A Fully Synthetic Transcriptional Platform for a Multicellular Eukaryote

Graphical Abstract

Highlights
- A fully synthetic transcriptional platform of engineered factors is created
- The pioneer factor Zelda is required to open chromatin at synthetic enhancers
- Synthetic enhancers encode transcription levels based on the number of binding sites
- Overlapping activator and repressor binding sites provide sharp expression boundaries

Authors
Justin Crocker, Albert Tsai, David L. Stern

Correspondence
crockerj@janelia.hhmi.org

In Brief
Crocker et al. build a fully synthetic transcriptional platform in Drosophila consisting of engineered transcription factor gradients and artificial enhancers. This synthetic platform confirms the need for pioneer factors to establish an active state and shows how overlapping activator and repressor binding sites can provide sharp expression boundaries.
A Fully Synthetic Transcriptional Platform for a Multicellular Eukaryote

Justin Crocker,1,2,* Albert Tsai,1 and David L. Stern1

1Janelia Research Campus, Howard Hughes Medical Institute, 19700 Helix Drive, Ashburn, VA 20147, USA
2Lead Contact
*Correspondence: crockerj@janelia.hhmi.org
http://dx.doi.org/10.1016/j.celrep.2016.12.025

SUMMARY

Regions of genomic DNA called enhancers encode binding sites for transcription factor proteins. Binding of activators and repressors increase and reduce transcription, respectively, but it is not understood how combinations of activators and repressors generate precise patterns of transcription during development. Here, we explore this problem using a fully synthetic transcriptional platform in Drosophila consisting of engineered transcription factor gradients and artificial enhancers. We found that binding sites for a transcription factor that makes DNA accessible are required together with binding sites for transcriptional activators to produce a functional enhancer. Only in this context can changes in the number of activator binding sites mediate quantitative control of transcription. Using an engineered transcriptional repressor gradient, we demonstrate that overlapping repressor and activator binding sites provide more robust repression and sharper expression boundaries than non-overlapping sites. This may explain why this common motif is observed in many developmental enhancers.

INTRODUCTION

Transcriptional enhancers in multicellular animals have been studied for about four decades, but we still have mainly a qualitative understanding of how they function. In brief, combinations of activating and repressing transcription factors act upon enhancers to drive specific patterns of expression (Stampfel et al., 2015). Natural enhancers have been studied experimentally usually by deleting individual transcription factor binding sites. These studies have therefore revealed specific sites required for proper enhancer function, but they have not necessarily identified all DNA sites that are sufficient to generate a functional enhancer. Similarly, genome-wide studies of the occupancy of transcription factors on DNA regions have provided correlational evidence for the role of transcription factors in enhancer function, but only for factors that were examined explicitly. Additionally, many transcription factor binding sites, as determined by occupancy assays, are not functional (Li et al., 2008). It would be useful to be able to test synthetic assemblages of transcription factor binding sites. However, although artificially concatenated arrays of activator binding sites typically drive expression, they do not always recapitulate the activators’ native expression domains (Erceg et al., 2014).

The construction of a synthetic system would allow comprehensive tests of alternative models of enhancer function, elucidating how specific DNA motifs and binding site architectures influence enhancer function. Indeed, one useful test of whether a biological phenomenon is understood is to build a working model of the system. However, attempts to build synthetic enhancers using binding sites for activators and repressors have largely failed (Johnson et al., 2008; Vincent et al., 2016). Here, we report a fully synthetic enhancer platform for the Drosophila blastoderm embryo and demonstrate the utility of this system.

RESULTS

Construction of a Synthetic Enhancer Platform

We reasoned that use of an exogenous transcription factor would allow study of the principles of enhancer architecture independently of the regulatory network that operates naturally in the Drosophila embryo. We therefore first engineered a gradient of transcription-activator like protein (TALEs) fused to a VP16 activator (Crocker and Stern, 2013) (TALEA) (Figure 1A). The gradient of TALEA protein was generated by driving TALEA expression with the hunchback promoter (Perry et al., 2010; Treisman and Desplan, 1989) (hb-TALEA), resulting in a smooth anterior-to-posterior RNA gradient (Figures 1C and 1D). The binding site for this TALEA, 5’-CCGGATGCTCCTT, is not present in the Drosophila genome and allowed construction of enhancers that would respond only to the TALEA (Figure 1B). Use of TALEs allows greater flexibility in the design of future experiments than other heterologous transcription factors, such as Gal4 and LexA, that are often used in Drosophila experiments, because TALEs with different DNA binding specificities can be generated easily.

