HHS Public Access

Author manuscript

Cell Rep. Author manuscript; available in PMC 2019 June 27.

Published in final edited form as:

Cell Rep. 2019 February 26; 26(9): 2407–2418.e5. doi:10.1016/j.celrep.2019.01.115.

Signal Integration by Shadow Enhancers and Enhancer Duplications Varies across the *Drosophila* Embryo

Clarissa Scholes[1,2], **Kelly M. Biette**[1], **Timothy T. Harden**[1], and **Angela H. DePace**[1,3,*]

[1]Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA

[2]Present address: Mammoth Biosciences, San Francisco, CA 94107, USA

[3]Lead Contact

SUMMARY

Transcription of developmental genes is controlled by multiple enhancers. Frequently, more than one enhancer can activate transcription from the same promoter in the same cells. How is regulatory information from multiple enhancers combined to determine the overall expression output? We measure nascent transcription driven by a pair of shadow enhancers, each enhancer of the pair separately, and each duplicated, using live imaging in *Drosophila* embryos. This set of constructs allows us to quantify the input-output function describing signal integration by two enhancers. We show that signal integration performed by these shadow enhancers and duplications varies across the expression pattern, implying that how their activities are combined depends on the transcriptional regulators bound to the enhancers in different parts of the embryo. Characterizing signal integration by multiple enhancers is a critical step in developing conceptual and computational models of gene expression at the locus level, where multiple enhancers control transcription together.

Graphical Abstract

*Correspondence: angela_depace@hms.harvard.edu.

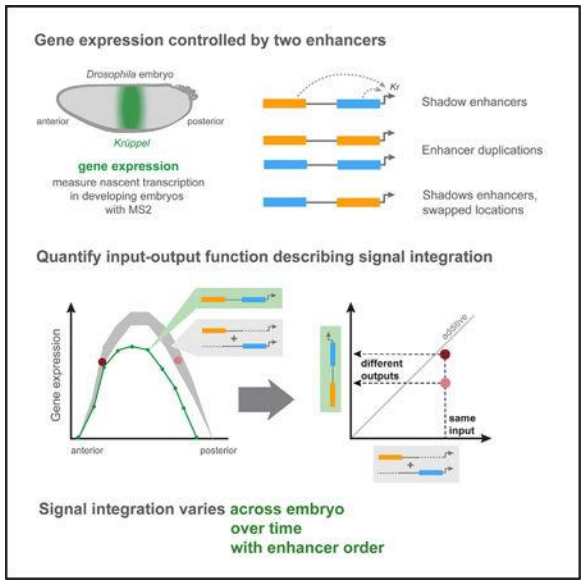
AUTHOR CONTRIBUTIONS C.S. and A.H.D. conceived the study, designed the experiments, and wrote the paper; C.S. conducted the experiments and analyzed the data; K.M.B. helped generate the transgenic fly lines; T.T.H. assisted in the image processing and analysis. All of the authors read and approved the final manuscript.

SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures and three tables and can be found with this article online at https://doi.org/10.1016/j.celrep.2019.01.115.

DECLARATION OF INTERESTS

The authors declare no competing interests.



In Brief

Scholes et al. examine how shadow enhancers and enhancer duplications drive expression from a single promoter. They show that ‘‘signal integration’’ varies across the expression pattern, implying that how enhancers’ activities combine depends on the transcriptional regulators they bind in different parts of the embryo.

INTRODUCTION

Developmental genes are controlled by many enhancers, some of which can drive overlapping patterns of gene expression in space and time (Prazak et al., 2010; Dunipace et al., 2011; Hong et al., 2008; Osterwalder et al., 2018; Frankel et al., 2010; Perry et al., 2011; Hoch et al., 1991; Hay et al., 2016; Dukler et al., 2016; Zeitlinger et al., 2007; Ghiasvand et al., 2011; Jeong et al., 2006; McBride et al., 2011; Lam et al., 2015). If enhancers are active in different cells and work independently of one another, their overall transcriptional output should be predicted by superimposing their activities. However, ‘‘shadow’’ enhancers, which drive spatiotemporally overlapping expression of the same gene (Hong et al., 2008), produce patterns and levels of expression that cannot be predicted from their separate activities (Prazak et al., 2010; Dunipace et al., 2011; Bothma et al., 2015). We do not yet understand how, when two or more enhancers are simultaneously active, their activities are integrated to determine the level and timing of gene expression.

Shadow enhancers are pervasive at developmental genes in vertebrates and invertebrates (Cannavò et al., 2016; Kvon et al., 2014; Barolo, 2012), where they are thought to improve the precision and robustness of expression (Barolo, 2012; Frankel, 2012; Frankel et al., 2010; Perry et al., 2010; Lam et al., 2015). While they drive similar expression patterns, shadow enhancers are not functionally identical and can collaborate to fine-tune the spatial or temporal boundaries of gene expression (Prazak et al., 2010; Perry et al., 2011; El-Sherif and Levine, 2016). Their partially redundant behavior may also enable shadow enhancers to

buffer against environmental or genetic perturbations; the deletion of a shadow enhancer may have no effect under normal developmental conditions and only be revealed under conditions of stress (Frankel et al., 2010; Perry et al., 2010). The mechanisms by which shadow enhancers may enable robust gene expression remain unclear. One suggestion is that having more than one enhancer increases the likelihood of gene activation at any given time (Perry et al., 2010); another is that shadow enhancers help to maintain the level of gene expression above a critical threshold (Ghiasvand et al., 2011; Lam et al., 2015). Deciphering the role of shadow enhancers in development will require elucidating the underlying molecular mechanisms that control how they act in combination.

While we know a great deal about the molecular mechanisms of transcription initiation and elongation (Fuda et al., 2009), we know much less about how enhancers regulate these processes, let alone how multiple enhancers control them through interactions with the same promoter. Enhancers contact the promoter to activate transcription, often from a genomic distance of many kilobases (Spitz, 2016). Recent live imaging experiments showed that transcription of a gene occurred only when the promoter and its distant enhancer were in close physical proximity (Chen et al., 2018). Enhancers also make contact with one another, as demonstrated by chromatin capture experiments (Ghavi-Helm et al., 2014), and this type of assay has shown that more than one enhancer is able to physically contact the promoter at a time on a single DNA template (Oudelaar et al., 2018). Reciprocally, live imaging reveals that a single enhancer can activate two promoters at a time (Fukaya et al., 2016). These studies paint a picture of ‘‘many-to-many’’ interactions between enhancers and promoters.

One route to unraveling the molecular mechanisms by which multiple enhancers interact with a single promoter is to examine the signal integration that they perform. For instance, a pair of enhancers together may drive expression that is equal to, greater than, or less than the sum of their separate activities (Bothma et al., 2015). We define the input-output function for a pair of concurrently active enhancers as the relationship between the expression level driven by an enhancer pair and the sum of expression driven by each enhancer alone. We can compare not only whether expression from a pair of enhancers deviates from additivity for a given level of expression driven by the two acting separately, but also the degree to which it does so. This input-output function describes the signal integration carried out by the enhancers and must reflect the underlying molecular mechanisms. Quantifying this function is a critical first step in developing conceptual and computational models of expression at the locus level, where multiple enhancers work together to dictate gene expression.

On the basis of live measurements of nascent transcription, shadow enhancers have been proposed to activate the promoter one at a time (Bothma et al., 2015). If this is the case, and enhancers activate the promoter frequently enough and for long enough, they will end up competing with one another; we refer to this as the ‘‘competition model.’’ As described by Bothma et al. (2015), this model predicts that when enhancers drive low expression, their activities combine additively, while when they drive high expression, their activities combine sub-additively. This assumes that increasing expression driven by a given enhancer reflects an increase in the fraction of time its cognate promoter spends actively transcribing (e.g., due to an increased frequency or duration of transcriptional bursts), as opposed to reflecting a change in the transcription initiation rate (i.e., an increase in burst amplitude), since the

latter would not alter the degree of competition between enhancers. Consistent with the competition model, Bothma et al. (2015) showed that the shadow enhancers of *hunchback* and *snail* act additively at low expression levels, but act sub-additively at high expression levels. However, the model cannot explain cases in which expression is greater than additive, as in their data for *knirps*. In this case, the authors speculate that super-additive expression may reflect cooperative interactions between these enhancers.

Krüppel is well suited to test the competition model because its central domain of expression in the *Drosophila* blastoderm embryo is driven by two shadow enhancers (Figures 1A and 1B), one of which drives low levels of expression and the other high levels (Figure 1E). Duplicating the lowly expressing and highly expressing enhancers is the minimal perturbation with which to probe the model's prediction that expression level, as a proxy for enhancer-promoter contact, determines the combined output from two enhancers. Switching the positions of the shadow enhancers provides a further test: under the competition model, there should be no effect of this perturbation on how the activities of the two enhancers combine. Finally, the shadow enhancers contain different complements of transcription factor (TF) binding sites (Wunderlich et al., 2015), while the duplicated enhancers are by definition identical in their binding site content. Comparing the behavior of the shadow enhancer pair to that of the duplicates lets us investigate whether binding the same versus different sets of TFs affects how the activities of two enhancers are combined.

To investigate how information from multiple active enhancers is integrated, we measured the input-output function for all combinations of the pair of shadow enhancers that drive *Krüppel* expression in *Drosophila melanogaster* blastoderm embryos. Our constructs include the *Krüppel* shadow enhancers in their endogenous locations, with their locations swapped with respect to the promoter, and duplications of each enhancer in the pair. Previous studies have measured expression from shadow enhancers in fixed embryos using *in situ* hybridization (Dunipace et al., 2011; Prazak et al., 2010; Wunderlich et al., 2015; Perry et al., 2011) and, more recently, in live embryos using the MS2 system to quantify nascent transcription (Bothma et al., 2015; El-Sherif and Levine, 2016). Live imaging of nascent transcription is particularly powerful for studying signal integration by multiple enhancers in *Drosophila* because its early development is rapid, with expression changing quickly in space and time. Since enhancer-promoter distance can affect expression (Fukaya et al., 2016; Quintero-Cadena and Sternberg, 2016; Figure 2), we compare the output from each two-enhancer construct to appropriately distance-controlled single enhancer constructs (Figure 1C). Eliminating this variable is essential to enabling our focus on the signal integration itself.

Our results demonstrate that signal integration by simultaneously active enhancers is not simply a constant function of their individual expression levels. For a given enhancer pair, the input-output function differs between the anterior and the posterior edges of the expression pattern, and in one case changes over time. By switching the order of the shadow enhancers, we also show that all relevant information is not contained within them; rather, additional information is contained in their relative positions to one another in the locus, which influences how their activities are integrated. Our results reinforce that molecular mechanisms to control gene expression must be considered at multiple levels, from

individual TFs, to collections of TFs in enhancers, to collections of enhancers in complex loci. In particular, the mechanisms operating at the locus level remain to be clearly elucidated.

