Ancestral Resurrection of the Drosophila S2E Enhancer Reveals Accessible Evolutionary Paths through Compensatory Change

Carlos Martinez,1,2* Joshua S. Rest,3 Ah-Ram Kim,1,2,4 Michael Ludwig,1,2 Martin Kreitman,1,2
Kevin White,1 and John Reinitz1,2,5,6
1Institute for Genomics and Systems Biology, University of Chicago
2Department of Ecology and Evolution, University of Chicago
3Department of Ecology and Evolution, Stony Brook University
4Department of Biochemistry and Cell Biology, Stony Brook University
5Department of Statistics, University of Chicago
6Department of Molecular Genetics and Cell Biology, University of Chicago

*Corresponding author: E-mail: martinezc@uchicago.edu.
Associate editor: Stuart Newfeld

Abstract
Upstream regulatory sequences that control gene expression evolve rapidly, yet the expression patterns and functions of most genes are typically conserved. To address this paradox, we have reconstructed computationally and resurrected in vivo the cis-regulatory regions of the ancestral Drosophila eve stripe 2 element and evaluated its evolution using a mathematical model of promoter function. Our feed-forward transcriptional model predicts gene expression patterns directly from enhancer sequence. We used this functional model along with phylogenetics to generate a set of possible ancestral eve stripe 2 sequences for the common ancestors of 1) D. simulans and D. sechellia; 2) D. melanogaster, D. simulans, and D. sechellia; and 3) D. erecta and D. yakuba. These ancestral sequences were synthesized and resurrected in vivo. Using a combination of quantitative and computational analysis, we find clear support for functional compensation between the binding sites for Bicoid, Giant, and Krüppel over the course of 40–60 My of Drosophila evolution. We show that this compensation is driven by a coupling interaction between Bicoid activation and repression at the anterior and posterior border necessary for proper placement of the anterior stripe 2 border. A multiplicity of mechanisms for binding site turnover exemplified by Bicoid, Giant, and Krüppel sites, explains how rapid sequence change may occur while maintaining the function of the cis-regulatory element.

Key words: evolution, transcription, modeling, enhancer.

Introduction
Our understanding of the evolution of Eukaryotic cis-regulatory sequences is weak relative to that of the evolution of protein coding sequences. This is a reflection of the fact that the cis-regulatory encoding is poorly understood but the genetic code for protein sequence is well understood. As a consequence, research in this area has tended to rely on sequence conservation, termed “phylogenetic footprinting,” of cis-regulatory modules (CRMs) across species as a proxy for functional importance (Cliften et al. 2003; Kellis et al. 2003; Berman et al. 2004). Phylogenetic footprints are thought to correspond to transcription factor binding sites (TFBSs) required for proper enhancer activity. Although this approach has been successful in the discovery of novel CRMs (Berman et al. 2004; Visel et al. 2008), there is mounting evidence that sequence conservation does not directly correlate with functional constraint (Ludwig et al. 1998; Li et al. 2008; Hare, Peterson, Iyer, et al. 2008; Swanson et al. 2011). The prevailing view is that the lack of a direct mapping between sequence conservation and functional constraint is the result of rapid binding site turnover (Dermitzakis et al. 2003; Doniger and Fay 2007; He, Holloway, et al. 2011). Models of CRM evolution have therefore focused on determining the mechanisms underlying the gain and loss of TFBSs in both functionally conserved (Durrett and Schmidt 2008; Lusk and Eisen 2010; Bullaughey 2011) and divergently expressed enhancers (Wittkopf and Kalay 2012). These efforts have benefited greatly from the rapid advancements of protein–DNA profiling techniques, such as ChIP-chip, ChIP-seq, SELEX, and MITOMI, that enable the characterization of in vivo and in vitro binding of TFs at a genomic level (Ren et al. 2000; Roulet et al. 2002; Robertson et al. 2007; Maerkl and Quake 2007). As TFs regulate transcription through multiple mechanisms, a fundamental challenge arises: How do we incorporate our knowledge of transcriptional regulation into evolutionary models of enhancers?

Transcription factors interact with DNA through short sequence motifs with a high degree of degeneracy. These motifs can be characterized by positional weight matrices (PWMs), which correlate the independent additive contributions of individual base pairs (bp) to the binding energy as a function of position and nucleotide (Berg and von Hippel 1987; Stormo 2000). Given a high quality PWM, the relative binding affinities of different motifs can be estimated with...
reasonable accuracy (Roider et al. 2007; Zhao et al. 2009; Zhao and Stormo 2011). Current evolutionary models of enhancers typically make use of this relatively simple mapping between sequence and affinity to describe the gain and loss of TFBSs as the result of neutral evolution, where the loss of affinity at one position is compensated by the gain at another (Dermitzakis et al. 2003; Durrett and Schmidt 2008; Lusk and Eisen 2010). The underlying assumption of these models is that the total number of sites of a given TF is what is under selection to satisfy a functional constraint. Although this assumption might be valid for relatively simple regulatory elements such as the heat-shock response element (HRE; He, Eichel, et al. 2011), studies have shown that many CRMs require additional mechanisms such as synergy, cooperativity, quenching, and direct repression for proper function (Simpson-Brose et al. 1994; Amost et al. 1996; Gray and Levine 1996; Hewitt et al. 1999; Lebrecht et al. 2005).

Indeed, efforts to create synthetic enhancers that mimic the expression of four Drosophila regulatory elements failed to recapitulate the native gene expression when assembled through the multimerization of known TFBSs (Johnson et al. 2008).