We synthesized a 252-bp DNA sequence that is transcriptionally silent in the early Drosophila embryo by starting with a random DNA sequence and systematically altering any motifs that resembled binding sites for known factors active in the early embryo. This sequence did not drive detectable expression in the blastoderm embryo (Figures 1E and 1F). The TALEA protein
gradient and silent enhancer are the foundational components of a synthetic regulatory network operating in parallel to the endogenous developmental networks.

The Synthetic Enhancer System Confirms the Role of Zelda as a Pioneer Factor

To test the ability of the TALEA to drive expression on its own, we introduced one, two, or three TALEA binding sites into the silent enhancer (Figure 1B). None of these enhancers drove detectable expression (Figures 1G and 1H). This was surprising, because introduction of the same three TALEA binding sites into a native enhancer drives strong ectopic expression (Figures S1C and S1D). This result indicated that additional regulatory information, other than the TALEA binding sites, is required for enhancer activity.

One candidate for this additional input in the early Drosophila embryo is the sequence-specific transcription factor protein Zelda. In Drosophila, Zelda is expressed ubiquitously just before most genes begin to be expressed in the blastoderm embryo and Zelda protein binds to many enhancers that are required to drive gene transcription in the early blastoderm embryo (Foo et al., 2014; Harrison et al., 2011; Li et al., 2014; Liang et al., 2008; Nien et al., 2011; Xu et al., 2014). Zelda activity is correlated with chromatin accessibility (Foo et al., 2014; Schulz et al., 2015; Sun et al., 2015), and Zelda appears to make enhancers accessible to transcription factors that drive specific patterns of gene expression (Foo et al., 2014; Li et al., 2014; Schulz et al., 2015; Xu et al., 2014). In particular, Xu et al. (2014) have demonstrated that Zelda binding sites enhance Bicoid binding and can convert silent enhancers containing Bicoid sites into Bicoid-responsive enhancers.

To test the hypothesis that Zelda is the missing element in our silent enhancers, we introduced a variable number of TALEA binding sites and a constant number of Zelda binding sites into the silent enhancer backbone (Figures 2A and 2B). We observed no activity from an enhancer with five Zelda sites in embryos not expressing the TALEA, indicating that Zelda sites alone are not sufficient to drive expression (Figures 2C and 2D). However, a single TALEA binding site together with five Zelda sites drove low levels of RNA expression in the anterior region of early embryos (Figures 2E–2H). Levels of expression were independent of the TALEA binding site location (compare Figures 2E and 2F with Figures 2G and 2H). Adding a second TALEA binding site increased the levels of expression in the anterior region (Figures 2I–2L) and activity remained independent of the location of the TALEA binding site within the 252-bp sequence. These results suggest that the precise arrangement of binding sites is not important for this engineered system, consistent with results from studies of many native enhancers (Arnosti and Kulkarni, 2005; Brown et al., 2007; Hare et al., 2008; Isley et al., 2013; Jin et al., 2013; Lusk and Eisen, 2010; Menoret et al., 2013; Rastegar et al., 2008) (compare Figures 2I and 2J with Figures 2K and 2L). Adding a third TALEA binding site further increased the levels of expression. We found that the levels of expression driven by the synthetic enhancer platform are similar to native gene expression...
Figure 2. Zelda Binding Sites Allow TALEA Binding Sites to Provide Quantitative Control of Gene Expression

(A) Expression patterns of the TALEA and of Zelda.

(B) Schematic of synthetic enhancers used to test the effect of systematically modifying the number of TALEA binding sites with five Zelda binding sites.

(C, E, G, I, K, and M) Stage 5 embryos stained for lacZ expression for enhancers with the indicated number of TALEA and Zelda binding sites.

(D, F, H, J, L, and N) Profiles of average expression levels across the bounding box of Figure 1C for the indicated genotype (n = 10 for each genotype).

(E–N) Enhancers with one (E–H), two (I–L), or three (M and N) TALEA binding sites.
(Figure S2) and that the synthetic enhancer can generate precise patterns of mRNA and protein expression (Figure S2).

