold that would separate the stripes: 0.1, 0.15, 0.15, 0.2, and 0.21 for T=2 through T=6 respectively. To compare the modeling of the reporter and the endogenous patterns, the ON set included all cells in the endogenous *eve* stripes 3 and 7 and the OFF set included all other cells. This OFF set is different from Ilsley et al., but this change does not have a large effect on the model prediction AUC scores in *bcd* RNAi embryos (Table S1).

Sensitivity analysis

For the sensitivity analysis (Fig. S5), for each TF, we scaled the concentration of the *bcd* RNAi atlas *in silico* and recomputed the model AUC scores.

Binding site predictions

For the *Kr* binding site analysis in Fig. S7, we predicted binding sites using PATSER (stormo.wustl.edu) with a position weight matrix derived from bacterial 1-hybrid data (54). Binding sites were visualized using InSite (cs.utah.edu/~miriah/projects).

Quantifying concordance between reporters and endogenous patterns

For each embryo, we used the pointcloud toolbox in Matlab to find pattern boundaries by creating 16 anterior-posterior line traces and finding the inflection point of each. Finding the boundary by using half the maximum value of the stripe peak identifies a very similar boundary to the inflection point. To find the peaks of the endogenous and reporter stripes, we took one line trace along the lateral part of the embryo using the pointcloud toolbox and found the local maxima.

Acknowledgements:

We thank Miki Fujioka for sharing the *eve* locus BAC reporter flies ahead of publication, Tara Lydiard-Martin for making the enhancer reporter fly lines, and Steve Small for the *sna::hb* flies; Garth Ilsley for developing the initial models and very helpful discussions; John Reinitz for the Hb antibody; Steve Small, Becky Ward, Garth Ilsley, Peter Combs, Alistair Boettiger and members of the DePace lab for extensive comments on the manuscript. This work was supported by the Harvard Herchel Smith Graduate Student Fellowship (MVS), Jane Coffin Childs Memorial Fund for Medical Research (ZBW), NIH K99HD073191 (ZBW), and NIH U01 GM103804-01A1 (AHD).

Figure Legends

Figure 1:

The linear and quadratic models formalize two alternative regulator sets for *eve* stripes 3 and 7. (A) The linear model includes repression (red) by Hb, *knirps (kni), giant (gt),* and *tailless (tll)* and activation (blue) by a constant term that represents spatially uniform factors. The quadratic model includes activation by a linear Hb term and repression by a quadratic Hb term, *kni*, *tll*, and uniform factors. (B) A schematic of the logistic regression framework. Logistic regression calculates the probability the target will be ON based on a linear combination of the concentrations of regulators (μ). We fit models in WT and use the perturbed regulator gene expression patterns to predict the perturbed *eve* patterns in *bcd* RNAi embryos.

Figure 2:

The *eve3+7* reporter pattern differs from the endogenous pattern. (A) The *eve* locus contains 5 annotated primary stripe enhancers. The endogenous pattern integrates the activity of the whole locus. The BAC reporter construct also integrates the activity of the whole locus, but the transcript is the same as the *eve3+7* reporter construct. The *eve3+7* reporter construct isolates the activity of the annotated enhancer sequence. (B) WT expression patterns are represented as line traces where anterior-posterior (A-P) position is plotted on the X-axis with expression level on the Y-axis for a lateral strip of the embryo. Endogenous *eve* pattern (gray), *eve3+7* reporter pattern (red). The reporter

pattern was manually scaled to match the level of the endogenous pattern. (C) Line traces in *bcd* RNAi embryos. (D) The boundaries of the endogenous pattern (gray), the *eve3+7* reporter pattern (red), and the BAC reporter pattern (blue) at T=3. All error bars are the standard error of the mean. The BAC reporter pattern is more faithful to the endogenous pattern than the *eve3+7* reporter pattern, especially in the anterior of *bcd* RNAi embryos (eve 3/7 ant). The endogenous pattern is shaded for visual clarity. (E) Peak positions of stripes 3 and 7, calculated from the line traces in B and C. The *eve3+7* reporter pattern shows better agreement to the endogenous pattern in WT than in *bcd* RNAi embryos. (F) Stripe widths, calculated from the inflection point of the line traces in B and C. The *eve3+7* reporter pattern is wider than the corresponding endogenous pattern.