RESULTS

To measure the dynamic signal integration performed by combinations of two enhancers across the *Drosophila* embryo, we generated a comprehensive set of constructs containing one or two copies of the blastoderm enhancers of *Krüppel* (Figure 1A). The constructs drive expression of an MS2 reporter cassette that enabled us to image nascent transcription in developing embryos (Garcia et al., 2013; Lucas et al., 2013) (see Method Details). *Krüppel's* enhancers, CD1 and CD2 (Hoch et al., 1990), lie within 4 kb upstream of the promoter and drive spatiotemporally overlapping patterns of expression in the blastoderm embryo (Hoch et al., 1990; Perry et al., 2011; El-Sherif and Levine, 2016) (Figure 1B). Since the distance between enhancer and promoter affects the expression levels (Fukaya et al., 2016), we made control constructs to measure expression from each enhancer in single copy at the alternate location (i.e., CD1 at the CD2 position and CD2 at the CD1 position, denoted CD1* and CD2*, respectively; Figure 1C). We followed the sequence definitions for the *Krüppel* enhancers established by Perry et al. (2011) (Table S1), which were based on the binding of the *Krüppel* regulators measured in chromatin immunoprecipitation (ChIP)-chip assays (Li et al., 2008). To maintain the endogenous distances between regulatory elements and to ensure that we did not accidentally introduce binding sites for the regulators of *Krüppel*, we replaced each enhancer in the single-enhancer constructs with an equivalent length of synthetic neutral sequence depleted of binding sites for TFs involved in patterning the blastoderm embryo (Estrada et al., 2016) (Method Details; Table S2).

Single Copies of Enhancers Drive Expression Influenced by Position Relative to Promoter

To determine the input-output functions for combinations of two enhancers—in other words, the expression driven by two enhancers together as a function of their individual activities—we must first accurately measure their individual activities. Our single-enhancer constructs control for quantitative variations in expression at different positions of the enhancer relative to the promoter, allowing us to subsequently focus on signal integration itself.

Krüppel expression driven by the pair of shadow enhancers (the CD1-CD2 construct) increases rapidly during nuclear cycle (nc)14 to a peak at ~18 min from the start of the cycle, before declining again (Figure 1E). The level of expression is set primarily by CD1, while CD2 turns on slightly later and more gradually to a lower peak level that, unlike CD1 expression, is maintained through the remainder of nc14 (Figure 1E). CD2 also drives a narrower pattern than CD1 (Figure 2G) and represses the activity of the distal CD1 enhancer in the posterior of the pattern (El-Sherif and Levine, 2016). We report the mean expression per actively transcribing nucleus (Figures 2A and 2E) and the fraction of nuclei that are transcribing (Figures 2B and 2F); together, these dictate the overall mean expression level driven by each enhancer. We note that the relative level of expression driven by CD1 differs between our results and a previous study, which may be due to differences in the sequence definitions of the enhancers and/or promoters between our two studies (see Discussion).

The expression driven by both *Krüppel* shadow enhancers is affected by their distance from the promoter, a property previously observed of other enhancers (Fukaya et al., 2016; Quintero-Cadena and Sternberg, 2016). Moving CD1 next to the promoter from its endogenous distance ~2.8 kb upstream reduces expression in the anterior of the *Krüppel* pattern while increasing it slightly in the posterior (Figures 2A, 2C, and 2D). This difference in behavior is particularly pronounced in the middle of nc14 and declines toward the end (Figure 2A). When CD2 is moved away from its endogenous position adjacent to the promoter, expression turns on later and to a lower level across the entire pattern (Figures 2E, 2F, and 2H). Given this lack of predictability in the effect of enhancer-promoter distance on gene expression, we carefully controlled for this variable in all of our subsequent experiments.

Signal Integration by *Krüppel* Enhancer Duplications and Shadow Enhancers Varies over the Anterior-Posterior Axis

We next examined how the expression driven by each pair compares to the sum of expression driven by the appropriate distance-controlled single enhancers (Figure 3). We used additive expression as the baseline for comparison since this is the predicted outcome if two enhancers independently activate transcription without competing with each other for the promoter (Bothma et al., 2015). Each combination of enhancers drives sub-additive expression (Figures 3A–3C).

To directly compare the signal integration performed by shadow enhancers and enhancer duplications, we plotted the expression level from each pair of enhancers (the ‘‘output’’ expression, on the y axis) against the sum of the expression driven by the relevant distance-controlled single enhancers (the ‘‘input’’ expression, on the x axis; Figure 3D). In the center of the *Krüppel* pattern, these pairs of enhancers follow a strikingly similar trend with respect to additivity over a wide range of expression levels (Figure 3D). Consistent with the predictions of the competition model, when the enhancers drive low expression levels, the output expression is close to additivity (exemplified by the duplication of CD2), while when they drive high levels, the output is sub-additive (exemplified by the CD1 duplication).

However, at the borders of the expression pattern, the input-output function differs between the anterior and posterior of the pattern and between pairs of enhancers (Figures 4A–4D). That is, for the same input level of expression (sum of expression from the two separate enhancers), their joint expression differs from additivity to different degrees. For example, CD1 and CD2 together drive expression that is close to additive in the anterior of the pattern (gray, Figure 4A), but less than additive in the posterior (black, Figure 4A). This behavior reflects a documented interaction between the two enhancers, in which CD2 represses the activity of CD1 in the posterior part of the pattern (El-Sherif and Levine, 2016).

The difference in the input-output function between the anterior and posterior of the *Krüppel* pattern also exists when the two enhancers involved are identical (Figures 4B and 4C). Consider the CD1 duplication: at a given level of input expression (on the x axis) it is possible to get different combined output expression levels (on the y axis). This trend is construct specific; the CD1 duplication is closer to additive in the posterior, whereas the shadow enhancer pair is closer to additive in the anterior (compare gray to gray in Figures

4A and 4B). Finally, the difference in the input-output function at the anterior and posterior borders of the pattern persists when the positions of the shadow enhancers are swapped in the reporter construct (compare Figures 4D to 4A). This variation in signal integration at the boundaries of the pattern likely results from different TFs controlling expression in the anterior versus posterior of the pattern (see Discussion).

We noticed that not only does the input-output function differ between the anterior and posterior of the *Krüppel* pattern, but in the case of the CD1 duplication, it also changes over time (Figure 4E). That is, for a given input expression level, the output level depends on the time in the nuclear cycle. Early in the cycle, expression from the CD1 duplication is approximately additive (light orange, Figure 4E), while later in the cycle it is much less than additive (dark orange). This effect was only apparent for the CD1 duplication and most pronounced in the middle to posterior of the pattern.

Repression of CD1 by CD2 Does Not Require CD2 to Be Adjacent to the Promoter

The CD2 enhancer refines the *Krüppel* expression pattern by inhibiting activity of the CD1 enhancer at the posterior boundary (Figure 5A; El-Sherif and Levine, 2016). We wondered whether the repressive activity of CD2 relies on either its promoter-proximal location or its positioning between CD1 and the promoter. To test this, we swapped the positions of the two enhancers and found that it does not: the CD2 enhancer is as effective at preventing activation by CD1 in the posterior of the *Krüppel* expression pattern from 2.4 kb upstream of the promoter and on the opposite side of CD1 from it (Figures 5B and 5C).

Regulatory Sequence Arrangement Influences Signal Integration by Two Enhancers

Finally, we asked whether the signal integration performed by the shadow enhancer pair depends only on the sequences of the enhancers, or whether the position of the enhancers relative to each other also influences how they interact. We swapped the locations of the pair of shadow enhancers and compared the input-output function between the endogenous and swapped arrangements. This analysis was possible because we measured the output of individual enhancers controlled for their distance from the promoter. These single enhancer controls enable us to separate the effect of individual enhancer location on expression from the effect of enhancer arrangement on the signal integration itself. The regulatory sequences in the endogenous and swapped constructs are essentially identical (Figure 1C; see Method Details). If the signal integration they perform is solely a function of the sequence of each enhancer, then we would expect their activities to combine in the same way, regardless of their arrangement. However, switching their positions affects the input-output function, as was evident when we compared CD1-CD2 and CD2*-CD1* constructs (Figure 5E). These data indicate that the organization of the gene locus influences not only the level and timing of expression driven by each enhancer alone (Figures 2A and 2E) but also how the two enhancers work together to determine the overall level of expression.

DISCUSSION

We quantified nascent transcription in developing *Drosophila* embryos to study how a promoter combines information from multiple concurrently active enhancers. Duplicating

strongly expressing and weakly expressing shadow enhancers of the *Krüppel* gene enabled us to test the idea that enhancers compete for access to the promoter. It also allowed us to investigate whether there are differences in how inputs from multiple enhancers are integrated, depending on whether they contain the same or different combinations of TF binding sites. Our data are consistent with the enhancers of *Krüppel* competing to activate the promoter one at a time, but they also indicate that the complement of TFs bound to the enhancers in different parts of the embryo and the arrangement of enhancers relative to one another influence the signal integration that the enhancers perform.

Sequence Context Matters in Quantitative Measurement of Gene Expression

Distance between Enhancer and Promoter Affects Expression—We found that the distance between enhancer and promoter quantitatively affects the expression level and the timing of *Krüppel* expression (Figure 2), as has been noted in previous work on other enhancers (Fukaya et al., 2016; Quintero-Cadena and Sternberg, 2016). This effect differs between the two enhancers we studied, and it is non-uniform across the anterior-posterior axis of the embryo. For instance, CD1 drives lower expression in the anterior of the pattern when it is positioned next to the promoter. This implies that *Krüppel* is more sensitive to the repressors that position its anterior boundary, Giant and Hunchback (Jaeger, 2011), when those repressors are bound close to the promoter. The unpredictable effect of moving an enhancer with respect to the promoter emphasized the need to carefully control for this variable when we subsequently measured signal integration by pairs of enhancers.

It is common for studies to use reporter constructs in which enhancers are placed immediately adjacent to the promoter or to delete enhancers from bacterial artificial chromosomes without replacing the deleted sequence, thereby altering the spacing between regulatory elements. Our results demonstrate the importance of controlling for the endogenous spacing between enhancers and promoters when making quantitative expression measurements. When introducing spacers, we used computationally designed synthetic sequences that lack predicted binding sites for TFs active in the early embryo (Estrada et al., 2016). We prefer this synthetic approach to using sequences from bacteria or phage, which is common, because even exogenous sequences contain numerous predicted binding sites for TFs that regulate blastoderm gene expression (Figure S2). However, it remains possible that even this ‘‘neutral sequence’’ has an effect on expression. In future work, it would be useful to test many different spacers to control for this possibility.