We performed a series of control experiments to confirm that the patterns of reporter gene expression resulted from binding of the TALEA to the synthetic enhancer. First, to test whether expression required the TALEA gradient, we constructed an alternative TALEA with a different binding sequence, 5′-AAGTTGTTGGTTGCT, driven by the Hb-promoter. This new TALEA drove expression from a new 252-bp sequence containing binding sites for this alternative activator in a pattern similar to the original TALEA (Figure S3). Second, to test whether expression from the synthetic enhancer resulted from binding of unknown transcription factors to the “silent” DNA sequence, we constructed an independent 252-bp silent DNA sequence both with and without TALEA binding sites. We found that these new sequences drove expression that was quantitatively equivalent to the original sequences (Figure S3), suggesting that the expression patterns we observed result from binding of the synthetic transcription factors.

Taken together, these results indicate that Zelda binding is required to enable an enhancer to respond quantitatively to a variable number of binding sites for an activator transcription factor in the early embryo. Our synthetic enhancers therefore behave like native enhancers that contain different numbers of Zelda binding sites (Driever et al., 1989; Gaudet and Mango, 2002; Stathopoulos et al., 2002).

To further test whether Zelda acts by making enhancers accessible to patterning transcription factors, as has been suggested by several previous studies (Foo et al., 2014; Li et al., 2014; Schulz et al., 2015; Xu et al., 2014), we systematically varied the number of Zelda motifs in synthetic enhancers containing three TALEA binding sites (3× TALEA) (Figures 3A and 3B). In embryos expressing the TALEA gradient, we did not detect any notable expression from enhancers containing one or two Zelda motifs (Figures 3C and 3D). However, enhancers containing three to five Zelda sites drove increased mRNA expression in a subset of cells in the anterior of the embryo, and the number of nuclei showing expression was correlated with the number of Zelda sites (Figures 3E–3J). To rule out the effect of position effects on the synthetic enhancers, we integrated the synthetic enhancers into two additional sites in the genome. We found that, in each case, a minimum of three Zelda binding sites was required for expression (Figure S3).

To test whether this pattern reflected stochastic transcription that is activated in different subsets of cells over time (Bothma et al., 2014; Chubb et al., 2006; Golding et al., 2005; Raj et al., 2006), we also examined patterns of expression for the proteins encoded by the reporter gene mRNA products, because the protein products perdure for much longer than the mRNA products (Figure S4). If transcription was temporally stochastic, then we would have expected more cells to express protein than mRNA. Instead, we observed very similar patterns of mRNA and protein expression, indicating that a subset of cells activated gene transcription from the synthetic enhancers and that these enhancers remained “on” for an extended time. Therefore, Zelda sites do not trigger transient stochastic expression, but instead mark enhancers in a subset of nuclei as available for binding of activator transcription factors.

These results are consistent with the hypothesis that Zelda marks enhancers as available for regulation and that other transcription factors control expression levels (Foo et al., 2014; Li et al., 2014; Schulz et al., 2015; Xu et al., 2014). To test this hypothesis, we segmented images to determine expression levels in each nucleus independently (Figure S5). In enhancers containing variable numbers of Zelda sites, we found that the levels of expression within each active nucleus are not different across enhancers, on average (Figures 3K and S5; ANOVA, $F_{(2,9)} = 1.76$, $p > 0.20$). In contrast, increasing the number of TALEA sites in synthetic enhancers increased levels of expression within active nuclei (Figures 3M and S5; ANOVA, $F_{(2,9)} = 6.01$, $p < 0.02$). Therefore, the number of Zelda sites alters the probability of transcription, whereas TALEA binding sites modulate the amplitude of expression.

The simplest proposed mechanism for Zelda activity is that Zelda makes DNA accessible to other transcription factors by displacing nucleosomes (Foo et al., 2014). We found that increasing the number of Zelda sites in synthetic enhancers increased DNA accessibility, as measured by DNase I digestion, even in the absence of TALEA expression (ANOVA, $F_{(3,16)} = 21.86$, $p < 0.001$) (Figure 3L). In contrast, increasing the number of TALEA binding sites in the context of a constant number of Zelda binding sites did not significantly alter DNA accessibility (ANOVA, $F_{(3,12)} = 1.65$, $p > 0.20$) (Figure 3N). These results agree with observations of native Drosophila enhancers (Foo et al., 2014) and confirm that binding of Zelda to enhancers increases local chromatin accessibility (Barozzi et al., 2014; Cirillo et al., 2002; Foo et al., 2014; Li et al., 2014; Schulz et al., 2015; Sherwood et al., 2014; Xu et al., 2014) (Figure 3O). Together, these results support the hypothesis that, in the blastoderm embryo, the regulatory state of an enhancer, ON versus OFF, is determined by Zelda binding and can be decoupled from the patterns and levels of expression driven by an enhancer (Figure 3O).