Figure 3:

In *bcd* RNAi embryos, the quadratic model more accurately predicts the endogenous pattern, and the linear model more accurately predicts the *eve3+7* reporter pattern. (A) The endogenous *eve* pattern in WT embryos. Cells with expression below an ON/OFF threshold (methods) are plotted in gray. For cells above this threshold, darker color indicates higher level. (B) The predictions of the linear and quadratic models in WT embryos. (C) Comparison of model predictions to the endogenous pattern in WT embryos. Green cells are true positives, purple cells are false positives, dark gray cells are false negatives, and light gray cells are true negatives. For visualization, the threshold is set to 80% sensitivity, but the AUC metric quantifies performance over all thresholds. (D) The endogenous *eve* pattern in *bcd* RNAi embryos. (E) The predictions of the linear (L) and quadratic (Q) models in *bcd* RNAi embryos. (F) Comparison of model predictions to the endogenous pattern in *bcd* RNAi embryos. The quadratic model more accurately predicts the endogenous pattern in *bcd* RNAi embryos. (G-L) Same as A-F, respectively, for the *eve3+7* reporter pattern. The linear model predicts the *eve3+7* reporter pattern more accurately in *bcd* RNAi embryos. Model parameters are in Table S1.

Figure 4:

In *hb* ventral misexpression *(sna::hb)* embryos, the quadratic model predicts the endogenous pattern while the linear model predicts the *eve3+7* reporter pattern. (A) The distribution of Hb in WT and *sna::hb* embryos, from a lateral view (left) and ventral view (right). The expression level of each regulator is shown for individual cells: cells with expression below an ON/OFF threshold (methods) are plotted in gray. For cells above this threshold, darker colors indicate higher levels. (B) Endogenous *eve* pattern. (C) The *eve3+7* reporter pattern. Both stripes retreat from ectopic Hb. (D) The *eve2+7* reporter pattern. Stripe 7 expands into the ectopic Hb domain. (E-F) Bottom (ventral) view of predictions of the quadratic model (E) and linear (F) models based on simulated *sna::hb* data. OFF cells are light pink and ON cells are red. Reproduced from Ilsley et al. 2013.

