Shadow enhancers suppress input transcription factor noise through distinct regulatory logic

Rachel Waymack[1], Alvaro Fletcher[2], German Enciso[1, 3], Zeba Wunderlich[1]

1. Department of Developmental and Cell Biology, University of California, Irvine, 92697
2. Mathematical, Computational, and Systems Biology Graduate Program, University of California, Irvine, 92697
3. Department of Mathematics, University of California, Irvine, 92697

Abstract

Shadow enhancers, groups of seemingly redundant enhancers, are found in a wide range of organisms and are critical for robust developmental patterning. However, their mechanism of action is unknown. We hypothesized that shadow enhancers drive consistent expression levels by buffering upstream noise through a separation of transcription factor (TF) inputs at the individual enhancers. By measuring transcriptional dynamics of several *Kruppel* shadow enhancer configurations in live *Drosophila* embryos, we showed individual member enhancers act largely independently. We found that TF fluctuations are an appreciable source of noise that the shadow enhancer pair can better buffer than duplicated enhancers. The shadow enhancer pair is uniquely able to maintain low levels of expression noise across a wide range of temperatures. A stochastic model demonstrated the separation of TF inputs is sufficient to explain these findings. Our results suggest the widespread use of shadow enhancers is partially due to their noise suppressing ability.

Introduction

The first evidence that transcription occurred in bursts, as opposed to as a smooth, continuous process, was observed in *Drosophila* embryos, where electron micrographs showed that even highly transcribed genes had regions of chromatin lacking associated transcripts between regions of densely associated nascent transcripts (Miller & McKnight, 1979). As visualization techniques have improved, it is increasingly clear that transcriptional bursting is the predominant mode of expression across organisms from bacteria to mammals (Dar, et al.,2012; Sanchez & Golding, 2013; Zenklusen, et al., 2008; Fukaya, et al., 2016). These bursts of transcriptional activity, separated by periods of relative silence, have important implications for cellular function, as mRNA numbers and fluctuations largely dictate these quantities at the protein level (Csardi, et al., 2015; Hansen, et al., 2018). Such fluctuations in regulatory proteins, like TFs and signaling molecules, can propagate down a gene regulatory network, significantly altering the expression levels or noise of downstream target genes (Blake, et al., 2003).

Given that protein levels fluctuate and that these fluctuations can cascade down regulatory networks, this raises the question of how organisms establish and maintain the precise levels of gene expression seen during development, where expression patterns can be reproducible down to half-nuclear distances in *Drosophila* embryos (Dubuis, et al., 2013; Gregor, et al., 2007). Many mechanisms that buffer against the noise inherent in gene expression or stemming from genetic or environmental variation have been observed (Lagha, et al., 2012;

Stapel, et al., 2017; Raj et al., 2010). For example, organisms use temporal and spatial averaging mechanisms and redundancy in genetic circuits to achieve the precision required for proper development (Stapel, et al., 2017; Raj, et al., 2010; Erdman, et al., 2009; Lagha, et al., 2012). Here, we propose that shadow enhancers may be another mechanism by which developmental systems manage noise (Barolo, S., 2012).

Shadow enhancers are groups of two or more enhancers that control the same target gene and drive overlapping spatiotemporal expression patterns (Barolo, S., 2012). Shadow enhancers are found in many organisms, from insects to plants to mammals, and are strongly associated with developmental genes (Cannavo, et al., 2016; Osterwalder, et al., 2018; Garnett, et al., 2012; Bomblies, et al., 1999). These seemingly redundant enhancers have been shown to be critical for proper gene expression in the face of both environmental and genetic perturbations, which may exacerbate fluctuations in upstream regulators (Frankel, et al., 2010; Osterwalder, et al. 2018; Perry, et al., 2010; Cheung & Ma, 2015, Chen, et al., 2015). However, shadow enhancers' precise mechanism of action is still unknown. Others have proposed that having multiple enhancers controlling the same promoter ensures a critical threshold of gene expression is reached, perhaps by reducing the effective "failure rate" of the promoter (Lam, et al., 2015; Perry, et al., 2011). An alternative, but not mutually exclusive, possibility is that shadow enhancers ensure precise expression by buffering noise in upstream regulators. Several studies suggest that individual enhancers of a shadow enhancer group tend to be controlled by different sets of TFs, which we call a "separation of inputs" (Wunderlich, et al., 2015; Cannavo, et al., 2016; Ghiasvand, et al., 2011). We hypothesize that this separation allows shadow enhancers to buffer against fluctuations in TF levels.

The *Drosophila* gap gene *Kruppel* (*Kr*) provides a useful system in which to address the mechanisms of action of shadow enhancers. During early embryogenesis, *Kr* is controlled by the activity of two enhancers, proximal and distal, that are each activated by different sets of TFs (Figure 1A; Wunderlich, et al., 2015). Here we focus on differences in activation, as the key repressors of *Kr*, *knirps* and *giant*, are likely to regulate both enhancers. *Kr* expression during this time is critical for thorax formation, and like the other gap genes in the *Drosophila* embryo, has quite low noise (Preiss, et al., 1985; Dubuis, et al., 2013). By measuring live mRNA dynamics, we can use the *Kr* system in *Drosophila* embryos to assess whether and how shadow enhancers act to buffer noise and identify the sources of noise in the developing embryo.

To test our hypothesis, we measured live mRNA dynamics driven by either single *Kr* enhancer, duplicated enhancers, or the shadow enhancer pair and compared the dynamics and noise associated with each. We showed that the individual *Kr* enhancers can act largely independently in the same nucleus while identical enhancers display correlated activity. We constructed a simple mathematical model to describe this system and found that TF fluctuations are necessary to reproduce the correlated activity of identical enhancers in the same nucleus. Using this model, we also found that the lower expression noise driven by the shadow enhancer pair compared to either duplicated enhancer is a natural consequence of the separation of TF inputs. Experimentally, we found the shadow enhancer pair achieves lower noise through decreases in both intrinsic and extrinsic sources of noise. Additionally, the shadow enhancer pair is uniquely able to maintain low levels of expression noise across a wide range of temperatures. We suggest that this noise suppression ability is one of the key features that explains the prevalence of shadow enhancers in developmental systems.

Results

Individual enhancers in the shadow enhancer pair act nearly independently within a nucleus

To test our hypothesis that the separation of inputs between *Kruppel's (Kr)* shadow enhancers provides them with noise-buffering capabilities, we needed to first test the ability of each enhancer to act independently. If variability in gene expression is driven primarily by fluctuations in upstream factors, the shadow enhancer pair, whose individual enhancers are controlled by different sets of TFs, could provide a form of noise buffering. Conversely, variability in upstream regulators may be low enough in the developing embryo that these fluctuations are not the primary driver of downstream expression noise. If this were the case, the separation of inputs is unlikely to be a key requirement of shadow enhancer function.

To investigate these possibilities, we measured and compared the correlation of allele activity in homozygous or heterozygous embryos that carry two reporter genes. *Proximal homozygotes* contained the proximal enhancer driving a reporter, inserted in the same location on both homologous chromosomes, and *distal homozygotes* similarly had the distal enhancer driving reporter expression on both homologous chromosomes (Figure 1B). We also made heterozygous embryos, called *shadow heterozygotes*, which had one proximal and one distal reporter, again in the same location on both homologous chromosomes. To measure live mRNA dynamics and correlations in allele activity, we used the MS2-MCP reporter system (Figure 1C, D). This system allows the visualization of mRNAs that contain the MS2 RNA sequence, which is bound by an MCP-GFP fusion protein (Bertrand, et al., 1998). In the developing embryo, only the site of nascent transcription is visible, as single transcripts are too dim, allowing us to measure the rate of transcription (Garcia, et al., 2013; Lucas, et al., 2013). In blastoderm-stage embryos with two MS2 reporter genes, we can observe two distinct foci of fluorescence corresponding to the two alleles (Figure 1D), in line with previous results that suggest there are low levels of transvection at this stage (Lim, et al., 2018; Fukaya & Levine, 2017). To confirm our ability to distinguish the two alleles, we imaged transcription in embryos hemizygous for our reporter constructs, which only show one spot of fluorescence per nucleus. Our counts of active transcription sites in homozygous embryos correspond well to the expected value calculated from hemizygous embryos (Supplemental Figure 1). Therefore, we are able to measure the correlation of allele activity, though we cannot identify which spot corresponds to which reporter.

We predicted that if variability in gene expression is driven by fluctuations in input TFs, we would observe lower correlations of allele activity in shadow heterozygotes than in either the proximal or distal homozygotes. However, if global factors affecting both enhancers dominate, there would be no difference in allele activity correlations. During the ~1 hour of nuclear cycle 14 (nc14) we found that allele activity is more than twice as correlated in both proximal and distal homozygotes than in shadow heterozygote embryos at 47-57% egg length, which encompasses the central region of *Kr* expression during this time period (Figure 1). This indicates not only that the individual member enhancers of the shadow enhancer pair can act largely independently in the same nucleus, but that differential TF inputs are the primary determinants of transcriptional bursts in this system. Notably, heterozygotes still show marginal allele correlation, indicating that some correlation is induced by either shared input TFs or factors that affect transcription globally. The independence of individual *Kr* enhancers allows for

the possibility that shadow enhancers can act to buffer noise by providing separate inputs to the same gene expression output.

Transcription factor fluctuations are required for the observed differences in the correlations of enhancer activity

To explore the conditions needed for the two *Kr* enhancers to act nearly independently within the same nucleus, we generated a simple model of enhancer-driven dynamics. We considered an enhancer E that interacts with a transcription factor T, which together bind to the promoter to form the active promoter-enhancer complex C (Figure 2A). When the promoter is bound by the enhancer, it drives the production of mRNA. Since the MS2 system only allows us to observe mRNA at the site of transcription, we modeled the diffusion of mRNA away from the transcription site as decay. The transcription factor T is produced in bursts of n molecules at a time, and it degrades linearly. For simplicity, the transcription factor T is an abstraction of the multiple activating TFs that interact with the enhancer, and T corresponds to a different set of TFs for the proximal and distal enhancer. This nonlinear model generalizes the linear model by Bothma et al. (Bothma et al., 2015) by explicitly taking into account the presence of TFs.

We estimated some model parameters directly from experimental data and others by fitting using simulated annealing. The mRNA degradation parameter α and production parameter r were measured directly from fluorescence data without any input from the model (see Methods for details). The remaining parameters were first estimated using mathematical analysis, then fine-tuned using simulated annealing. We found separate parameter sets for the proximal and distal enhancers that, when used to simulate transcription, fit the experimentally measured characteristics of the transcriptional traces, including transcription burst size, frequency, and duration, as well as the total mRNA produced (Supplementary Figure 2).

We hypothesized that a model that lacks fluctuations in the input TFs could not recapitulate the high correlation of transcriptional activity in homozygotes versus the low correlation in heterozygotes. To test this hypothesis, we generated another model of TF production. We call our original model described above *bursting TFs*. The other model is one in which TF numbers are constant over time, which we call *constant TFs* and is equivalent to the model in (Bothma et al., 2015). If the difference in transcription correlation between homozygotes and heterozygotes is due to fluctuating numbers of TFs, we expected that the bursting TFs model will recapitulate this behavior, while the constant TFs model will not. However, if the constant TFs model is also able to recapitulate the observed difference in correlations, then the correlations are likely a consequence of the identical enhancers simply being regulated by the same set of TFs.

For each model, we used the 10 best parameter sets to simulate transcriptional activity in homozygote and heterozygote embryos and analyzed the resulting allele correlations. We found that the bursting TFs model always produced results in which both homozygote allele correlations are significantly higher than the heterozygote, as observed experimentally (Figure 2B). None of the best fitting parameter sets for the constant TF model were able to produce the experimentally-observed behavior and always resulted in similar correlations for the homozygote and heterozygote embryos (Figure 2C). Therefore, in our minimalist model of enhancer-driven transcription, the presence of TF fluctuations is required for the observed differences in allele correlation. These results also demonstrate the advantage of using a single generic TF for each

enhancer. By abstracting away TF interactions, we reduced the complexity and number of parameters in the model, which allowed us to explore the relationship between TF production and allele correlation.

The shadow enhancer pair drives less noisy expression than enhancer duplications

Since the individual *Kr* enhancers can act independently, we wanted to further test whether this separation of inputs enables the shadow enhancer pair to provide more stable gene expression output. We compared the noise in expression driven by the shadow enhancer pair to that driven by two copies of either the distal or proximal enhancer (Figure 3). If the shadow enhancer pair drives lower noise, this observation, combined with our finding of enhancer independence, strongly suggests that the shadow enhancer pair reduces variability and mediates robustness by buffering fluctuations in upstream regulators. Alternatively, if duplicated enhancers drive similar levels of expression noise, this suggests that a separation of inputs is not critical for shadow enhancer's function and that shadow enhancers mediate robustness through a different mechanism, such as ensuring a critical threshold of expression is met (Lam, et al., 2015; Perry, et al., 2011).

We tracked the transcriptional activity in embryos expressing MS2 under the control of the shadow enhancer pair, a duplicated proximal enhancer, or a duplicated distal enhancer (Figure 3). To measure noise associated with each enhancer, we used these traces to calculate the coefficient of variation (CV) of transcriptional activity across nc14. CV is the standard deviation divided by the mean and provides a unitless measure of noise to allow comparisons among our enhancer constructs. We then grouped these CV values by the AP position of the transcriptional spots and found the average CV at each position for each enhancer construct. All of the enhancer constructs display the lowest expression noise at the egg length of their peak expression (Figure 3A), in agreement with previous findings of an inverse relationship between mean expression and noise levels (Dar et al., 2016; Supplemental Figure 3). The shadow enhancer pair's expression noise is almost 30% or 15% lower, respectively, than that of the duplicated proximal or distal enhancers in their positions of maximum expression.

If the primary function of shadow enhancers is only to ensure a critical threshold of expression is reached, we would not expect to also see the lower expression noise associated with the shadow enhancer pair compared to either duplicated enhancer. Furthermore, this decreased expression noise is not simply a consequence of higher expression levels, as the shadow enhancer pair produces less mRNA than the duplicated distal enhancer during nc14 (Figure 3B). The lower expression noise associated with the shadow enhancer pair suggests that it is less susceptible to fluctuations in upstream TFs than multiple identical enhancers.

The separation of input TFs is sufficient to explain the low noise driven by the shadow enhancer pair

To explore which factors drive the difference in CVs between the duplicated and shadow enhancer constructs, we extended our model to have a single promoter controlled by two enhancers (Figure 4A). To do so, we assumed that either or both enhancers can be looped to the promoter and drive mRNA production. The rate of mRNA production when both enhancers are looped is the sum of the rates driven by the individual enhancers. We assumed that some parameters, e.g. the TF production rates and mRNA decay rate, are the same as the single

enhancer case. We allowed the parameters describing the promoter-enhancer looping dynamics (the k_{on} and k_{off} values) to differ, depending on the enhancer's position in the construct relative to the promoter and whether another enhancer is present. To fit the k_{on} and k_{off} values, we used the medians of the 10 best single enhancer parameter sets as a starting point and performed simulated annealing to refine them.

This approach allowed us to examine how the model parameters that describe promoter-enhancer looping dynamics change when two enhancers are controlling the same promoter. We compared the k_{off} and k_{on} values for each enhancer in the two enhancer constructs to their values from the single enhancer model. We generally found that k_{off} values increased and k_{on} values decreased (Figure 4B). The effect is most pronounced in the duplicated distal enhancer, with large changes to the k_{off} and k_{on} values for the enhancer in the position far from the promoter (position 2). Given that our model assumes that enhancers act additively and only allows for changes in the k_{off} and k_{on} values, these observed effects may indicate that either the presence of a second enhancer interferes with promoter-enhancer looping or that the promoter can be saturated. Our model cannot distinguish between these two possibilities, but these observations are consistent with our (Supplementary Figure 4) and previous results indicating that the *Kr* enhancers act sub-additively (Scholes, et al., 2019). Additionally, the dramatic changes in k_{off} and k_{on} values in the duplicated distal enhancer are consistent with a previous assertion that enhancer sub-additivity is most pronounced in cases of strong enhancers (Bothma et al., 2015).

We used these models to simulate transcription and predict the resulting CVs from the duplicated enhancer and shadow enhancer constructs. In line with experimental data, we found the model predicts that the shadow enhancer construct drives lower noise than the duplicated distal or duplicated proximal enhancer constructs in the middle of the embryo. This is particularly notable, as we did not explicitly fit our model to reproduce the experimentally observed CVs. There is only one fundamental difference between the shadow and duplicated enhancer models, namely the use of separate TF inputs for the shadow enhancers. Therefore, we can conclude that the separation of input TFs is sufficient to explain the low noise driven by the shadow enhancer construct.

The shadow enhancer pair buffers against intrinsic and extrinsic sources of noise

To further validate that the more stable expression driven by the shadow enhancer pair is due to its separation of inputs, we compared the extrinsic and intrinsic noise associated with the shadow enhancer pair to that associated with either single or duplicated enhancers. To do so, we measured the transcriptional dynamics of embryos with two identical reporters in each nucleus and calculated noise sources using the approach of Elowitz, et al. (Elowitz, et al., 2002). Intrinsic noise corresponds to sources of noise, such as TF binding and unbinding, that affect each allele separately. It is quantified by the degree to which the activities of the two reporters in a single nucleus differ. Extrinsic noise corresponds to global sources of noise, such as TF levels, that affect both alleles simultaneously. It is measured by the degree to which the activities of the two reporters change together. Intrinsic and extrinsic noise are defined such that, when squared, their sum is equal to total noise^2, which corresponds to the CV^2 of the two identical alleles in each nucleus in our system (see Methods). Because our data do not meet one key assumption needed to measure extrinsic and intrinsic noise with the two reporter approach (see Discussion; Supplementary Figure 5), we use the terms inter-allele noise and covariance in place of intrinsic and extrinsic noise.

Based on our separation of inputs hypothesis and CV data, we expected the total noise associated with the shadow enhancer pair to be lower than that associated with the duplicated enhancers. We predicted that the shadow enhancer pair will mediate lower total expression noise through lower covariance, as the two member enhancers are regulated by different TFs. Given the complexity of predicting inter-allele noise from first principles (see Supplementary Note), we predicted that constructs with two enhancers will have lower inter-allele noise than single enhancer constructs, but did not have a strong prediction regarding the relative inter-allele noise among the different two-enhancer constructs. Comparisons of noise between the single and duplicated enhancer constructs would further allow us to discern whether reductions in noise are generally associated with two-enhancer constructs or whether this is a particular feature of the shadow enhancer pair.

