Modeling the zebrafish segmentation clock’s gene regulatory network constrained by expression data suggests evolutionary transitions between oscillating and non-oscillating transcription

Jamie Schwendinger-Schreck, Yuan Kang, Scott A. Holley

Department of Molecular, Cellular and Developmental Biology. Yale University, New Haven, CT 06520.
Running Title: Simulated Annealing and the Clock

Key words: gene regulatory network, simulated annealing, segmentation clock, zebrafish, \textit{her/hes}

Corresponding author: Scott Holley, P.O. Box 208103; 203-432-3230; scott.holley@yale.edu
ABSTRACT

During segmentation of vertebrate embryos, somites form in accordance with a periodic pattern established by the segmentation clock. In the zebrafish (Danio rerio), the segmentation clock includes six hairy/enhancer of split-related (her/hes) genes, five of which oscillate due to negative autofeedback. The non-oscillating gene hes6 forms the hub of a network of ten Her/Hes protein dimers, which includes seven DNA binding dimers and four weak or non-DNA binding dimers. The balance of dimer species is critical for segmentation clock function, and loss of function studies suggest that the her genes have both unique and redundant functions within the clock. However, the precise regulatory interactions underlying the negative feedback loop are unknown. Here, we combine quantitative experimental data, in silico modeling and a global optimization algorithm to identify a gene regulatory network (GRN) designed to fit measured transcriptional responses to gene knockdown. Surprisingly, we find that hes6, the clock gene that does not oscillate, responds to negative feedback. Consistent with prior in silico analyses, we find that variation in transcription, translation, and degradation rates can mediate the gain and loss of oscillatory behavior for genes regulated by negative feedback. Extending our study, we found that transcription of the non-oscillating Fgf pathway gene sef responds to her/hes perturbation similarly to oscillating her genes. These observations suggest a more extensive underlying regulatory similarity between the zebrafish segmentation clock and the mouse and chick segmentation clocks, which exhibit oscillations of her/hes genes as well as numerous other Notch, Fgf and Wnt pathway genes.
INTRODUCTION

Segmentation of the anterior-posterior axis in vertebrate embryos results in the formation of somites, which are paired metameric structures within the paraxial mesoderm that lie on either side of the notochord and neural tube. Somites form sequentially, from anterior to posterior, during a process called somitogenesis. The precise temporal and spatial control of somitogenesis was first explained by the “clock and wavefront” model (COOKE and ZEEMAN 1976). This model proposes that segment formation results from the interaction of two distinct components: a posterior-progressing wavefront that coincides with axis elongation and genetic oscillations (the segmentation clock). Somite border formation occurs only when the wavefront encounters cells in an appropriate phase of the clock. Thus, the clock gates wavefront activity, resulting in the creation of discrete, repeated tissue borders. The wavefront consists of Fgf and Wnt signals that originate in the tailbud, forming a posterior to anterior gradient that recedes as the tail elongates (AULEHLA et al. 2007; DELFINI et al. 2005; DUBRULLE et al. 2001; DUBRULLE and POURQUIÉ 2004; NAICHE et al. 2011; SAWADA et al. 2001). The segmentation clock is comprised of the hairy/enhancer of split-related (her/hes) family of Notch target genes (BESSHÖ et al. 2001; HOLLEY et al. 2000; PALMEIRIM et al. 1997). Components of the Fgf and Wnt signaling pathways oscillate in mouse and chick, suggesting a role for Fgf and
Wnt in the clock, but the her/hes genes are the only conserved oscillatory clock components in mouse, chick and zebrafish (KROL et al. 2011). The anole lizard and alligator segmentation clocks share some features with the zebrafish segmentation clock along with some characteristics found in the mouse or chick suggesting that there is substantial variation in the segmentation clock network among vertebrate species (ECKALBAR et al. 2012).