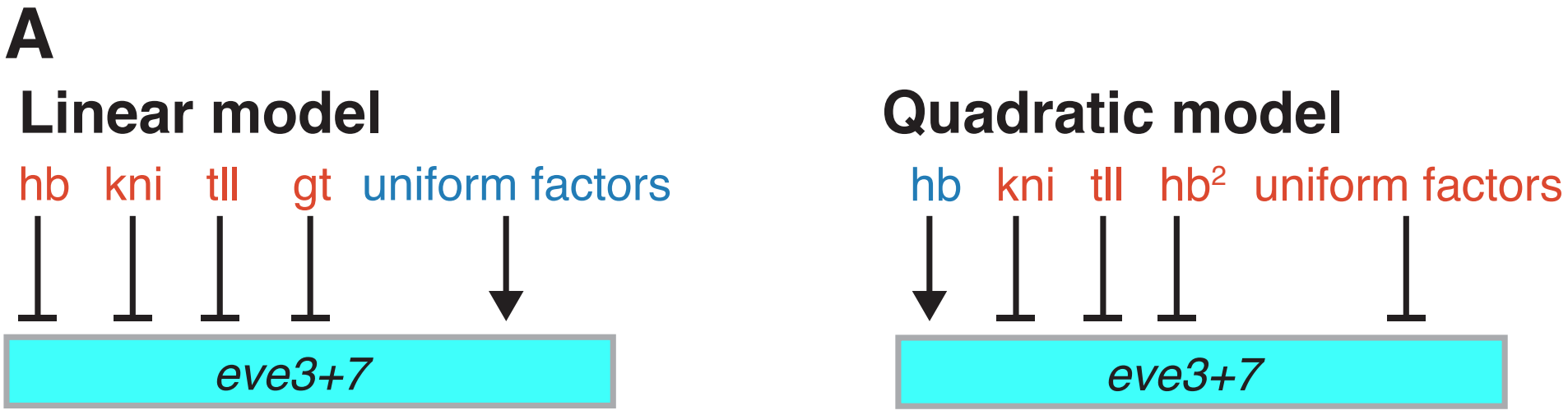
References

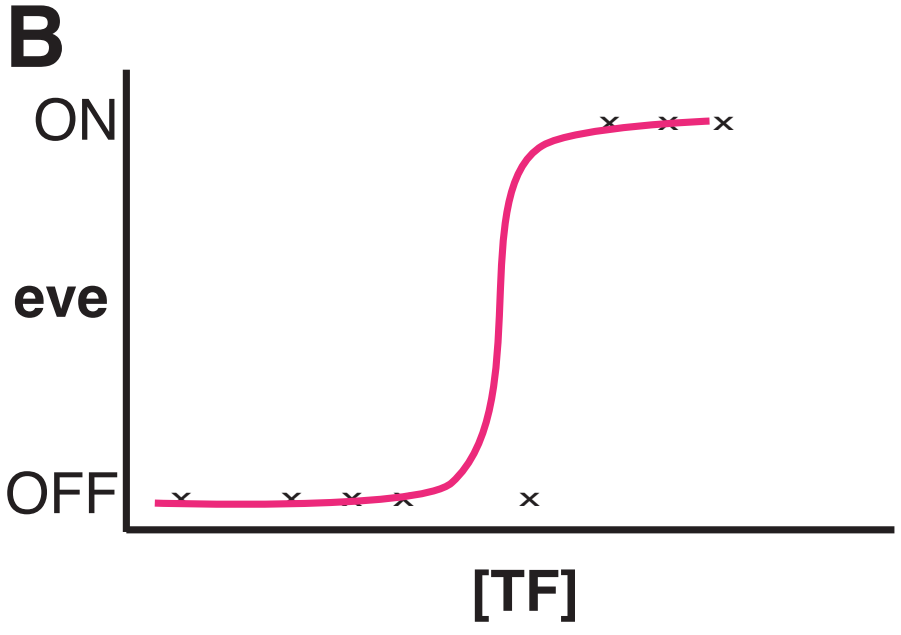
1. Struhl G, Johnston P, Lawrence PA (1992) Control of Drosophila body pattern by the hunchback morphogen gradient. *Cell* 69:237–249.
2. Sauer F, Jäckle H (1993) Dimerization and the control of transcription by Krüppel. *Nature* 364:454–457.
3. Deng Z, Cao P, Wan MM, Sui G (2010) Yin Yang 1: a multifaceted protein beyond a transcription factor. *Transcription* 1:81–84.
4. Di Stefano B, Sardina JL, van Oevelen C, Collombet S, Kallin EM, Vicent GP, Lu J, Thieffry D, Beato M, Graf T (2014) C/EBPalpha poises B cells for rapid reprogramming into induced pluripotent stem cells. *Nature* 506:235–239.
5. Shore D, Nasmyth K (1987) Purification and cloning of a DNA binding protein from yeast that binds to both silencer and activator elements. *Cell* 51:721–732.
6. Small S, Kraut R, Hoey T, Warrior R, Levine MS (1991) Transcriptional regulation of a pair-rule stripe in Drosophila. *Genes & Development* 5:827–839.
7. Arnosti DN, Barolo S, Levine MS, Small S (1996) The eve stripe 2 enhancer employs multiple modes of transcriptional synergy. *Development (Cambridge, England)* 122:205–214.
8. Small S, Blair A, Levine MS (1996) Regulation of two pair-rule stripes by a single enhancer in the Drosophila embryo. *Developmental biology* 175:314–324.
9. Schulz C, Tautz D (1994) Autonomous concentration-dependent activation and repression of Kruppel by hunchback in the Drosophila embryo. *Development* 120:3043–3049.
10. Zuo P, Stanojevic D, Colgan J, Han K, Levine MS, Manley JL (1991) Activation and repression of transcription by the gap proteins hunchback and Kruppel in cultured Drosophila cells. *Genes & Development* 5:254–264.
11. Levine M, Cattoglio C, Tjian R (2014) Looping back to leap forward: transcription enters a new era. *Cell* 157:13–25.
12. Fujioka M, Emi-Sarker Y, Yusibova GL, Goto T, Jaynes JB (1999) Analysis of an even-skipped rescue transgene reveals both composite and discrete neuronal and early blastoderm enhancers, and multi-stripe positioning by gap gene repressor gradients. *Development (Cambridge, England)* 126:2527–2538.
13. Ilsley GR, Fisher J, Apweiler R, DePace AH, Luscombe NM (2013) Cellular resolution models for even skipped regulation in the entire Drosophila embryo. *Elife* 2:e00522.
14. Papatsenko D, Levine MS (2008) Dual regulation by the Hunchback gradient in the Drosophila embryo. *Proc Natl Acad Sci U S A* 105:2901–2906.
15. Goto T, Macdonald P, Maniatis T (1989) Early and late periodic patterns of< i> even skipped</i> expression are controlled by distinct regulatory elements that respond to different spatial cues. *Cell* 57:413–422.
16. Harding K, Hoey T, Warrior R, Levine MS (1989) Autoregulatory and gap gene response elements of the even-skipped promoter of Drosophila. *The EMBO journal* 8:1205–1212.
17. Barolo S (2012) Shadow enhancers: frequently asked questions about distributed cis-regulatory information and enhancer redundancy. *Bioessays* 34:135–141.