Neither the duplicated proximal nor distal enhancers drive significantly lower total noise than the corresponding single enhancers, indicating that the addition of an identical enhancer is not sufficient to reduce expression noise in this system (Figure 5A). The shadow enhancer pair drives lower total expression noise than either single or duplicated enhancer, consistent with the temporal CV data in Figure 3. The median total expression noise associated with the duplicated distal and duplicated proximal enhancers is 1.4 or 2.4 times higher, respectively, than that associated with the shadow enhancer pair (Figure 5A). Note that for measurements of noise, our distal construct places the enhancer at the endogenous spacing from the promoter, as we wanted to control for positional effects on expression and noise (Scholes, et al., 2019; Supplemental Figure 6).

In line with our expectations, the shadow enhancer pair has significantly lower covariance levels than either single or duplicated enhancers (Figure 5B). The shadow enhancer pair also has lower inter-allele noise than all of the other constructs, though these differences are only marginally significant ($p = 0.13$) when compared to the duplicated distal enhancer. Covariance makes a larger contribution to the total noise for the duplicated distal enhancer and the shadow enhancer pair, while inter-allele noise is the larger source of noise for the single distal enhancer and the single or duplicated proximal enhancers (Figure 5B).

The lower total noise and covariance of the shadow enhancer pair support our hypothesis that, by separating regulation of the member enhancers, the shadow enhancer pair can buffer against upstream fluctuations. The lower inter-allele noise associated with the shadow enhancer pair warrants further investigation. A simple theoretical approach predicts that two enhancer constructs will have lower inter-allele noise (see Supplementary Note). Given that this is not universally observed in our data, this suggests that there is still much to discover about how inter-allele noise changes as additional enhancers control a gene's transcription.

The shadow enhancer pair drives low noise at several temperatures

We showed the *Kr* shadow enhancer pair drives expression with lower total noise than either single or duplicated enhancer, yet previous studies have generally found individual member enhancers of a shadow enhancer set are dispensable under ideal conditions (Frankel, et al., 2010; Perry et al., 2011; Osterwalder, et al., 2018). However, in the face of environmental or genetic stress, the full shadow enhancer group is necessary for proper development (Frankel et al., 2010; Osterwalder, et al., 2018; Perry, et al., 2011). We therefore decided to investigate whether temperature stress causes significant increases in expression noise and whether the shadow enhancer pair or duplicated enhancers can buffer potential increases in noise.

Similar to our findings at ambient temperature (26.5°C), the shadow enhancer pair drives lower total noise than all other tested enhancer constructs at 32°C (Figure 6B). At 32°C, the duplicated distal and duplicated proximal enhancers display 35% or 52%, respectively, higher total noise than the shadow enhancer pair. At 17°C, the shadow enhancer pair has approximately 46% lower total noise than either the single or duplicated proximal enhancer, 21% lower total noise than the single distal enhancer, and is not significantly different than the duplicated distal enhancer (Figure 6A). As seen by the variety of shapes in the temperature response curves (Figure 6C), temperature perturbations have enhancer-specific effects, suggesting input TFs may differ in their response to temperature change. The low noise driven by the shadow enhancer pair across conditions is consistent with previous studies showing shadow enhancers are required for robust gene expression at elevated and lowered temperatures (Frankel, et al., 2010; Perry, et al., 2010).

Discussion

Fluctuations in the levels of transcripts and proteins are an unavoidable challenge to precise developmental patterning (Raser & O'Shea, 2005; Arias & Hayward, 2006; Hansen, et al., 2018). Given that shadow enhancers are common and necessary for robust gene expression (Osterwalder, et al., 2018; Frankel, et al., 2010; Perry, et al., 2010), we proposed that shadow enhancers may function to buffer the effects of fluctuations in the levels of key developmental TFs. To address this, we have, for the first time, extensively characterized the noise associated with shadow enhancers critical for patterning the early *Drosophila* embryo. By tracking biallelic transcription in living embryos, we tested the hypothesis that shadow enhancers buffer noise through a separation of TF inputs to the individual member enhancers. Our results show that TF fluctuations play a significant role in transcriptional noise and that a shadow enhancer pair is better able to buffer both extrinsic and intrinsic sources of noise than duplicated enhancers. Using a simple mathematical model, we found that fluctuations in TF levels are required to reproduce the observed correlations between reporter activity and that the low noise driven by the shadow enhancer pair is a natural consequence of the separation of TF inputs to the member enhancers. Lastly, we showed that a shadow enhancer pair is uniquely able to buffer expression noise across a wide range of temperatures. Together, these results support the hypothesis that the separation of inputs of shadow enhancers allow them to buffer input TF noise and therefore drive more robust gene expression patterns during development.

Temporal fluctuations in transcription factor levels drive expression noise in the embryo

When measured in fixed embryos, the TFs used in *Drosophila* embryonic development show remarkably precise expression patterns, displaying errors smaller than the width of a single nucleus (Dubuis, et al., 2013; Gregor, et al., 2007; Little 2013; He, et al., 2008). It therefore was unclear whether fluctuations in these regulators play a significant role in transcriptional noise in the developing embryo. By measuring the temporal dynamics of the individual *Kr* enhancers, each of which is controlled by different transcriptional activators, we show that TF fluctuations do significantly contribute to the noise in transcriptional output of a single enhancer. Within a nucleus, expression controlled by the two different *Kr* enhancers is far less correlated than expression driven by two copies of the same enhancer, indicating that TF inputs, as opposed to more global factors, are the primary regulators of transcriptional bursting in this system.

Given that individual *Kr* enhancers are influenced by fluctuations in input TFs, it may seem puzzling that endogenous *Kr* expression patterns are rather reproducible (Little 2013). Previous work has cited the role of spatial and temporal averaging, which buffers noisy nascent transcriptional dynamics to generate more precise expression levels. Shadow enhancers operate upstream of this averaging, driving less noisy nascent transcription than either single enhancers or enhancer duplications.

A stochastic model underscores importance of transcription factor fluctuations

We developed a stochastic mathematical model of *Kr* enhancer dynamics and mRNA production that recapitulates our main experimental results. This model is based on that by (Bothma, et al., 2015), but it is expanded to include the dynamics of a TF that regulates each enhancer. We placed a strong emphasis on the simplicity of this model, e.g. by using a single abstract TF for each enhancer. This choice both avoids a combinatorial explosion of parameters and makes the model results and parameters easier to interpret. One of the most notable features of the model is that it recreates the differences in noise between shadow and duplicated enhancer constructs without any additional fitting, indicating that these differences are a direct result of the separation of input TFs to the proximal and distal enhancers.

Future versions of this model can include refinements. For example, in the current model, we do not include the influence of repressiveTFs or fluctuations that affect transcription globally. The absence of these features may partially explain the non-zero correlation experimentally observed in the shadow heterozygote embryos. Future experiments and models can also be designed to identify the mechanism of enhancer non-additivity: changes in promoter-enhancer looping, saturation of the promoter, or other mechanisms.

Noise source decomposition suggests competition between reporters

In our investigation of sources of noise, we decomposed total noise into extrinsic and intrinsic components as in (Elowitz, et al., 2002). In that study, the authors showed that the activity of one reporter did not inhibit expression of the other reporter, and therefore their calculations assume no negative covariance between the reporters' expression output. In our system, we found a small amount of negative covariance between the activity of two alleles in the same nucleus (Supplemental Figure 5). For this reason, we called our measurements covariance and inter-allele noise. The negative covariance we observe indicates that activity at one allele can sometimes interfere with activity at the other allele, suggesting competition for limited amounts of a factor necessary for reporter visualization. The two possible limiting factors are MCP-GFP or an endogenous factor required for transcription. If MCP-GFP were limiting, we would expect to see the highest levels of negative covariance at the center of the embryo, where the highest number of transcripts are produced and bound by MCP-GFP. Since the fraction of nuclei with negative covariance is highest at the edges of the expression domain (Supplementary Figure 5), the limiting resource is likely not MCP-GFP, but instead a spatially-patterned endogenous factor, like a TF.

Currently, the field largely assumes that adding reporters does not appreciably affect expression of other genes. However, sequestering TFs within repetitive regions of DNA can impact gene expression (Liu, et al., 2007; Janssen, et al., 2000), and a few case studies show that reporters can affect endogenous gene expression (Laboulaye, et al., 2018; Thompson & Gasson, 2001). If TF competition is responsible for the observed negative covariance between reporters, a

389 closer examination of the effects of transgenic reporters on the endogenous system is warranted.
390 In addition, TF competition may be a feature, not a bug, of developmental gene expression
391 control, as modeling has indicated that molecular competition can decrease expression noise and
392 correlate expression of multiple targets (Yuan, et al., 2018).

393 *Additional functions of shadow enhancers and outlook*

394 There are likely several features of shadow enhancers selected by evolution outside of their
395 noise-suppression capabilities. Preger-Ben Noon, et al. recently showed that all shadow
396 enhancers of *shavenbaby*, a developmental TF gene in *Drosophila*, drive expression patterns in
397 tissues and times outside of their previously-characterized domains in the larval cuticle (Preger-
398 Ben Noon, et al., 2018). This suggests that shadow enhancers, while seemingly redundant at one
399 developmental stage, may play separate, non-redundant roles in other stages or tissues. In several
400 other cases, both members of a shadow enhancer pair are required for the precise expression
401 pattern generated by the endogenous locus (El-Sherif & Levine, 2016; Perry, et al., 2012;
402 Dunipace, et al., 2011; Perry, et al., 2011; Yan, et al., 2017). In the case of *Kr*, the early
403 embryonic enhancers drive observable levels of expression in additional tissues and time points,
404 but these expression patterns overlap those driven by additional, generally stronger, enhancers,
405 suggesting that the primary role of the proximal and distal enhancers is in early embryonic
406 patterning (Hoch, et al., 1990). In addition, the endogenous expression domain of *Kr* is best
407 recapitulated by the pair of shadow enhancers (El-Sherif & Levine, 2016). Therefore, while we
408 cannot rule out the possibility that the proximal and distal enhancers perform separate functions
409 at later stages, it seems that their primary function, and evolutionary substrate, is controlling
410 *Kruppel* expression pattern and noise levels during early embryonic development.

411 Here, we have investigated the details of shadow enhancer function for a particular
412 system, and we expect that some key observations may generalize to many sets of shadow
413 enhancers. Shadow enhancers seem to be a general feature of developmental systems (Cannavo,
414 et al., 2016; Osterwalder, et al., 2018), but the diversity among them has yet to be specifically
415 addressed. While we worked with a pair of shadow enhancers with clearly separated TF
416 activators, shadow enhancers can come in much larger groups and with varying degrees of TF
417 input separation between the individual enhancers (Cannavo, et al., 2016; Osterwalder, et a.,
418 2018). To discern how expression dynamics and noise driven by shadow enhancers depend on
419 their degree of TF input separation, we are investigating these characteristics in additional sets of
420 shadow enhancers with varying degrees of differential TF regulation. Our current experimental
421 data and computational results, combined with that gathered from additional shadow enhancers
422 will inform fuller models of how developmental systems ensure precision and robustness.

423

424 **Materials and Methods**

425 *Generation of transgenic reporter fly lines*

426 The single, duplicated, or shadow enhancers were each cloned into the pBphi vector, upstream of
427 the *Kruppel* promoter, 24 MS2 repeats, and a *yellow* reporter gene as in (Fukaya, et al., 2016).
428 We defined the proximal enhancer as chromosome 2R:25224832-25226417, the distal enhancer
429 as chromosome 2R:25222618-25223777, and the promoter as chromosome 2R:25226611-
430 25226951, using the *Drosophila melanogaster* dm6 release coordinates. The precise sequences

for each reporter construct are given in Supplementary File 1. For the allele correlation experiments, each enhancer was cloned 192 bp upstream of the *Kr* promoter, separated by the endogenous sequence found between the proximal enhancer and the promoter. For transcriptional noise experiments, the distal enhancer was placed at its endogenous spacing, 2835 bp upstream of the promoter, and the proximal enhancer sequence was replaced by a region of the lambda genome that is predicted to have few relevant TF binding sites. In the shadow enhancer pair or duplicated enhancer constructs, the two enhancers were separated by the sequence separating the proximal and distal enhancers in the endogenous locus.

Using phiC31-mediated integration, each reporter construct was integrated into the same site on the second chromosomes by injection into yw; PBac{y[+]-attP-3B}VK00002 (BDRC stock #9723) embryos by BestGene Inc. (Chino Hills, CA). To produce embryos with biallelic expression of the MS2 reporter, female flies expressing RFP-tagged histones and GFP-tagged MCP (yw; His-RFP/Cyo; MCP-GFP/TM3.Sb) were crossed with males containing one of the enhancer-MS2 reporter constructs. Virgin female F1 offspring were then mated with males of the same parental genotype, except in the case of shadow heterozygous flies, which were mated with males containing the other single enhancer-MS2 reporter.

Sample preparation and image acquisition

Live embryos were collected prior to nc14, dechorionated, mounted on a permeable membrane, immersed in Halocarbon 27 oil, and put under a glass coverslip as in (Garcia, et al., 2013). Individual embryos were then imaged on a Nikon A1R point scanning confocal microscope using a 60X/1.4 N.A. oil immersion objective and laser settings of 40uW for 488nm and 35uW for 561nm. To track transcription, 21 slice Z-stacks, at 0.5um steps, were taken throughout the length of nc14 at roughly 30 second intervals. To identify the Z-stack's position in the embryo, the whole embryo was imaged after the end of nc14 at 20x using the same laser power settings. Later in the analysis, each transcriptional spot's location is described as falling into one of 42 AP bins, with the first bin at the anterior of the embryo. Unless otherwise indicated, embryos were imaged at ambient temperature, which was on average 26.5°C. To image at other temperatures, embryos were either heated or cooled using the Bioscience Tools (Highland, CA) heating-cooling stage and accompanying water-cooling unit.

Calculation of transcription parameters

For every spot of transcription imaged, the fluorescence traces across the time of nc14 were first subject to smoothing by the LOWESS method with a span of 10%. The resulting smoothed traces were used to measure transcriptional parameters and noise. Traces consisting of fewer than three time frames were removed from calculations. To calculate transcription parameters, we used the smoothed traces to determine if the promoter was active or inactive at each time point. A promoter was called active if the slope of its trace (change in fluorescence) between that point and the next was greater than or equal to the instantaneous fluorescence value calculated for one mRNA molecule (F_{RNAP} , described below). Once called active, the promoter is considered active until the slope of the fluorescence trace becomes less than or equal to the negative instantaneous fluorescence value of one mRNA molecule, at which point it is called inactive until another active point is reached. The instantaneous fluorescence of a single mRNA was chosen as the threshold because we reasoned that an increase in fluorescence greater than or equal to that of a single transcript is indicative of an actively producing promoter, while a decrease in fluorescence greater than that associated with a single transcript indicates transcripts

are primarily dissociating from, not being produced, at this locus. Visual inspection of fluorescence traces agreed well with the burst calling produced by this method (Supplemental Figure 7).

Using these traces and promoter states, we measured burst size, frequency and duration. Burst size is defined as the integrated area under the curve of each transcriptional burst. Duration is defined as the amount of time occurring between the frame a promoter is determined active and the frame it is next determined inactive. Frequency is defined as the number of bursts occurring in the period of time from the first time the promoter is called active until 50 minutes into nc14 or the movie ends, whichever is first. The time of first activity was used for frequency calculations because the different enhancer constructs showed different characteristic times to first transcriptional burst during nc14. For these, and all other measurements, we control for position of the transcription trace by first individually analyzing the trace and then using all the traces in each AP bin to calculate summary statistics of the transcriptional dynamics and noise values at that AP position.

Conversion of integrated fluorescence to mRNA molecules

To put our results in physiologically relevant units, we calibrated our fluorescence measurements in terms of mRNA molecules. As in (Lammers, et al., 2018), for our microscope, we determined a calibration factor, α, between our MS2 signal integrated over nc13, F_{MS2},and the number of mRNAs generated by a single allele from the same reporter construct in the same time interval, N_{FISH}, using the *hunchback* P2 enhancer reporter construct (Garcia et al., 2013). Using this conversion factor, we can calculate the integrated fluorescence of a single mRNA (F_1) as well as the instantaneous fluorescence of an mRNA molecule (F_{RNAP}). With our microscope, F_{RNAP} is 379 AU/RNAP and F_1 is 1338 AU/RNAP·min. With these values, we are able to convert both integrated and instantaneous fluorescence into total mRNAs produced and number of nascent mRNAs present at a single time point, by dividing by F_1 and F_{RNAP}, respectively.

Calculation of noise metrics

To calculate the temporal CV each transcriptional spot *i*, we used the formula:

$$\mathrm{CV}(i) = \mathrm{mean}(m^i(t))/\mathrm{standard\ deviation}(m^i(t))$$

where $m^i(t)$ is the fluorescence of spot *i* and time *t.*

We also decomposed the total noise experienced in each nucleus to inter-allele noise and co-variance, analogous to the approach of (Elowitz, et al., 2002).

Inter-allele noise is calculated one nucleus at a time. It is the mean square difference between the fluorescence of the two alleles in a single nucleus:

$$\eta_{IA}^2 = \frac{\langle (m_1(t) - m_2(t))^2 \rangle}{2\langle m_1(t) \rangle \langle m_2(t) \rangle}$$

where $m_1(t)$ is the fluorescence of one allele in the nucleus at time *t*, and $m_2(t)$ is the fluorescence of the other allele in the same nucleus and the angled brackets indicate the mean across the time of nc14.

Covariance is the covariance of the activity of the two alleles in the same nucleus across the time of nc14:

514 $$\eta_{CV}^2 = \frac{\langle m_1(t)m_2(t)\rangle - \langle m_1(t)\rangle\langle m_2(t)\rangle}{\langle m_1(t)\rangle\langle m_2(t)\rangle}$$

515 The inter-allele and covariance values are defined such that they sum to give the total
516 transcriptional noise displayed by the two alleles in a single nucleus.

517 $$\eta_{tot}^2 = \frac{\langle m_1(t)^2 + m_2(t)^2\rangle - 2\langle m_1(t)\rangle\langle m_2(t)\rangle}{2\langle m_1(t)\rangle\langle m_2(t)\rangle}$$

518 This total noise value is equal to the coefficient of variation of the expression of the two alleles
519 in a single nucleus across the time of nc14.