**Overlapping Activator and Repressor Binding Sites Provide Sharper Boundaries Than Non-overlapping Sites**

With this confirmation of the utility of our synthetic enhancer system for testing models of transcription factor function in enhancers, we next examined a classical problem in developmental biology, the use of broadly distributed gradients of transcription factors to produce sharp boundaries of gene expression (Driever and Nüsslein-Volhard, 1988; Turing, 1990; Wolpert, 1969). The mechanisms that generate precise patterns of gene expression are not fully understood (Lagha et al., 2012; Little et al., 2013), and some authors have proposed that binding site competition, whereby activators and repressors compete to bind to the same DNA sites, might produce sharp boundaries of gene expression (Rushlow et al., 2001; Saller and Bienz, 2001; Small et al., 1991; Stanojevic et al., 1991). Consistent with this hypothesis, overlapping activator and repressor binding sites are a common feature in transcriptional enhancers (Cheng et al., 2013; Makeev et al., 2003; Papatsenko et al., 2009; Stanojevic et al., 1991). However, there have been no experimental tests of this hypothesis in embryos (Payankaulam et al., 2010). Our synthetic enhancer provides an ideal platform for testing this hypothesis.
Figure 3. Increasing the Number of ZELDA Sites Increases the Probability That an Enhancer Will Be Active in a Cell
(A) Expression of the synthetic TALEA and Zelda.
(B) Schematic of synthetic enhancers used to test the effect of varying the number of Zelda binding sites.
(C and D) Enhancers with zero, one, or two Zelda binding sites.
(E–J) Enhancers with three (E and F), four (G and H), or five (I and J) Zelda binding sites.
(Figure 1E for the indicated genotype (n = 10 for each genotype).
(K) Cell-by-cell quantification of the staining intensities in all cells displaying expression for enhancers with the indicated number of Zelda binding sites.
(N) The effect of the number of Zelda binding sites on DNase I sensitivity, each with three TALEA binding sites. N = 5 samples of embryos per genotype.
(M) Cell-by-cell quantification of the staining intensities in cells displaying expression for enhancers with the indicated number of TALEA binding sites, each with five Zelda binding sites.
(O) Heuristic model of the synthetic enhancer activity. Zelda opens chromatin and allows binding by transcription factors that modulate expression amplitude.
To test the role of overlapping activator and repressor binding sites, we first generated orthogonal gradients of an activator and a repressor. We started with the anterior-posterior gradient of the TALEA described above and added an orthogonal ventral-dorsal gradient of a TALE fused to a Hairy repression domain (TALER) (Figure 4A). The TALER protein gradient was generated by driving TALER expression with the snail promoter (Ip et al., 1992) (sna-TALER). This results in a smooth ventral-dorsal gradient (Figures 4C and 4D). We then created two synthetic enhancers, one with three activator and three repressor binding sites in an alternating tandem array and one where activator and repressor sites shared exactly the same three binding sites (Figure 4B). For the enhancer with tandem binding sites, the TALER targeted the sequence 5′-AAGTTGTGGTTTGTCT. For the enhancer with overlapping sites, the TALER and TALEA both targeted the sequence 5′-CCGGATGCTCCTCTT. To test for potential differential affinity of the two binding sites, we targeted these two sites separately with a TALEA and found that they drove indistinguishable patterns of expression (Figure S3). Therefore, the two binding sites arranged in a tandem array appear to have similar affinity for the TALEEs and therefore...
provides a useful comparison with the enhancer containing overlapping binding sites.

We observed that both the tandem and overlapping enhancers generated repression of reporter gene expression in the region in which the TALER was expressed (Figures 4E–4G). However, the enhancer with overlapping sites generated stronger reduction in reporter gene expression at the highest levels of TALER expression and a steeper transition from high to low levels of expression along the TALER gradient, compared with the enhancer containing tandem binding sites (Figures 4F and 4G).