18. Staller MV, Fowlkes CC, Bragdon MD, Wunderlich Z, DePace AH (2014) A gene expression atlas of a *bicoid*-depleted *Drosophila* embryo reveals early canalization of cell fate. Submitted. Available from BioRXiv dx.doi.org/10.1101/004788.
19. Luengo Hendriks CL, Keränen SVE, Fowlkes CC, Simirenko L, Weber GH, DePace AH, Henriquez C, Kaszuba DW, Hamann B, Eisen MB *et al.* (2006) Three-dimensional morphology and gene expression in the Drosophila blastoderm at cellular resolution I: data acquisition pipeline. *Genome Biology* 7:R123.
20. Keränen SVE, Fowlkes CC, Luengo Hendriks CL, Sudar D, Knowles DW, Malik J, Biggin MD (2006) Three-dimensional morphology and gene expression in the Drosophila blastoderm at cellular resolution II: dynamics. *Genome Biology* 7:R124.
21. Fowlkes CC, Hendriks CLL, Keränen SVE, Weber GH, Rübel O, Huang M-Y, Chatoor S, DePace AH, Simirenko L, Henriquez C *et al.* (2008) A quantitative spatiotemporal atlas of gene expression in the Drosophila blastoderm. *Cell* 133:364–374.
22. Swets JA (1988) Measuring the accuracy of diagnostic systems. *Science* 240:1285–1293.
23. Clyde DE, Corado MSG, Wu X, Pare A, Papatsenko D, Small S (2003) A self-organizing system of repressor gradients establishes segmental complexity in Drosophila. *Nature* 426:849–853.
24. Small S, Blair A, Levine M (1992) Regulation of even-skipped stripe 2 in the Drosophila embryo. *EMBO J* 11:4047–4057.
25. Janssens H, Hou S, Jaeger J, Kim AR, Myasnikova E, Sharp D, Reinitz J (2006) Quantitative and predictive model of transcriptional control of the Drosophila melanogaster even skipped gene. *Nat Genet* 38:1159–1165.
26. Hare EE, Peterson BK, Iyer VN, Meier R, Eisen MB (2008) Sepsid even-skipped Enhancers Are Functionally Conserved in Drosophila Despite Lack of Sequence Conservation. *PLoS Genetics* 4:e1000106.
27. Peterson BK, Hare EE, Iyer VN, Storage S, Conner L, Papaj DR, Kurashima R, Jang E, Eisen MB (2009) Big genomes facilitate the comparative identification of regulatory elements. *PLoS ONE* 4:e4688.
28. Stanojevic D, Small S, Levine MS (1991) Regulation of a segmentation stripe by overlapping activators and repressors in the Drosophila embryo. *Science* 254:1385–1387.
29. Venken KJ, Bellen HJ (2012) Genome-wide manipulations of Drosophila melanogaster with transposons, Flp recombinase, and PhiC31 integrase. *Methods Mol Biol* 859:203–228.
30. Ren X, Sun J, Housden BE, Hu Y, Roesel C, Lin S, Liu LP, Yang Z, Mao D, Sun L *et al.* (2013) Optimized gene editing technology for Drosophila melanogaster using germ line-specific Cas9. *Proc Natl Acad Sci U S A* 110:19012–19017.
31. Crocker J, Stern DL (2013) TALE-mediated modulation of transcriptional enhancers in vivo. *Nat Methods* 10:762–767.
32. Struffi P, Corado M, Kaplan L, Yu D, Rushlow C, Small S (2011) Combinatorial activation and concentration-dependent repression of the Drosophila even skipped stripe 3+7 enhancer. *Development (Cambridge, England)* 138:4291–4299.
33. Li X-Y, MacArthur S, Bourgon R, Nix D, Pollard DA, Iyer VN, Hechmer A, Simirenko L, Stapleton M, Hendriks CLL *et al.* (2008) Transcription Factors Bind Thousands of Active and Inactive Regions in the Drosophila Blastoderm. *PLoS biology* 6:e27.