520 *Statistical methods*

521 To determine any significant differences in total noise, covariance, or inter-allele noise values
522 between the different enhancer constructs, we performed Kruskal-Wallis tests with the
523 Bonferroni multiple comparison correction.

524 *Description of the single enhancer model and associated parameters*

525 We constructed a model of enhancer-driven transcription based on the following chemical
526 reaction network,

527 528 529 $$T + E \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} C \xrightarrow{r} C + R$$

530

531 $$\begin{aligned} \emptyset &\xrightarrow{\beta_1} nT \\ T &\xrightarrow{\beta_{-1}} \emptyset \\ R &\xrightarrow{\alpha} \emptyset \end{aligned}$$

532 where E is an enhancer that interacts with a transcription factor T, which together bind to the
533 promoter at a rate k_{on} to form the active promoter-enhancer complex C. When the promoter is in
534 this active form, it leads to the production of mRNA denoted by R, which degrades by diffusion
535 from the gene locus at a rate α. Transcription is interrupted whenever the complex C
536 disassociates spontaneously at a rate k_{off}. In the bursting TFs model, the transcription factor T
537 appears at a rate β_1 and degrades at a rate β_{-1}. To recapitulate *Kruppel* expression patterns, the
538 value of β_1 was assumed to be given by

539 (1) $$f(x) = c\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

540

541 where x is the percentage along the length of the egg and c is a scaling constant. Since *Kruppel*
542 activity peaks near the center of the egg, we chose $\mu = 50$, while c and σ were fitted along with
543 the other parameters. Lastly, n was assumed to be fixed across the length of the egg.

544 We also generated a constant TF model, which is an adaptation of the model in (Bothma et al.,
545 2015). This model implicitly assumes that TF numbers are constant and, therefore, are
546 incorporated into the value of k_{on} as described by the reactions

$$E \underset{k_{\text{off}}}{\overset{Tk_{\text{on}}}{\rightleftharpoons}} C \xrightarrow{r} C + R$$

$$R \xrightarrow{\alpha} \emptyset$$

In this case, the value for T was fitted for each bin in a similar way to β_1, i.e. the constant number of TFs was assumed to be described by equation (1) (values were rounded to the nearest integer).

To simulate the transcriptional traces, we implemented a stochastic approach. Individual chemical events such as enhancer-promoter looping take place at random times and are influenced by transcription factor numbers. Individual trajectories of chemical species over time were calculated using the Gillespie algorithm (Gillespie, 1976), and these trajectories are comparable to the experimentally measured transcriptional traces. Since the enhancer is either bound or not bound to the promoter, we imposed the constraint that $C + E = 1$ when simulating model dynamics.

Estimation of model parameters from experimental data

To yield a starting estimate for the k_{on} and k_{off} parameters, we defined the start and end of a burst as the time when the reactions $E \xrightarrow{k_{on}} C$ and $C \xrightarrow{k_{off}} E$ occur, respectively. The length of the ith burst was defined as the range of $[b_i,p_i]$ where b_i corresponds to the time of the i^{th} instance of the reaction $E \xrightarrow{k_{on}} C$ and p_i to the time of the i^{th} instance of the reaction $C \xrightarrow{k_{off}} E$. The time between the i^{th} burst and the $i + 1^{\text{th}}$ burst is $[p_i,b_{i+1}]$. The Gillespie algorithm dictates that the time spent in any given state is determined by an exponentially distributed random variable with a rate parameter equal to the product of two parts: the sum of rate constants of the outgoing reactions, and the number of possible reactions. If the enhancer is either bound or unbound, we have that $C = 1$ or $E = 1$, respectively. Therefore, by letting t_b be the average time between bursts and t_d be the average duration of a burst, we can write

$$t_b = \lim_{M\to\infty} \frac{1}{M} \sum_{j=1}^{M} \left(\frac{1}{N-1} \sum_{i=1}^{N-1} (b_{i+1_j} - p_{i_j}) \right) = \frac{1}{k_{\text{on}} ET} \approx \frac{1}{k^{\text{on}}},$$

and

$$t_d = \lim_{M\to\infty} \frac{1}{M} \sum_{j=1}^{M} \left(\frac{1}{N} \sum_{i=1}^{N} (p_{i_j} - b_{i_j}) \right) = \frac{1}{k_{\text{off}} C} = \frac{1}{k_{\text{off}}},$$

where N is the number of bursts for spot j, b_{ij} and p_{ij} denote the beginning and end of burst i in spot j respectively, and M denotes the total number of spots in the egg. The right-hand sides are given by the expected value of the exponential distribution and the assumption that, on average, T is close to 1. While this may not be the case for T, the assumption provides a convenient upper bound for the average time between bursts, which is likely not to have a much smaller value for a lower bound. (A low enough value of t_b would imply nearly constant fluorescence intensity instead of bursts.) Finally, the average duration of a burst t_d can be calculated directly from the

data and used to obtain k_{off} by calculating $1/t_d$. Similarly, the average time between bursts t_b is readily available from the data giving us $k_{on} \approx 1/t_b$.

We were able to directly estimate mRNA production and degradation rates from the experimental data. To estimate α, we focused on periods of mRNA decay; i.e. periods where no active transcription is taking place and are thus described by

$$R' = -\alpha R,$$

which in turn can be solved to be

$$R = ce^{-t\alpha}, \quad (2)$$

where c is a constant of integration. Taking the derivative of equation 2 yields

$$R'(t) = -\alpha c e^{-t\alpha}, \quad (3)$$

which corresponds to the slope of the decaying burst. We define the interval of decay of the i^{th} burst as $[p_i, b_{i+1}]$. For some point $t_0 \in (p_i, b_{i+1})$, let $R_0 = R(t_0) = ce^{-t_0\alpha}$. Solving this expression for c gives that $c = R_0 e^{t_0\alpha}$. Substituting for c in equation 3 evaluated at t_0 results in $R'(t_0) = -\alpha R_0 e^{t_0\alpha} e^{-t_0\alpha}$ $= -\alpha R_0$. Then, it follows that

$$\alpha = -\frac{R'(t_0)}{R_0}. \quad (4)$$

In other words, the rate of decay of mRNA fluorescence can be calculated from any trace by taking the ratio of the slope during burst decay and its intensity at a given time $t_0 \in (p_i, b_{i+1})$.

Adjacent measurements of fluorescence intensity from the single enhancer systems were used to approximate the slope at each point in the traces. Then, equation 4 was applied to each point. A histogram of all calculated values was generated (Supplemental Figure 8). In this figure, there was a clear peak, which provided us with an estimate of $\alpha \approx 1.95$.

The estimation of r was done for periods of active transcription, which are also accompanied by simultaneous mRNA decay. By noting that $C = 1$ during mRNA transcription, we can approximate these periods as the zeroth order process

$$\emptyset \underset{\alpha}{\overset{r}{\rightleftharpoons}} R$$

The differential equation associated with this system is given by

$$R' = r - \alpha R, \quad (5)$$

and has steady state $R^* = r/\alpha$. Equation 5 can be solved explicitly for R by choosing

$$R(t) = \frac{r}{\alpha} + ce^{-t\alpha},$$

where c is a constant of integration. For two adjacent measurements at times t_1 and t_2 we can write their respective measured amounts of mRNA as

$$R_1 = \frac{r}{\alpha} + c_1 e^{-t_1\alpha}, \quad (6)$$

and

$$R_2 = \frac{r}{\alpha} + c_2 e^{-t_2 \alpha} \quad (7)$$

Solving for c_1 and c_2 gives

$$c_1 = (R_1 - \frac{r}{\alpha}) e^{t_1 \alpha}$$

$$c_2 = (R_2 - \frac{r}{\alpha}) e^{t_2 \alpha}.$$

The short-term fluctuations of mRNA from R_1 to R_2 between two adjacent discrete time points in the stochastic system can be approximated by equations 6 and 7. This implies that

$$(R_1 - \frac{r}{\alpha}) e^{t_1 \alpha} = (R_2 - \frac{r}{\alpha}) e^{t_2 \alpha},$$

which in turn gives

$$r = \alpha \frac{R_1 - R_2 e^{\alpha \Delta t}}{1 - e^{\alpha \Delta t}}.$$

Therefore, the estimation of r can be computed given two adjacent measurements of fluorescence and the time between them. Finally, we used a similar approach as done with α to calculate values of r from fluorescence data. However, unlike α, r was calculated for each bin to account for differences in transcriptional efficiency across the length of the embryo.

Parameter fitting with simulated annealing

Simulations and parameter fitting were done with MATLAB®. Optimization in fitting was done by minimizing the sum of squared errors (SSE) between the normalized vectors of burst properties and allele correlations of the experimental and simulated data. In particular, a vector y of experimental data was created by concatenating the following vectors: burst size, integrated fluorescence, frequency, duration, and allele correlation across the length of the embryo. The vector y was subsequently normalized by dividing each burst property by the largest element in their respective vectors (except correlation which by definition is unitless between -1 and 1). A vector x was created in an analogous fashion to y but using simulated instead of experimental data. However, x was normalized using the same elements that were used to normalize y. Then, the discrepancy between the experimental and simulated data was measured by

$$SSE = \sum_{i=1}^{n} (y_i - x_i)^2 .$$

We used a high-performance computing cluster to compute 200 independent runs of parameter fitting with simulated annealing for each model variant. The algorithm requires an initial guess of the parameter set P_0, an initial temperature Γ_0, a final temperature Γ', the number of iterations per temperature N, and a cooling factor μ. Then, each iteration is as follows:

(1) If the current iteration i is such that $i > N$, then update the current temperature $\Gamma_k = \mu^k \Gamma_0$ to $\mu^{k+1} \Gamma_0$ and set $i = 0$. Otherwise, set i to $i + 1$.

(2) Check if $\Gamma_k < \Gamma$'. If so, return the current parameter set P_j and terminate.

(3) Choose a parameter randomly from P_j and multiply it by a value sampled from a normal distribution with a mean equal to 1. The standard deviation of such distribution should be continuously updated to be Γ_k. The result of this step is the newly generated parameter set P_{j+1}.
(4) Calculate ΔE as the difference in SSE between the data generated by P_j and that generated by P_{j+1}. Update P_j to P_{j+1} if $\Delta E < 0$ or with probability $p < e^{\Delta E/\Gamma k}$ where p is a uniformly distributed random number.
(5) Repeat all steps until termination.

To generate our results, we chose $\Gamma_0 = 1$, $\Gamma' = \Gamma_0/10$, $N = 30$, and $\mu = 0.8$. We observed an improvement in the quality of the fittings by using analysis-derived parameter values as initial guesses instead of values given through random sampling. The sampled space ranged from 10^{-3} to 10^3 for all parameters, except n, which was sampled from 10^0 to 10^2, and σ, which was randomly chosen to be an integer between 1 and 20. Equal numbers of parameter values were sampled at each order of magnitude. The analysis in the section above was used to estimate the parameters in P_0. Parameters that were not estimated in the previous section were given the following initial guesses: $n = 10$, $\beta_{-1} = 1$, $\sigma = 6$, and $c = 40$. Initial guesses for c and σ were based on the experimental observation that there is little transcription outside of 20-80% egg length. Based on this observation, simulations were limited to this egg length range, as well. For the constant TFs model, both analysis-derived and random initial parameter values were used to maximize the likelihood of finding any parameter set capable of recapitulating the observed allele correlation.

Generation of simulated experimental data

Parameter sets resulting from fitting were sorted in ascending order based on their sum of squared errors, and the 10 lowest error parameter sets are what we called the 10 best parameter sets. For all figures, we simulated 80 spots per bin and simulated each bin 5 times to generate error bars. Data for the distal enhancer at the proximal location was used to reproduce simulated allele correlations in all cases.

Gillespie simulations update the counts of each chemical species at random time intervals. However, for ease of parameter fitting and to better recapitulate the experiments, we generated data in two distinct timescales: one consisting of 30 second intervals after which mRNA counts were recorded, and another consisting of random time intervals generated by the algorithm after which chemical counts were updated. The former one was used for all parameter fitting rounds and generation of figures.

Description of two enhancer model, parameter estimation, and fitting

To explore two enhancer systems, we expanded our previous model to include an additional enhancer. First, we considered duplicated enhancer systems, which consist of either two proximal or two distal enhancers. Enhancers were denoted by E_1 and E_2, which correspond to two identical enhancers that exist in different locations relative to the promoter. They are activated by the same transcription factors as described by the reactions

$$T + E_1 \underset{k_{\mathrm{off}_1}}{\overset{k_{\mathrm{on}_1}}{\rightleftharpoons}} C_1 \xrightarrow{r_1} C_1 + R$$
$$T + E_2 \underset{k_{\mathrm{off}_2}}{\overset{k_{\mathrm{on}_2}}{\rightleftharpoons}} C_2 \xrightarrow{r_2} C_2 + R$$
$$\emptyset \xrightarrow{\beta_1} nT$$
$$T \xrightarrow{\beta_{-1}} \emptyset$$
$$R \xrightarrow{\alpha} \emptyset$$

Without loss of generality, we used E_1 to denote the enhancer at the proximal location and E_2 to denote the enhancer at the distal location. This model describes independent enhancer dynamics; i.e. the behavior of one enhancer does not affect the behavior of the other, and, as such, both enhancers can be simultaneously looped to the promoter. Consequently, to account for potential enhancer interference or competition for the promoter, we assumed distinct k_{on} and k_{off} values for each enhancer in the duplicated enhancer constructs. We also used distinct values of r for each distal enhancer in the duplicated distal construct since fluorescence data was available for this enhancer at the proximal and endogenous location. For proximal enhancers, we assume $r_1 = r_2$.

To describe the dynamics of the shadow enhancer pair, we denoted the activators for E_1 (the proximal enhancer) and E_2 (the distal enhancer) by T_1 and T_2, respectively:

$$T_1 + E_1 \underset{k_{\mathrm{off}_1}}{\overset{k_{\mathrm{on}_1}}{\rightleftharpoons}} C_1 \xrightarrow{r_1} C_1 + R$$
$$T_2 + E_2 \underset{k_{\mathrm{off}_2}}{\overset{k_{\mathrm{on}_2}}{\rightleftharpoons}} C_2 \xrightarrow{r_2} C_2 + R$$
$$\emptyset \xrightarrow{\beta_1} n_1 T_1$$
$$\emptyset \xrightarrow{\gamma_1} n_2 T_2$$
$$T_1 \xrightarrow{\beta_{-1}} \emptyset$$
$$T_2 \xrightarrow{\gamma_{-1}} \emptyset$$
$$R \xrightarrow{\alpha} \emptyset$$

The production rate of T_2, γ_1, was calculated in the same way as production rate of T_1, β_1, but differed in the values of c and σ. The two enhancer models were also used to calculate allele correlation between homozygotes and heterozygotes because a distinction between the mRNA produced by C_1 and C_2 was made. This approach works because, e.g., when considering the homozygote embryos, each single enhancer resides in the same nucleus and is therefore affected by the same fluctuating TF numbers. In the duplicated enhancer model, each enhancer E_1 or E_2 is affected by the same fluctuations in the number of transcription factor T. An analogous logic applies to the heterozygotes.

To fit the two enhancer models to experimental data, we retained several parameters from the single enhancer models. Parameters r and α were directly calculated from the data, and, as such, did not vary across models. We assume that parameters concerning transcription factors (β_1, β_{-1}, γ_1, γ_{-1}, n_1, and n_2) are not affected by the presence of an additional enhancer. Therefore, in our model, only k_{on} and k_{off} are free to change. To fit the values of k_{on1}, k_{on2}, k_{off1}, and k_{off2}, we set the

718 other model parameters to the median values of the 10 best parameter sets in the respective
719 single enhancer model. We then used a similar simulating annealing approach to fit the k_{on} and
720 k_{off} values. We used the resulting values to simulate transcriptional traces and to calculate the
721 predicted CV values shown in Figure 4.

722

723 **Acknowledgements**

724 The authors would like to thank Hernan Garcia for providing advice on MS2 imaging, the MCP-
725 GFP fly line, and a plasmid containing the MS2 reporter construct, Nick Lammers for assistance
726 in the calibration of MS2 signal to mRNA counts, Clarissa Scholes for assistance in image
727 processing, Lily Li for assistance in data analysis and comments on the text, and Thomas
728 Schilling and Rahul Warrior for useful suggestions for the project.

729

730 **Competing Interests**

731 The authors have no competing interests.

732

733 **Funding**

734 This work was supported by NIH grants R00-HD073191, R01-095246, and a Hellman
735 Foundation Fellowship to Z.W, NIH grant T32-EB009418 to A. F., an ARCS Foundation
736 Fellowship to R.W., and NSF grant DMS1763272 and Simons Foundation grant 594598 from
737 the University of California, Irvine, NSF-Simons Center for Multiscale Cell Fate Research to
738 Z.W and G.E.

739

740 **References**

741 Arias, Alfonso Martinez, and Penelope Hayward. 2006. "Filtering Transcriptional Noise
742 during Development: Concepts and Mechanisms." *Nature Reviews Genetics* 7 (1): 34–
743 44. https://doi.org/10.1038/nrg1750.
744 Barolo, Scott. 2012. "Shadow Enhancers: Frequently Asked Questions about Distributed
745 Cis-Regulatory Information and Enhancer Redundancy." *BioEssays : News and*
746 *Reviews in Molecular, Cellular and Developmental Biology* 34 (2): 135–41.
747 https://doi.org/10.1002/bies.201100121.
748 Bertrand, Edouard, Pascal Chartrand, Matthias Schaefer, Shailesh M. Shenoy, Robert H.
749 Singer, and Roy M. Long. 1998. "Localization of ASH1 MRNA Particles in Living
750 Yeast." *Molecular Cell* 2 (4): 437–45. https://doi.org/10.1016/S1097-2765(00)80143-
751 4.
752 Biggin, Mark D. 2011. "Animal Transcription Networks as Highly Connected, Quantitative
753 Continua." *Developmental Cell* 21 (4): 611–26.
754 https://doi.org/10.1016/J.DEVCEL.2011.09.008.

Blake, William J., Mads KÆrn, Charles R. Cantor, and J. J. Collins. 2003. "Noise in Eukaryotic Gene Expression." *Nature* 422 (6932): 633–37. https://doi.org/10.1038/nature01546.