Differences in Relative Expression Level Compared to Previous Studies May Stem from Differences in Enhancer and Promoter Definitions—In our constructs, CD1 and CD2 drive different relative levels of expression compared to constructs reported by two previous studies, including one from our own lab (El-Sherif and Levine, 2016; Wunderlich et al., 2015). Both studies reported lower peak expression from CD1 alone, a level that is similar to that driven by their CD2-only constructs.

The constructs used in both previous studies differed from those in the present work. Wunderlich et al. (2015) used the minimal *even-skipped* promoter and did not control for the distance between CD1 and the promoter. Meanwhile, El-Sherif and Levine (2016) used

different sequence definitions of the enhancers and a minimal core *Krüppel* promoter; in addition, the CD1-only construct was not controlled for the distance between the CD1 enhancer and the promoter. When we compared the predicted binding sites across all of our constructs and those of El-Sherif and Levine (2016), we saw that their CD1-only construct alone does not contain a consensus motif for Zelda close to the *Krüppel* core promoter. Zelda is an activator of CD1, and this site is likely to be bound *in vivo* since its predicted location coincides with a strong Zelda ChIP sequencing peak (Harrison et al., 2011). We suspect that this may account for the weaker expression from their CD1-only construct compared to ours.

CD2 Represses CD1 through an Unknown Long-Range Mechanism

The CD2 enhancer represses CD1 activity in the posterior of the *Krüppel* pattern (El-Sherif and Levine, 2016; Figure 5A). We found that this repression does not depend on CD2 being next to the promoter or being positioned between CD1 and the promoter: CD2 represses CD1 just as effectively when the locations of the enhancer are swapped (Figures 5B and 5C). The posterior border of *Krüppel* expression is set by Giant and Knirps, both of which are short-range repressors (Arnosti et al., 1996; Strunk et al., 2001). Short-range repressors act over distances of <~100 bp to quench the activity of nearby activators or inhibit other enhancers by directly repressing the promoter from within this range (Arnosti et al., 1996; Hewitt et al., 1999). Even in its endogenous position, CD2 is ~200 bp away from the transcription start site, and when the positions of the enhancers are swapped, it is ~2.4 kb away, meaning that the short-range repressive action of Knirps or Giant cannot account for its inhibition of CD1. Repression between the *Krüppel* enhancers must therefore either be mediated by an unknown long-range repressor or depend on the previously undetected long-range repressive activity of Knirps and/or Giant.

The Effect of Enhancer Order Suggests a Role for Topology in Combinatorial Control at the Locus Level

The constructs containing the shadow enhancer pair in their endogenous and swapped locations comprise the same regulatory sequences, and yet the signal integration they perform to determine the overall gene expression output is different (Figure 5E). Topology affects gene expression at the genome scale, organized by DNA-binding looping factors whose position relative to one another is critical (Guo et al., 2015). Furthermore, gene expression is affected when topological interactions are forced through the introduction of looping factors (Deng et al., 2014; Bartman et al., 2016). We hypothesize that looping factors may also influence gene expression at the level of single loci. In our constructs, these looping factors could bind either in the enhancers themselves or in the spacer sequences that replaced them in the single-enhancer constructs, although we deliberately designed the spacers to lack binding sites for many regulatory proteins (see Method Details and Table S2). Predicted binding sites for multiple *Drosophila* looping proteins are present in the enhancers, the promoter, and the endogenous sequence between them (see Figure S3); investigating the role of these sites in the integration of enhancer activities at the locus level is an exciting direction for future work.

Signal Integration in the Center of the *Krüppel* Pattern Supports a Model of Enhancer Competition

In the center of the *Krüppel* pattern, the enhancer pairs combine additively when their expression is low and sub-additively when it is high (Figure 3D). This is consistent with high expression levels being the result of an increase in the fraction of time an enhancer spends activating the promoter (Bothma et al., 2015), which would lead to competition between the enhancers if only one can activate the promoter at a time.

The concentration of the repressors that set the boundaries of *Krüppel* expression is low in the center of the pattern, meaning that expression there is dictated primarily by activators. An interesting feature of the *Krüppel* shadow enhancers is that they are controlled by different activators: CD1 is sensitive to Zelda and Bicoid (Wunderlich et al., 2015; Hoch et al., 1991), while CD2 responds to Stat92E and (perhaps indirectly) to Hunchback (Wunderlich et al., 2015; Jaeger, 2011). We wondered whether the use of distinct activators by CD1 and CD2 affects how their activities combine. If this were the case, then we would expect the shadow enhancer pair to display a different input-output function to the two sets of duplicated enhancers. Instead, the input-output functions for the shadow enhancers and duplications are very similar over a wide range of expression levels in the center of the pattern (Figure 3D), suggesting that the activator's identities do not influence the input-output function in this case.

Explaining Differences in Signal Integration across the *Krüppel* Expression Pattern

The competition model as laid out by Bothma et al. (2015) cannot account for the combined output of two enhancers across the *Krüppel* expression domain as a whole. Specifically, having quantified expression across the entire *Krüppel* pattern, we observed significant differences in the input-output function at its anterior and posterior borders and over time, both within and between constructs (Figure 4). That is, there is not a single relation between the level of expression that the enhancers drive separately and together.

How might different output levels be generated from the same input levels of expression driven by two enhancers? In the model described by Bothma et al. (2015), high expression results from an increase in the fraction of time that an enhancer engages with the promoter; this leads to sub-additive behavior at high expression levels due to competition between enhancers that can only activate the promoter one at a time. Achieving different output expression levels for a given input level under this model would require that expression level also be altered by changing the rate of transcription initiation, not just the fraction of time that an enhancer activates the promoter (Figure S4 and STAR Methods). However, this runs counter to recent *Drosophila* studies that show that enhancers alter the frequency of transcriptional bursts (changing the fraction of time that the promoter is active), rather than burst amplitude (i.e., the transcription initiation rate during active periods) (Fukaya et al., 2016; Lammers et al., 2018; Zoller et al., 2018).

A possible explanation for the differences in input-output function that we observe across the pattern is that different TFs are present in each location, and these may differentially influence transcriptional kinetics based on their biochemical functions. This explanation is

supported by the behavior of the duplications, since the input-output function differs across the pattern, despite both copies of the enhancer presumably binding the same TFs in a given cell. The anterior and posterior borders of *Krüppel* are determined by different combinations of transcriptional repressors. The anterior boundary is set by Giant and Hunchback and the posterior by Giant and Knirps (Jaeger, 2011; Kraut and Levine, 1991; Harding and Levine, 1988; Jäckle et al., 1986). Whether CD1 and CD2 respond differently to these repressors has not been directly tested. However, the expression pattern of CD2 does not extend as far to the posterior as that driven by CD1, and so it seems likely that CD2 is sensitive to Knirps (whose expression overlaps the posterior of the pattern of *Krüppel*), while CD1 is not. TFs have diverse and context-dependent biochemical roles in transcription (Fuda et al., 2009; Duarte et al., 2016). In *Drosophila*, repressors can act on different steps in transcription when bound to shadow enhancers of the *sloppy-paired* gene (Hang and Gergen, 2017). A similar scenario may exist for *Krüppel*, with different combinations of repressors at the anterior and posterior of the pattern acting on distinct processes in transcription; the mechanism(s) by which each enhancer activates or represses transcription could in turn affect how the activities of two enhancers combine quantitatively.

Outlook

Given the complexity of enhancer function over space and time, dynamic models of transcription are required to decipher the underlying molecular mechanisms. Models are most helpful when they can be confronted with relevant experimental data. While great strides have been made in measuring the dynamic production of mRNA using live imaging, measuring the underlying steps of transcription outside of bulk biochemical assays remains highly challenging. Single-molecule approaches in bacteria have been very successful at linking models to mechanism through measurements of the underlying steps and their rates. Our results indicate that such measurements will also be transformative for animal systems, where the underlying biochemical functions and rates likely influence not only how individual enhancers operate but also how they collaborate with one another.

STAR★METHODS

CONTACT FOR REAGENT AND RESOURCE SHARING

Please direct requests for information and resources to the Lead Contact, Angela H. DePace (angela_depace@hms.harvard.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Drosophila melanogaster transgenic lines (see Key Resources Table) were raised at 25°C. For live imaging of blastoderm-stage embryos, virgin females of line yw; His2Av-mRFP1; MCP-NoNLS-eGFP (Garcia et al., 2013) were crossed to males bearing each of our transgenic reporter constructs (transgenics were generated in Bloomington Stock BL8622 by BestGene Inc.).

METHOD DETAILS

Cloning and transgenesis—We constructed a reporter gene in the pBOY vector backbone (Hare et al., 2008) using isothermal assembly (Gibson et al., 2009), into which we

cloned the *Krüppel* (*Kr*) regulatory sequence out to ~4.5 kb upstream of the transcription start site. The reporter consisted of the *D. melanogaster Krüppel* core promoter and its surrounding sequence from the 3′ end of the CD2 enhancer to the beginning of the second exon of *Krüppel*; a 1.5 kb cassette encoding 24 MS2 stem loops (from Addgene 31865, pCR4–24XMS2SL-stable plasmid); 3 kb of the *lacZ* gene; and the alpha-tubulin 3′ UTR. We commercially synthesized the 4.5 kb of *Krüppel* regulatory sequence that contains the proximal and distal blastoderm enhancers (using GenScript's gene synthesis service) in order to insert unique restriction sites flanking each enhancer, enabling us to replace each sequence in a modular fashion. We then ligation-cloned this sequence into our pBOY construct, directly upstream of the *Krüppel* promoter. For subsequent manipulations of the two *Krüppel* enhancers, we used the sequence definitions from Perry et al. (2011); the sequences and genome coordinates for the *Krüppel* enhancers we used are listed in Table S1.

In order to maintain the endogenous spacing between enhancers and promoter in the single-enhancer constructs we replaced each with an equivalent length of non-regulatory sequence. We computationally designed these sequences be depleted of binding sites for transcription factors active in patterning the blastoderm embryo, sites for architectural binding proteins, and core promoter sequences (Table S1). To do so we used the online binding site removal tool, SiteOut (Estrada et al., 2016), beginning with a randomly-generated sequence of the appropriate GC content for *Drosophila* intergenic DNA (40.6%). SiteOut locates binding sites on the basis of their Position Weight Matrices using PATSER (http://genetics.wustl.edu/gslab) and a p value threshold that we set at It then removes sites iteratively using a Monte Carlo algorithm while maintaining the GC content specified. We commercially synthesized a length of binding site-depleted sequence, and from it PCR-amplified unique sections of the appropriate lengths to replace the CD1 and CD2 enhancers (1160bp and 1586bp, respectively; Table S1). These we cloned into our reporter plasmid in place of the enhancers using isothermal assembly (Gibson et al., 2009), which leaves scar-less junctions.