The her/hes genes form both hetero- and homodimers and negatively regulate their own transcription (BESSHO et al. 2003; BREND and HOLLEY 2009a; GAJEWSKI et al. 2003; GIUDICELLI et al. 2007; HANISCH et al. 2013; HIRATA et al. 2002; HOLLEY et al. 2002; OATES and Ho 2002; SCHRÖTER et al. 2012; TROFKA et al. 2012). Due to a short half-life, this negative feedback results in repeated cycles of expression and repression of the her/hes genes within the presomitic mesoderm (PSM), the frequency of which correlates with the rate of somite formation (AY et al. 2013; DELAUNE et al. 2012; HARIMA et al. 2013; HIRATA et al. 2004; ISHII et al. 2008; LEWIS 2003; MONK 2003; SCHRÖTER and OATES 2010; TAKASHIMA et al. 2011). Within the zebrafish segmentation clock, the repertoire of her/hes genes is expanded, and at least six members of the her/hes family play a role in clock oscillations—her1, her7, her11, her12, hes6 and her15 (HENRY et al. 2002; HOLLEY et al. 2002; KAWAMURA et al. 2005; OATES and Ho 2002; SHANKARAN et al. 2007; SIEGER et al. 2006; SIEGER et al. 2004). We previously determined the dimerization partners of her1, her7, her11, her12, hes6 and her15. Between these six proteins, ten dimers form in vivo, only seven of which
can strongly bind DNA. We identified Hes6 as the hub of the protein dimer network, as it participates in half of the dimer species (TROFFKA et al. 2012). hes6 is also unique in that it is the only her/hes clock gene that does not oscillate during segmentation. Rather, hes6 forms a posterior to anterior gradient in the PSM that is promoted by Fgf signaling (KAWAMURA et al. 2005). However Hes6 protein levels do oscillate due to their heterodimerization with the oscillating Her proteins (SCHRÖTER et al. 2012).

Loss of function studies suggest that the her genes have both unique and redundant functions within the clock. For example, though Her7 can bind DNA as a dimer with Hes6, Her7 primarily functions by modulating the topology of the dimer network by sequestering Hes6 and limiting the molecules available to dimerize with other Her proteins (SCHRÖTER et al. 2012; TROFFKA et al. 2012). Several models of zebrafish segmentation have been published and provide insights into the parameters most important to maintaining oscillations and identify ranges of parameter values that enable sustained oscillations of the appropriate frequency and amplitude (CAMPANELLI and GEDEON 2010; CINQUIN 2007; LEWIS 2003; MONK 2003; TERRY et al. 2011). More recent models of zebrafish segmentation have also taken into account the specific Her/Hes dimers that form in vivo (AY et al. 2013; SCHRÖTER et al. 2012). However, we still do not understand the precise transcriptional regulation of each her/hes gene, or the role of her11, her12, or her15 in the clock. Therefore, we aimed to generate a model of zebrafish segmentation that includes all six well-characterized her/hes
genes, uses experimentally determined dimer species, and allows for differential regulation of each her/hes gene. Previous methods to model networks of transcription factors included data on binding affinities, cell-specific protein concentrations and chromatin accessibility (KAPLAN et al. 2011). However, quantifying these parameters can be challenging. Furthermore, transcription factor binding may not always lead to significant changes in transcription (BIGGIN 2011). We aimed to sidestep some of these complications by generating a model of the segmentation clock that depends on quantitative measurements of transcriptional output. We measured changes in transcription rate of her1, her7, her11, her12, her15 and hes6 using qPCR in response to gene perturbation of those same genes by morpholino-mediated knockdown. We used simulated annealing, a global optimization method, to search a large parameter space for the network that best matches our experimental data. We performed a series of in silico experiments to validate this approach but ultimately found that the problem was too under-constrained to identify the in vivo network. Thus, as a case study, we chose a gene regulatory network (GRN) that appeared 5 times among the 78 networks produced by the simulated annealing. This GRN matches 55.6% of the qPCR data within 2 standard deviations. Further, the average log fold difference between simulated and experimental qPCR data is 0.102 for the training data, while novel triple knockdowns were predicted with an average difference of 0.134. In comparison, a set of 100 randomly generated networks not fit using simulated annealing have an average log fold difference between simulated and qPCR values ranging from 0.13 to 0.44. Further, this
network recapitulates known synergistic interactions between her1+her7 morpholino and her1+hes6 morpholino. However, a number of inconsistencies between the predicted GRN and in vivo behavior of the network suggest we lack sufficient information to determine the actual in vivo network.