34. Dunipace L, Ozdemir A, Stathopoulos A (2011) Complex interactions between cis-regulatory modules in native conformation are critical for Drosophila snail expression. *Development* 138:4075–4084.
35. Perry MW, Boettiger AN, Bothma JP, Levine M (2010) Shadow enhancers foster robustness of Drosophila gastrulation. *Curr Biol* 20:1562–1567.
36. Frankel N, Davis GK, Vargas D, Wang S, Payre F, Stern DL (2010) Phenotypic robustness conferred by apparently redundant transcriptional enhancers. *Nature* 466:490–493.
37. Boettiger AN, Levine MS (2009) Synchronous and Stochastic Patterns of Gene Activation in the Drosophila Embryo. *Science* 325:471–473.
38. Dunipace L, Saunders A, Ashe HL, Stathopoulos A (2013) Autoregulatory feedback controls sequential action of cis-regulatory modules at the brinker locus. *Dev Cell* 26:536–543.
39. Lynch M (2007) The frailty of adaptive hypotheses for the origins of organismal complexity. *Proceedings of the National Academy of Sciences of the United States of America* 104 Suppl 1:8597–8604.
40. Yan R, Small S, Desplan C, Dearolf CR, Darnell JE (1996) Identification of a Stat gene that functions in Drosophila development. *Cell* 84:421–430.
41. Andrioli LPM, Vasisht V, Theodosopoulou E, Oberstein AL, Small S (2002) Anterior repression of a Drosophila stripe enhancer requires three position-specific mechanisms. *Development (Cambridge, England)* 129:4931–4940.
42. Simpson-Brose M, Treisman J, Desplan C (1994) Synergy between the hunchback and bicoid morphogens is required for anterior patterning in Drosophila. *Cell* 78:855–865.
43. Kim A-R, Martinez C, Ionides J, Ramos AF, Ludwig MZ, Ogawa N, Sharp DH, Reinitz J (2013) Rearrangements of 2.5 kilobases of noncoding DNA from the Drosophila even-skipped locus define predictive rules of genomic cis-regulatory logic. *PLoS Genetics* 9:e1003243.
44. Bieler J, Pozzorini C, Naef F (2011) Whole-Embryo Modeling of Early Segmentation in Drosophila Identifies Robust and Fragile Expression Domains. *Biophysj* 101:287–296.
45. Kim Y, Coppey M, Grossman R, Ajuria L, Jiménez G, Paroush Z, Shvartsman SY (2010) MAPK Substrate Competition Integrates Patterning Signals in the< i> Drosophila Embryo. *Current Biology* 20:446–451.
46. Gunawardena J (2014) Models in biology:'accurate descriptions of our pathetic thinking'. *BMC Biology*
47. Wunderlich Z, DePace AH (2011) Modeling transcriptional networks in Drosophila development at multiple scales. *Curr Opin Genet Dev* 21:711–718.
48. Maeda RK, Karch F (2011) Gene expression in time and space: additive vs hierarchical organization of cis-regulatory regions. *Curr Opin Genet Dev* 21:187–193.
49. Fowlkes CC, Eckenrode KB, Bragdon MD, Meyer M, Wunderlich Z, Simirenko L, Luengo Hendriks CL, Keranen SV, Henriquez C, Knowles DW *et al.* (2011) A conserved developmental patterning network produces quantitatively different output in multiple species of Drosophila. *PLoS Genet* 7:e1002346.