To clarify the mechanisms that may be acting to generate these differences between the two enhancers, we constructed simple steady-state models of activators and repressors binding to enhancers with either tandem arrays or overlapping binding sites (Figures 4H–4L and S6). We assumed that activators and repressors compete to bind to overlapping binding sites. The model of overlapping binding sites predicts a similar pattern of reporter gene expression as we observed empirically, with an early and sharp reduction in reporter gene expression across the repressor gradient and a strong reduction in reporter gene expression at the highest levels of repressor concentration (Figures 4F and 4I). Notably, in real embryos, reporter gene expression in the region of highest repression was indistinguishable from background (Figure 4F). This indicates that there is virtually no binding of activators to this enhancer at the highest repressor concentrations. This observation agrees with the model, which predicts that when repressors entirely outcompete activators, the reporter gene expression should drop to background levels (Figures 4K and 4L).

To explore the results for the tandem enhancer, we built a series of models in which activators and repressors can enhance or inhibit activity of factors bound at neighboring sites through various mechanisms (Figure S6). We fixed the apparent affinities of the activator and repressor using the experimental results from the enhancer containing overlapping binding sites. The apparent affinities account for differences in the absolute concentrations between the factors and any additional interactions that may affect their binding and transcriptional activity. All models make the same qualitative prediction that the enhancer should display incomplete repression in the region of highest repressor concentration (Figure S6), because in all models activator proteins remain bound to the enhancer. This pattern is consistent with our experimental observations (Figures 4K and 4L).

An additional salient experimental result was most consistent with one of the models of tandem sites. We observed that the enhancer with tandem binding sites displayed the first signs of reduced expression at higher repressor concentrations than the enhancer with overlapping sites (e.g., approximately at 15% of ventral/dorsal axis) and that the slope of the reduction in expression across the repressor gradient was more shallow than the slope for the enhancer with overlapping sites (Figures 4K and 4L). These two results were most consistent with a model where repressors bond to sites flanking an activator site prevented binding of, or suppressed activity of, the TALEA at the intervening activator binding site (Figures 4K and 4L). The neighboring sites in our tandem arrays should be separated sufficiently to prevent direct competition. We therefore hypothesize that the repression domain on the TALER is responsible for this novel activity. The mechanism underlying this repressor activity remains to be investigated. Additionally, our model required a much higher apparent affinity for the repressor than for the activator to achieve complete transcriptional shutdown with overlapping binding sites. This may reflect a real activity difference between the activation and repression domains we used. It will be valuable to learn the mechanism of this repressor-activator interaction because tandem activator and repressor binding sites are observed in many native enhancers (Fakhouri et al., 2010; Gray and Levine, 1996; Payankaulam et al., 2010; Small et al., 1991).

**DISCUSSION**

Disentangling regulatory networks in multicellular eukaryotic development has proven challenging because native enhancers usually contain activator and repressor binding sites for multiple factors that each exert nuanced, context-dependent control of enhancer activity (Crocker et al., 2008). Drawing from our experience exploring the activity of engineering TALEs in developing embryos (Crocker and Stern, 2013; Crocker et al., 2016) and dissecting native enhancer elements (Crocker et al., 2015), we have constructed a simple yet functional synthetic enhancer platform in Drosophila blastoderm embryos. We have thus extended techniques from cellular synthetic biology (Amit et al., 2011; Atkinson et al., 2003; Basu et al., 2005; Elowitz and Leibler, 2000; Endy, 2005; Friedland et al., 2009; Garcia and Phillips, 2011; Gardner et al., 2000; Mukherji and van Oudenaarden, 2009) to organismal systems. Our system provides clean tests of hypotheses of regulatory function, as we demonstrate for the function of Zelda and the role of overlapping binding sites (Driever et al., 1989; Gaudet and Mango, 2002; Stathopoulos et al., 2002). In particular, our results comparing overlapping with tandem arrays of repressor and activator binding sites show how overlapping binding sites can create well-defined expression boundaries during development.

The specific design of our engineered enhancer raises several caveats. First, previous studies suggest that enhancers rarely contain three or more Zelda binding sites (Xu et al., 2014). Second, some enhancers clearly do not require Zelda activity. For example, the binary UAS-Gal4 expression system drives high levels of expression in Drosophila melanogaster, and these constructs do not contain Zelda binding sites. One explanation for our results is that TALE proteins bind poorly to nucleosomal DNA. Indeed, GAL4 can bind to nucleosomal templates, and different transcription factors vary in their ability to bind to nucleosomal templates (Taylor et al., 1991). This variability may be important to the function of different transcription factors, and our engineered system provides a novel platform for examining these phenomena in vivo. Finally, we used strong activation and repression domains to test our engineered system. It is possible that DNA accessibility plays a more important role for our assays than during native developmental gene expression. It will be possible to use this system to test different activation and repression domains and their context-dependent activity on transcription (Stampfel et al., 2015).