Despite these limitations, an analysis of our simulated annealing results enables more in depth study of previous assumptions about the her/hes network. In particular, we find that hes6 responds to negative feedback by the Her/Hes dimers, and given equivalent biochemical parameters between all her/hes genes, our GRN model predicts that hes6 oscillates. Since hes6 has not been observed to oscillate in vivo, we varied an array of biochemical parameters including transcription, translation, and degradation rates of hes6, to identify parameter sets that reduced the amplitude of hes6 oscillations in our computational model. Prior in silico analyses of the segmentation clock found that varying these parameters can determine whether a given gene oscillates and offered an explanation for the observed differences between the zebrafish, mouse and chick segmentation clocks (Lewis 2003). Indeed, we found that protein degradation, transcription, and translation rates were highly constrained in models that showed minimized hes6 oscillations. To extend this analysis, we examined the transcriptional response of non-oscillatory Wnt and Fgf target genes in the zebrafish and found that sef, but not axin2, respond to perturbation of the her/hes somitogenesis clock in a manner significantly correlated with that of oscillating her genes. These observations suggest a more extensive underlying regulatory similarity between the zebrafish segmentation clock and the mouse and chick
segmentation clocks, which exhibit oscillations of her/hes genes as well as numerous other Notch, Fgf and Wnt pathway genes.

MATERIALS AND METHODS

Zebrfish maintenance: Zebrafish care and breeding followed standard protocols (NÜSSLIN-VOLHARD and DAHM 2002) as approved by Yale IACUC. The wild-type strains used were Tübingen and Tupfel long fin.

Morpholino injection and luciferase assay: Percentage knockdown was calculated using a modified luciferase assay as previously described (TROFKA et al. 2012) with the following modifications: all fish were raised at 28°C. Morpholino sequences and concentrations used can be found in Table S1. For in situ hybridization, embryos were injected with morpholino at the one-cell stage and raised at 28°C until 80% epiboly, then moved to 18°C overnight. Raising embryos at lower temperatures increases the penetrance of defects due to loss or reduction of hes6 (KAWAMURA et al. 2005; SCHRÖTER and OATES 2010; TROFKA et al. 2012). We raised all embryos at 18°C overnight so as to minimize variation within experimental replicates. Embryos were fixed at the 8-10-somite stage and processed using standard protocols as previously described (HOLLEY et al. 2000). Probes for her1 and hes6 were made from plasmids, while probes for ripply1 and sef were transcribed directly from RT-PCR products by adding a T7 promoter in the antisense orientation. Fluorescent in situ hybridization was performed as
described in (Brend and Holley 2009b) and imaged using a Zeiss LSM 350 confocal microscope. For DIC images, embryos were mounted in methylcellulose at the ~15-somite stage and imaged using a Zeiss Axioskop 2 mot plus widefield microscope. Three experimental replicates were performed for luciferase assays and in situ hybridization.

**Quantitative Real-Time PCR:** Injected or un-injected control embryos were raised at 28°C until 80% epiboly, then kept at 18°C overnight until the 8-10-somite stage. Embryos were then manually dissected in Mark’s Modified Ringer buffer with 10mM HEPES. Approximately 50 PSMs were pooled and stored at -80°C until processing. RNA was extracted using an RNeasy Micro Kit (Qiagen) and cDNA was reverse transcribed using the High Capacity RT kit (AB). qPCR was performed as previously described (Stulberg *et al.* 2012). For analysis of *her/hes* genes, five replicates were performed for all single injections, and three replicates for all double injections. Three replicates were performed for all treatments for *sef* and *axin2* analysis. Primer sequences are provided in Table S1. Primers were designed to include one exonic sequence and one intronic sequence, so only nascent transcripts were measured. Values shown are fold change (Figure 2B) or normalized log fold change (Figure 3A), with fold change calculated as $2^{-\Delta\Delta C_t}$. Because we used morpholinos to knock down protein levels, embryos were mosaic. Therefore, we normalized the qPCR data to help account for mosaicism, so that individual gene responses for each experimental
replicate were measured relative to the average response of all genes (which may be higher in embryos with a greater proportion of morphant cells).

\[ qPCR\text{value} = \frac{\log\text{FoldChange}(x_i)}{\left(\sum_{i=1}^{6} \log\text{FoldChange}(x_i)\right) / 6} \]

Correlation of log fold change qPCR values was calculated in Matlab using Pearson’s linear correlation coefficient, with p-values calculated by Student’s t-test. Average qPCR values with standard error are provided in File S1.