Bomblies, Kirsten, Nicole Dagenais, and Detlef Weigel. 1999. "Redundant Enhancers Mediate Transcriptional Repression of AGAMOUS by APETALA2." *Developmental Biology* 216 (1): 260–64. https://doi.org/10.1006/DBIO.1999.9504.

Bothma, Jacques P, Hernan G Garcia, Samuel Ng, Michael W Perry, Thomas Gregor, and Michael Levine. 2015. "Enhancer Additivity and Non-Additivity Are Determined by Enhancer Strength in the Drosophila Embryo." https://doi.org/10.7554/eLife.07956.001.

Bothma, Jacques P., Matthew R. Norstad, Simon Alamos, and Hernan G. Garcia. 2018. "LlamaTags: A Versatile Tool to Image Transcription Factor Dynamics in Live Embryos." *Cell* 173 (7): 1810-1822.e16. https://doi.org/10.1016/j.cell.2018.03.069.

Cannavò, Enrico, Pierre Khoueiry, David A. Garfield, Paul Geeleher, Thomas Zichner, E. Hilary Gustafson, Lucia Ciglar, Jan O. Korbel, and Eileen E.M. Furlong. 2016. "Shadow Enhancers Are Pervasive Features of Developmental Regulatory Networks." *Current Biology* 26 (1): 38–51. https://doi.org/10.1016/j.cub.2015.11.034.

Chen, Jun, Viola Nolte, and Christian Schlötterer. 2015. "Temperature Stress Mediates Decanalization and Dominance of Gene Expression in Drosophila Melanogaster." *PLoS Genetics* 11 (2): e1004883. https://doi.org/10.1371/journal.pgen.1004883.

Cheung, David, and Jun Ma. 2015. "Probing the Impact of Temperature on Molecular Events in a Developmental System." *Scientific Reports* 5. https://doi.org/10.1038/srep13124.

Csárdi, Gábor, Alexander Franks, David S Choi, Edoardo M Airoldi, and D Allan Drummond. 2015. "Accounting for Experimental Noise Reveals That MRNA Levels, Amplified by Post-Transcriptional Processes, Largely Determine Steady-State Protein Levels in Yeast." *PLoS Genetics* 11 (5): e1005206. https://doi.org/10.1371/journal.pgen.1005206.

Dar, Roy D., Sydney M. Shaffer, Abhyudai Singh, Brandon S. Razooky, Michael L. Simpson, Arjun Raj, and Leor S. Weinberger. 2016. "Transcriptional Bursting Explains the Noise–Versus–Mean Relationship in MRNA and Protein Levels." Edited by Ashok Chauhan. *PLOS ONE* 11 (7): e0158298. https://doi.org/10.1371/journal.pone.0158298.

Dar, Roy D, Brandon S Razooky, Abhyudai Singh, Thomas V Trimeloni, James M McCollum, Chris D Cox, Michael L Simpson, and Leor S Weinberger. 2012. "Transcriptional Burst Frequency and Burst Size Are Equally Modulated across the Human Genome." *Proceedings of the National Academy of Sciences of the United States of America* 109 (43): 17454–59. https://doi.org/10.1073/pnas.1213530109.

Dubuis, Julien O, Reba Samanta, and Thomas Gregor. 2013. "Accurate Measurements of Dynamics and Reproducibility in Small Genetic Networks." *Molecular Systems Biology* 9: 639. https://doi.org/10.1038/msb.2012.72.

Dunipace, Leslie, Anil Ozdemir, and Angelike Stathopoulos. 2011. "Complex Interactions between Cis-Regulatory Modules in Native Conformation Are Critical for Drosophila Snail Expression." *Development* 138: 4566. https://doi.org/10.1242/dev.074377.

El-Sherif, Ezzat, and Michael Levine. 2016. "Shadow Enhancers Mediate Dynamic Shifts of Gap Gene Expression in the Drosophila Embryo." *Current Biology*. Vol. 26. https://doi.org/10.1016/j.cub.2016.02.054.

Elowitz, Michael B., Arnold J. Levine, Eric D. Siggia, and Peter S. Swain. 2002. "Stochastic Gene Expression in a Single Cell." *Science* 297 (5584).

Erdmann, Thorsten, Martin Howard, and Pieter Rein ten Wolde. 2009. "Role of Spatial Averaging in the Precision of Gene Expression Patterns." *Physical Review Letters* 103 (25): 258101. https://doi.org/10.1103/PhysRevLett.103.258101.

Frankel, Nicolás, Gregory K Davis, Diego Vargas, Shu Wang, François Payre, and David L Stern. 2010. "Phenotypic Robustness Conferred by Apparently Redundant Transcriptional Enhancers." *Nature* 466 (7305): 490–93. https://doi.org/10.1038/nature09158.

Fukaya, Takashi, and Michael Levine. 2017. "Transvection." *Current Biology* 27 (19): R1047–49. https://doi.org/10.1016/J.CUB.2017.08.001.

Fukaya, Takashi, Bomyi Lim, and Michael Levine. 2016. "Enhancer Control of Transcriptional Bursting." *Cell* 166 (2): 358–68. https://doi.org/10.1016/j.cell.2016.05.025.

Garcia, Hernan G, Mikhail Tikhonov, Albert Lin, and Thomas Gregor. 2013. "Quantitative Imaging of Transcription in Living Drosophila Embryos Links Polymerase Activity to Patterning." *Current Biology : CB* 23 (21): 2140–45. https://doi.org/10.1016/j.cub.2013.08.054.

Garnett, Aaron T, Tyler A Square, and Daniel M Medeiros. 2012. "BMP, Wnt and FGF Signals Are Integrated through Evolutionarily Conserved Enhancers to Achieve Robust Expression of Pax3 and Zic Genes at the Zebrafish Neural Plate Border." *Development (Cambridge, England)* 139 (22): 4220–31. https://doi.org/10.1242/dev.081497.

Ghiasvand, Noor M, Dellaney D Rudolph, Mohammad Mashayekhi, Joseph A Brzezinski, Daniel Goldman, Tom Glaser, and Tom Glaser. 2011. "Deletion of a Remote Enhancer near ATOH7 Disrupts Retinal Neurogenesis, Causing NCRNA Disease." *Nature Neuroscience* 14 (5): 578–86. https://doi.org/10.1038/nn.2798.

Gregor, Thomas, David W Tank, Eric F Wieschaus, and William Bialek. 2007. "Probing the Limits to Positional Information." *Cell* 130 (1): 153–64. https://doi.org/10.1016/j.cell.2007.05.025.

Hansen, Maike M.K., Ravi V. Desai, Michael L. Simpson, and Leor S. Weinberger. 2018. "Cytoplasmic Amplification of Transcriptional Noise Generates Substantial Cell-to-Cell Variability." *Cell Systems* 7 (4): 384-397.e6. https://doi.org/10.1016/J.CELS.2018.08.002.

He, Feng, Ying Wen, Jingyuan Deng, Xiaodong Lin, Long Jason Lu, Renjie Jiao, and Jun Ma. 2008. "Probing Intrinsic Properties of a Robust Morphogen Gradient in Drosophila." *Developmental Cell* 15 (4): 558–67. https://doi.org/10.1016/J.DEVCEL.2008.09.004.

Hoch, Michael, Christian Schroder1, Eveline Seifert2, and Herbert Jackle. 1990. "Cis-Acting Control Elements for Kruppel Expression in the Drosophila Embryo." *The EMBO Journal* 9 (8): 2587–95. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC552291/pdf/emboj00235-0229.pdf.

Janssen, Sam, Olivier Cuvier, Martin Müller, and Ulrich K Laemmli. 2000. "Specific Gain- and Loss-of-Function Phenotypes Induced by Satellite-Specific DNA-Binding Drugs Fed to Drosophila Melanogaster." *Molecular Cell* 6 (5): 1013–24. https://doi.org/10.1016/S1097-2765(00)00100-3.

Laboulaye, Mallory A., Xin Duan, Mu Qiao, Irene E. Whitney, and Joshua R. Sanes. 2018. "Mapping Transgene Insertion Sites Reveals Complex Interactions Between Mouse Transgenes and Neighboring Endogenous Genes." *Frontiers in Molecular Neuroscience* 11 (October): 385. https://doi.org/10.3389/fnmol.2018.00385.

Lagha, Mounia, Jacques P Bothma, and Michael Levine. 2012. "Mechanisms of Transcriptional Precision in Animal Development." *Trends in Genetics : TIG* 28 (8): 409–16. https://doi.org/10.1016/j.tig.2012.03.006.

Lam, Daniel D, Flavio S J de Souza, Sofia Nasif, Miho Yamashita, Rodrigo López-Leal, Veronica Otero-Corchon, Kana Meece, et al. 2015. "Partially Redundant Enhancers Cooperatively Maintain Mammalian Pomc Expression above a Critical Functional Threshold." *PLoS Genetics* 11 (2): e1004935. https://doi.org/10.1371/journal.pgen.1004935.

Lim, Bomyi, Tyler Heist, Michael Levine, and Takashi Fukaya. 2018. "Visualization of Transvection in Living Drosophila Embryos." *Molecular Cell* 70 (2): 287-296.e6. https://doi.org/10.1016/J.MOLCEL.2018.02.029.

Little, Shawn C, Mikhail Tikhonov, and Thomas Gregor. 2013. "Precise Developmental Gene Expression Arises from Globally Stochastic Transcriptional Activity." *Cell* 154 (4): 789–800. https://doi.org/10.1016/j.cell.2013.07.025.

Liu, Xiaowei, Bo Wu, Jaroslaw Szary, Eric M Kofoed, and Fred Schaufele. 2007. "Functional Sequestration of Transcription Factor Activity by Repetitive DNA." *The Journal of Biological Chemistry* 282 (29): 20868–76. https://doi.org/10.1074/jbc.M702547200.

Lucas, Tanguy, Teresa Ferraro, Baptiste Roelens, Jose De Las Heras Chanes, Aleksandra M. Walczak, Mathieu Coppey, and Nathalie Dostatni. 2013. "Live Imaging of Bicoid-Dependent Transcription in Drosophila Embryos." *Current Biology* 23 (21): 2135–39. https://doi.org/10.1016/J.CUB.2013.08.053.

Miller Jr., Oscar L., and Steven L. McKnight. 1979. "Post-Replicative Nonribosomal Transcription Units in D. Melanogaster Embryos." *Cell* 17 (3): 551–63. https://doi.org/10.1016/0092-8674(79)90263-0.

Osterwalder, Marco, Iros Barozzi, Virginie Tissières, Yoko Fukuda-Yuzawa, Brandon J. Mannion, Sarah Y. Afzal, Elizabeth A. Lee, et al. 2018. "Enhancer Redundancy Provides Phenotypic Robustness in Mammalian Development." *Nature* 554 (7691): 239–43. https://doi.org/10.1038/nature25461.

Perry, Michael W., Alistair N. Boettiger, Jacques P. Bothma, and Michael Levine. 2010. "Shadow Enhancers Foster Robustness of Drosophila Gastrulation." *Current Biology* 20 (17): 1562–67. https://doi.org/10.1016/J.CUB.2010.07.043.

Perry, Michael W., Alistair N. Boettiger, and Michael Levine. 2011. "Multiple Enhancers Ensure Precision of Gap Gene-Expression Patterns in the Drosophila Embryo." *Proceedings of the National Academy of Sciences* 108 (33): 13570–75. https://doi.org/10.1073/PNAS.1109873108.

888 Perry, Michael W., Jacques P. Bothma, Ryan D. Luu, and Michael Levine. 2012. "Precision
889 of Hunchback Expression in the Drosophila Embryo." Current Biology 22 (23): 2247–
890 52. https://doi.org/10.1016/J.CUB.2012.09.051.
891 Preger-Ben Noon, Ella, Gonzalo Sabarís, Daniela M Ortiz, Jonathan Sager, Anna
892 Liebowitz, David L Stern, and Nicolás Frankel. 2018. "Comprehensive Analysis of a
893 Cis-Regulatory Region Reveals Pleiotropy in Enhancer Function." *Cell Reports* 22
894 (11): 3021–31. https://doi.org/10.1016/j.celrep.2018.02.073.
895 Preiss, A, U B Rosenberg, A Kienlin, E Seifert, and H Jäckle. n.d. "Molecular Genetics of
896 Krüppel, a Gene Required for Segmentation of the Drosophila Embryo." *Nature* 313
897 (5997): 27–32. Accessed January 11, 2019.
898 http://www.ncbi.nlm.nih.gov/pubmed/3917552.
899 Raj, Arjun, Scott A Rifkin, Erik Andersen, and Alexander van Oudenaarden. 2010.
900 "Variability in Gene Expression Underlies Incomplete Penetrance." *Nature* 463
901 (7283): 913–18. https://doi.org/10.1038/nature08781.
902 Raser, Jonathan M, and Erin K O'Shea. 2005. "Noise in Gene Expression: Origins,
903 Consequences, and Control." *Science (New York, N.Y.)* 309 (5743): 2010–13.
904 https://doi.org/10.1126/science.1105891.
905 Sanchez, Alvaro, and Ido Golding. 2013. "Genetic Determinants and Cellular Constraints in
906 Noisy Gene Expression." *Science (New York, N.Y.)* 342 (6163): 1188–93.
907 https://doi.org/10.1126/science.1242975.
908 Scholes, Clarissa, Kelly M Biette, Timothy T Harden, and Angela H DePace. 2019. "Signal
909 Integration by Shadow Enhancers and Enhancer Duplications Varies across the
910 Drosophila Embryo." *Cell Reports* 26 (9): 2407-2418.e5.
911 https://doi.org/10.1016/j.celrep.2019.01.115.
912 Stapel, L Carine, Christoph Zechner, and Nadine L Vastenhouw. 2017. "Uniform Gene
913 Expression in Embryos Is Achieved by Temporal Averaging of Transcription Noise."
914 *Genes & Development* 31 (16): 1635–40. https://doi.org/10.1101/gad.302935.117.
915 Thompson, A, and M J Gasson. 2001. "Location Effects of a Reporter Gene on Expression
916 Levels and on Native Protein Synthesis in Lactococcus Lactis and Saccharomyces
917 Cerevisiae." *Applied and Environmental Microbiology* 67 (8): 3434–39.
918 https://doi.org/10.1128/AEM.67.8.3434-3439.2001.
919 Wunderlich, Zeba, Meghan D J Bragdon, Ben J Vincent, Jonathan A White, Javier Estrada,
920 and Angela H DePace. 2015. "Krüppel Expression Levels Are Maintained through
921 Compensatory Evolution of Shadow Enhancers." *Cell Reports* 12 (11): 1740–47.
922 https://doi.org/10.1016/j.celrep.2015.08.021.
923 Yan, Jenny, Caitlin Anderson, Kayla Viets, Sang Tran, Gregory Goldberg, Stephen Small,
924 and Robert J Johnston. 2017. "Regulatory Logic Driving Stable Levels of Defective
925 Proventriculus Expression during Terminal Photoreceptor Specification in Flies."
926 *Development (Cambridge, England)* 144 (5): 844–55.
927 https://doi.org/10.1242/dev.144030.
928 Yuan, Ye, Lei Wei, Tao Hu, Shuailin Li, Tianrun Cheng, Jinzhi Lei, Zhen Xie, Michael Q.
929 Zhang, and Xiaowo Wang. 2018. "Quantitative Understanding of Molecular
930 Competition as a Hidden Layer of Gene Regulatory Network." *BioRxiv*, February,
931 258129. https://doi.org/10.1101/258129.

Zenklusen, Daniel, Daniel R Larson, and Robert H Singer. 2008. “Single-RNA Counting Reveals Alternative Modes of Gene Expression in Yeast.” *Nature Structural & Molecular Biology* 15 (12): 1263–71. https://doi.org/10.1038/nsmb.1514.

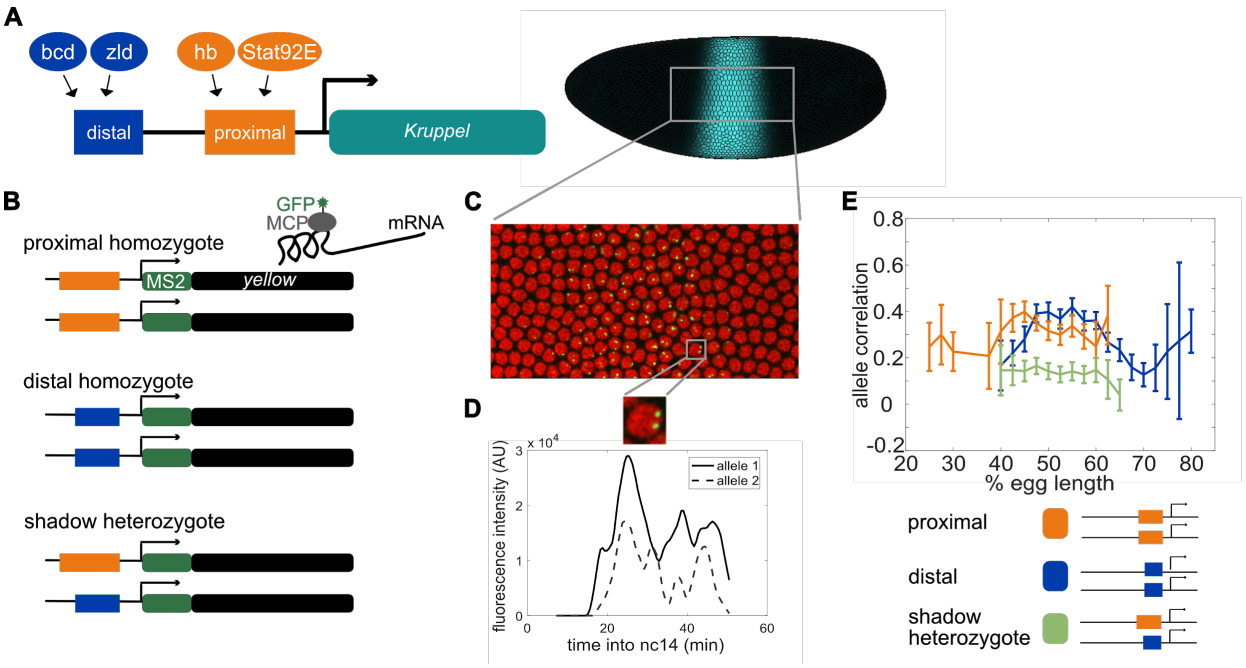


Figure 1: Dual allele imaging shows the individual *Kruppel* enhancers drive largely independent transcriptional dynamics. A. Schematic of the endogenous *Kruppel* locus with distal (blue) and proximal (orange) shadow enhancers driving *Kr* (teal) expression in the central region of the embryo. Known transcriptional activators of the two enhancers are shown. **B**. Schematics of single enhancer reporter constructs driving expression of MS2 and a *yellow* reporter. When transcribed, the MS2 sequence forms stem loops that are bound by GFP-tagged MCP expressed in the embryos. Proximal embryos have expression on each allele controlled by the 1.5kb proximal enhancer at its endogenous spacing from the *Kr* promoter, while distal embryos have expression on each allele controlled by the 1.1kb distal enhancer at the same spacing from the *Kr* promoter. Shadow heterozygote embryos have expression on one allele controlled by the proximal enhancer and expression on the other allele controlled by the distal enhancer. **C**. Still frame from live imaging experiment where nuclei are red circles and active sites of transcription are green spots. MCP-GFP is visible as spots above background at sites of nascent transcription (Garcia, et al., 2013). **D.** The fluorescence of each allele in individual nuclei can be tracked across time as a measure of transcriptional activity. Graph shows a representative trace of transcriptional activity of the two alleles in a single nucleus across the time of nc14. These traces are used to calculate the correlation of allele activity in each nucleus. Correlation values are grouped by position of the nucleus along the egg length and averaged across all imaged nuclei in all embryos of each construct. **E**. Graph of average correlation between the two alleles in each nucleus as a function of egg length. 0% egg length corresponds to the anterior end. Error bars indicate 95% confidence intervals. The shadow heterozygotes have much lower allele correlation than either homozygote, demonstrating that the individual shadow enhancers drive nearly independent transcriptional activity and that upstream fluctuations in regulators are a significant driver of transcriptional bursts. The total number of nuclei used in calculations for each construct by AP bin are given in Supplementary Table 2.