We sequence-verified the enhancers and promoter of all reporter constructs prior to injection, and checked the length of the MS2 cassette by restriction digest. The pBOY backbone contains an attB site for phiC31-mediated site-specific recombination (Fish et al., 2007) and a mini-white gene for transformant selection. For each construct, BestGene Inc. (Chino Hills, CA) injected midi-prepped DNA into 200 embryos of Bloomington Stock BL8622, which contains the attP2 landing site on chromosome 3L (Markstein et al., 2008). All constructs are integrated into this same attP2 landing site. After the constructs were successfully integrated into the fly genome, we prepared genomic DNA, PCR-amplified the transgene and repeated the sequencing and restriction digest verification of the reporters. Doing so revealed that a single construct, the CD1 duplication, may have four extra MS2 stem loops (giving it 56 rather than 48 copies of GFP if all stem loops are fully bound by MCP-GFP). This does not change the main conclusion that we draw from analysis of the CD1 duplication, which is that the input-output function for the two copies of CD1 is different in the anterior and the posterior of the expression pattern. Conservatively correcting for the possible extra loops simply shifts the data in Figure 4B down without changing their positions relative to one another (data not shown). Note that in the construct containing the shadow enhancers in their swapped location (CD2*-CD1*), the proximal enhancer (which is

endogenously located ~200bp upstream of the promoter) is at 2.4kb upstream, compared to 2.8 kb in the corresponding single-enhancer control.

Live imaging: embryo preparation and data acquisition—Virgin females of the line yw; His2Av-mRFP1; MCP-NoNLS-eGFP (Garcia et al., 2013) were crossed to males bearing each of our transgenic reporter constructs. At two hours after egg deposition, embryos were dechorionated in freshly-made 50% bleach for 1 minute and then mounted in halocarbon 27 oil between a semi-permeable membrane (Biofoile, *In Vitro* Systems and Services) and a coverslip (No. 1.5, 18 3 18 mm).

Live imaging was carried out at 20 s intervals for the first 50 minutes of nuclear cycle 14 (nc14, beginning ~2hr post-egg deposition). Each time-lapse series was registered to its position on the anterior-posterior axis by cross-correlation with a 20x image of the full embryo. Our field of view was narrower than the expression pattern of some of our constructs; we therefore report mean fluorescence data from multiple embryos for each transgenic line, with an n of at least 3 for each anterior-posterior bin.

Live imaging was carried out using a Yokagawa CSU-22 spinning disk confocal with Borealis modification (Spectral Applied Research) with a Hamamatsu ORCA-R2 cooled CCD camera, mounted on an inverted Nikon Ti microscope. The MCP-eGFP and Histone-mRFP1 were imaged using a 488nm solid state laser with 525/50 nm emission filter, and 561 nm solid state laser with 620/60 nm emission filter, respectively. Laser lines were selected using an AOTF, and a 405/488/568/647 multi-band pass dichromatic mirror (Chroma) was used. Time-lapsed images were acquired using a Nikon Plan Apo 60× 1.4 NA oil immersion objective. Low-magnification images used to register the movie on the anterior-posterior axis of the embryo were taken with a Nikon Plan Apo 20× 0.75 NA objective. Acquisition was controlled using MetaMorph software (Molecular Devices). To enable quantitative comparison of expression levels between embryos, we normalized illumination intensity at the start of each imaging session by measuring the power of 488 nm laser exiting the objective lens, using an EXFO X-Cite XR2100 power meter and adjusting the laser AOTF to reach a set target value. A 10mm z stack of 21 images spaced 0.5 um apart was taken every 20 s using a Prior NanoScanZ piezo-electric focus stage insert, an exposure time of 80 ms, and 2×2 camera binning, resulting in a pixel size of 212 nm. Flatfield images were taken under identical imaging conditions to those described above, using a slide of concentrated fluorescien sodium salt solution (Model and Blank, 2008). Using the full CCD chip with 60x magnification did not cover the full extent of the *Krüppel* pattern; we report mean fluorescence data from multiple embryos for each transgenic line, with at least 3-fold coverage across the pattern from different embryos.

QUANTIFICATION AND STATISTICAL ANALYSIS

Analysis of live imaging data—Imaging data were processed in MATLAB (Mathworks) as described in Garcia et al., 2013; the scripts are available on GitHub from the Garcia lab at UC Berkeley. Briefly, the histone-RFP image stacks were maximum-projected at each time point, and the nuclei segmented. MCP-GFP images were corrected for uneven laser illumination by subtracting the camera offset (a flat value of 200 added to every pixel) and

then dividing by an offset-subtracted, normalized flat-field image. Spots of transcription were located in each z-slice of each time point using a difference of Gaussians filter and associated in z to a given site of transcription (or 'particle'). Each particle was then associated with its closest respective nucleus; in cases where more than one particle is detected in the vicinity of a nucleus, the brightest particle alone was kept. A 2D Gaussian was fitted to the brightest z-spot in each particle to determine the offset, which was used as an estimate of the local background fluorescence. The fluorescence of a particle was calculated by integrating the fluorescence over a fixed area of 9 pixels in diameter centered on the 2D Gaussian in the image in the z stack containing the highest intensity value pixels in this region, and subtracting the estimated local background. This background fluorescence estimate dominates the imaging error (Garcia et al., 2013). Spots for which there was no peak in fluorescence in the z axis in a given time point (i.e., where part of the spot was cut off) were discarded.

A single 20x image of each embryo was generated by automatically stitching together two 512 3 675 pixel images (pixel size 600 nm) of the anterior and posterior of the embryo in MATLAB (Figures S1A and S1B). The anterior, posterior, dorsal and ventral poles of the embryo were manually assigned and the full embryo image used to align the 60x imaging region by cross-correlation using the histone-RFP channel. Each particle of fluorescence was then assigned its anterior-posterior and dorso-ventral coordinates and parsed into anterior-posterior bins of 2.5% embryo length (Figure S1C).

Expression data for a given transgenic construct are reported as the mean fluorescence per actively transcribing nucleus, across multiple embryos, in the anterior-posterior bin in question. A nucleus was considered actively transcribing at a given time point if a spot of GFP fluorescence was detected in it (see above). Error bars represent standard error of the mean across multiple embryos. Number of embryos across which data were averaged can be found in the figure legends. Where multiple panels of a figure show different types of data for the same construct, the number of embryos used is reported only for the first panel in which that construct appears.

The borders of the *Krüppel* expression pattern were calculated by fitting a Gaussian to the expression pattern (the mean fluorescence in 'on' nuclei) for a given construct at each time point (Figure S1E). The anterior (posterior) border was determined to be the position of the Gaussian's peak minus (plus) the half width of the Gaussian at half max (HWHM).

The initial transcription rate was calculated for individual particle traces by finding the maximum derivative of the time series prior to the initial peak in its expression in nc14 (Figure S1D). The mean transcription rate for a construct was determined by averaging that for all the particles in a given anterior-posterior bin across all embryos of that construct.

Exploring input-output relationships under a model of enhancer competition —Previous work laid out a scheme for two enhancers activating a promoter (Bothma et al., 2015), based on a two-state model of transcription. We first briefly summarize this model and then discuss the input-output functions that can arise from two enhancers independently activating the same promoter when the underlying parameters are systematically varied.

Model for two enhancers acting on a promoter—Let's consider first a promoter that is activated by a single enhancer. The enhancer engages with and disengages from the promoter with rate constants k_{on} and k_{off}, respectively (Figure S4A) and the ratio of these rates determines the ''strength'' of interaction between enhancer and promoter (Bothma et al., 2015). When the enhancer and promoter are engaged, transcription initiates at a constant rate, r. The promoter therefore exists in one of two states: unoccupied, which we denote as *OFF*; and enhancer-engaged, which we denote as *ON*.

The overall rate of mRNA production is a function of the probability that the enhancer is engaged with the promoter (i.e., the probability that the promoter is in the *ON* state, P(*ON*)) and the rate of initiation from the engaged state (r). The mRNA production rate is therefore given by

$$\frac{d(mRNA)}{dt} = r * P(ON).$$

Since the promoter can only be in the *OFF* or *ON* states, the occupancy of the states must sum to 1. Given this, and assuming that the temporal evolution of the state occupancies is at steady state, the rate of mRNA production can be written as

$$\frac{d(mRNA)}{dt} = r * \frac{k_{on}}{k_{on} + k_{off}}$$

(see Bothma et al., 2015 for fuller explanation). To simplify the above equation, we can plug in the association constant, $K_a = k_{on}/k_{off}$, which gives

$$\frac{d(mRNA)}{dt} = r * \frac{K_a}{1 + K_a}.$$

Now, if two enhancers can independently activate the promoter, we need to calculate the occupancy of the states when the promoter is engaged with Enhancer 1 (ON_1) and that when it is engaged with Enhancer 2 (ON_2). By calculating the temporal evolution of the occupancy of states ON_1 and ON_2, the rate of mRNA production from the two enhancers in the same construct (E1,2), acting independently, is

$$\frac{d(mRNA(E1,2))}{dt} = r_1 * P(ON_1) + r_2 * P(ON_2),$$

where r_1 and r_2 are the transcription initiation rates from states ON_1 and ON_2, respectively. As was the case for the single-enhancer example, we can now apply the constraint that the occupancy of the three states (ON_1, ON_2 and *OFF*) must sum to 1, and assume that the temporal evolution of the occupancies of these states is at steady state. By doing so, we can calculate the rate of mRNA production due to Enhancer 1 and that due to Enhancer 2 (see Bothma et al., 2015), and substitute these results into the above equation to write it as

$$\frac{d(mRNA(E1,2))}{dt} = \frac{r_1 k^1_{on} k^2_{off} + r_2 k^2_{on} k^1_{off}}{k^1_{on} k^2_{off} + k^2_{on} k^1_{off} + k^1_{off} k^2_{off}}.$$

Recognizing that $K^1_a = k^1_{on}/k^1_{off}$ and similarly $K^2_a = k^2_{on}/k^2_{off}$, the above equation can be rewritten in simplified form,

$$\frac{d(mRNA(E1,2))}{dt} = \frac{r_1 K^1_a + r_2 K^2_a}{1 + K^1_a + K^2_a}.$$

Similarly, the rate of mRNA production from each enhancer on its own can be written as follows, for Enhancer 1,

$$\frac{d(mRNA(E1))}{dt} = \frac{r_1 K^1_a}{1 + K^1_a},$$

and for Enhancer 2,

$$\frac{d(mRNA(E2))}{dt} = \frac{r_2 K^2_a}{1 + K^2_a}.$$

Exploring parameter space—We showed that the input-output function varies across the anterior-posterior axis for a given two-enhancer construct, and at the same anterior-posterior position between constructs (Figures 4 and 5). That is, for the same summed level of 'input' expression from two enhancers, their combined 'output' depends both on the enhancer pair and on the position along the anterior-posterior axis. How might this difference in input-output function arise under the model outlined above?