Each additional regulatory mechanism introduces new potential avenues through which evolution can explore the enhancer sequence space by compensatory changes (Bullaughey 2011). As a consequence, CRMs controlled through multiple mechanisms might well evolve faster than those with a simpler cis-regulatory logic. This is consistent with a recent study showing that the HREs of fly and human heat-shock genes can correctly induce transcription in C. elegans, whereas most Drosophila CRMs fail to function in this context (He, Eichel, et al. 2011). In turn, phylogenetic comparisons of the C. elegans, Drosophila, and human HREs show strong sequence conservation of binding sites for the heat-shock transcription factor, which is thought to be the main determinant of the heat-shock response (He, Eichel, et al. 2011). These studies illustrate how simple regulatory “on–off” response elements such as the HRE can remain conserved at both the functional and binding site level across the 1,000–1,200 My separating Chordates, Arthropods, and Nematodes (Wang et al. 1999). In contrast, the CRMs of the Drosophila even-skipped (eve) become essentially unalignable within 100 My while retaining functional conservation (Hare, Peterson, Iyer, et al. 2008). Thus, models of CRMs need to incorporate all relevant regulatory mechanisms in order to correctly characterize their evolution, including the contribution of compensation, redundancy, and selection.

A major obstacle in trying to incorporate cis regulatory information in models of enhancer evolution is that incorporating these regulatory elements into a precise quantitative model is a complex task. The approach used here was first proposed in 2003 (Reinitz et al. 2003); it and other approaches have been applied to the Drosophila blastoderm with varying degrees of success (Janssens et al. 2006; Segal et al. 2008; Fakhouri et al. 2010; He et al. 2010; Isley et al. 2013; Kim et al. 2013; Samee and Sinha 2013). We have previously shown that a theoretical model of transcriptional control is capable of accurately fitting the expression patterns of the proximal 1.7 kb control region of the D. melanogaster eve gene (Janssens et al. 2006). This region contains the eve stripe 2 element (S2E) responsible for driving eve stripe 2 expression in the blastoderm embryo (Goto et al. 1989; Harding et al. 1989). Functional analysis of the S2E across multiple species has uncovered the existence of a stabilizing selection mechanism acting to minimize functional divergence (Ludwig et al. 1998, 2000). Functional conservation of the S2E can be seen despite an almost complete lack of sequence conservation (Hare, Peterson, Eisen 2008; Hare, Peterson, Iyer 2008). In this work, we use our theoretical model to study how the S2E sequence diverges while its function is conserved. By combining ancestral sequence reconstruction with model-based functional constraints in transcriptional expression, we show how the S2E enhancer has evolved both through compensatory and noncompensatory mechanisms that allow the maintenance of the correct expression pattern. To validate this finding, putative S2E sequences for several ancestral sequences were synthesized and tested in vivo using site-specific reporter constructs. The combination of both phylogenetics and transcriptional modeling shows that the S2E enhancer has preserved proper boundary and expression levels by compensatory evolution of cis-regulatory sites that bind Bicoid (Bcd), Krüppel (Kr), and Giant (Gt).

Results

In this study, we seek to understand functional conservation in the face of sequence divergence. For that reason, we hypothesized at the outset of our investigation that both the trans-environment and the expression pattern of S2E have been conserved over the course of the evolution of these lineages. We therefore assumed for our model that each of a set of eight homologous S2E sequences from D. melanogaster (mel), D. simulans (sim), D. sechelia (sec), D. yakuba (yak), D. erecta (ere), D. ananasse (ana), D. pseudoobscura (psd), and D. mojavensis (moj) would be expressed identically if placed into a D. melanogaster background. We then took the experimentally observed eve–lacZ expression driven by the mel eve minimal stripe 2 element (MSE2) and used it as a conserved expression data set for training. The assumption of trans-environment conservation was necessary in light of the limited amount of data on quantitative expression in organisms other than melanogaster. It is also a reasonable assumption, as evidenced by the fact that two CRMs from Sepsid flies express stripe 2 in D. melanogaster, and these systems are separated by about 100 My (Hare, Peterson, Iyer, et al. 2008). Moreover, it is known that S2E’s from psd and yak can functionally rescue a knockout in mel. At the same time, it is clear that our assumption does not hold for one of the studied species because the ere S2E, when expressed in a mel S2E knockout, cannot rescue stripe 2 function (Ludwig et al. 2005). Despite this exception, we used this initial data set to analyze how expression has been conserved using a quantitative model of transcriptional regulation.
Transcriptional Model

We developed a theoretical model of *eve* stripe 2 transcriptional regulation capable of predicting the expression pattern driven by homologous S2E sequences, given TF concentrations. The input to the model is the presence, affinity, order, and spacing of the TFBSs in the S2E, and the output is the spatial and temporal expression pattern driven by them. The locations and affinities of the TFBSs of the main regulators of *eve* stripe 2 were determined using high quality PWMs (supplementary materials and methods, Supplementary Material online). We note that while some TFs appear not to conform to a simple model of independent additive contributions to the binding energy, a recent quantitative study has shown that a standard PWMs is sufficient to model-binding specificities of most TFs (Zhao and Stormo 2011). The predicted binding sites and affinities together with quantitated TF concentration profiles (Janssens et al. 2006; Surkova, Kosman, et al. 2008; Pisarev et al. 2008) were then used as inputs to the model. The quantitated TFs profiles were obtained from the FlyEx database and are discussed in the following paragraphs. Model output is determined by the successive application of rules describing binding competition, short-range repression, and activation as previously described (Janssens et al. 2006), with the addition of the mechanism of coactivation by which Hunchback (Hb) switches from being a repressor to an activator when bound in proximity to Bcd (Small et al. 1991, 1993; Simpson-Brose et al. 1994).