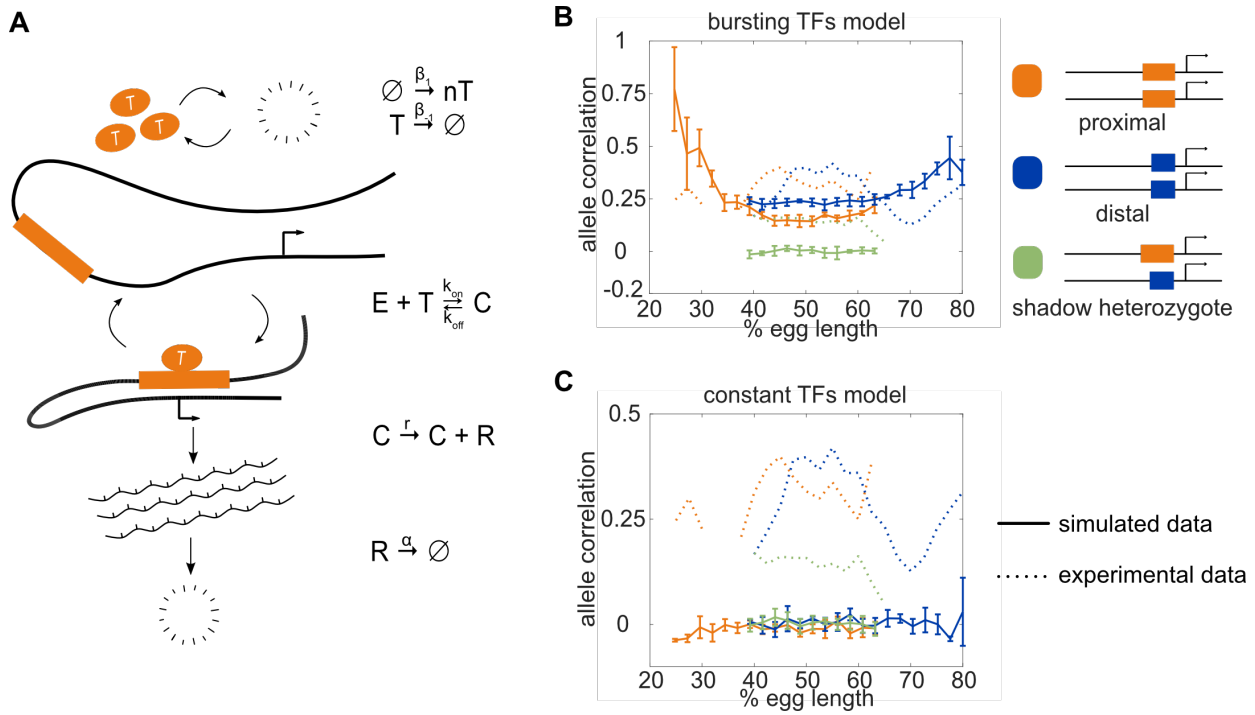


Figure 2: Model of enhancer-driven dynamics demonstrates TF fluctuations are required for correlated reporter activity. To investigate the factors required for the observed correlated behavior of identical enhancers and largely independent behavior of the individual enhancers, we developed a simple stochastic model of enhancer-driven transcription. **A**. Schematic of model of transcription driven by a single enhancer (the *bursting TFs* model). For each enhancer, we assume there is a single activating TF, T, that appears in bursts of size *n* molecules at a rate β_1, which varies by the position in the embryo. TFs degrade linearly at rate β_{-1}. When present, T can bind the enhancer, E, to form a transcriptionally active complex, C, at a rate k_{on} and dissociates at rate k_{off}. This complex then produces mRNA at an experimentally determined rate *r* that degrades at an experimentally determined rate, α. **B**. The bursting TFs model is able to recapitulate the experimentally observed pattern of allele correlation. We plot the correlation between the two alleles in a nucleus as a function of egg length. Simulated data is created using the lowest energy parameter set for each enhancer. The data shown is the average of five simulated embryos that have 80 transcriptional spots per AP bin. In B and C simulated data are shown by solid lines, experimental data are shown by dotted lines. **C**. The constant TFs model fails to recapitulate the experimentally observed pattern of allele correlation. Without TF fluctuations, both heterozygous and homozygous embryos display independent allele activity. Error bars in B and C represent 95% confidence intervals.

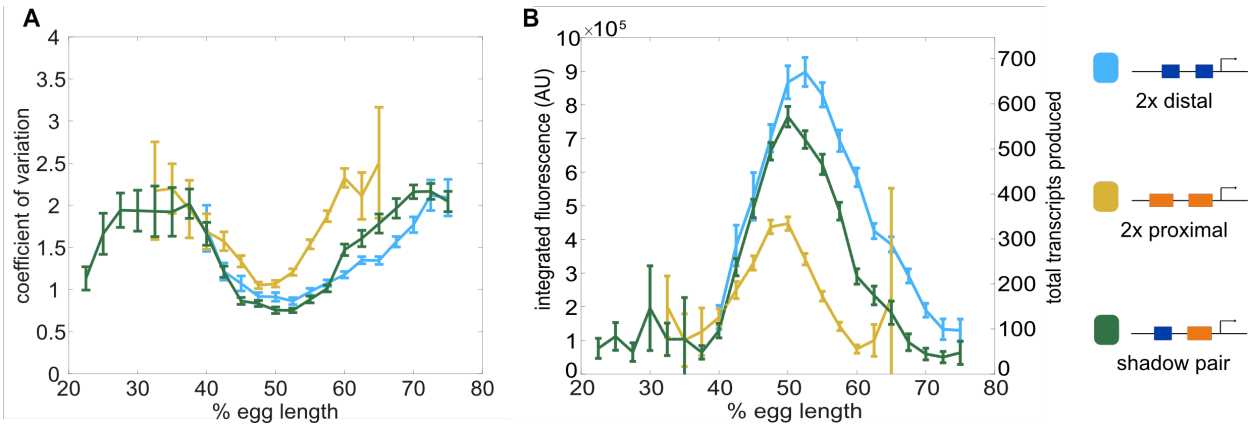


Figure 3: Shadow enhancer pair produces lower expression noise than duplicated enhancers. To investigate whether the shadow enhancer pair drives less noisy expression, we calculate the coefficient of variation (CV) associated with the shadow enhancer pair or either duplicated enhancer across time of nc14. **A**. The shadow enhancer pair displays lower temporal expression noise than either duplicated enhancer. Graph is mean coefficient of variation of fluorescence traces across time as a function of embryo position. **B**. The shadow enhancer pair shows the lowest expression noise, but not the highest expression levels, indicating that the lower noise is not simply a function of higher expression. Graph is average total expression during nc14 as a function of embryo position. Error bars in A and B represent 95% confidence intervals. Total number of transcriptional spots used for graphs are given in Supplementary Table 1 by construct and AP bin.

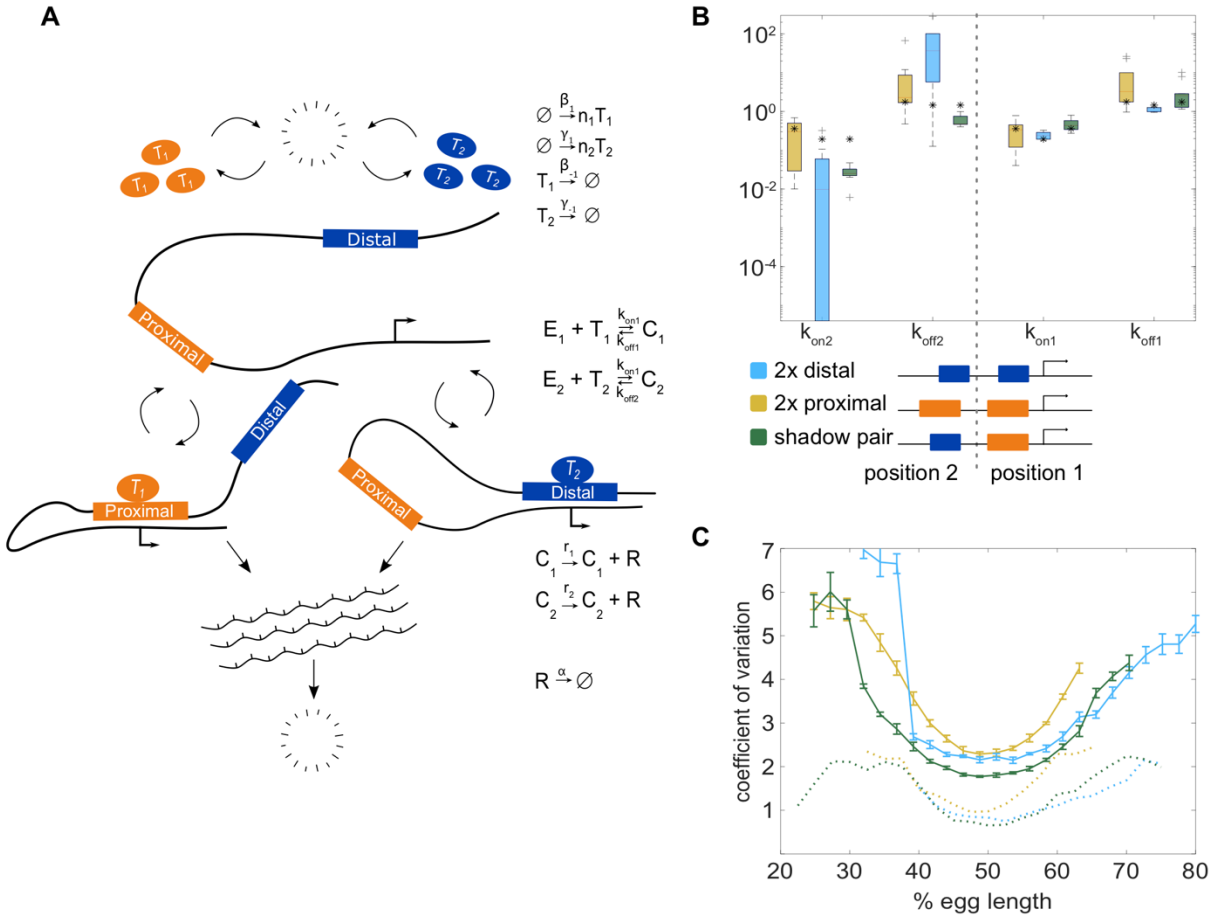


Figure 4: The two enhancer model recapitulates low expression noise associated with the shadow enhancer pair. To assess whether the separation of input TFs mediates the lower expression noise driven by the shadow enhancer pair, we expanded our model to incorporate two enhancers driving transcription. **A**. Schematic of the two enhancer model. We assume that when two enhancers control a single promoter, either or both can loop to the promoter and drive transcription. We defined model parameters as in Figure 2, and only allowed the k_{on} and k_{off} values to vary from the single enhancer model. **B**. To understand the effect of adding a second enhancer, we examined how the k_{on} and k_{off} values vary from those in the single enhancer model. We plotted the distribution of the values for k_{on} and k_{off} for each enhancer in the three different constructs measured. The distribution shows the values derived from the 10 best-fitting parameter sets, and the black star in each column indicates the k_{on} or k_{off} value from the corresponding single enhancer model. In general, the k_{off} values increased relative to the single enhancer model, and the k_{on} values decreased, indicating that the presence of a second enhancer inhibits the activity of the first. **C.** Graph of average coefficient of variation of simulated or experimental transcriptional traces as a function of egg length. The model is able to recapitulate the lower expression noise seen with the shadow enhancer pair with no additional fitting, indicating that the separation of TF inputs to the two enhancers is sufficient to explain this observation. Simulated data are shown in solid lines, experimental data are shown in dotted lines.

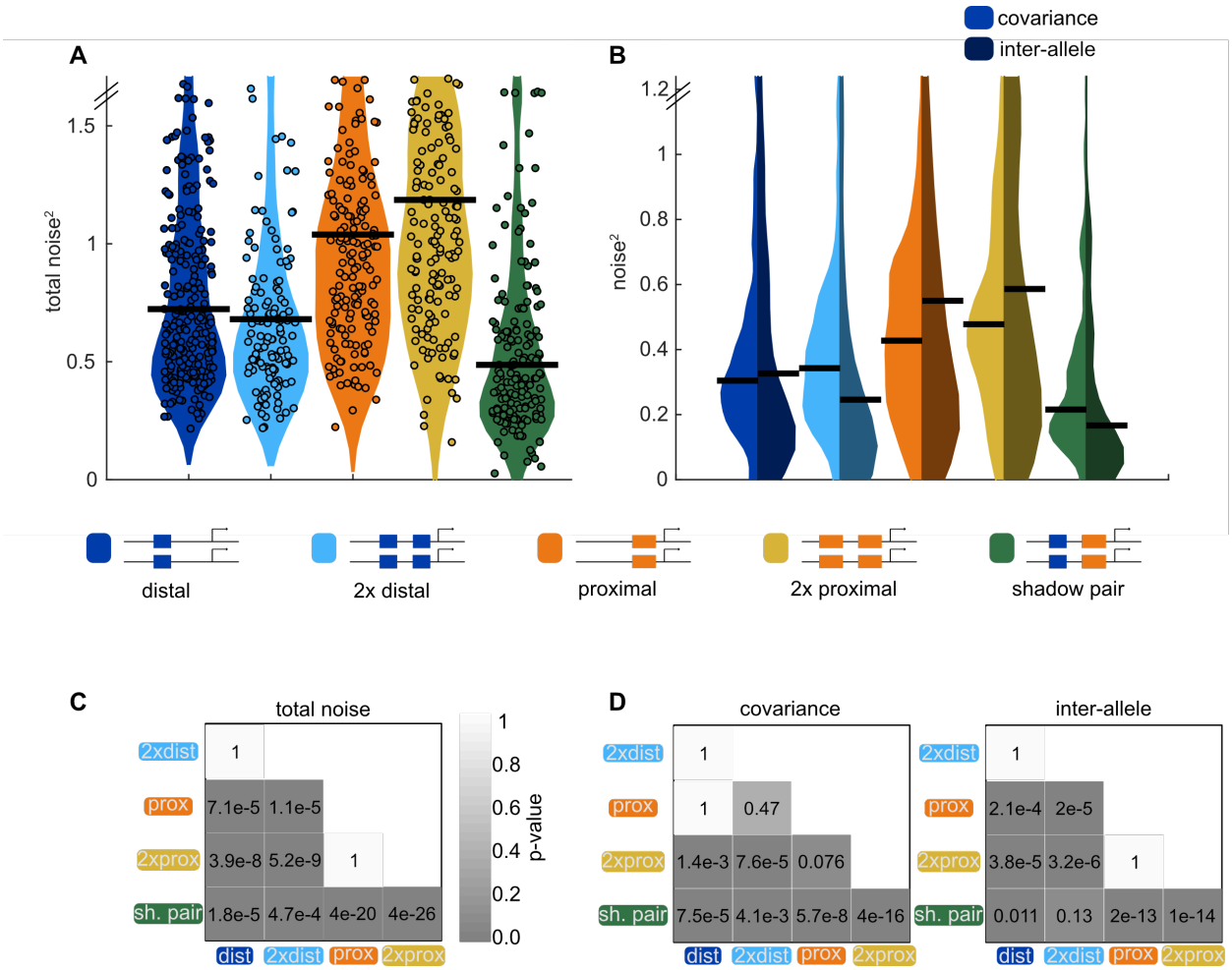


Figure 5: Shadow enhancer pair achieves lower total noise by buffering global and allele-specific sources of noise. To determine how the shadow enhancer pair produces lower expression noise, we calculated the total noise associated with each enhancer construct and decomposed this into the contributions of covariance and inter-allele noise. Covariance is a measure of how the activities of the two alleles in a nucleus change together and is indicative of global sources of noise. Inter-allele noise is a measure of how the activities of the two alleles differ and is indicative of allele-specific sources of noise. **A.** The shadow enhancer pair has lower total noise than single or duplicated enhancers. Circles are total noise values for individual nuclei in AP bin of peak expression for the given enhancer construct. Horizontal line represents the median. The y-axis is limited to 75th percentile of the proximal enhancer, which has the largest noise values. The shadow enhancer pair has significantly lower total noise than all other constructs. **B.** The shadow enhancer pair displays significantly lower covariance than either single or duplicated enhancer and significantly lower inter-allele noise than both single enhancers and the duplicated proximal enhancer. The left half of each violin plot shows the distribution of covariance values of nuclei in the AP bin of peak expression, while the right half shows the distribution of inter-allele noise values. Horizontal lines represent median. The y-axis is again limited to the 75th percentile of enhancer with the largest noise values, which is duplicated proximal. The lower covariance and inter-allele noise associated with the shadow enhancer pair indicates it is better able to buffer both global and allele-specific sources of noise.