Under this model, the rate of transcription driven by a single enhancer could vary as a result of changes in association between enhancer and promoter (K_a) and/or the rate of initiation from the promoter *ON* state (r). Because we are interested in the signal integration performed by two enhancers acting on the same promoter, we will specifically look at the relationship between the rate of transcription driven by both enhancers together in the same construct compared to that from the two enhancers in different constructs. To do so, we vary the rates K^1_a, K^2_a, r_1 and r_2 over three orders of magnitude (between 0.1 – 10 a.u.), and in Figure S3 we plot $d(mRNA(E1,2))/dt$ against $d(mRNA(E1))/dt + d(mRNA(E2))/dt$, as we did in Figures 3D and 4 and 5D in the main text.

To explore this parameter space in a systematic fashion, we reasoned that for duplicated enhancers which bind the same transcription factors, K^1_a and K^2_a (the association constants for Enhancer 1 and Enhancer 2, which in this case are identical) may vary in proportion to one

another as the concentrations of their input transcription factors change over space and time in the embryo (and ditto for r_1 and r_2). On the other hand, for shadow enhancers, which respond to different activating and repressing transcription factors, K_a^1 and K_a^2 may vary independently of one another (likewise for r_1 and r_2). Therefore, in Figure S4 we specified that the values for K_a and/or r for each enhancer either:

a. remain constant;

b. *vary proportionally* to one another (i.e., $\boldsymbol{r_1}$ varies in proportion to $\boldsymbol{r_2}$; K_a^1 varies in proportion to K_a^2);

c. *vary independently* of one another (i.e., $\boldsymbol{r_1}$ varies independent of $\boldsymbol{r_2}$; K_a^1 varies independent of K_a^2).

The parameters used to create the plots in Figure S4 are listed in Table S3.

We are interested in the conditions required to deviate from a single input-output relationship between the level of expression driven by two enhancers separately and that driven by both together in the same construct. Deviation from a single input-output relationship can be seen in Figure S3B, panels iv-viii. Figure S4B shows that there are two broad ways to deviate from a single input-output relationship under this model:

a. if the association constants for the two enhancers (K_a's) vary in proportion – as might be expected from a pair of duplicated enhancers, which respond to the same transcription factors – then the rates of transcription initiation (r's) must also vary (compare iv and v to iii).

b. the rate of transcription initiation (r) can remain constant, but the association constants for the two enhancers (K_a's) would have to vary independently to achieve different output expression (on the y axis) for a given input level (on the x axis). It seems somewhat unlikely that the association constants of duplicated enhancers would vary completely independently.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This work was generously supported by a National Science Foundation (NSF) Career Award (IOS-1452557, to A.H.D.), NIH grant 4U01GM103804–04 (to A.H.D.), the Giovanni Armenise-Harvard Foundation (to A.H.D.), and a Harvard Graduate Society Merit Fellowship (to C.S.). We would like to thank Hernan Garcia for many helpful discussions, for teaching us MS2 imaging in fly embryos and for sharing the computational pipeline used to process the imaging data. We thank Armando Reimer and Jacques Bothma for advice on imaging and image processing. We carried out all of the imaging at the Nikon Imaging Center at Harvard Medical School, where Anna Jost and Jennifer Waters provided invaluable guidance on microscopy. We thank members of the DePace lab for their thoughtful input throughout this work.

REFERENCES

Arnosti DN, Gray S, Barolo S, Zhou J, and Levine M (1996). The gap protein knirps mediates both quenching and direct repression in the Drosophila embryo. EMBO J 15, 3659–3666. [PubMed: 8670869]

Barolo S (2012). Shadow enhancers: frequently asked questions about distributed cis-regulatory information and enhancer redundancy. BioEssays 34, 135–141. [PubMed: 22083793]

Bartman CR, Hsu SC, Hsiung CC, Raj A, and Blobel GA (2016). Enhancer Regulation of Transcriptional Bursting Parameters Revealed by Forced Chromatin Looping. Mol. Cell 62, 237–247. [PubMed: 27067601]

Bothma JP, Garcia HG, Ng S, Perry MW, Gregor T, and Levine M (2015). Enhancer additivity and non-additivity are determined by enhancer strength in the Drosophila embryo. eLife 4, e07956.

Cannavò E, Khoueiry P, Garfield DA, Geeleher P, Zichner T, Gustafson EH, Ciglar L, Korbel JO, and Furlong EEM (2016). Shadow Enhancers Are Pervasive Features of Developmental Regulatory Networks. Curr. Biol 26, 38–51. [PubMed: 26687625]

Chen H, Levo M, Barinov L, Fujioka M, Jaynes JB, and Gregor T (2018). Dynamic interplay between enhancer-promoter topology and gene activity. Nat. Genet 50, 1296–1303. [PubMed: 30038397]

Deng W, Rupon JW, Krivega I, Breda L, Motta I, Jahn KS, Reik A, Gregory PD, Rivella S, Dean A, and Blobel GA (2014). Reactivation of developmentally silenced globin genes by forced chromatin looping. Cell 158, 849–860. [PubMed: 25126789]

Duarte FM, Fuda NJ, Mahat DB, Core LJ, Guertin MJ, and Lis JT (2016). Transcription factors GAF and HSF act at distinct regulatory steps to modulate stress-induced gene activation. Genes Dev 30, 1731–1746. [PubMed: 27492368]

Dukler N, Gulko B, Huang Y-F, and Siepel A (2016). Is a super-enhancer greater than the sum of its parts? Nat. Genet 49, 2–3. [PubMed: 28029159]

Dunipace L, Ozdemir A, and Stathopoulos A (2011). Complex interactions between cis-regulatory modules in native conformation are critical for Drosophila snail expression. Development 138, 4075–4084. [PubMed: 21813571]

El-Sherif E, and Levine M (2016). Shadow Enhancers Mediate Dynamic Shifts of Gap Gene Expression in the Drosophila Embryo. Curr. Biol 26, 1164–1169. [PubMed: 27112292]

Estrada J, Ruiz-Herrero T, Scholes C, Wunderlich Z, and DePace AH (2016). SiteOut: An Online Tool to Design Binding Site-Free DNA Sequences. PLoS One 11, e0151740. [PubMed: 26987123]

Fish MP, Groth AC, Calos MP, and Nusse R (2007). Creating transgenic Drosophila by microinjecting the site-specific phiC31 integrase mRNA and a transgene-containing donor plasmid. Nat. Protoc 2, 2325–2331. [PubMed: 17947973]

Frankel N (2012). Multiple layers of complexity in cis-regulatory regions of developmental genes. Dev. Dyn 241, 1857–1866. [PubMed: 22972751]

Frankel N, Davis GK, Vargas D, Wang S, Payre F, and Stern DL (2010). Phenotypic robustness conferred by apparently redundant transcriptional enhancers. Nature 466, 490–493. [PubMed: 20512118]

Fuda NJ, Ardehali MB, and Lis JT (2009). Defining mechanisms that regulate RNA polymerase II transcription in vivo. Nature 461, 186–192. [PubMed: 19741698]

Fukaya T, Lim B, and Levine M (2016). Enhancer Control of Transcriptional Bursting. Cell 166, 358–368. [PubMed: 27293191]

Garcia HG, Tikhonov M, Lin A, and Gregor T (2013). Quantitative imaging of transcription in living Drosophila embryos links polymerase activity to patterning. Curr. Biol 23, 2140–2145. [PubMed: 24139738]

Ghavi-Helm Y, Klein FA, Pakozdi T, Ciglar L, Noordermeer D, Huber W, and Furlong EEM (2014). Enhancer loops appear stable during development and are associated with paused polymerase. Nature 512, 96–100. [PubMed: 25043061]

Ghiasvand NM, Rudolph DD, Mashayekhi M, Brzezinski JA 4th, Goldman D, and Glaser T (2011). Deletion of a remote enhancer near ATOH7 disrupts retinal neurogenesis, causing NCRNA disease. Nat. Neurosci 14, 578–586. [PubMed: 21441919]

Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA 3rd, and Smith HO (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods 6, 343–345. [PubMed: 19363495]

Guo Y, Xu Q, Canzio D, Shou J, Li J, Gorkin DU, Jung I, Wu H, Zhai Y, Tang Y, et al. (2015). CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function. Cell 162, 900–910. [PubMed: 26276636]

Hang S, and Gergen JP (2017). Different modes of enhancer-specific regulation by Runt and Even-skipped during Drosophila segmentation. Mol. Biol. Cell 28, 681–691. [PubMed: 28077616]

Harding K, and Levine M (1988). Gap genes define the limits of antennapedia and bithorax gene expression during early development in Drosophila. EMBO J 7, 205–214. [PubMed: 2896123]

Hare EE, Peterson BK, Iyer VN, Meier R, and Eisen MB (2008). Sepsid even-skipped enhancers are functionally conserved in Drosophila despite lack of sequence conservation. PLoS Genet 4, e1000106. [PubMed: 18584029]

Harrison MM, Li X-Y, Kaplan T, Botchan MR, and Eisen MB (2011). Zelda Binding in the Early Drosophila Melanogaster Embryo Marks Regions Subsequently Activated at the Maternal-to-Zygotic Transition. PLoS Genet 7, e1002266. [PubMed: 22028662]

Hay D, Hughes JR, Babbs C, Davies JOJ, Graham BJ, Hanssen L, Kassouf MT, Marieke Oudelaar AM, Sharpe JA, Suciu MC, et al. (2016). Genetic dissection of the a-globin super-enhancer in vivo. Nat. Genet 48, 895–903. [PubMed: 27376235]

Hertz GZ, and Stormo GD (1999). Identifying DNA and protein patterns with statistically significant alignments of multiple sequences. Bioinformatics 15, 563–577. [PubMed: 10487864]

Hewitt GF, Strunk BS, Margulies C, Priputin T, Wang XD, Amey R, Pabst BA, Kosman D, Reinitz J, and Arnosti DN (1999). Transcriptional repression by the Drosophila giant protein: cis element positioning provides an alternative means of interpreting an effector gradient. Development 126, 1201–1210. [PubMed: 10021339]

Hoch M, Schröder C, Seifert E, and Jäckle H (1990). cis-acting control elements for Krüppel expression in the Drosophila embryo. EMBO J 9, 2587–2595. [PubMed: 2114978]

Hoch M, Seifert E, and Jäckle H (1991). Gene expression mediated by cisacting sequences of the Krüppel gene in response to the Drosophila morphogens bicoid and hunchback. EMBO J 10, 2267–2278. [PubMed: 2065664]

Hong J-W, Hendrix DA, and Levine MS (2008). Shadow enhancers as a source of evolutionary novelty. Science 321, 1314. [PubMed: 18772429]

Jäckle H, Tautz D, Schuh R, Seifert E, and Lehmann R (1986). CrossRegulatory Interactions among the Gap Genes of Drosophila. Nature 324, 668–670.