For each predicted binding site, the fractional occupancy is calculated using the equilibrium chemical equations as in Reinitz and Vaisnys (1990). Competitive binding, where two or more binding sites for different TFs overlap, is modeled for a given site by adding the binding affinity multiplied by the TF concentration to the partition function for each overlapping site (supplementary eqs. S3 and S4, Supplementary Material online). Next, we calculate the degree of Hb coactivation for Hb bound at each site, defined as the percentage of the time in which Hb is in a coactivated state (i.e., behaves as an activator) (supplementary eq. S5, Supplementary Material online). Coactivation is assumed to occur by protein–protein interactions between bound Hb and nearby bound Bcd (less then 150 bp), where the interaction efficiency is a piecewise linear function of distance (maximum efficiency at distances less then 100 bp, 0% efficient at more then 150 bp, and linear interpolation in between; supplementary eqs. S6 and S7, Supplementary Material online). We next calculate the effect of short-range repression (also known as quenching) (Hewitt et al. 1999). We represent short range repression phenomenologically as a reduction in effective fractional occupancy of nearby activator binding sites, where the quenching efficiency decreases with distance in the same manner as coactivation (supplementary eq. S8, Supplementary Material online). Finally, the model calculates the reduction in the energy barrier of transcriptional initiation by assuming that bound activators and coactivated Hb interact with the holoenzyme via indirect protein–protein interactions to stabilize the transcriptional complex at the promoter. The efficiency of activation is independent of distance and is a specific property of the activator or coactivated transcription factor.

We model, in one dimension, a region of the blastoderm with anteroposterior (A-P) positions ranging from 35.5% to 93.5% of the EL (0% EL at the anterior pole), which corresponds to the segmental anlage. We do not consider the embryo regions anterior to 35.5% EL or posterior to 93.5% EL in the model because we do not have quantitative data on the expression of TFs that act in the terminal regions. We consider a period of 32.5 min which extends from 6.5 to 40 min after the 13th nuclear division. In our data, this period is divided into five temporal classes of 6.5 min each; these are time classes 2–6 (T2–T6) described elsewhere (Surkova, Kosman, et al. 2008). These data represent 58 nuclei. Data for each nucleus contain the level of mRNA observed, together with the concentrations of the controlling TFs Bicoid (Bcd), Hunchback (Hb), Caudal (Cad), Kruppel (Kr), Giant (Gt), Knirps (Kni), and Tailless (Tll). Because the lifetime of the reporter RNA is very short, its concentration is proportional to the transcription rate (Janssens et al. 2006).

We constructed a model of conserved expression as follows. We first measured the expression of a *eve-lacZ* reporter transgene of the *D. melanogaster* MSE2 in an OreR genetic background for the same embryonic time classes and nuclei for which TF concentrations were measured. A total of 36 embryos between time classes T2–T6 were imaged and their expression profiles quantitated. We then assumed that the homologous S2E sequences from eight species would drive the identical expression if placed in this same melanogaster context. The model described earlier contains 31 free parameters. We fit a model for the eight sequences to the *eve-lacZ* reporter expression pattern. Over 5 time classes and 58 nuclei were used in the model fit, this amounts to determining 31 parameters from 2,320 observations. Optimization was carried out by simulated annealing to find the minimum of the sum of the squared difference between the observed and model-calculated expression data for the 8 sequences. This parametrized model produces calculated expression patterns for homologous S2E sequences that overall most closely match the *mel* MSE2 expression profile. The quality of the solution compared with data is well described by calculating the sum of the squared difference between the observed MSE2 expression and the predicted expression for all S2Es. The sum of the squared difference between expression levels in the model and the data. A solution was deemed a good fit if the rms score was below 7.0 and had no obvious defects as judged by visual inspection. We obtained ten very similar model fits from our data with rms scores ranging from 5.24 to 5.48, and selected the lowest scoring one, with an rms of 5.24, for further study (fig. 1). This fit has a proportional error of about 5%, although its accuracy varies from species to species.

Our use of a least squares cost function was not based on a likelihood model, but we can find a rough measure of the significance of our model fit by comparing the quality of fits to expression data with fits to random data. We used two types of random data: data in which the stripe 2 profile was randomly shifted and data in which RNA levels were
randomly shuffled with respect to TF data. For randomly
shifted stripes, the summed squared deviation of the model
from the shifted test data was always at least 7.9 times larger
than that obtained from actual biological data. It is interesting
to note that such fits (supplementary fig. S1, Supplementary
Material online) either produces a stripe of greatly reduced
amplitude in the correct location, or in other cases are com-
pletely incapable of producing a stripe in desired location, and
instead express elsewhere. Fits to randomly shuffled data
failed to express, producing summed squared scores no smal-
ler than 9.6 times that of the biological data (supplementary
fig. S1, Supplementary Material online).

As an additional quality control, we checked that our op-
timization algorithm found the unique true minimum of the
cost function as follows. First, the model resulting from the fit
was used to calculate the mRNA expression patterns for each
construct. Then, using the calculated mRNA as input, the
model was reoptimized to see whether the initial parameter
set was recovered. The model with the best fit had an rms
score of 0.0274, which is very close to the minimum rms score
possible of 0. By comparing the initial model parameters with
the reoptimized model, it can be seen that majority of the
parameters recovered were within 10% of the initial param-
eter set (supplementary table S2, Supplementary Material
online).

Experimental Resurrection of Ancestors

Because we are interested in inferring the modes of functional
evolution of the enhancer, it is necessary to determine the
branch-specific changes that have occurred in the phylogeny
of the sequence. To do this, we must first reconstruct the
ancestral state of the S2E sequences. This is because it is not
possible to determine whether compensatory evolution has
occurred using extant sequences alone because the changes
could be the result of convergent evolution. For a rapidly
evolving segment of noncoding DNA, however, ancestral
sequence reconstruction is challenging. Ancestral reconstruc-
tion using Bayesian inference techniques often finds multiple
sequences with similar posterior probabilities. To refine the
selection of an ancestral sequence, we took advantage of our
transcription model to calculate an rms score for each one.
Following our assumption of functional conservation, we
then chose as a plausible ancestral sequence, one with both
a low posterior probability and a low rms. If, upon in vivo
testing with a transgenic reporter, the sequence could drive
stripe 2 expression as the model predicted, we could then use
it to infer the branch-specific changes and apply the tran-
scription model to analyze their functional effects on
expression.