C. *p*-value table of Kruskal-Wallis pairwise comparison of the total noise values of each enhancer construct. *p*-value gradient legend applies to C and D. **D.** *p*-value table of Kruskal-Wallis pairwise comparison of covariance (on left) and inter-allele noise (on right) values for each enhancer construct. Bonferroni multiple comparison corrections were used for *p*-values in C and D. Total number of nuclei used in noise calculations are given in Supplementary Table 2.

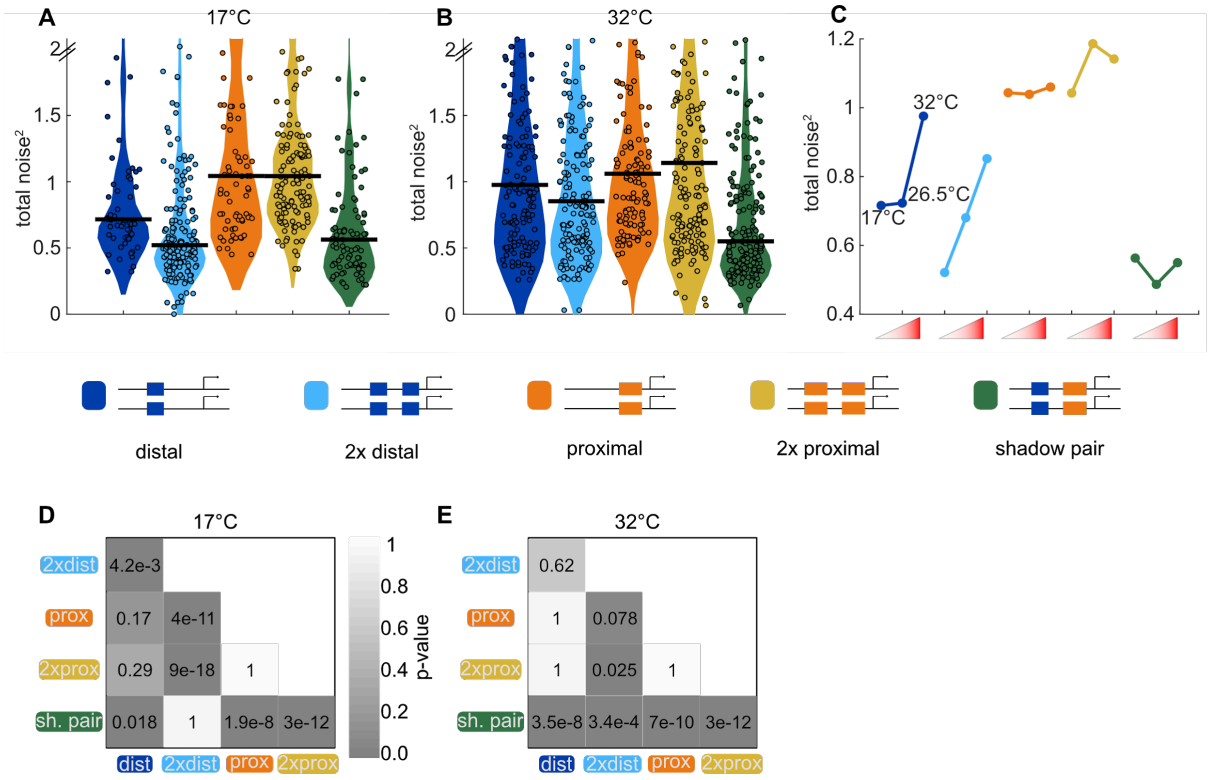
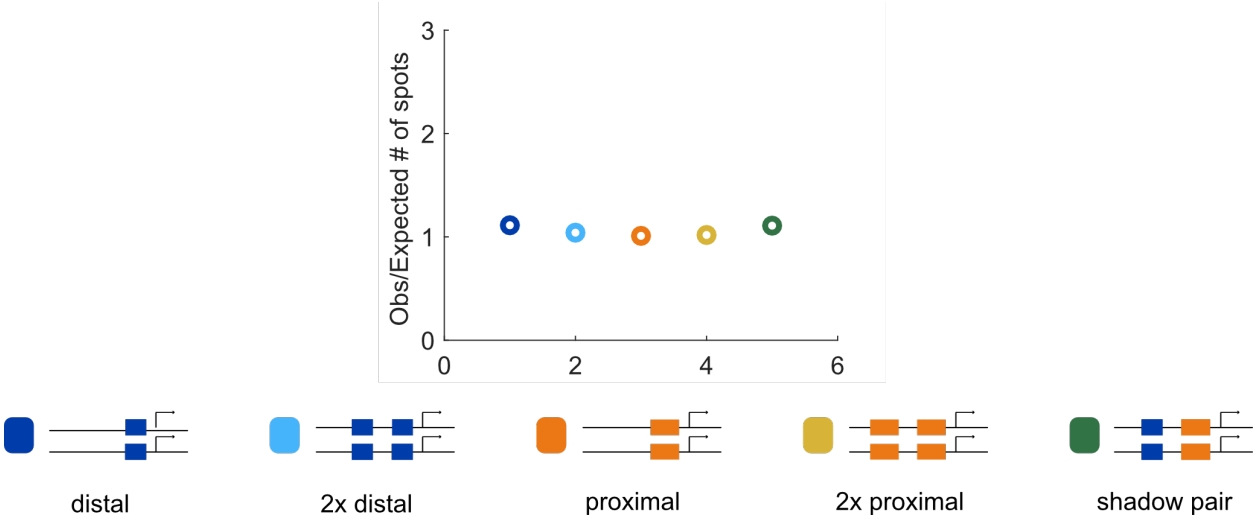
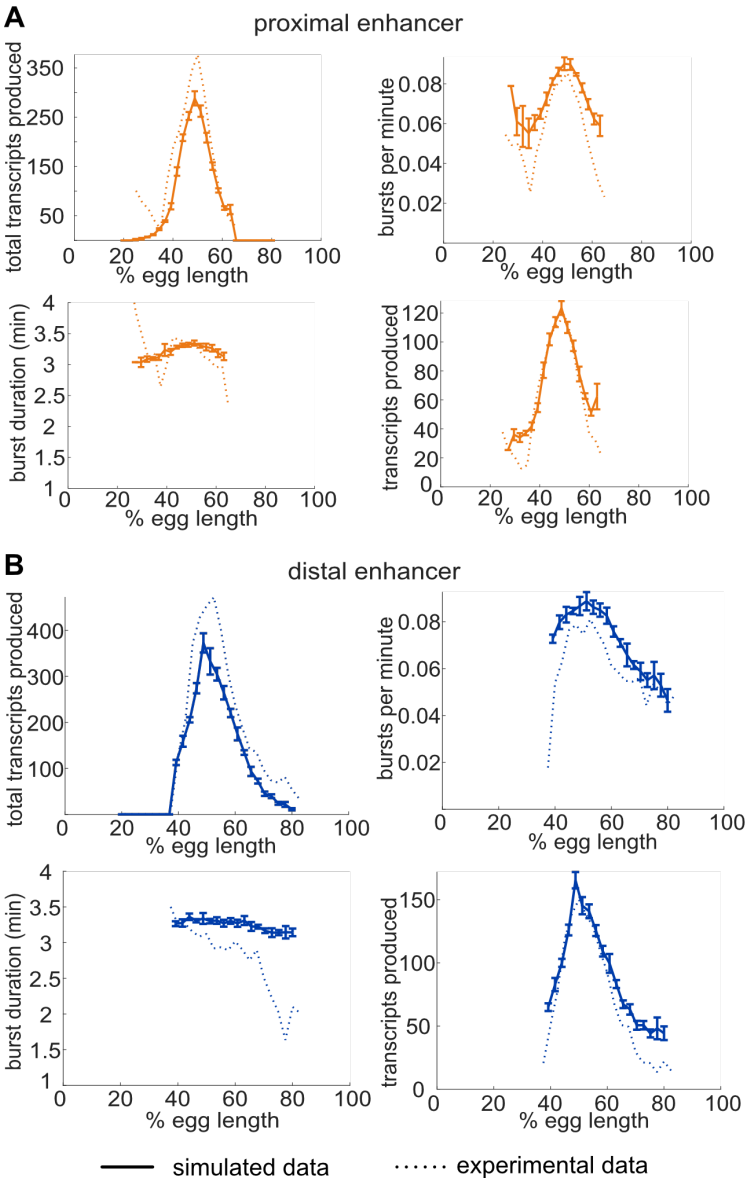


Figure 6: Shadow enhancer pair maintains lower total noise across temperature perturbations. To test the ability of each enhancer construct to buffer temperature perturbations, we measured the total expression noise associated with each for embryos imaged at 17°C or 32°C. **A**. The shadow enhancer pair displays significantly lower total noise than the single or duplicated proximal enhancer and the single distal enhancer at 17°C. Circles are total noise values for individual nuclei in AP bin of peak expression for the given enhancer construct and horizontal bars represent medians. The y-axis is limited to 75th percentile of construct with highest total noise at 17°C (single proximal). **B**. The shadow enhancer pair has significantly lower total noise than all other constructs at 32°C. The y-axis is limited to 75th percentile of the enhancer construct with highest total noise at 32°C (duplicated proximal). **C.** Temperature changes have different effects on the total noise associated with the different enhancers. The median total noise value at the AP bin of peak expression at the three measured temperatures is shown for each enhancer construct. Within each enhancer, the median total noise values are shown left to right for 17°C, 26.5°C, and 32°C. **D**. *p*-value table of Kruskal-Wallis pairwise comparison of the total noise values of each enhancer construct at 17°C. *p*-value gradient legend applies to D and E. **E**. *p*-value table of Kruskal-Wallis pairwise comparison of the total noise values of each enhancer construct at 32°C.

Supplemental Figures:



Supplemental Figure 1: Correspondence of observed and expected number of spots. To ensure that we can accurately measure two spots of expression in the embryo, we compared the number of transcriptional spots seen in embryos hemizygous or homozygous for each construct. Our rationale was that in the absence of transvection, the number of transcriptional spots in homozygous embryos should be twice the number in embryos expressing the reporter on only one allele. The number of transcriptional spots tracked during nc14 in the AP bin of maximum expression was counted for all embryos imaged for each homozygous and hemizygous construct. The graph shows the average of this value for homozygous embryos, divided by double the value observed in the corresponding hemizygous construct. Assuming no transvection occurs, this value should be close to 1. The ratio of observed to expected number of spots is close to 1 for all of our enhancer constructs, indicating we are reliably able to track the two individual spots of transcription in single nuclei.



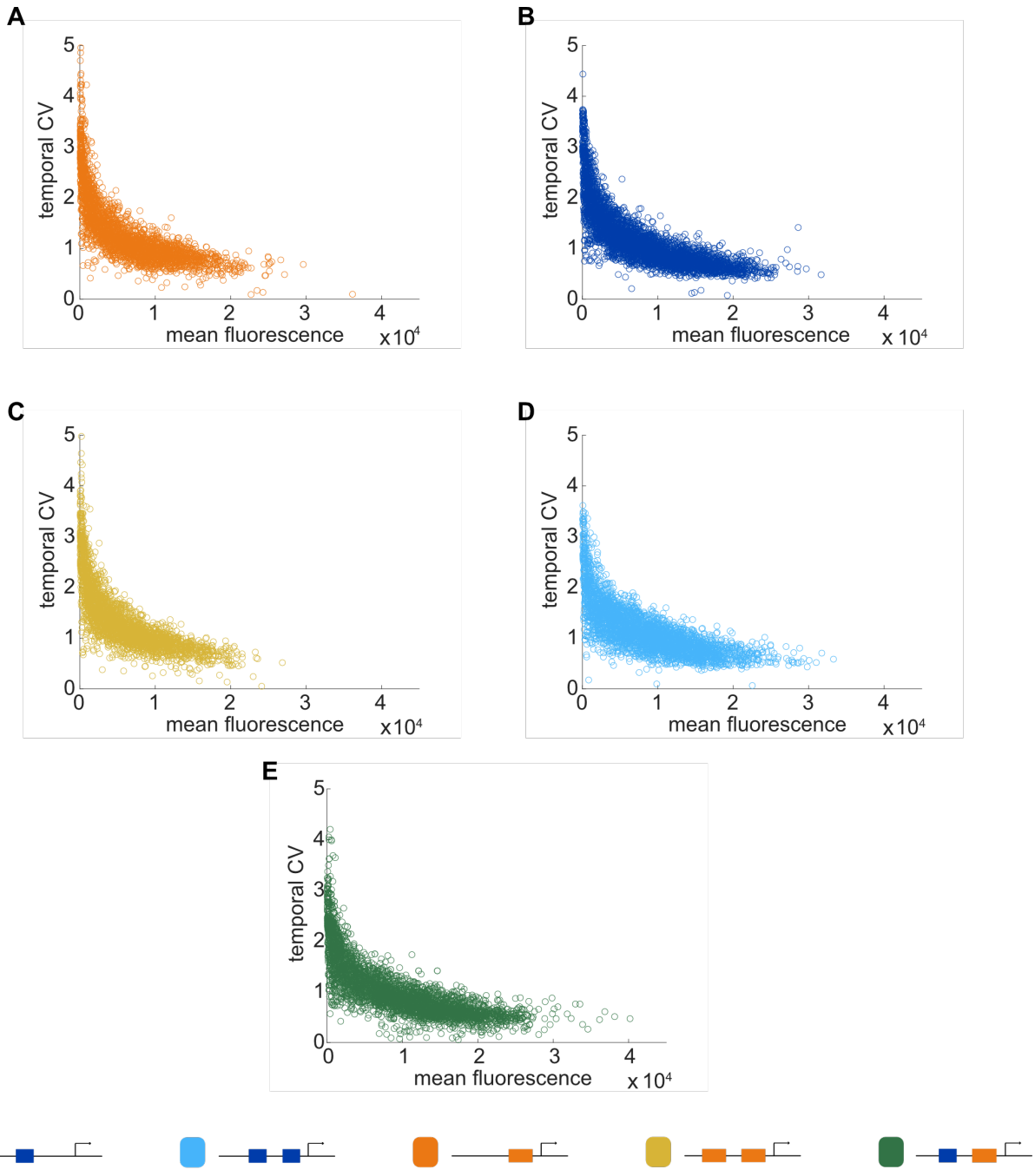
C

	median	CV
β_{-1}	2.7	2.4
k_{off}	1.8	0.40
k_{on}	0.36	0.90
n	4	2.0
c	6.8	0.75
σ	8.1	0.08

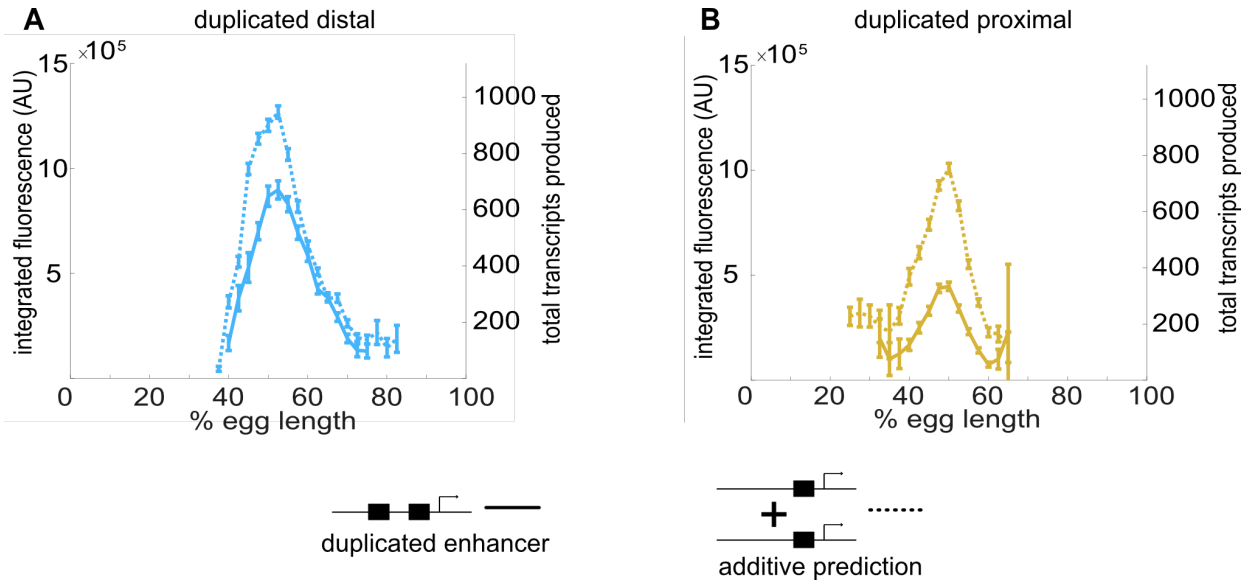
D

	median	CV
β_{-1}	3.9	1.2
k_{off}	1.5	0.13
k_{on}	0.19	1.3
n	12.5	1.3
c	9.2	0.34
σ	12.5	0.13