**Computational Methods**

**Model:** A system of delay differential equations similar to those of (Cinquín 2007) describing transcription, translation, dimerization and repression of *her1, her7, her11, her12, hes6* and *her15* was generated using Matlab. Codes are available in Files S3-S5.

\[ \dot{h}_i^m = -\delta_i^m h_i^m + \beta_i^m \times f(h_{i,j}) \]

\[ \dot{h}_i = -\delta h_i + \beta(h_i^m - h_i^m^*) - k_{i,j}^+ h_j h_i^* + k_{i,j}^- h_{i,j} \]

\[ h_{i,j}^p = -\delta^p h_{i,j} + k_{i,j}^+ h_j h_i - k_{i,j}^- h_{i,j} \]

Repression, \( f(h_{i,j}) \), was described using a modified Hill function with either multiplicative logic (Bintu et al. 2005):

\[ f(h_{i,j}) = \frac{1}{\prod 1 + S_k^{i,j} \left( \frac{h_j^{m^*}}{h_i} \right)^3} \]

or additive logic (Balaskas et al. 2012):
Additive logic was used for simulations that identified the GRN described here.

\[ f(h_{i,j}) = \frac{1}{1 + \sum S_k^{i,j} \left( \frac{h_{i,j}}{\theta} \right)^3} \]

\( h_i^m \) represents mRNA for gene \( i \), where \( i \in \{ \text{her1, her7, her11, her12, hes6, her15} \} \). \( h_i \) is the protein for gene \( i \), and \( h_{i,j} \) is the dimer made up of genes \( i \) and \( j \), where only experimentally observed dimers are allowed: Her1/Her1, Her7/Hes6, Her11/Her11, Her12/Her12, Her12/Hes6, Her15/Hes6, Her15/Her15, Her1/Hes6, Her7/Her15, and Hes6/Hes6. Note that only the first seven dimers can be involved in repression, as Her1/Hes6, Her7/Her15, and Hes6/Hes6 do not strongly bind DNA. The mRNA degradation rate for gene \( i \) is given by \( \delta_i^m \), for protein degradation the value is \( \delta_i \), and for dimer degradation it is \( \delta_{i,j} \). Maximum transcription rate is \( \beta_i^m \) and translation rate is \( \beta_i \). Dimerization association and dissociation rates are denoted as \( k_{i,j}^+ \) and \( k_{i,j}^- \), respectively. Delays from mRNA and protein processing are denoted by \( \tau^m \) and \( \tau \). \( S_k^{i,j} \) represents the strength of repression of dimer \( h_{i,j} \) on the target \( h_k \), while \( r_{i,j}^{k} \) is a scaling factor. The value of \( S_k^{i,j} \) for each dimer/transcriptional target pair was randomly selected from a discrete distribution \( \{0, S_{\text{weak}}, S_{\text{strong}}\} \) to generate a connectivity network. The parameter \( n \) denotes the relative difference between strong and weak dimers. For example, in the initial parameter set \( n=1.26 \), so \( S_{\text{weak}}=162.7 \) and \( S_{\text{strong}}=205 \).

All biochemical parameters used in identification of the GRN presented here are
provided in Table S2. We ran several tests of simulated annealing using various assumptions. In addition to testing multiplicative and additive logics of repression, we changed the values of \( n \) and \( S_{\text{weak}} \) to 2.0 and 1627, respectively (for a total of eight tests). In general, we found that additive logic modestly outperformed multiplicative logic, and that our initial estimate of dimer repressor strength (\( S_{\text{weak}}=1627 \)) was too high. There was no discernable advantage to having a greater distinction between weak and strong dimers, as the value of \( n \) did not greatly impact overall performance. Morpholino knockdown was simulated by reducing translation rate in accordance with luciferase assay measurements (Figure S1). Since morpholino injection results in mosaic embryos, we focused on oscillations within a single cell in silico, so that the protein knockdown level would be equivalent to that measured using the luciferase assay. We kept the dynamic model as simple as possible, using only those physical interactions that have been experimentally defined. We kept activation rates constant, and focused only on Her/Hes feedback. Initial parameter values were selected based on a previously published model (CINQUIN 2007) and have already been tuned to enable oscillations of appropriate frequency and amplitude. Since relative values for each her/hes gene have not been measured, we further simplified the parameter set by making all biochemical rates equivalent between her/hes genes so that differences in expression would be due only to differences in network connectivity. For each randomly generated network, all knockdown conditions were simulated for 300 minutes and qPCR data was calculated as log\(_{10}\) fold change in \( x \), with:
where \( t_f = 300 \) minutes. Using this model, we then fit the “black box” of Her/Hes repressor-DNA interactions to experimental data with simulated annealing, a global optimization method. While in vivo, transcription factors bind with a wide range of occupancy levels (BIGGIN 2011), here we discretized repressor binding to be “weak,” “strong,” or “unbound.” We found that using continuous variables to describe repression greatly reduced the efficiency of simulated annealing, while using only binary decisions for binding (bound versus unbound) reduced the performance of output GRNs. Matlab script is available in File S4.