Fig. 1. Model fit to eight homologous S2E sequences using MSE2
expression as reference shows good fit to all sequences. Shown are
the calculated expression patterns for the homologous S2E enhancers
from Drosophila mel, D. sim, D. sec, D. sim, D. yak, D. ere, D. ana, D. pse,
and D. moj at different stages during nuclear division cycle 14A. The
interphase of nuclear division cycle 14A is approximately 50-min long
and has been partitioned into eight time classes (T1–T8) each about
6 min in length as previously described (Surkova, Kosman, et al. 2008).
Only time classes T2–T6 were used in the model fit. The dotted line
 corresponds to the observed expression pattern of a P-element
transformant line expressing eve–lacZ under control of the D. mel MSE2.
The observed D. mel MSE2 pattern was used as a reference in the model
fit. The rms score between the model calculated pattern for all species
and the reference D. mel MSE2 expression was 5.24. Figure inset in each
time class shows a representative in situ hybridization and the dashed
box shows the embryo region being modeled (35.5–93.5% EL).
We proceeded to generate the putative ancestral S2E sequences as follows: First, we aligned homologous S2E sequences at conserved 5' and 3' blocks using ProbConsMorph, a hidden Markov-based software that uses PWMs for improved alignment of CRM (Kim et al. 2009) sequences. Bayesian inference was then used to sample the posterior distribution of the ancestral sequences at select internal nodes of the Drosophila phylogenetic tree. A collection of ancestral S2E sequences was inferred for the common ancestors of 1) D. sim and D. sec (sim–sec); 2) D. mel, D. sim, and D. sec (mel–sim–sec); and 3) D. ere and D. yak (ere–yak) (fig. 2).

Using the optimized transcription model, the rms score between the predicted expression pattern of a sampled sequence and the observed eve stripe 2 pattern was calculated. By plotting the posterior probability of alternative ancestral sequences versus the calculated rms scores, we find that there is increasing variance in the expression scores of sampled sequences with longer time of divergence (fig. 3, compare top, middle and bottom panels; $P$ value $\leq 2.2 \times 10^{-16}$, Levene's test for homogeneity of variance). However, no linear correlation was observed between the posterior probability and the rms score (fig. 3; Pearson correlation coefficient ($r$) of $-0.0618$, $-0.1705$, and $-0.0065$ for sim–sec, mel–sim–sec, and ere–yak sampled sequences). This implies that the...

**Fig. 2.** Phylogenetic relationships of all Drosophila species used in the initial model fit and alignment. Red circles denote the internal nodes of the Drosophila phylogenetic tree where the Bayesian inference reconstruction of ancestral eve S2E sequences was carried out. The length of the branches in the tree is scaled according to divergence times. The last common ancestor between mel and moj was 62.9 Ma, between mel and wil 62.2 Ma, between mel and obs 54.9 Ma, between mel and ana 44.2 Ma. Divergence times for the inferred ancestral S2E sequences were 0.93, 5.4, and 10.4 My for the sim–sec, mel–sim–sec, and ere–yak ancestors, respectively. Divergence times are genomic mutation-clock based estimates (Tamura et al. 2004).

**Fig. 3.** Scatter plot of the phylogenetic posterior probability and the predicted rms score of sampled S2E sequences show no correlation. Top, middle, and bottom panels show the scatter plots for the sampled S2E sequences from the sim–sec ($n = 118$), mel–sim–sec ($n = 250$), and ere–yak ($n = 6,234$) internal nodes of the Drosophila phylogenetic tree. Each black dot represents a single sampled sequence. The vertical axis corresponds to the posterior probability of each ancestral sequence based on phylogenetic inference. The rms score between the predicted expression pattern given by the transcriptional model and the reference MSE2 pattern of the eve–lacZ reporter line. The value of the Pearson product–moment correlation coefficient ($r$) is shown in each panel. No correlation between the posterior probabilities of the sequence and the corresponding model rms was observed. The dotted line marks the viability threshold, defined as the maximum calculated rms score for all eight homologous sequences used in the model fit (rms = 5.68). The red star (arrow) to the left of the viability threshold represents the putative ancestral S2E selected for synthesis and experimental validation. The putative ancestral S2E for each internal node corresponds to the sequence with the highest posterior probability from the subset of sampled sequences with an rms below the viability threshold. In the case of the ere–yak internal node, two additional sequences were selected as controls. The blue triangle (arrow) represents a sampled sequence with a high posterior probability and high rms (ere–yak/Hpp-Hrms). The green triangle (arrow) represents a sampled sequence with a low posterior probability and low rms (ere–yak/Lpp-Lrms). Each panel contains a figure inset showing a part of the Drosophila phylogenetic tree with a red star over the sampled internal node.
Bayesian posterior probability and the predicted model rms are independent of each other.

We determined the rms score at the threshold of embryo viability by assuming that all current extant S2E sequences are viable. Therefore, the viability threshold was set as the maximum model rms score from the set of sequences used in the fit (supplementary table S3, Supplementary Material online). This threshold ensures that the candidate ancestral sequences have a predicted expression pattern close to that of the reference eve stripe 2. The most likely ancestral sequence for each node sampled was determined by selecting the sequence with the highest posterior probability that lies below the threshold. Because of the much greater dispersion in the rms scores of the more distant ere–yak ancestral sequences, two additional sequences were selected as controls for our methodology: an S2E with a high ancestral reconstruction posterior probability and high rms score (ere–yak/Hpp–Hrms), and an S2E sequence with a low ancestral reconstruction posterior probability and a low rms score (ere–yak/Lpp–Lrms).