Our synthetic system will allow deeper investigation into how different combinations of protein domains contribute to
enhancer activity than is possible using native enhancers alone. It is possible to imagine extending this system to build more sophisticated synthetic regulatory systems that could be engineered to test the roles of specific features of regulatory architecture during development.

**EXPERIMENTAL PROCEDURES**

**Construction of TALE Plasmids**

TALE constructs were based on the VP64 TALEA construct (Crocker and Stern, 2013) and were assembled using the Golden Gate method (Cermak et al., 2011). The TALE binding domain was previously characterized in a plant base (Matys et al., 2003). Enhancer sequences were subcloned into pLacZattB that resembled binding sites for known factors based on the TRANSFAC database (Matys et al., 2003). Enhancer sequences were subcloned into pLacZattB (Crocker et al., 2015).

**Fly Strains and Crosses**

*D. melanogaster* strains were maintained under standard laboratory conditions. Transgenic TALE constructs were created by Rainbow Transgenic Flies and were integrated at the attP2 landing site.

**Embryo Manipulations**

Embryos were raised at 25°C and fixed and stained according to standard protocols. Briefly, anti-DIG RNA probes were used against lacZ. Antibody staining was subsequently carried out against the DIG-antigen, according to standard procedures. LacZ protein was detected using an anti-β-Gal antibody (1:1000; Promega). Detection of primary antibodies was done using secondary antibodies labeled with Alexa Fluor dyes (1:500; Invitrogen).

**Microscopy**

Each series of experiments to measure transcript levels was performed entirely in parallel. Embryo collections, fixations, and hybridizations, and image acquisition and processing were performed side-by-side in identical conditions. Confocal exposures were identical for each series and were set to not exceed the 255 maximum level. Series of images were acquired over a 1-day time frame, to minimize any signal loss or aberration. Confocal images were obtained on a Leica DM5500 Q Microscope with an ACS APO 20×/0.60 IMM CORR lens and Leica Microsystems LAS AP software. Sum projections of confocal stacks were assembled, embryos were scaled to match sizes, background was subtracted using a 50-pixel rolling-ball radius, and plot profiles of fluorescence intensity were analyzed using ImageJ software (https://imagej.nih.gov/ij/). Data from the plot profiles were further analyzed in MATLAB. Expression levels of the nuclei in Figure 4 were obtained by segmenting the nuclei based on DAPI expression and measuring the average level of expression within each nucleus. The expression levels were then further analyzed in MATLAB.

**DNase I Sensitivity**

DNase I digestion was performed on 1.5- to 3-hr-old embryos as described previously (Foo et al., 2014; Thomas et al., 2011), with some modifications. Four biological replicates were performed for each DNase I digestion experiment. PCR experiments were performed on the isolated nuclei with primers for the synthetic enhancer and control regions with the following set of common synthetic and negative control (Neg) primers:

- Synthetic Enhancer, 5'- CGGATGTCCTCCTTTTTCCCA
- Synthetic Enhancer, 3'- [T7]ggGGTCCCCAGCAGTTAAGCT
- Neg Enhancer, 5'- TGCGTACCAGATAGAGCCCA
- Neg Enhancer, 3'- [T7]ggCTGGCTTTGGAAAACCCC

Each set of 3'-primers contained a T7 promoter, 5'-GAAATATACGACT CACTATA. Samples were subjected to six rounds of amplification. The PCR products were cleaned with a Qiagen PCR purification and then were added to a MEGASHortscript T7 Transcription kit (Thermo Fisher Scientific) for a 12-hr linear DNA amplification (Shankaranarayanan et al., 2011). The resulting RNA products were run on a denaturing gel, and the fluorescence intensity was quantified. Fluorescence values were normalized and DNase I hypersensitivity values were calculated as described previously (Foo et al., 2014).