Jaeger J (2011). The gap gene network. Cell. Mol. Life Sci 68, 243–274. [PubMed: 20927566]

Jeong Y, El-Jaick K, Roessler E, Muenke M, and Epstein DJ (2006). A functional screen for sonic hedgehog regulatory elements across a 1 Mb interval identifies long-range ventral forebrain enhancers. Development 133, 761–772. [PubMed: 16407397]

Kraut R, and Levine M (1991). Mutually repressive interactions between the gap genes giant and Krüppel define middle body regions of the Drosophila embryo. Development 111, 611–621. [PubMed: 1893878]

Kvon EZ, Kazmar T, Stampfel G, Yáñez-Cuna JO, Pagani M, Schernhuber K, Dickson BJ, and Stark A (2014). Genome-scale functional characterization of Drosophila developmental enhancers in vivo. Nature 512, 91–95. [PubMed: 24896182]

Lam DD, de Souza FSJ, Nasif S, Yamashita M, López-Leal R, OteroCorchon V, Meece K, Sampath H, Mercer AJ, Wardlaw SL, et al. (2015). Partially redundant enhancers cooperatively maintain Mammalian pomc expression above a critical functional threshold. PLoS Genet 11, e1004935. [PubMed: 25671638]

Lammers NC, Galstyan V, Reimer A, Medin SA, Wiggins CH, and Garcia HG (2018). Binary Transcriptional Control of Pattern Formation in Development. bioRxiv 10.1101/335919.

Li X-Y, MacArthur S, Bourgon R, Nix D, Pollard DA, Iyer VN, Hechmer A, Simirenko L, Stapleton M, Luengo Hendriks CL, et al. (2008). Transcription factors bind thousands of active and inactive regions in the Drosophila blastoderm. PLoS Biol 6, e27. [PubMed: 18271625]

Lucas T, Ferraro T, Roelens B, De Las Heras Chanes J, Walczak AM, Coppey M, and Dostatni N (2013). Live imaging of bicoid-dependent transcription in Drosophila embryos. Curr. Biol 23, 2135–2139. [PubMed: 24139736]

Markstein M, Pitsouli C, Villalta C, Celniker SE, and Perrimon N (2008). Exploiting position effects and the gypsy retrovirus insulator to engineer precisely expressed transgenes. Nat. Genet 40, 476–483. [PubMed: 18311141]

McBride DJ, Buckle A, van Heyningen V, and Kleinjan DA (2011). DNaseI hypersensitivity and ultraconservation reveal novel, interdependent longrange enhancers at the complex Pax6 cis-regulatory region. PLoS One 6, e28616. [PubMed: 22220192]

Model MA, and Blank JL (2008). Concentrated dyes as a source of two-dimensional fluorescent field for characterization of a confocal microscope. J. Microsc 229, 12–16. [PubMed: 18173639]

Osterwalder M, Barozzi I, Tissières V, Fukuda-Yuzawa Y, Mannion BJ, Afzal SY, Lee EA, Zhu Y, Plajzer-Frick I, Pickle CS, et al. (2018). Enhancer redundancy provides phenotypic robustness in mammalian development. Nature 554, 239–243. [PubMed: 29420474]

Oudelaar AM, Davies JOJ, Hanssen LLP, Telenius JM, Schwessinger R, Liu Y, Brown JM, Downes DJ, Chiariello AM, Bianco S, et al. (2018). Single-Cell Chromatin Interactions Reveal Regulatory Hubs in Dynamic Compartmentalized Domains. bioRxiv 10.1101/307405.

Perry MW, Boettiger AN, Bothma JP, and Levine M (2010). Shadow enhancers foster robustness of Drosophila gastrulation. Curr. Biol 20, 1562–1567. [PubMed: 20797865]

Perry MW, Boettiger AN, and Levine M (2011). Multiple enhancers ensure precision of gap gene-expression patterns in the Drosophila embryo. Proc. Natl. Acad. Sci. USA 108, 13570–13575. [PubMed: 21825127]

Prazak L, Fujioka M, and Gergen JP (2010). Non-additive interactions involving two distinct elements mediate sloppy-paired regulation by pair-rule transcription factors. Dev. Biol 344, 1048–1059. [PubMed: 20435028]

Quintero-Cadena P, and Sternberg PW (2016). Enhancer Sharing Promotes Neighborhoods of Transcriptional Regulation Across Eukaryotes. G3 (Bethesda) 6, 4167–4174. [PubMed: 27799341]

Spitz F (2016). Gene regulation at a distance: from remote enhancers to 3D regulatory ensembles. Semin. Cell Dev. Biol 57, 57–67. [PubMed: 27364700]

Strunk B, Struffi P, Wright K, Pabst B, Thomas J, Qin L, and Arnosti DN (2001). Role of CtBP in transcriptional repression by the Drosophila giant protein. Dev Biol 239, 229–240. [PubMed: 11784031]

Wunderlich Z, Bragdon MDJ, Vincent BJ, White JA, Estrada J, and DePace AH (2015). *Krüppel* Expression Levels Are Maintained through Compensatory Evolution of Shadow Enhancers. Cell Rep 12, 1740–1747. [PubMed: 26344774]

Zeitlinger J, Zinzen RP, Stark A, Kellis M, Zhang H, Young RA, and Levine M (2007). Whole-genome ChIP-chip analysis of Dorsal, Twist, and Snail suggests integration of diverse patterning processes in the Drosophila embryo. Genes Dev 21, 385–390. [PubMed: 17322397]

Zoller B, Little SC, and Gregor T (2018). Diverse Spatial Expression Patterns Emerge from Unified Kinetics of Transcriptional Bursting. Cell 175, 835–847.e25. [PubMed: 30340044]

Highlights

- Measured dynamic expression from *Krüppel* shadow enhancers and their duplicates
- The input-output function of the pairs depends on the position in the embryo
- The input-output function of the shadow pair depends on the enhancers' relative positions
- Competition between enhancers to activate the promoter cannot fully explain these data

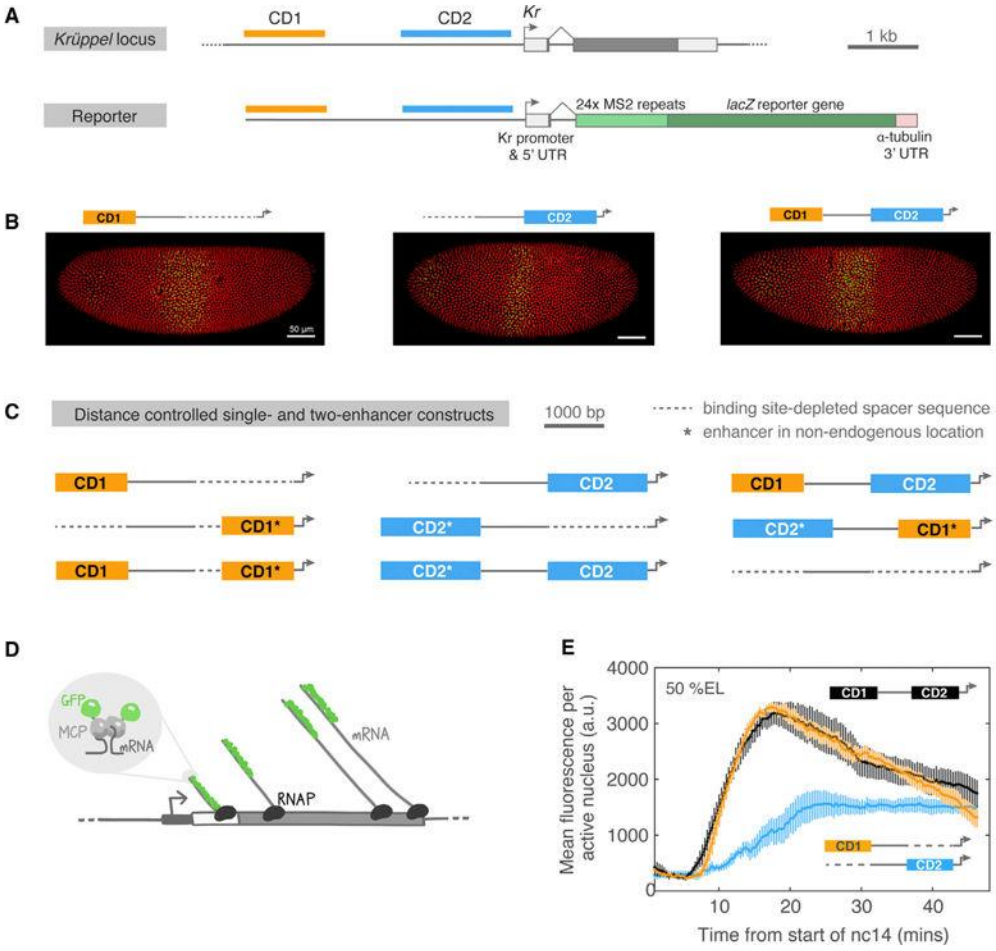


Figure 1. Constructs for Live Imaging of Nascent Transcription

(A) *Krüppel* (*Kr*) locus showing the two blastoderm enhancers, CD1 (orange) and CD2 (blue). The *Krüppel* reporter used here drives expression of an MS2 cassette followed by a *lacZ* reporter gene and α-tubulin 3′ UTR, under the control of the endogenous *Krüppel* promoter and 5′ UTR.

(B) The blastoderm enhancers of *Krüppel* drive overlapping expression patterns in the central region of the embryo during nuclear cycle 14 (nc14). Representative maximum projection images showing nascent transcription (MS2 bacteriophage coat protein [MCP]-GFP; green) and nuclei (histone-red fluorescent protein [RFP]; red) in mid-nc14 (anterior left, dorsal up).

(C) Reporter constructs used to investigate the behavior of enhancer duplications. * indicates an enhancer is at its non-endogenous location relative to the promoter. Dotted lines represent synthetic ‘‘neutral’’ sequences, computationally designed to be depleted for binding sites of transcription factors active in patterning the blastoderm embryo. These sequences were used to maintain the endogenous spacing of the enhancers and promoter in the reporter constructs.

(D) Live imaging of nascent transcription using the MS2 system. A sequence encoding 243 MS2 repeats was incorporated into the 5′ region of the reporter gene. When it is transcribed, the sequence forms RNA stem loops that are bound by a constitutively expressed MCP-GFP

fusion protein. As a train of polymerases transcribes through the reporter, GFP builds up at the locus and nascent transcripts become visible. Once a transcript dissociates from the gene and diffuses, it is no longer discernible.