The selected ancestral sequences were synthesized and were, together with four wild type (mel, ere, yak, and pse) S2E sequences placed 42 bp upstream of a eve–lacZ attB reporter construct (fig. 4A). We then generated transgenic fly lines carrying individual reporter constructs by site-specific integration in mel embryos. Embryos carrying the eve–lacZ reporter were imaged during nuclear division cycle 14A. The dashed line represents the observed expression data. The number of embryos used per time class and reporter line are listed in supplementary table S6, Supplementary Material online. The solid red line corresponds to the initial model fit. The solid blue line corresponds to the extended model fit.

**Compensatory Evolution of S2E**

As the model predictions for the sim–sec and mel–sim–sec ancestral reconstructions were a good fit to the observed...
data, we initially sought to learn from the model about any functionally compensatory evolution that occurred in the cis S2E enhancer within the mel, sim, and sec clade. Compensatory evolution in cis was determined by calculating the predicted changes in transcription factor occupancy and activity along the branches of the Drosophila phylogenetic tree.

The central idea of the transcriptional model is that activators recruit adapter factor proteins that in turn lower the energy barrier for transcriptional initiation. By separating the contributions of the different transcription factors toward adapter factor recruitment, we can analyze the different cis-regulatory elements and describe their different expression patterns as a balance between activating and repressive components. Because repression is represented in the model by prevention of activation, we define the repressive contribution of Gt as the decrease in adapter factor recruitment from the peak of stripe 2 expression to the anterior border position that is attributable to Gt. Analogously, we define the repressor contribution of Kr as the decrease in adapter factor recruitment from the peak of stripe 2 expression to the posterior border position that is attributable to Kr. Gt and Kr repressive contributions are then calculated using supplementary equations (S18) and (S19) (Supplementary Material online), respectively. In the case of activation, we define the activator contribution as the maximum number of adapter factors recruited in the absence of repression (supplementary eq. S13, Supplementary Material online). For Bcd, we calculate the activation contribution using supplementary equation (S14) (Supplementary Material online). We do the analogous calculation in the case of Hb activation.

Comparison of total Bcd and Hb activation during the course of S2E evolution shows a loss of activation along the branch leading from mel–sim–sec to mel (fig. 5A). This loss is a consequence of a mutation in the mel lineage leading to a decrease in the binding affinity of a Bcd site and the creation of a new overlapping Kr site, further reducing its occupancy (fig. 5D, red arrows). Concomitant with the loss of Bcd activation, the use of supplementary equation (S18) (Supplementary Material online) revealed a decrease in the repressive contribution of Gt in the formation of the anterior eve stripe 2 border (fig. 5B). The reduction in Gt repression is due to the loss of a Gt site at the 3’ end of the S2E (fig. 5D, black arrow). Although there was an increase in the total amount of Kr bound along this branch, the net functional effect of this change as measured by equation (S19) (Supplementary Material online) on the repression contribution of Kr at the posterior boundary of stripe 2 was negligible (fig. 5B and C). An analogous evolutionary analysis carried out along the branches of the mel, sim, sec clade showed that total Bcd and Hb activation levels changed in tandem with Gt repression (fig. 5E). In each case, an increase or decrease in total Bcd and Hb activation is matched with a respective increase of decrease in Gt repression at the anterior border.

To expand our analysis, we used the newly observed expression data—representing a total of 558 imaged and quantitated embryos—to fit a new implementation of the transcriptional model (described in Kim et al. 2013). In contrast to the initial model, which was fit using the same data set for eight homologous S2E sequences, the extended model was fit to the measured expression during time classes T1–T7 of four wild type S2E sequences (mel, ere, yak, and pse) and five putative ancestral S2E sequences (sim–sec, mel–sim–sec, ere–yak, ere–yak/Hpp–Hrms, and ere–yak/Lpp–Lrms). In addition to quenching and co-activation, the new model also incorporates direct repression (Arnosti et al. 1996; Gray and Levine 1996) as well as cooperativity which have been shown to be important for Bcd binding (Ma et al. 1996; Burz and Hanes 2001).

The expanded model had 52 free parameters (supplementary table S4, Supplementary Material online) and was fit to a set of 3,654 mRNA data points (7 time classes, 58 nuclei, and 9 eve–lacZ reporter constructs). As with the initial model, we confirmed that our algorithm finds a unique minimum (supplementary table S5, Supplementary Material online). We confirmed the significance of the fit by attempting to fit a randomly displaced stripe or completely randomized data, as we did for the initial fit. Fits to randomly displaced stripes failed, showing the same qualitative behavior as randomly displaced stripes in the initial model, with summed squared deviations over ten trials which were larger than those from observed data by factors ranging from 4 to 7.5. Similar results were observed for completely randomized data.

The expanded model showed a good fit of rms 1.82 to the observed expression for all reporter constructs (fig. 4B). We then used this model to determine the functional compensatory changes that occur in a set of 13 homologous S2E sequences spanning 40–60 My of evolutionary distance (Tamura et al. 2004) (fig. 6). The results showed a statistically significant linear correlation between the level of Bcd activation and Gt repression (r = 0.709, P value ≤ 0.0045). Because the S2E sequences for mel, sim, and sec are very similar to one another, we re-determined this correlation with sim and sec removed. Removal of these sequences still resulted in a statistically significant correlation (r = 0.713, P value ≤ 0.0092). In this larger data set, we also observed a statistically significant correlation between Bcd activation and Kr repression (r = 0.704, P value ≤ 0.0049) that was not seen in the smaller mel, sim, sec clade. Note that the expanded model predicts an expression pattern at T6 consistent with eve stripe 2 expression for all 13 homologous S2E sequences, nine of which were not used in the fit (supplementary fig. S3, Supplementary Material online). Some of the predicted patterns from the distantly related takahashii (tak), piciornis (pic), ficushila (fic), and grimshawi (grl) show marked derepression of the 1–2 interstripe, a phenotype suggestive of incomplete Gt repression (Frasch and Levine 1987; Manu et al. 2013). If such derepression in fact occurs in a mel background, it would suggest Gt is expressed at higher levels in these four species. We also carried out an analogous analysis of Hb activation, the results of which are shown in figure 7. The results show a very weak negative correlation between Hb and Gt (r = −0.629, P value ≤ 0.01588), and no statistically significant correlation between Hb and Kr activities (r = −0.4180, P value ≤ 0.1369).
Discussion