**Enhancer Modeling**

Total transcription output of the synthetic enhancer was modeled assuming that the system is in steady state. For overlapping binding sites available to an activator or a repressor, transcriptional repression occurs because binding of a repressor prevents an activator from accessing the same site. In this case, each binding site introduces the following term:

\[ 1 + K_A A + K_R R \]

The elements in the term describe the relative probabilities that the site is, respectively, unbound (1), bound by an activator (K_A A), or bound by a repressor (K_R R), and A and R are the relative concentrations of the activator and repressor normalized so that the maximum concentration is 1. K_A and K_R are the apparent affinities of the activator and the repressor to the site and include all molecular mechanisms that influence activity, including DNA affinity. Note that, because the concentrations of the activator and repressor are relative, the apparent binding affinities also include adjustments for their absolute concentrations and any additional interactions that may modify the activity of either factor. With three overlapping binding sites, the population with no activator bound is the following:

\[ P_{0,0}(A, R) = 1 + 3 K_R R + 3 K_A^2 R^2 + K_A^3 R^3 \]

The population with one activator bound is the following:

\[ P_{0,1}(A, R) = (3 + 6 K_A^2 R^2 + K_R R) K_A A \]

The population with two activators bound is the following:

\[ P_{0,2}(A, R) = (3 + 3 K_R R) K_A^2 A^2 \]

Finally, the population with three activators bound is the following:

\[ P_{0,3}(A, R) = K_A^3 A^3 \]

Note that the sum of all populations is the following:

\[ P_{0,0}(A, R) + P_{0,1}(A, R) + P_{0,2}(A, R) + P_{0,3}(A, R) = (1 + K_A A + K_R R)^2 \]

Assuming that each activator additively contributes one unit of transcriptional activity and that the maximum transcriptional output is 3, the total transcriptional output is the following:

\[ T_{\text{max}}(A, R) = \frac{P_{0,1}(A, R) + 2 P_{0,2}(A, R) + 3 P_{0,3}(A, R)}{P_{0,0}(A, R) + P_{0,1}(A, R) + P_{0,2}(A, R) + P_{0,3}(A, R)} \]

With separate binding sites for activators and repressors, the sum of the relative probabilities of activator site being unbound or bound are described by the following:

\[ 1 + K_A A \]

The sum of the relative probabilities of a repressor site being unbound or bound are the following:

\[ 1 + K_R R \]

With six total alternating activator and repressor sites in tandem, the model that best describes the experimental results assumes that having two bound
repressor sites flanking an activator site precludes the activator site in question from functioning. After removing configurations prohibited by the above rule, the relative populations for no activator bound is the following:

\[ P_{1,0}(A, R) = 1 + 3 K_{d, R} + 3 K_{d, R}^{2} + K_{d, R}^{3}. \]

The term for one activator bound is the following:

\[ P_{1,1}(A, R) = (3 + 9 K_{d, R} + 7 K_{d, R}^{2} + 6 K_{d, R}^{3}) K_{d, A}. \]

The term for two activators bound is the following:

\[ P_{1,2}(A, R) = (3 + 9 K_{d, R} + 5 K_{d, R}^{2}) K_{d, A}^{2}. \]

The term for three activators bound is the following:

\[ P_{1,3}(A, R) = (1 + 3 K_{d, R} + K_{d, R}^{2}) K_{d, A}^{3}. \]

The total transcriptional output is the following:

\[ TS_{X_{A Tam}}(A, R) = P_{1,0}(A, R) + 2 P_{1,1}(A, R) + 3 P_{1,2}(A, R) + 4 P_{1,3}(A, R). \]

The panels in Figures 6H–IJ were generated using Mathematica (Wolfram) with the following parameters: \( K_{d, A} = 5 \) and \( K_{d, R} = 500 \).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures and six figures and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2016.12.025.

AUTHOR CONTRIBUTIONS

J.C. conceived of, designed, and executed the experiments and analyzed the data, with mentorship of D.L.S. A.T. led the modeling analyses. J.C., A.T., and D.L.S. conceived of, designed, and executed the experiments and analyzed the data, with mentorship of D.L.S. A.T. led the modeling analyses. J.C., A.T., and D.L.S. revised and wrote the manuscript.

ACKNOWLEDGMENTS

We thank T. Shirangi for critical insight into the experimental results; Colby Starker and Daniel Voytas for kindly providing the TALE binding domain; G. Isley, R. Mann, and E. Preger-Ben Noon for valuable discussions; and several anonymous reviewers for helpful comments. A.T. is a Damon Runyon Fellow supported by the Damon Runyon Cancer Research Foundation (DRG-2220-15).

Received: November 9, 2015
Revised: December 14, 2016
Accepted: December 7, 2016
Published: January 3, 2017

REFERENCES

Establishment of regions of genomic activity during the zygotic genome activation