50. Wunderlich Z, Bragdon MD, Eckenrode KB, Lydiard-Martin T, Pearl-Waserman S, DePace AH (2012) Dissecting sources of quantitative gene expression pattern divergence between Drosophila species. *Mol Syst Biol* 8:604.
51. Staller MV, Yan D, Randklev S, Bragdon MD, Wunderlich ZB, Tao R, Perkins LA, DePace AH, Perrimon N (2013) Depleting gene activities in early Drosophila embryos with the "maternal-Gal4-shRNA" system. *Genetics* 193:51–61.
52. Groth AC, Fish M, Nusse R, Calos MP (2004) Construction of transgenic Drosophila by using the site-specific integrase from phage phiC31. *Genetics* 166:1775–1782.
53. Pisarev A, Poustelnikova E, Samsonova M, Reinitz J (2009) FlyEx, the quantitative atlas on segmentation gene expression at cellular resolution. *Nucleic Acids Res* 37:D560–D566.
54. Noyes MB, Meng X, Wakabayashi A, Sinha S, Brodsky MH, Wolfe SA (2008) A systematic characterization of factors that regulate Drosophila segmentation via a bacterial one-hybrid system. *Nucleic Acids Res* 36:2547–2560.

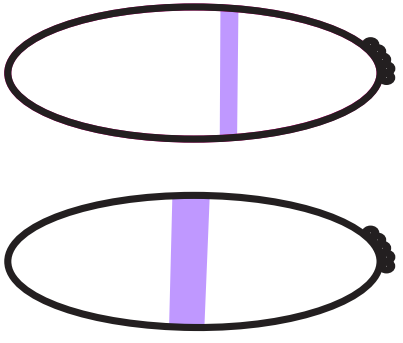




$$P\,(\text{eve} = \text{ON}) = \frac{1}{1 + e^{-\mu}}$$

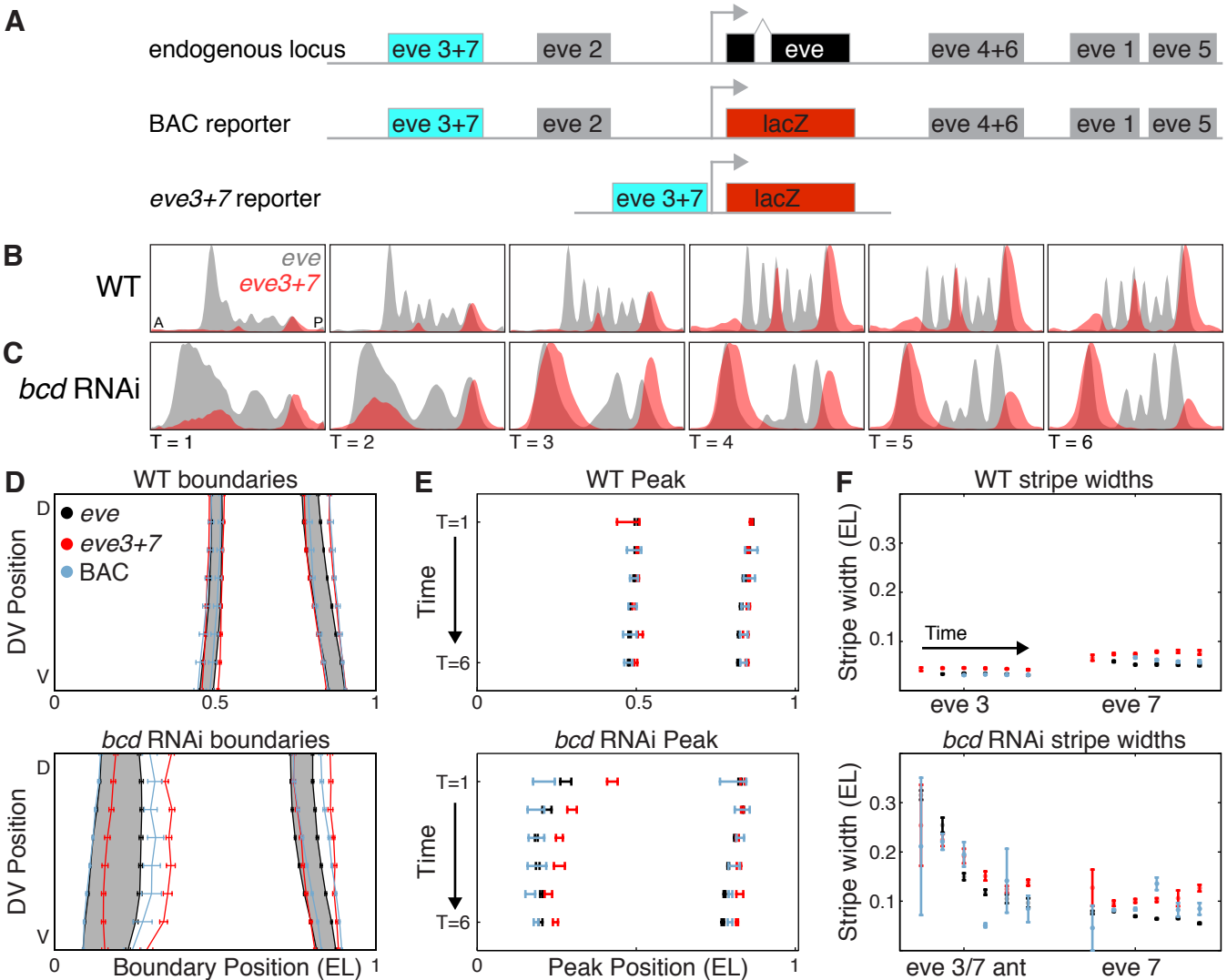
Fit in WT

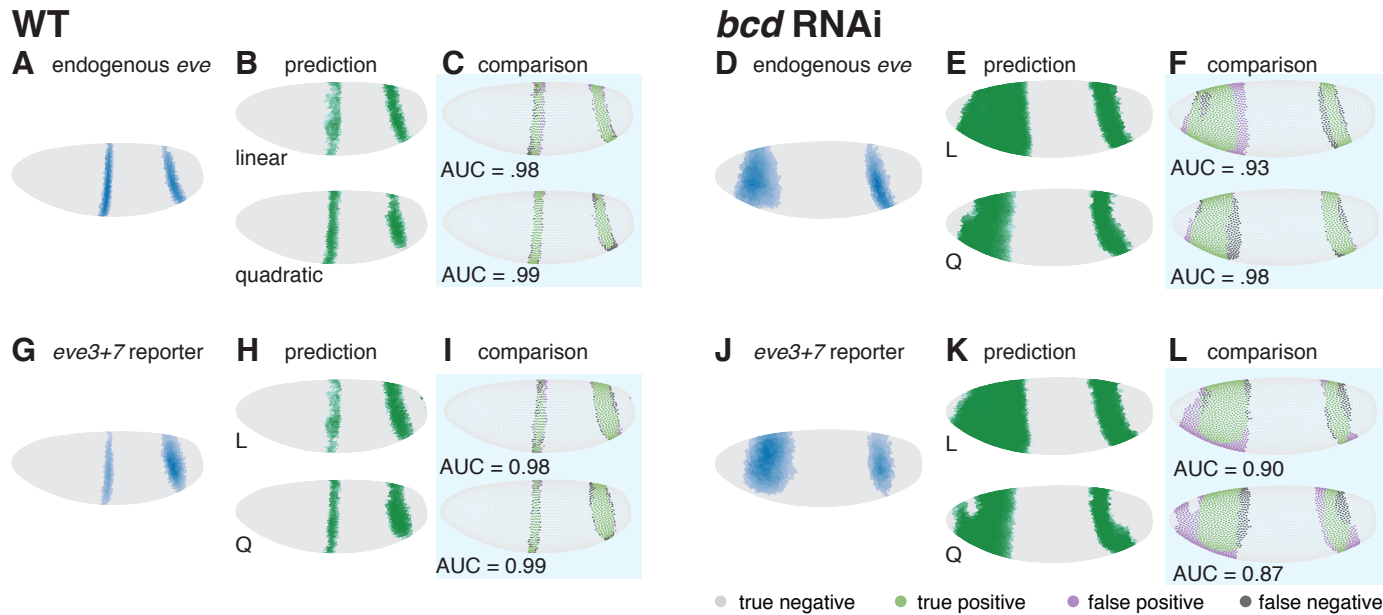
Predict *bcd* RNAi

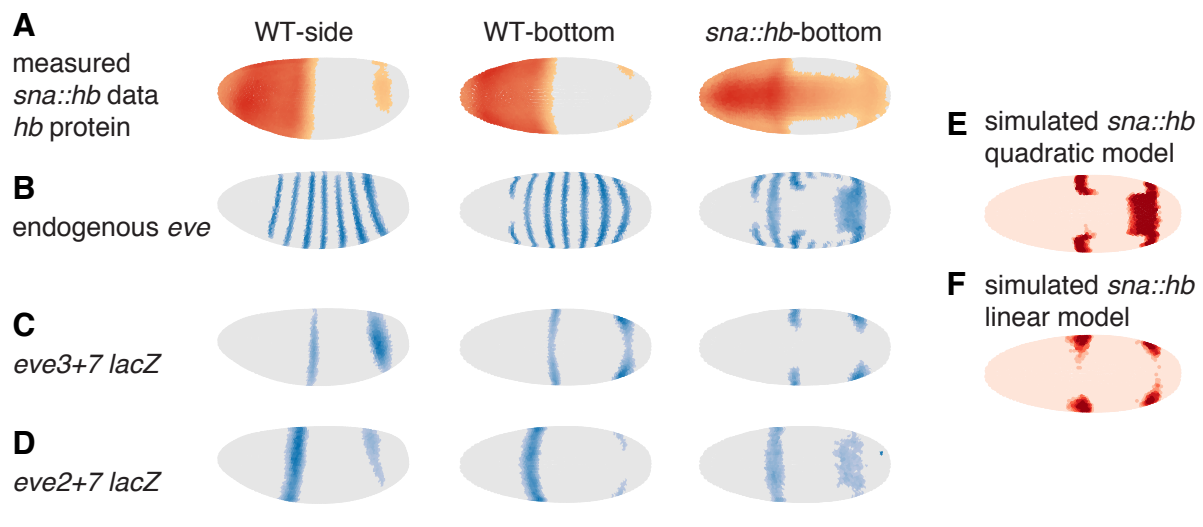


$$\mu = \beta_{uniform} + \beta_{hb}[hb] + \beta_{kni}[kni] + \beta_{tll}[tll] + \beta_{gt}[gt]$$ Linear model

$$\mu = \beta_{uniform} + \beta_{hb}[hb] + \beta_{kni}[kni] + \beta_{tll}[tll] + \beta_{hb2}[hb^2]$$ Quadratic model


A
endogenous locus
eve 3+7
eve 2
eve
eve 4+6
eve 1
eve 5
BAC reporter
eve 3+7
eve 2
lacZ
eve 4+6
eve 1
eve 5
eve3+7 reporter
eve 3+7
lacZ
B
WT
eve
eve3+7
A
P
C
bcd RNAi
T = 1
T = 2
T = 3
T = 4
T = 5
T = 6
D
WT boundaries