(E) *Krüppel* shadow enhancers drive differing levels and dynamics of expression over time in nuclear cycle 14 (CD1, orange; CD2, blue; CD1-CD2, black). Plot shows mean fluorescence per actively transcribing nucleus in a single anterior-posterior bin of 2.5% embryo length (EL) covering between 50% and 52.5% EL (see also Figure S1). Error bars are SEMs across multiple embryos. Each time point is separated by 20 s. CD1, n = 5 embryos; CD2, n = 6; CD1-CD2, n = 6.

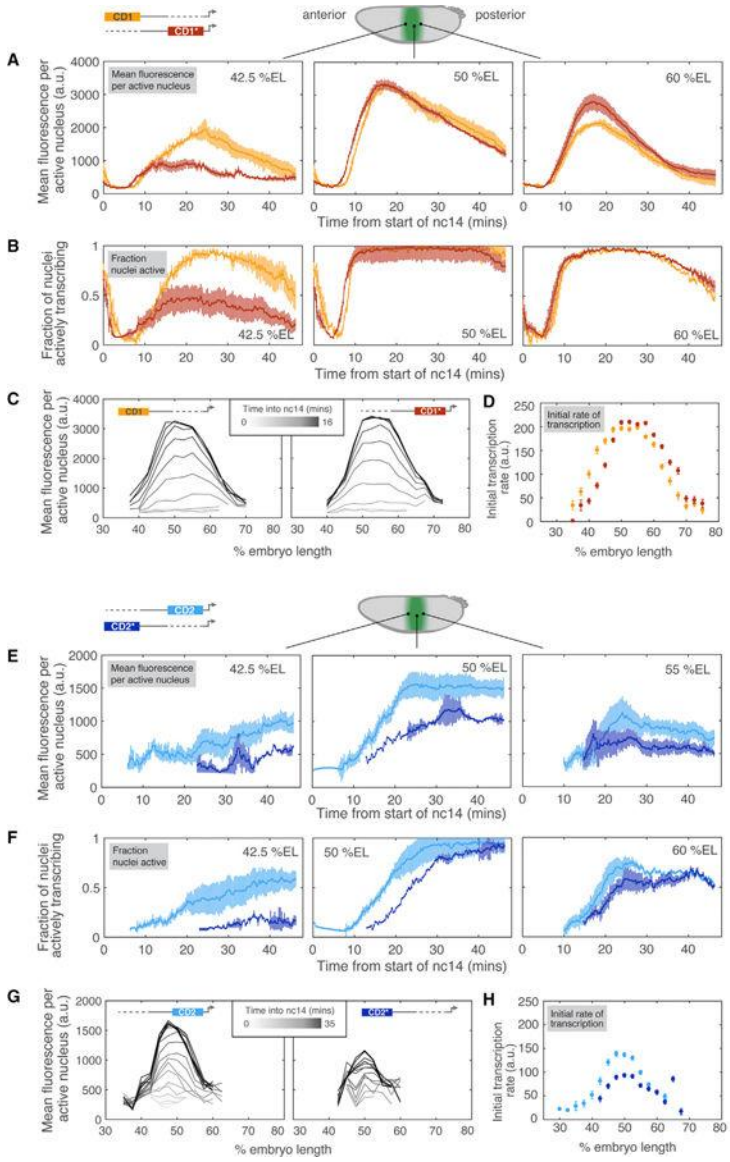


Figure 2. Single Copies of Enhancers Drive Expression Influenced by Position Relative to Promoter

(A) Comparing mean expression over time in nuclear cycle 14 (nc14) driven by the CD1 enhancer at its endogenous position (CD1, orange) and at the promoter (CD1*, red). Shown is mean fluorescence per actively transcribing nucleus (a.u.) within bins in the anterior, middle, and posterior of the expression pattern. Each bin is 2.5% embryo length (EL) in width. Error bars show the SEMs, and data points are 20 s apart. CD1, n = 3 embryos; CD1*, n = 6.

(B) Fraction of total nuclei in the indicated anterior-posterior bin that are actively driving transcription, over time, in nc14. Nuclei are counted as active at a given time point if a spot of MS2 fluorescence is detected in them. The number of nuclei in a given bin in nc14 is ~30. Error bars show the SEMs across multiple embryos.

(C) Evolution of the expression pattern driven by CD1 (left) and CD1* (right) in the first 16 min of nc14. Line traces show the mean fluorescence in active nuclei across the pattern and show time points ~100 s apart.

(D) Initial transcription rate across the patterns driven by CD1 and CD1*, estimated by finding the maximum derivative of the initial rise in fluorescence for each transcriptionally active nucleus; data points show means across all of the active nuclei across a number of embryos for each anterior-posterior bin, ± SE.

(E) Mean fluorescence per active nucleus (a.u.) driven by the CD2 enhancer at its endogenous position (blue) and upstream of the promoter (CD2*, indigo) across nc14. CD2, n = 6 embryos; CD2*, n = 4.

(F) Fraction of actively transcribing nuclei driven by CD2 and CD2* shown over time in nc14.

(G) Evolution of the expression pattern driven by CD2 (left) and CD2* (right) during the first 35 min of nc14. The window of time shown is longer than in the equivalent plots in (C) because the CD2 enhancer takes longer to start driving expression and to reach its peak level.

(H) Initial transcription rate across the expression patterns driven by CD2 and CD2* (see D).

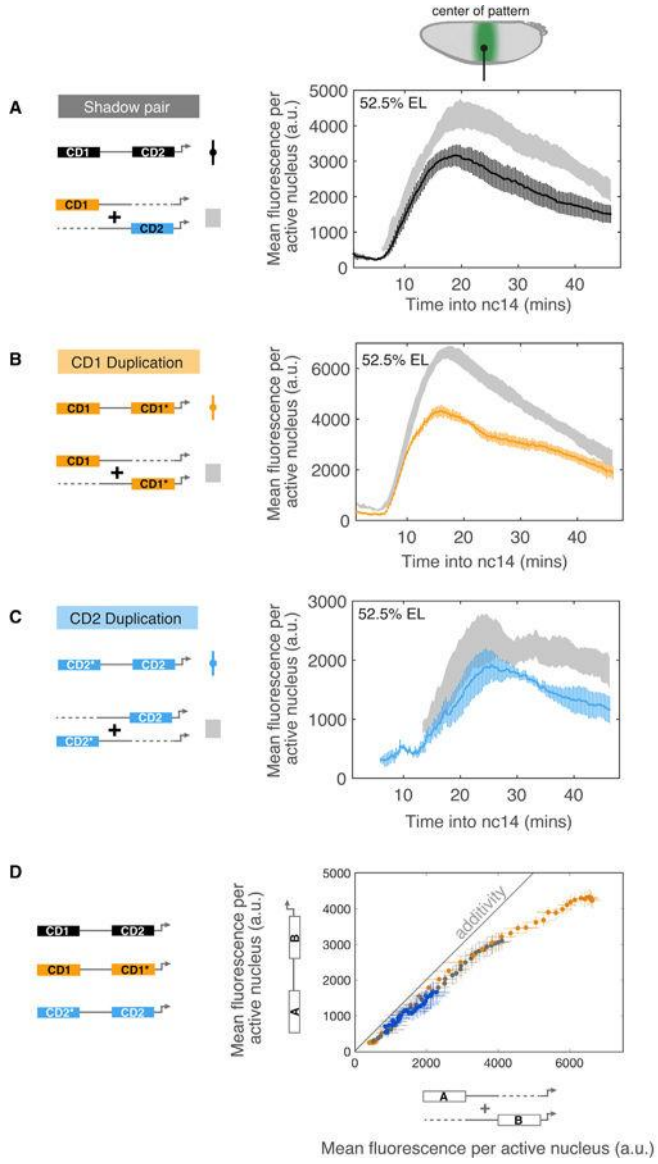


Figure 3. Duplications and Shadow Enhancers of *Krüppel* Drive Largely Sub-additive Expression

(A) Mean expression over time in nuclear cycle 14 (nc14) driven by the shadow enhancer pair (black), compared to the sum of expression driven by each distance-controlled individual enhancer (gray). Mean fluorescence per active nucleus ± SE at approximately the center of the expression pattern, in the bin at 52.5% embryo length (EL). CD1, n = 3 embryos; CD2, n = 6; CD1-CD2, n = 6.

(B) Comparison of expression from the duplicated CD1 construct (orange) compared to the sum of expression from CD1 and CD1* alone (gray). CD1, n = 3 embryos; CD1*, n = 6; CD1-CD1*, n = 3.

(C) Comparison of expression from the duplicated CD2 construct (blue) to the projected sum of expression driven by CD2* and CD2 (gray). CD2, n = 6 embryos; CD2*, n = 4; CD2*-CD2, n = 5.

(D) In the center of the expression pattern, the two-enhancer constructs (shadow enhancer pair, CD1 duplication and CD2 duplication) follow the same relation to additivity over a wide range of expression levels. This holds over ~10% of embryo length in the anterior-middle of the pattern (not shown). Plot shows comparison of expression driven by each two-enhancer construct to the sum of expression driven by the constituent enhancers (A and B). Time points shown are from the period during which mean expression is increasing to a peak: 3–18 min from the start of nc14 (shadow enhancer pair and CD1 duplication) and 10–25 min (CD2 duplication). Shown is the mean fluorescence per active nucleus, ± SE, from the middle of the expression patterns (bin at 52.5% embryo length). The SE on the x axis is the summed error in the fluorescence driven by each constituent enhancer. The gray line indicates additivity.