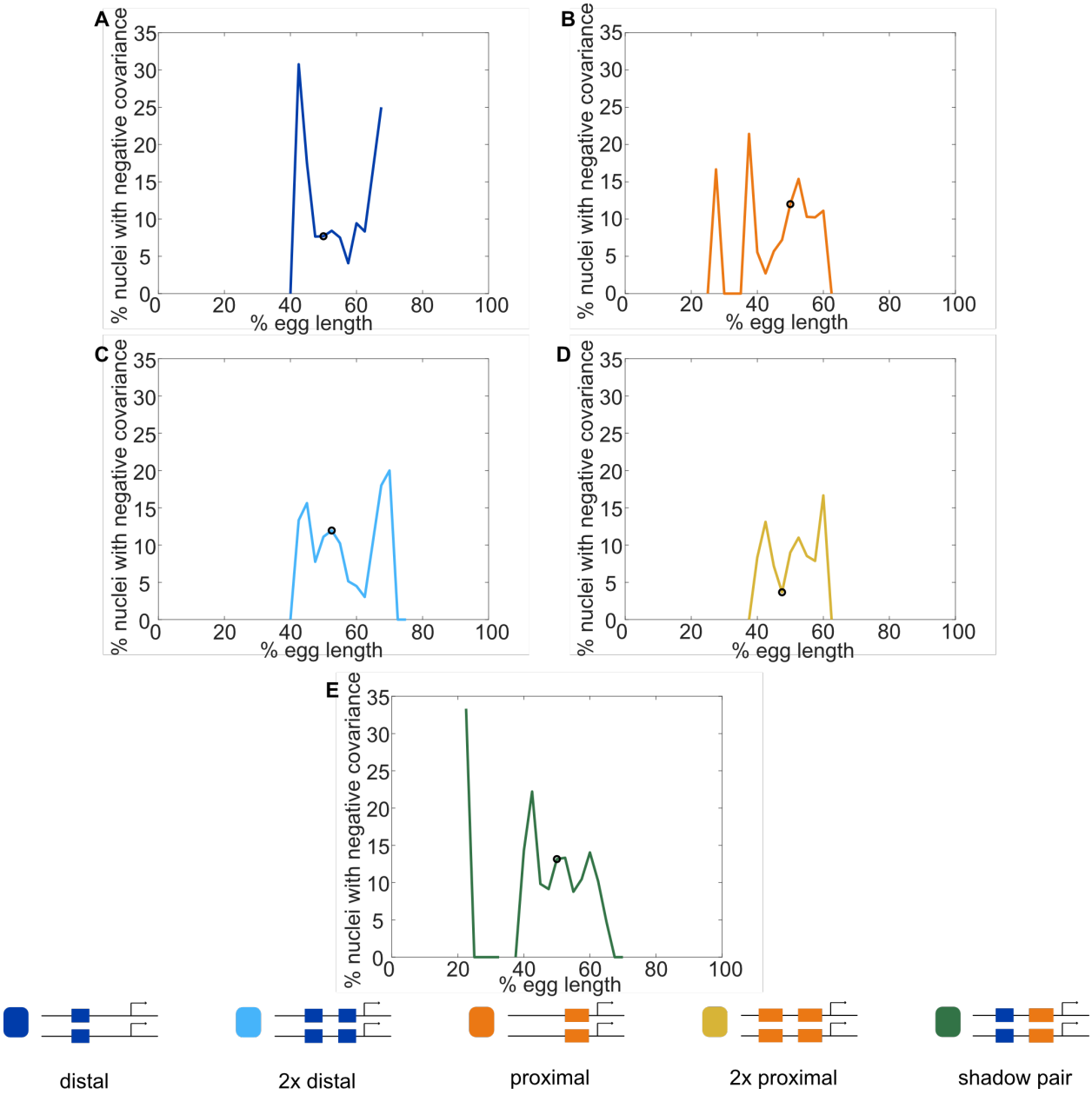
Supplemental Figure 2: Single enhancer models recreate observed transcriptional bursting properties. To investigate whether our model is accurately simulating our experimental system, we compared the transcriptional burst properties produced by model simulations of transcription to those observed experimentally (see Supplementary Figure 7 for description of burst properties). **A**. Graphs of average values of transcriptional burst properties, total mRNA produced during nc14, burst frequency, burst duration, and burst size associated with the proximal enhancer as a function of egg length. In A and B, simulated data are represented with solid lines and experimental data are shown with dotted lines. **B**. Graphs of average values of transcriptional burst properties as in A, associated with the distal enhancer. For both the proximal and distal enhancers, our model is largely able to recapitulate the experimentally observed transcriptional burst properties associated with each enhancer. **C**. The median and CV values of the model parameters for the proximal enhancer in the top 10 performing parameter sets. **D**. The median and CV values of the model parameters for the distal enhancer in the top 10 performing parameter sets. Explanations of model parameters are given in the Methods.



Supplemental Figure 3: Temporal CV as a function of mean fluorescence. To investigate the relationship between our noise measurement of temporal CV and the mean activity of each construct, we plotted the temporal CV of each transcription spot as a function of its mean fluorescence. **A**. Distal; **B**. Proximal; **C**. 2x Proximal; **D**. 2x Distal; **E**. Shadow pair. With all constructs, we find the general trend that CV decreases with increasing average expression, flattening out at a baseline noise level specific to each enhancer construct.

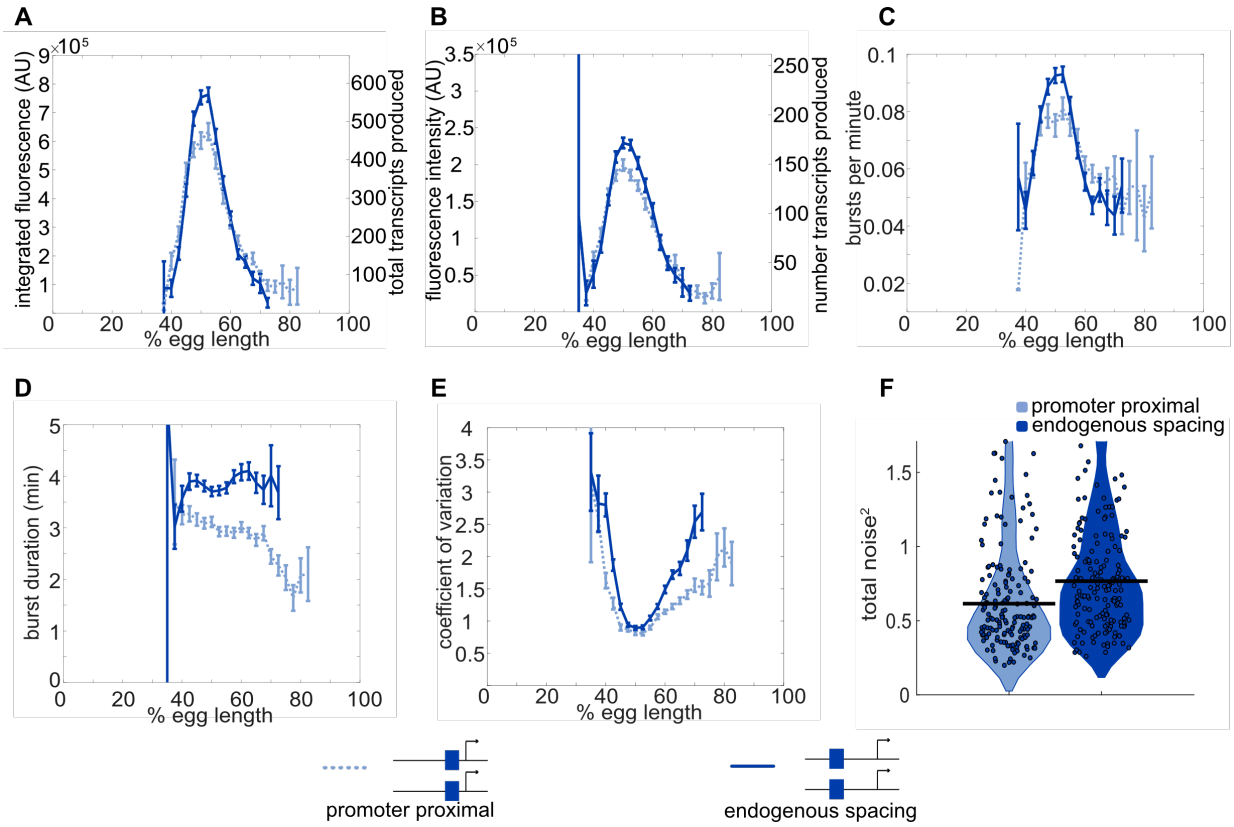


Supplemental Figure 4: Individual *Kr* enhancers display sub-additive behavior. To assess the way input from two enhancers is integrated at the Kr promoter, we compared the experimentally observed mRNA production of duplicated enhancers to that predicted from additive behavior of the single enhancers. **A**. The duplicated distal enhancer displays sub-additive behavior. The solid line is the experimentally observed total mRNA produced by the duplicated distal enhancer during nc14 as a function of egg length and the dotted line is that expected by doubling the total mRNA produced by the single distal enhancer. **B**. The duplicated proximal enhancer also acts sub-additively. The solid line is the experimentally observed total mRNA produced by the proximal enhancer during nc14 as a function of egg length and the dotted line is that expected by doubling the total mRNA produced by the single proximal enhancer. These results, along with the observation that k_{off} values increased and k_{on} values decreased in our model with the addition of a second enhancer, suggests that the *Kr* enhancers compete with each other for interactions with the promoter.



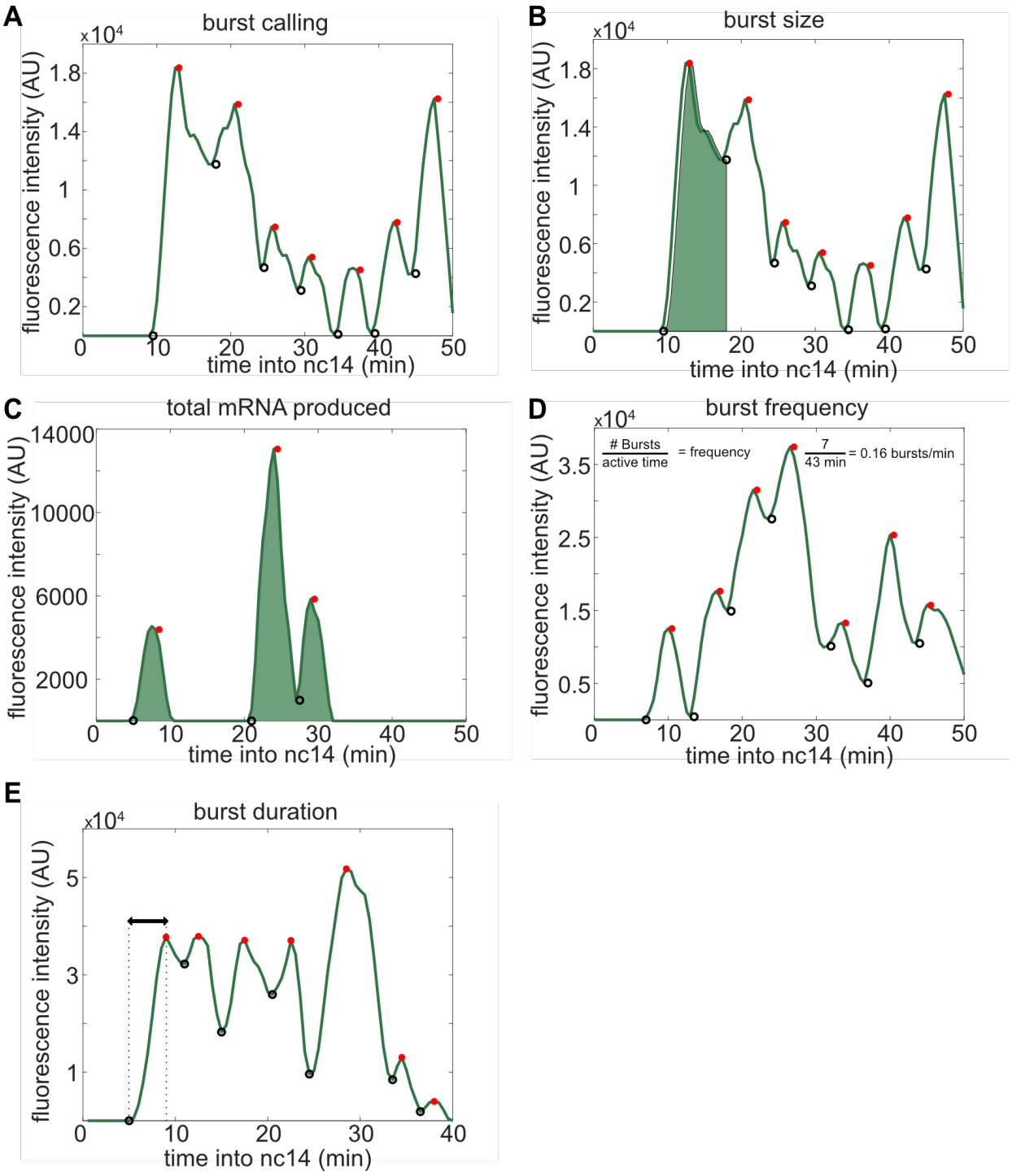
Supplemental Figure 5: Fraction of nuclei with negative covariance of allele activity. To identify the likely cause of the observed negative covariance between allele activity in some nuclei, we calculated the fraction of nuclei displaying negative covariance out of all nuclei that had active reporter transcription. Graphs show the fraction of transcribing nuclei with negative covariance as a function of egg length for each reporter construct, with a black circle indicating the position along the embryo of maximal expression for that construct. **A**. Distal; **B**. Proximal; **C**. 2x Proximal; **D**. 2x Distal; **E**. Shadow pair. Note that for all constructs, the highest rates of negative covariance are outside of the region of maximal reporter expression. MCP-GFP is expressed uniformly along the length of the embryo and we would therefore expect if MCP-GFP were the limiting factor, we would see the highest rates of negative covariance in the center of the expression pattern, where the highest number of transcripts are produced. Instead, the highest

rates of negative covariance are seen at the edges of the *Kr* expression pattern, suggesting a spatially patterned factor, such as a TF, may be what is limiting.



Supplemental Figure 6: Position-dependent effects on distal enhancer. To best mimic the endogenous system, we looked at expression driven by the distal enhancer at its endogenous spacing from the promoter for our noise calculations. In this construct we replaced the sequence of the proximal enhancer with sequence of the same length from the lambda phage genome predicted to have low number of *Drosophila* TF binding sites. This increased distance from the promoter had observable effects on the transcriptional dynamics and noise associated with the distal enhancer. **A**. Comparison of total transcriptional expression mediated by the distal enhancer at its endogenous spacing or proximal to the promoter. The distal enhancer at its endogenous spacing, shown as the solid line, produces significantly more total mRNA in the center region of expression than the distal enhancer proximal to the promoter, shown as the dotted line. **B**. Comparison of the average number of transcripts produced per transcriptional burst by each distal enhancer configuration as a function of egg length. **C**. Average burst frequency associated with either distal enhancer configuration as a function of egg length. **D**. Average burst duration associated with either distal enhancer configuration as a function of egg length. **E**. Coefficient of variation of transcriptional activity across nc14 for each distal enhancer configuration as a function of egg length. **F**. Total expression noise associated with either distal enhancer configuration at the AP bin of that construct's peak expression. The total noise distribution for the distal enhancer proximal to the promoter is on the left and that for the distal enhancer at its endogenous spacing from the promoter is on the right. The distal enhancer at its endogenous spacing displays significantly higher total noise ($p = 0.018$) than the distal enhancer proximal to the promoter. Each circle represents the total noise of an individual nucleus and the

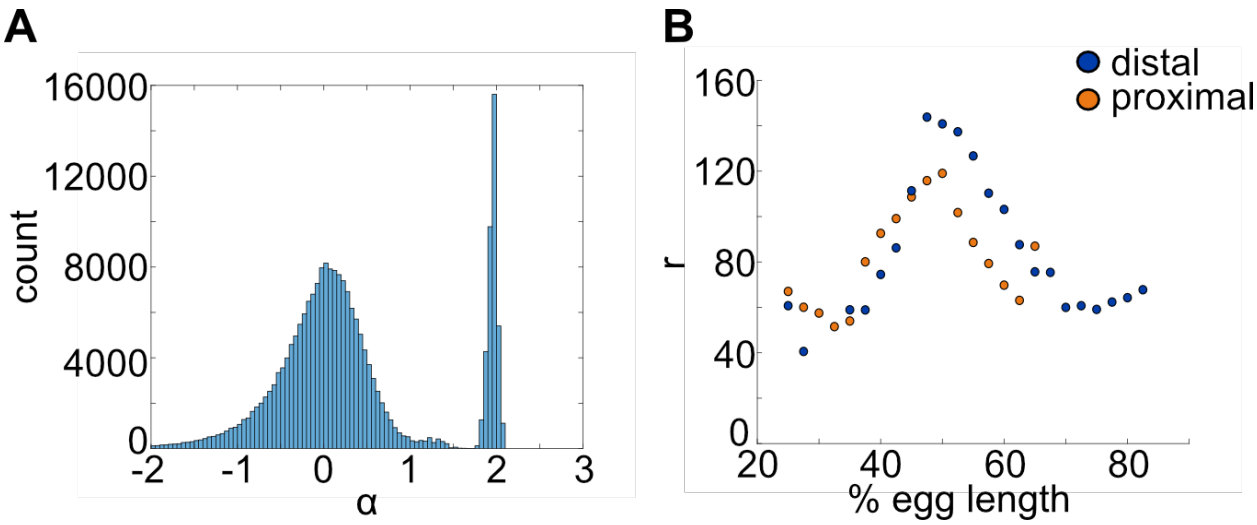
1147 horizontal bar marks the median total noise value. Y-axis limited to the 75th percentile of the
1148 construct with the highest total noise values (distal promoter at endogenous spacing).



Supplemental Figure 7: Visual inspection of burst calling algorithm

To extract the bursting parameters examined (burst size, frequency, and duration), individual fluorescence traces were first smoothed using the LOWESS method with a span of 0.1. Our burst calling algorithm then determined the periods of promoter activity or inactivity based on the slope of the fluorescence trace. **A**. Representative fluorescence trace of a single spot across the

time of nc14. Black open circles indicate time points where the promoter is called “on”, red filled circles indicate time points where the promoter is called “off”. **B**. Same trace as in A with shading representing the area under the curve used to calculate the size of the first burst. This area is calculated using the trapz function in MATLAB and is done for each burst, from the time point the promoter is called “on” until the next time it is called “on”. C-E show additional representative fluorescence traces of single transcriptional spots across the time of nc14. **C**. A trace with shading representing the area under the entire curve during nc14 used to calculate the total amount of mRNA produced. This area is calculated using the trapz function in MATLAB and is done from the time the promoter is first called active until 50 minutes into nc14 or the movie ends, whichever comes first. **D**. Burst frequency is calculated by dividing the number of bursts that occur from the time the promoter is first called active until 50 minutes into nc14 or the movie ends, whichever comes first. **E**. Burst duration is defined as the amount of time between when the promoter is called active and it is next called inactive.



Supplemental Figure 8: mRNA production and decay rates can be directly estimated from experimental data. The mRNA degradation parameter α and production parameter *r* were measured directly from fluorescence data without any input from the model. **A**. To estimate α, we used adjacent measurements of fluorescence intensity to approximate the slope at each point in the fluorescence traces. These values are compared with an exponential rate of mRNA decay (see Methods) and the resulting predicted values are shown in the histogram. Periods of mRNA production have negative α values and periods of decay have positive values. The histogram shows a distinct peak for $\alpha > 0$, which provided us with an estimate of $\alpha \approx 1.95$. **B.** A similar computational approach was used to calculate values of *r* from fluorescence data (see Methods). We calculated different values of *r* for each bin to account for differences in transcriptional efficiency across the length of the embryo due to factors that are not explicitly included in the model. For example, different combinations of TF bound to the enhancer may give rise to different mRNA production rates. Different values of *r* were found for the proximal and distal enhancers. Notice that distal *r* values shown correspond to the distal enhancer at the proximal location.