eve
eve3+7
BAC
DV Position
D
V
0
0.5
1
bcd RNAi boundaries
DV Position
D
V
0
Boundary Position (EL)
1
E
WT Peak
Time
T=1
T=6
0
0.5
1
bcd RNAi Peak
Time
T=1
T=6
0
Peak Position (EL)
1
F
WT stripe widths
Stripe width (EL)
0.3
0.2
0.1
Time
eve 3
eve 7
bcd RNAi stripe widths
Stripe width (EL)
0.3
0.2
0.1
eve 3/7 ant
eve 7


WT
A endogenous eve
B prediction
linear
quadratic
C comparison
AUC = .98
AUC = .99
bcd RNAi
D endogenous eve
E prediction
L
Q
F comparison
AUC = .93
AUC = .98
G eve3+7 reporter
H prediction
L
Q
I comparison
AUC = 0.98
AUC = 0.99
J eve3+7 reporter
K prediction
L
Q
L comparison
AUC = 0.90
AUC = 0.87
true negative
true positive
false positive
false negative


A
measured
sna::hb data
hb protein
WT-side
WT-bottom
sna::hb-bottom
B
endogenous eve
C
eve3+7 lacZ
D
eve2+7 lacZ
E
simulated sna::hb
quadratic model
F
simulated sna::hb
linear model