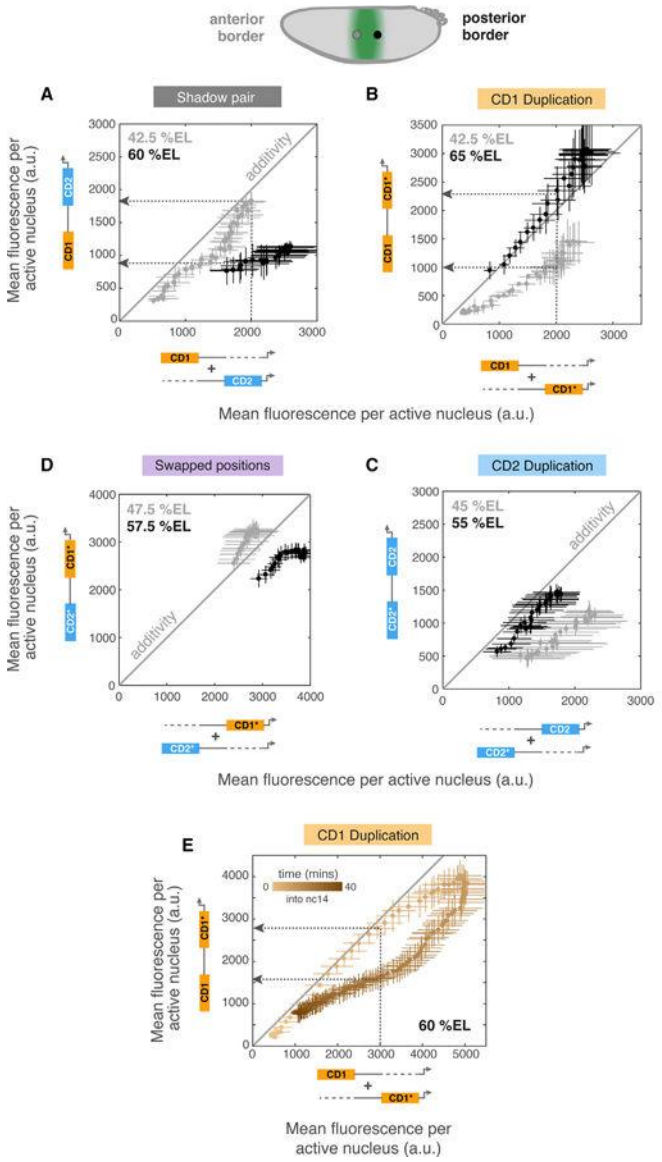


Figure 4. Signal Integration by Two Enhancers Varies over the Anterior-Posterior Axis and over Time

We compare mean fluorescence per actively transcribing nucleus driven by each combination of two enhancers to the sum of the constituent enhancers' expression at the anterior (gray) and posterior (black) borders of the pattern. Each data point is a different time point (at 20-s intervals) over the first 20 min of nuclear cycle 14. Note that we only plot data for time points at which all three relevant constructs (the single enhancers and both enhancers together) drive expression. Error bars indicate SEM.

(A) The shadow enhancer pair drives approximately additive expression in the anterior (at 42.5% embryo length, gray), but sub-additive expression in the posterior of the pattern (at 60% embryo length, black). The dotted arrows highlight how the same summed expression level from the two enhancers alone (on the x axis, where CD1 + CD2 = 2,000 a.u.) yields different expression level outputs at different anterior-posterior positions in the embryo.

(B) Expression from the CD1 duplication is approximately additive in the posterior of the pattern (black) and sub-additive in the anterior (gray; note that this is the opposite trend to that displayed by the shadow pair). For a given summed level of expression from CD1 and CD1* on the x axis, the duplication yields different expression levels at different anterior-posterior positions (dotted arrows).

(C) Expression from the CD2 duplication compared to the sum of expression driven by CD2 and CD2* separately. Data shown in (C) are from the first 25 min of nc14.

(D) When the positions of the shadow enhancers are swapped (CD2*-CD1* construct), expression is additive (or even super-additive) in the anterior (gray), while it is less than additive in the posterior of the pattern (black), following the same trend as when the two enhancers are in their endogenous arrangement.

(E) Expression from the CD1 duplication (shown at 60% embryo length) is approximately additive during the early part of nc14 (light orange), but is much less than additive later in the cycle (dark orange) for the same summed level of expression (dotted arrows highlights an example at CD1 + CD1* = 3,000 a.u.).

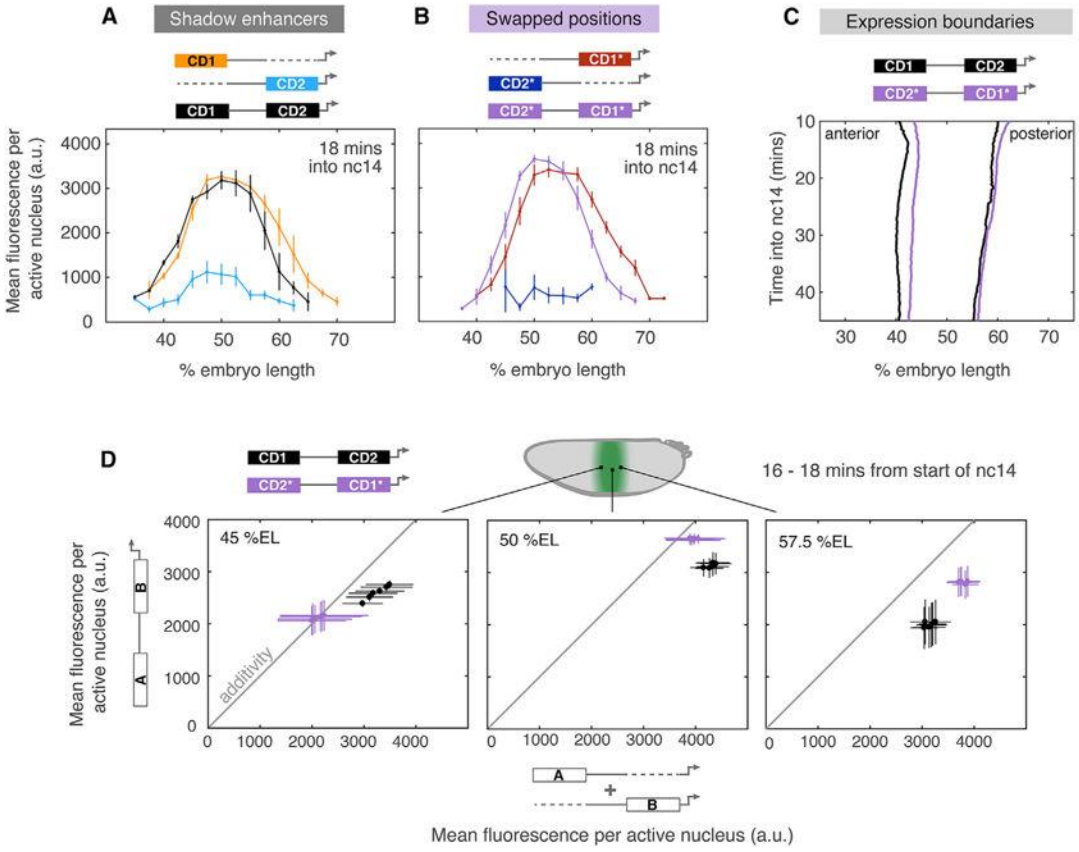


Figure 5. Regulatory Sequence Arrangement Influences Signal Integration by Two Enhancers

(A) When both shadow enhancers drive expression together (CD1-CD2, black), the CD2 enhancer (blue) represses the activity of CD1 (orange) in the posterior of the *Krüppel* pattern. Plotted is mean expression per active nucleus across the *Krüppel* expression pattern at a single time point, 18 min into nuclear cycle 14 (nc14). Error bars are SEMs across multiple embryos.

(B) When the positions of the shadow enhancers are swapped (CD2*-CD1*, purple), the CD2* enhancer (indigo) still represses the activity of CD1* (red) effectively in the posterior of the pattern.

(C) Comparison of the anterior and posterior boundary positions driven by the shadow enhancer pair in their endogenous (black) and swapped (purple) locations over time in nc14. The posterior boundary remains the same irrespective of the order of the enhancers. However, the anterior border driven by the construct containing the enhancers in their swapped positions (CD2*-CD1*) is shifted backward by ~2.5% embryo length relative to that driven by the endogenous construct (CD1-CD2). Anterior and posterior boundaries were defined by fitting a Gaussian to the expression pattern at each time point and finding the positions at half-maximum expression (see Figure S1).

(D) Comparison of the input-output function for the shadow enhancers in their endogenous (black) and swapped positions (purple) across the *Krüppel* pattern. Mean fluorescence from each two-enhancer construct (y axis) is plotted against the sum of the mean expression from the appropriate single-enhancer constructs (x axis). Data are shown from six time points

between 16 and 18 min from the start of nc14. Error bars are SEMs across multiple embryos.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and Virus Strains		
DH5alpha competent cells	NEB	Cat # C2988J
Experimental Models: Organisms/Strains		
D. melanogaster: yw; His2Av-mRFP1; MCP-NoNLS-eGFP. Used for live imaging of nascent transcription.	Hernan Garcia lab, UC Berkeley	Garcia et al., 2013
All *Drosophila melanogaster* transgenic fly lines below generated in Bloomington Stock BL8622.	BestGene	See below
CD1 – CD2 (both enhancers)	This manuscript	AHD stock# 0694
CD1 – ___ (distal enhancer only)	This manuscript	AHD stock# 0695
___ – CD2 (proximal enhancer only)	This manuscript	AHD stock# 0696
CD2 – ___ (proximal enhancer at distal location)	This manuscript	AHD stock# 0697
CD2 – CD2 (proximal enhancer duplication)	This manuscript	AHD stock# 0698
___ – CD1 (distal enhancer at proximal location)	This manuscript	AHD stock# 0700
CD1 – CD1 (distal enhancer duplication)	This manuscript	AHD stock# 0699
CD2 – CD1 (enhancers in switched positions)	This manuscript	AHD stock# 0701
___ – (promoter only – data not shown)	This manuscript	AHD stock# 0702
Oligonucleotides		
‘‘Neutral’’ spacers sequences used to replace each enhancer (see Table S1 for sequences)	Genscript gene synthesis	https://www.genscript.com
Recombinant DNA		
pBOY plasmid	Hare et al., 2008	N/A
pCR4–24XMS2SL-stable plasmid. Source of the 24xMS2 cassette for the reporter gene.	Robert Singer lab, via Addgene	Addgene plasmid # 31865
CD1 - CD2: pBOY-mKr1-MS2-LacZ	This manuscript	Quartzy ID# C00539
CD1 - ___: pBOY-mKr2-MS2-LacZ	This manuscript	Quartzy ID# C00540
___ - CD2: pBOY-mKr3-MS2-LacZ	This manuscript	Quartzy ID# C00541
CD2 - ___: pBOY-mKr4-MS2-LacZ	This manuscript	Quartzy ID# C00542
CD2 - CD2: pBOY-mKr5-MS2-LacZ	This manuscript	Quartzy ID# C00543
___ - CD1: pBOY-mKr6-MS2-LacZ	This manuscript	Quartzy ID# C00544
CD1 - CD1: pBOY-mKr7-MS2-LacZ	This manuscript	Quartzy ID# C00545
CD2 - CD1: pBOY-mKr8-MS2-LacZ	This manuscript	Quartzy ID# C00547
Software and Algorithms		
MATLAB	MathWorks	N/A
MetaMorph	Molecular Devices	N/A
SiteOut	Estrada et al., 2016	https://depace.med.harvard.edu/siteout/
PATSER	Hertz and Stormo, 1999	http://genetics.wustl.edu/gslab
inSite	N/A	https://www.cs.utah.edu/~miriah/insite/

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Other		
Semi-permeable membrane for mounting embryos to conduct live imaging	*In Vitro* Systems and Services	N/A