Adaptive simulated annealing (ASA) (INGBER 2012) was used to search the energy space of all possible network configurations for the best possible fit by minimizing the energy score. The ASA algorithm is weakly ergodic. However, ergodicity is not a rigorous concept when referring to sampling of a very large parameter space as in our study due to practical limitations of computational cost (INGBER 2012). Energy score was calculated as the sum over all 21 knockdowns and all six genes of the absolute value of the difference between simulated and experimental qPCR values using the Matlab code “MO_cost_func.m”, available as File S6. The Matlab-ASA interface ASAMIN (MOINS 2002; SAKATA 1999) was used to call ASA. Adaptive simulated annealing uses an exponential temperature gradient that is specific for each parameter, with the reannealing schedule based on the sensitivity of each parameter at the current energy minimum (INGBER 1993). Essentially, this means that the stringency with which networks are
evaluated becomes greater over time. Default parameters for ASA were used, with the following exceptions: the maximum number of accepted states allowed is 1000; the maximum number of generated states allowed is 10,000; integer parameters are included in derivative and reannealing calculations. Biochemical parameters were consistent from posterior to anterior PSM, but equations describing transcription of $her1$, $her7$, $her11$, $her12$, $hes6$ and $her15$ were distinct between posterior and anterior PSM. Only $her1$, $her7$, $her12$, $hes6$ and $her15$ were actively transcribed in the posterior PSM, and only $her1$, $her7$ and $her11$ were actively transcribed in the anterior PSM. Allowed values for $S_{\text{weak}}$ and $S_{\text{strong}}$, which represent the total possible strength of repression, were scaled to compensate for the smaller number of dimers available in the anterior versus posterior PSM. The transition from posterior to anterior occurs halfway through the simulation at $t = \frac{1}{2} t_f$.

Oscillation constraints were added to the simulated annealing algorithm using a power spectrum, the code for which is available as File S7: “getAvgSpectrum.m”. The power spectrum analyzes the dynamics of $her1$ or $her7$ mRNA over a long time span, and can be used to determine whether the gene oscillates, and at what frequency these oscillations occur. Briefly, oscillations for $her1$ and $her7$ were detected using a custom Matlab code that calculates the normalized average power spectrum across a simulation run for 3000 minutes. The long running time is required to generate enough data points to accurately calculate the power spectrum. Since activation is constant in our model, undamped oscillations would hypothetically continue forever. Networks
that fail to generate a high enough power spectrum are discarded. For $her1+her7$ MO and $her1+hes6$ MO synergy conditions, calculation of the power spectrum was used to penalize networks that maintained oscillations of $her1$ or $her7$ under these conditions. Penalties were added to the energy score and scaled with power spectrum values, such that

$$\text{penalty} = 5 \times 10^{-4} \left( \text{Power}_{her1}^{\text{anterior}} + \text{Power}_{her1}^{\text{posterior}} + \text{Power}_{her7}^{\text{anterior}} + \text{Power}_{her7}^{\text{posterior}} \right)$$

so that networks with larger oscillations were penalized more than those with small amplitude or damped oscillations.

One hundred simulations were run when testing experimental results, but simulations were limited by available CPU time with approximately 90 networks completing successfully. Simulations were run in parallel for 168 hours. All predictions were visually inspected to ensure that oscillations of $her1$ and $her7$ were fully damped under $her1+her7$ and $her1+hes6$ knockdown conditions. GRN predictions were then ranked by energy score and Jensen-Shannon divergence to provide two separate metrics for evaluation. Jensen-Shannon divergence values were calculated as follows. Normalized qPCR values were binned into two groups: -1 if qPCR values were less than zero, and +1 if values were equal to or greater than zero. A “state” was defined as the value of all six genes within any given knockdown. Given two bins and six genes, this meant 64 states were possible. The distribution of states for experimental results was calculated using all qPCR measurements and compared to the distribution calculated in silico for each network prediction using custom Matlab script, “JSdivergence.m” (File S8).
After identification of a putative GRN, we explored the transition between oscillatory and non-oscillatory behavior of \textit{hes}6 by searching the biochemical parameter space. Here, the negative feedback network was fixed in the simulation, and adaptive simulated annealing was used as before to continuously vary eight biochemical parameters: \( \{\delta^m_{\text{hes}6}, \delta^m_{\text{Hex}6/6}, \delta^m_{\text{Hex}6/6}, \beta^m_{\text{Hex}6/6}, \beta^m_{\text{Hex}6/6}, \beta^m_{\text{Hex}6/6}, k^-_{\text{Hex}6/6}, n\} \).