Our results show that a model of stripe 2 transcriptional regulation, constrained by expression data from \textit{mel}, accurately predicts the functional conservation of expression patterns of reconstructed common ancestral sequences of the \textit{mel}, \textit{sim}, and \textit{sec} clade. We also demonstrated that an expanded transcriptional model incorporating cooperative interactions and direct repression can accurately describe the observed expression from all reporter constructs. Our model describes evolutionary changes in binding sites that result from CRM sequence substitutions, but more significantly it provides a representation of the evolutionary changes in
functional interactions between bound TFs that give rise to the conserved expression patterns observed.

The approach used here affords new confidence in the reconstruction of ancestral sequences by Bayesian methods and has increased our understanding of evolutionary constraint. The correct prediction of the expression of \textit{mel–sim–sec} and \textit{sim–sec} ancestral sequences obtained by MCMC demonstrates that such Bayesian reconstruction is, at least in this case, neutral with respect to expression. This point can only be established by a model with an explicit and quantitative connection between DNA sequence and expression. With respect to evolutionary constraint, these results

**Fig. 6.** Functional analysis of homologous S2E sequences shows compensatory evolution between Bcd, Gt, and Kr activities. (Top) Gt repression contribution versus Bicoid activation. (Bottom) Kr repression contribution versus Bicoid activation. Activation and repression contributions are calculated as described in figure 5. Dashed lines correspond to a linear regression fit of the repressive contribution as a function of Bcd activation. Each dot corresponds to a homologous S2E sequence. The 13 homologous S2E sequences tested correspond to the following \textit{Drosophila} species: \textit{Drosophila} species: \textit{melanogaster} (\textit{mel}), \textit{simulans} (\textit{sim}), \textit{sechellia} (\textit{sec}), \textit{erecta} (\textit{ere}), \textit{yakuba} (\textit{yak}), \textit{eugracilis} (\textit{eug}), \textit{takahashii} (\textit{tak}), \textit{kikkawai} (\textit{kik}), \textit{mojavensis} (\textit{moj}), \textit{grimshawi} (\textit{gri}), \textit{ficusphila} (\textit{fic}), \textit{picticornis} (\textit{pic}), and \textit{pseudoobscura} (\textit{pse}).

**Fig. 7.** Functional analysis of homologous S2E sequences shows no correlation between Hb, Gt, and Kr activities. Hb activation is calculated at the anterior of stripe 2. (Top) Gt repression contribution versus Hb activation. (Bottom) Kr repression contributions versus Bicoid activation. Activation and repression contributions are calculated as in figure 5. The homologous S2E sequences tested are the same as in figure 6.
permit comparative functional analysis between mel, sim, and sec. The extended model can predict the expression of homologous S2E sequences that were not included in the training set (supplementary fig. S3, Supplementary Material online), and analysis of this model provides functional insight (fig. 6) into a broad set of extant and resurrected enhancers. A limitation of our initial model was shown by our inability to correctly predict expression driven by ere–yak ancestral sequence. Nevertheless, it is reasonable to expect that wider phylogenetic expression data sets and more precise models of transcriptional control will ameliorate this problem.

The analysis of the resurrected mel–sim–sec ancestor and the extant mel, sim, and sec S2E ancestral enhancers reported here indicate that there has been evolutionary compensation between Bcd activation and Gt repression. This observation can be understood in terms of well-known properties of the Drosophila segmentation system. Experimentally, it is known that in the absence of Gt in trans, eve expression spreads anteriorly to the 1–2 interstripe with fatal results (Frasch and Levine 1987; Petschek et al. 1987), and inactivation of all MSE2 Gt sites in cis abolishes the anterior border of reporter expression. Theoretical analysis has demonstrated that the border location is determined by a balance of activation and repression (Reinitz and Sharp 1995; Janssens et al. 2006), and hence it is reasonable to expect that an evolutionary change in Gt repression should be balanced by a change in Bcd activation of the opposite sign. We have provided additional validation for this result by using an expanded model fit to the observed data to show that Bcd-Gt functional compensation has occurred throughout 40–60 My of Drosophila evolution.

In contrast, while the occupancy of Kr evolved on the lineage leading from the mel–sim–sec ancestor, our initial analysis indicated it had little or no effect on the degree of repression at the posterior border of eve stripe 2 (fig. 5). However, the incorporation of additional homologous S2E sequences in our analysis allowed us to see that Bcd activation and Kr repression are also functionally coupled. This suggests that Bcd-Kr functional compensation occurs over longer evolutionary time frames than is the case for Bcd-Gt. It is also possible that the lower concentration of Bcd at the posterior border of stripe 2 reduces the importance of Kr sites for proper regulation, so that changes in Kr binding may be subject to less extensive purifying selection than changes in Gt.

We failed to predict the correct expression pattern of the putative ancestral ere–yak S2E sequences. The failure of our original model can be attributed to the inclusion of the ere S2E sequence in our initial fit without corresponding quantitative measurements. Because the expanded model incorporates the actual quantitatively measured eve–lacZ driven by the ere S2E sequence as well as the putative ancestral constructs, we can infer whether changes in the trans environment in D. ere explain the lack of T6 S2E expression driven by the ere S2E. As can be seen in supplementary figure S4, Supplementary Material online, we can rescue in silico the correct stripe 2 expression in ere if we increase the Bcd concentration at stripe 2 by 90%. Because rescaling the Bcd gradient shifts the expression domains of Hb, Kr, Gt, Kni, and eve patterns to the posterior, partially compensating for the increase in Bcd (Driejer and Nüsslein-Volhard 1988; Houchmandzadeh et al. 2002), a larger rescaling would be required overall to attain a 90% increase at stripe 2. It is possible therefore that Bcd levels have increased in the lineage leading to D. ere and that the S2E sequence has evolved to compensate for this difference. In fact, native eve stripes in ere are shifted posteriorly with respect to those of mel (Ludwig et al. 2005), consistent with what would be expected with an increase in Bcd levels (Driejer and Nüsslein-Volhard 1988).