Additional Supplementary Materials

Supplementary Note: A note describing the theoretical estimates of inter-allele noise in single and two enhancer constructs.

Supplementary Table 1: Number of total single alleles tracked for each construct.

Each row corresponds to a construct, named in column 42, and columns 1-41 correspond to that AP bin of the embryo. The value in each cell in columns 1-41 is the number of single transcriptional spots used in calculations of burst size, frequency, and duration and CV in that AP bin for the given construct. The value in column 43 is the total number of independently imaged embryos for that construct.

Supplementary Table 2: Number of nuclei tracked for each construct.

Each row corresponds to a construct, named in column 42, and columns 1-41 correspond to that AP bin of the embryo. The value in each cell in columns 1-41 is the number of nuclei used for correlation and total noise/covariance/inter-allele noise calculations in that AP bin for the given construct. The value in column 43 is the total number of independently imaged embryos for that construct.

Supplementary File 1: The sequences of all the enhancer constructs generated in this paper.

Supplementary Note

To make a prediction about the expected change in inter-allele noise between single and two enhancer reporter constructs, we used the theory put forth in (Sánchez and Kondev, 2008; Sanchez et al., 2011). This formalism can be used to calculate the expected mean and variance of the transcriptional output of a promoter, given the possible states of the promoter, transition rates between states, and the rate of transcription resulting from each state. In these papers, the authors apply their formalism to different promoter architectures. Here, we generate a simpler model, in which we abstract away the individual transcription factor (TF) binding configurations, which would be numerous and poorly parametrized, and simply define states by whether an enhancer is looped to the promoter and activating transcription. Since these models do not account for fluctuations that would contribute to extrinsic noise, e.g. fluctuations in TF or RNA polymerase levels, they can predict the dependence of intrinsic noise on enhancer arrangement.

To apply this model to our system, we use theses parameters:

γ	degradation rate of mRNA
p	production rate of mRNA
k	on rate for enhancer-promoter looping
l	off rate for enhancer-promoter looping

Below, we describe several models that represent different configurations of either one or two enhancers controlling a single promoter and provide the variables, as defined in (Sanchez et al., 2011), needed to calculate the coefficient of *intrinsic* variation (CV) associated with each model. Briefly, $\boldsymbol{R}$ and $\boldsymbol{r}$ describe the production rates of mRNA in the different promoter-enhancer staets, and $\boldsymbol{K}$ contains the transition rates in and out of states. Key assumptions are that the parameters describing this system are independent of both the position of the enhancer relative to the promoter and the presence of a second enhancer controlling the same promoter. We chose to make these simplifying assumptions to give the reader a general sense of the expected behavior of noise when adding an additional enhancer, since the possible behaviors are nearly infinite with the removal of these simplifying assumptions.

Model 1: Single enhancer

In this model, there is a single enhancer controlling one promoter.

States	Rate of mRNA production
Enhancer-promoter unlooped	0
Enhancer-promoter looped	p

$\boldsymbol{K} = [-k \quad l; \; k \quad -l]$
$\boldsymbol{R} = [0 \; 0; \; 0 \; p]$
$\boldsymbol{r} = [0 \; p]$

Model 2: OR model

In this model, there are two enhancers controlling one promoter, transcription is activated if either enhancer is looped, and both enhancers can't be bound at the same time.

States	Rate of mRNA production
Enhancers-promoter unlooped	0
Enhancer 1-promoter looped	p
Enhancer 2-promoter looped	p

$\boldsymbol{K} = [-2k \;\; l \;\; l; \; k \;\; -l \;\; 0; \; k \;\; 0 \;\; -l]$
$\boldsymbol{R} = [0\ 0\ 0;\ 0\ p\ 0;\ 0\ 0\ p]$
$\boldsymbol{r} = [0\ p\ p]$

Model 3: Additive model

In this model, there are two enhancers controlling one promoter, transcription is activated if either enhancer is looped, and, if both enhancers are bound, transcription occurs at twice the rate of single enhancer looping states.

States	Rate of mRNA production
Enhancers-promoter unlooped	0
Enhancer 1-promoter looped	p
Enhancer 2-promoter looped	p
Both enhancers looped	$2p$

$\boldsymbol{K} = [-2k \;\; l \;\; l \;\; 0; \; k \;\; -k-l \;\; 0 \;\; l; \; k \;\; 0 \;\; -k-l \;\; l; \; 0 \;\; k \;\; k \; -2l]$
$\boldsymbol{R} = [0\ 0\ 0\ 0;\ 0\ p\ 0\ 0;\ 0\ 0\ p\ 0;\ 0\ 0\ 0\ 2p]$
$\boldsymbol{r} = [0\ p\ p\ 2p]$

Model 4: Synergistic model

In this model, there are two enhancers controlling one promoter, transcription is activated if either enhancer is looped, and, if both enhancers are bound, transcription occurs at three times the rate of single enhancer looping states.

States	Rate of mRNA production
Enhancers-promoter unlooped	0
Enhancer 1-promoter looped	p
Enhancer 2-promoter looped	p
Both enhancers looped	$3p$

$\boldsymbol{K} = [-2k \;\; l \;\; l \;\; 0; \; k \;\; -k-l \;\; 0 \;\; l; \; k \;\; 0 \;\; -k-l \;\; l; \; 0 \;\; k \;\; k \; -2l]$
$\boldsymbol{R} = [0\ 0\ 0\ 0;\ 0\ p\ 0\ 0;\ 0\ 0\ p\ 0;\ 0\ 0\ 0\ 3p]$
$\boldsymbol{r} = [0\ p\ p\ 3p]$

Model 5: XOR model

In this model, there are two enhancers controlling one promoter, transcription is activated if either enhancer is looped, and, if both enhancers are bound, no transcription occurs.

States	Rate of mRNA production
Enhancers-promoter unlooped	0
Enhancer 1-promoter looped	p
Enhancer 2-promoter looped	p
Both enhancers looped	0

$\boldsymbol{K} = [-2k \;\; l \;\; l \;\; 0;\; k \;\; -k-l \;\; 0 \;\; l;\; k \;\; 0 \;\; -k-l \;\; l;\; 0 \;\; k \;\; k \;\; -2l]$
$\boldsymbol{R} = [0\ 0\ 0\ 0;\ 0\ p\ 0\ 0;\ 0\ 0\ p\ 0;\ 0\ 0\ 0\ 0]$
$\boldsymbol{r} = [0\ p\ p\ 0]$

To explore the behavior of CV in these different models, we use several approaches.

Figure 1: CV decreases upon the addition of a second enhancer.

Here we plot the mean expression level versus CV for the five models above and one set of parameters, $k = l = 1$, $p = 1$, $\gamma = 0.1$. The single enhancer model (dark purple) drives the highest CV, indicating that, under the assumptions of our models, adding an additional enhancer generally lowers intrinsic noise. Except for XOR model (yellow), all other models produce more mRNA than the single enhancer model. The other colors are: blue, OR model; green, additive model; brown, synergistic model.

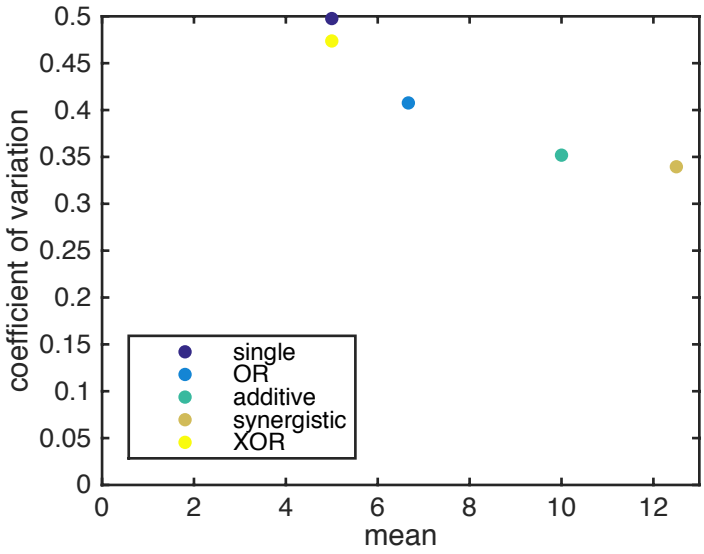
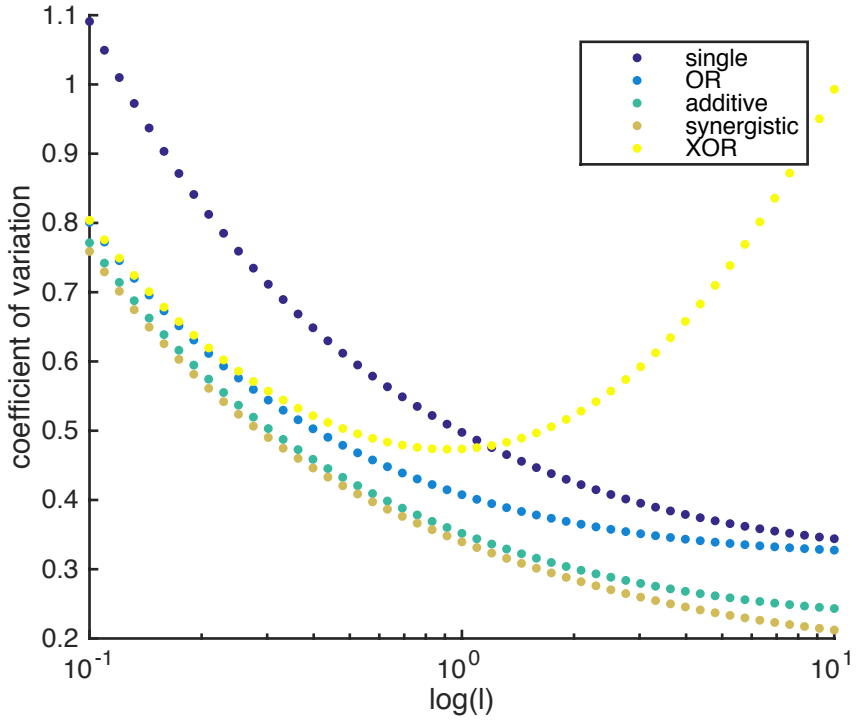


Figure 2: In most cases, two enhancer models drive lower noise than the single enhancer model.
Here we plot the CV as a function of *l*, the rate of promoter-enhancer dissociation, for the five models above and vary *l* from 0.1 to 10 on a logarithmic scale with $k = 1, p = 1, \gamma = 0.1$. With the exception of the XOR model with low *l*, the single enhancer model drives a higher CV than the models with two enhancers for the same value of *l*.



The results above show that, under the simplifying assumptions that the production rates and on-off rates of enhancers are independent of the position and number of enhancers, the addition of a second enhancer generally lowers the predicted intrinsic noise. In our experimental data (Figure 5), we only observe a significant decrease in interallele noise for the shadow enhancer pair compared to the single distal or single proximal enhancer. Duplications of either the proximal or distal enhancer do not have significantly lower noise than their respective single enhancer constructs. Therefore, we expect that the simple addition of an identical enhancer likely does not fulfill the simplifying parameter assumptions used here and suggests that further investigation is needed to understand the complexity of the relationship between interallele noise and the numbers of enhancers controlling a promoter.

References

Sanchez A, Garcia HG, Jones D, Phillips R, Kondev J. 2011. Effect of promoter architecture on the cell-to-cell variability in gene expression. *PLoS Comput Biol* **7**: e1001100. 10.1371/journal.pcbi.1001100.

Sánchez A, Kondev J. 2008. Transcriptional control of noise in gene expression. *Proc Natl Acad Sci U S A* **105**: 5081-5086. 10.1073/pnas.0707904105.

AP Bins (anterior to posterior)																																										Construct Name	No of Embryos
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	249	623	1038	1616	1710	1874	1929	1530	1551	1246	809	597	261	145	56	22	23	15	0	0	0	0	0	0	0	KrDist	11
0	0	0	0	0	0	0	0	0	0	53	54	70	37	18	115	181	479	996	1673	1613	1270	973	542	190	62	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	KrProx	11
0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	20	269	668	1219	1625	2083	2121	1779	1476	1019	521	155	108	120	34	0	0	0	0	0	0	0	0	0	0	0	0	0	KrBothSep	12
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	16	115	196	403	368	439	410	364	309	276	196	58	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	KrDistEmpty	4
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	29	78	185	202	194	152	91	18	6	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	KrProxEmpty	3
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	76	249	586	1368	1280	1755	1775	1583	1508	1191	1013	717	294	94	83	0	0	0	0	0	0	0	0	0	0	0	0	KrDistDuplicN	8
0	0	0	0	0	0	0	0	0	0	0	0	0	20	12	20	114	521	1180	1551	1586	1471	1031	609	171	42	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	KrProxDuplic	10
0	0	0	0	0	0	0	0	0	25	33	36	43	27	19	41	181	1023	1477	2012	1842	2472	1550	1250	755	443	222	48	23	12	9	0	0	0	0	0	0	0	0	0	0	0	KrBoth	11
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	21	94	181	342	385	398	340	294	242	95	4	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	KrBothEmpty	4
0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	27	88	398	664	800	1017	1190	708	580	436	287	291	166	135	117	34	1	0	0	0	0	0	0	0	0	0	0	0	KrDist32C	6
0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	109	226	593	958	1133	1092	1150	799	446	132	13	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	KrProx32C	7
0	0	0	0	0	0	0	0	0	0	0	9	12	55	31	246	676	1219	1789	2051	2025	1842	1599	888	592	267	79	67	36	10	0	0	0	0	0	0	0	0	0	0	0	0	KrBoth32C	13
0	0	0	0	0	0	0	0	0	51	58	36	32	20	57	126	266	640	1199	1475	1455	1317	1012	710	307	31	13	0	2	0	1	0	0	0	0	0	0	0	0	0	0	0	Kr2xProx32C	12
0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	2	3	20	115	231	326	240	226	106	21	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Kr2xProxEmpty	5
0	0	0	0	0	0	0	0	2	1	1	0	0	0	3	56	219	728	1373	1745	2056	2368	2094	1851	1375	1199	821	140	55	0	0	0	0	0	0	0	0	0	0	0	0	KrDist17C	11
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	35	373	701	1517	1349	1884	1475	1339	1278	815	707	438	308	97	0	0	0	0	0	0	0	0	0	0	0	0	Kr2xDist32C	8
0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	134	261	392	753	1157	1102	1108	1038	922	677	296	43	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	KrBoth17C	5
0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	90	680	1269	1815	1752	1975	1829	1497	1151	1030	870	328	91	12	0	0	0	0	0	0	0	0	0	0	0	Kr2xDist17C	8
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	85	664	703	691	606	431	312	158	61	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	KrProx17C	3
0	0	0	0	0	0	0	0	0	0	0	0	0	37	8	6	14	211	645	956	1079	922	778	445	43	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Kr2xProx17C	7
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	53	259	391	478	628	595	524	409	269	158	59	34	21	0	0	0	0	0	0	0	0	0	0	0	0	Kr2xDistEmpty	5
0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	11	56	565	1548	2340	2852	2894	2063	1453	817	577	303	133	35	11	0	0	0	0	0	0	0	0	0	0	0	0	KrEndogDist	13
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	13	43	577	794	1409	1470	1467	1317	1169	825	699	486	178	70	1	0	0	0	0	0	0	0	0	0	0	0	0	KrEndogDist32C	8
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	384	577	609	762	888	908	849	683	538	364	162	96	61	32	14	29	0	0	0	0	0	0	0	0	0	KrEndogDist17C	5

AP Bins (anterior to posterior)																																										Construct Name	No of Embryos
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	28	80	116	176	193	194	220	185	213	170	107	73	25	15	2	2	2	0	0	0	0	0	0	0	0	KrDist	11
0	0	0	0	0	0	0	0	0	0	8	6	4	1	1	14	18	74	141	222	200	169	136	88	18	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	KrProx	11
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	30	64	129	126	159	176	175	178	165	133	100	40	11	7	0	0	0	0	0	0	0	0	0	0	0	KrDistDuplicN	8
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	12	61	167	191	200	200	164	89	18	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	KrProxDuplic	10
0	0	0	0	0	0	0	0	0	3	3	2	1	2	0	2	21	126	173	219	190	240	171	153	121	59	21	8	5	0	1	0	0	0	0	0	0	0	0	0	0	0	KrBoth	11
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	8	46	78	88	104	118	79	74	62	44	48	21	10	7	1	0	0	0	0	0	0	0	0	0	0	0	KrDist32C	6
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	17	31	85	120	133	118	144	126	66	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	KrProx32C	7
0	0	0	0	0	0	0	0	0	0	0	0	0	3	2	26	77	140	187	209	205	185	184	134	110	60	8	2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	KrBoth32C	13
0	0	0	0	0	0	0	0	0	4	7	2	4	1	6	20	38	104	193	191	179	169	143	103	45	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Kr2xProx32C	12
0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	32	97	143	190	208	242	214	215	185	187	127	20	8	0	0	0	0	0	0	0	0	0	0	0	0	0	KrDist17C	11
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	48	83	172	130	185	159	158	166	128	108	73	51	17	0	0	0	0	0	0	0	0	0	0	0	0	Kr2xDist32C	8
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	16	33	37	73	103	95	97	102	99	92	49	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	KrBoth17C	5
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	69	146	203	182	203	199	185	150	147	118	46	7	0	0	0	0	0	0	0	0	0	0	0	0	0	Kr2xDist17C	8
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14	73	79	79	72	72	61	34	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	KrProx17C	3
0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	0	0	27	111	139	143	134	127	68	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Kr2xProx17C	7
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	52	137	170	182	178	173	172	106	72	24	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	KrEndogDist	8
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	4	70	85	152	160	168	168	164	124	103	68	24	7	0	0	0	0	0	0	0	0	0	0	0	0	0	KrEndogDist32C	8
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	49	60	62	79	93	87	89	77	68	54	8	8	2	0	0	0	0	0	0	0	0	0	0	0	0	KrEndogDist17C	4