The ancestral sequence reconstruction of sim–sec and mel–sim–sec S2Es sheds light on the dynamics of enhancer evolution. Evolutionary changes in the sequences occur along individual lineages as they diverge from a common ancestor. Mechanistically, the effect of each mutation is contingent on the order of the previous changes that have occurred on a particular branch. Compensatory evolution is one example where the functional analysis used here is beneficial. By comparing the evolutionary changes between the ancestral and the extant sequences, the model provides mechanistic insight into the functional constraints that create correlated changes. We can see from figure 5E that there has been an increase in total activation going from mel–sim–sec to sim–sec compensated by an increase in Gt repression. This increase was then followed by a small decrease in total activation and Gt repression in the branch from sim–sec to sec (fig. 5E). Because the time of divergence of mel, sim, and sec is on the order of 1 My (Tamura et al. 2004) this indicates that compensatory evolution can occur within a very short time frame.

The approach presented here provides a refined perspective on the evolution of cis-regulatory regions. One previous study involved analysis of the affinity of sites on a TF by TF basis (He, Holloway, et al. 2011), while another imposed a functional constraint by requiring the sequence to maintain a minimum number of sites for each TF in the enhancer (Lusk and Eisen 2010). Yet another model investigated the rate at which turnover could occur under compensatory evolution, but only considered compensation between homotypic sites (Durrett and Schmidt 2008). The model used here goes further by incorporating a richer set of regulatory mechanisms providing an explicit map between the genotype of the enhancer and the expression phenotype which it drives. Given the opportunistic nature of evolution and the multiple mechanisms by which S2E evolution has occurred, this study highlights the need for realistic models of transcription to understand CRM evolution. The approach we have taken in this study where we have combined phylogenetic, functional modeling, and gene synthesis techniques, produces a deeper understanding of both the forces driving the evolution of CRM and also transcriptional regulatory mechanisms.

Materials and Methods

Alignment and Ancestral Reconstruction of eve S2E

Sequence alignment was done using the ProbConsMorph multiple alignment tool (Kim et al. 2009). ProbConsMorph takes as input the PWMs of transcription factors and a phylogenetic tree in Newick format. A ProbConsMorph alignment
was created using eight homologous eve S2E sequences from D. mel, D. sim, D. sec, D. yak, D. ere, D. ana, D. pse, and D. moj. Sequences used in alignment, selected ancestral sequences, and Newick tree are listed in the supplementary information S1 (Supplementary Material online).

We reconstructed the ancestral sequence at each internal node using Bayesian inference in MrBayes 3.1.2 (Huelsenbeck and Ronquist 2001). Sequence at each internal node was reconstructed in a separate analysis by enforcing a constraint only on the resolution of that internal node, based on accepted relationships between the species. The reconstruction was conducted under a general time reversible model with a gamma distribution and four rate categories. Simultaneously, we used a simple binary model to reconstruct the presence or absence of a gap at each position, under the assumption that the absence of a gap in all taxa cannot be observed. Reconstruction at each node was run for 5 million Markov chain Monte Carlo generations with six chains and sampled every 100 generations after a 100,000 generation burn in. We recorded the sequence from each generation based on the maximum of the gap or nucleotide probabilities. We then summed up the proportion each unique ancestral sequence was observed over the 5 million generations as an estimate of its phylogenetic probability.

Transcriptional Model and Functional Analysis
A detailed description of the transcriptional model, activation, and repression contributions is given in supplementary materials and methods (Supplementary Material online).

Plasmid Construction and Site-Specific Integration
Site-specific integration into D. mel embryos was achieved through recombimase-mediated cassette exchange (RMCE) (Bateman et al. 2006). The attB donor vector piB-GFP vector (gift from J.R. Bateman [Bateman et al. 2006]) was modified to carry a S2E enhancer flanked 5’ by Sall and 3’ by HindIII placed at −42 bp from the +1 transcriptional start site of a eve–lacZ fusion driven by the eve basal promoter (fig. 4A). Briefly, the piB-GFP vector was cut with BamHI and XbaI to remove the mGFP transcript. A BamHI/XbaI fragment from a eve–lacZ pCaSpeR plasmid (Small et al. 1993) was then ligated into the piB vector backbone resulting in the piB-lacZ vector carrying the lacZ coding region. A HindIII/BamHI polymerase chain reaction (PCR) fragment containing the eve promoter (from −42 to +1 bp), the 100 bp 5' UTR and the first 22 codons of the eve gene was then inserted between HindIII and BamHI sites in piB-lacZ resulting in the piB-eve-lacZ vector. Wild-type and reconstructed ancestral S2E sequences were then inserted as Sall and HindIII PCR fragments resulting in the piB-S2E-eve-lacZ donor vector. Plasmids carrying the mel, ere, yak, and pse wild type S2E sequences are as described (Ludwig et al. 1998). Plasmids containing the reconstructed ancestral S2E sequences were ordered from GENEART AG (http://www.lifetechnologies.com/us/en/home/life-science/cloning/c-misc/geneart.html, last accessed January 27, 2014).

Site-specific integration by RMCE of the S2E-containing donor vector was done using a fly line containing target attB sites on chromosome arm 3R (Bloomington stock 25091). Embryo injection and transformation was performed by Rainbow Transgenic Flies, Inc. (http://www.rainbowgene.com, last accessed January 27, 2014). Correct reporter gene insertion was verified by genomic extraction, PCR amplification, and sequencing of the S2E–eve–lacZ reporter gene in all transformant lines.

In Situ Hybridization
Staged embryos (2–4 h) bearing the wild type and reconstructed ancestral enhancers eve-lacZ fusion gene were collected, fixed, and stained for eve–lacZ mRNA by in situ hybridization and for Eve protein by immunostaining as previously described (Janssens et al. 2006). To visualize the embryo nuclei, embryos were stained and mounted as in Kim et al. (2013).

Quantitative Expression Data
Embryos were imaged using a Vti Infinity three multipoint confocal system (Visitech International, Sunderland, UK) mounted on a Zeiss AxioPlan 2 microscope (Carl Zeiss, Inc., USA). A backthinned EMCCD camera (Hamamatsu Photonics UK Ltd, Hertfordshire, UK) having a resolution of 512 × 512 pixels and 16-bit depth was employed for fluorescence detection. Fluorophores were excited with three different laser wavelengths (491, 561, and 642 nm) and the detection was done using a triple band (488, 568, and 647 nm) dichroic emission filter. The corresponding detection channels were 488 nm for nuclei, 568 nm for Eve protein, and 647 nm for eve–lacZ mRNA. All embryos were imaged with the same microscope settings. Microscope settings were standardized using an eve–lacZ fly line (M32C) bearing a fusion of the eve MSE3 and MSE2 enhancers (Kim et al. 2013). This fly line expresses eve–lacZ mRNA at levels greater then any other line tested in this work. Camera exposure times for the nuclei, Eve protein, and eve–lacZ mRNA channels were 1.5, 0.7, and 0.85 s, respectively. Embryos were classified temporally as belonging to either C13, or one of eight time classes (T1–T8), each about 6.5-min long, in cycle 14A (C14A), as described (Surkova, Kosman, et al. 2008). Image analysis and segmentation was done as previously described (Janssens et al. 2006).

Initial eve stripe 2 Expression Data
The quantitative expression data for eve stripe 2 pattern used in the initial model fit was obtained from embryos carrying a P-element eve–lacZ reporter driven by the minimal stripe 2 element (line 1511B, gift from M. Levine) as previously described (Kim et al. 2013).

Transcription Factor Data Set
Concentration profiles for Bcd, Hb, Cad, Kr, Gt, Kni, and TII were obtained from the FlyEx database (http://flyex.uchicago.edu/flyex/index.jsp, last accessed January 27, 2014) as previously described (Poustelnikova et al. 2004; Pisarev et al. 2005;
Generation and Selection of PWMs

A family of PWMs for Bcd, Hb, Gt, Kr, Kni, Tll, and Cad were created by running MEME (Bailey et al. 2006) on a set of SELEX-derived binding sites for each ligand. The set of binding site oligonucleotides were obtained from the Berkeley Drosophila Transcription Network Project using a SAGE like SELEX protocol (Roulet et al. 2002; Orgawa and Biggin 2012). Because the SELEX PWMs for Kni and Tll performed poorly in tests of false positive and false negative binding site detection, PWMs derived from a set of experimentally characterized binding sites for Kni (27) and Tll (20) were used instead (Rajewsky et al. 2002). For the extended model fit, the Bcd and Hb PWMs were replaced by MITOMI-derived PWMs described in He, Holloway, et al. (2011), the Kni PWM was replaced with a bacterial one-hybrid obtained from the Fly Factor Survey Database (http://pgfe.umassmed.edu/ffs/, last accessed January 27, 2014). Additional PWMs for Slp, Zld, and DSTAT were used for the extended model and described in the supplementary information S1 (Supplementary Material online). All PWMs used in this work are listed in supplementary materials and methods (Supplementary Material online).

Computation and Optimization

The model equations were implemented in C as previously described (Janssens et al. 2006). The code version was 1.15 and can be downloaded from http://flyex.uchicago.edu/newlab/download.shtml, last accessed January 27, 2014. Parameters were determined by minimizing the summed squared difference between the model output and the observed data. Because the concentration of ligand factors important for the transcriptional regulation at the terminal pole regions of the embryos have not been determined, model fits and calculations are carried out only from 35.5% to 93.5% egg length (EL) along the embryo A-P axis. Only time classes T2–T6 in cycle 14A were used in the model fit. The total number of data points used in each model fit was 232 (58 nuclei and 5 time classes). Optimization was performed using the Lam-simulated annealing schedule (Lam and Delosme 1988, Reinitz and Sharp 1995). Model parameters are listed in supplementary tables S1 and S4 (Supplementary Material online). The total model root mean square (rms) for all S2E sequences used in the initial model fit was 5.24. The calculated model rms scores for the S2E sequences from individual Drosophila species were as follows: 4.98 (D. mel), 5.33 (D. sim), 4.96 (D. sec), 5.32 (D. ere), 5.64 (D. yak), 5.68 (D. ana), 5.36 (D. pse), and 4.55 (D. moj). Because the reference eve stripe 2 pattern used to optimize the initial model was obtained from a P-element eve–lacZ reporter its level of expression depends on where it lands in the genome due to position effect. To directly compare the initial model calculated expression with the observed site-specific reporter constructs, a scaling factor was applied. The scaling factor is multiplied to the model calculated expression such that it matches the measured expression of the mel S2E reporter construct. The scaling factor was determined as 1/3.4.

Supplementary Material

Supplementary information S1, figures S1–S4, and tables S1–S6 are available at Molecular Biology and Evolution online (http://mbe.oxfordjournals.org/).

Acknowledgments

The authors thank Manu, K. Barr, B. He, and J. Gavin-Smyth for valuable advice and discussions, and also thank two anonymous reviewers for extremely useful comments on the manuscript. This work was supported by the University of Chicago and the National Institute of Health grants RO1 OD010936 (formerly RO1 RR07801), RO1 GM70444, and NIH P50 GM081892.

References