Parameter sets were searched for wild-type simulations exhibiting \textit{hes}6 oscillations with amplitude less than half that of \textit{her}1 oscillations, and where the minimum value of \textit{hes}6 mRNA in the posterior tailbud never reached zero. The range for each parameter and average parameter values from the parameter search are given in Table S2. We identified which parameters were more constrained when \textit{hes}6 oscillations were reduced using the coefficient of variation (C.V.), which is the standard deviation of each parameter value divided by the mean. For comparison between all parameter sets and those with minimized \textit{hes}6 oscillations, p-values were calculated using Student’s t-test.

\textbf{Validation:} For validation of SA, we used five different synthetic networks. Multiple synthetic GRNs were produced randomly, and those networks that resulted in oscillations were selected for use in validating simulated annealing. Four of these five exhibited continuous oscillations under wild type conditions, while one synthetic network resulted in damped oscillations. In silico qPCR data was calculated as above. Error was introduced into in silico qPCR values by adding noise chosen from a normal distribution \( N(\mu, \sigma^2) \) with \( (\mu, \sigma^2) = (0, \text{std}) \), where \textit{std} is the standard deviation from experiment for each of 126
measurements. Simulated annealing was run as for experimental networks and results for ~50 runs for each network are shown in Figure 3B. Accuracy of each prediction was calculated as the proportion of network connections (out of 36 total) that match the input synthetic network. Each GRN is plotted by Jensen-Shannon divergence versus energy score, which together sort the networks reasonably well according to accuracy, which is denoted by color (Figure 3B).

RESULTS

Design of computational model of her/hes negative feedback loop: Six her/hes genes: her1, her7, her11, her12, her15 and hes6 have been well-characterized and shown to be involved in the segmentation clock (Gajewski et al. 2003; Henry et al. 2002; Holley et al. 2000; Holley et al. 2002; Kawamura et al. 2005; Oates and Ho 2002; Shankaran et al. 2007; Sieger et al. 2006; Sieger et al. 2004). Mutant and morpholino knockdown analyses suggest that while there is a degree of functional redundancy between the her/hes genes, they also have distinct roles during segmentation. In particular, the process of segmentation fails specifically after loss of both her1 and her7 (Henry et al. 2002; Schröter et al. 2012), or both her1 and hes6 (Schröter et al. 2012; Schröter and Oates 2010; Sieger et al. 2006). We previously identified seven dimers that form in vivo and bind DNA: Her1/Her1, Her7/Hes6, Her11/Her11, Her12/Her12, Her12/Hes6, Hes6/Her15 and Her15/Her15. In addition, three dimers form in vivo but do not exhibit significant DNA binding: Her1/Hes6,
Her7/Her15 and Hes6/Hes6 (TROFKA et al. 2012). The her/hes genes have unique expression patterns with only her1 and her7 oscillating in both the posterior and anterior presomitic mesoderm (PSM) (GAJEWSKI et al. 2003; HOLLEY et al. 2000; OATES and HO 2002). Oscillations of her11 are restricted to the anterior PSM (SIEGER et al. 2004), while oscillations of her12 and her15 are only in the posterior PSM (SHANKARAN et al. 2007). hes6 expression does not oscillate, and is expressed in a gradient in the posterior PSM (KAWAMURA et al. 2005).

In order to predict a GRN for the her/hes genes, we wrote a mathematical model of negative feedback using delay differential equations adapted from (CINQUIN 2007). We included all six her/hes genes and all experimentally observed dimer species. We explicitly modeled the processes of translation, dimer binding and unbinding, and degradation (Figure 1A). We modeled transcription using a modified Hill function to account for repression by the seven different DNA-binding Her/Hes dimers. We assumed the dimers bind DNA independently and used a Hill coefficient of three, as previous models have indicated this level of cooperativity provides robust oscillations (CINQUIN 2007; TERRY et al. 2011). In order to simplify the model, we focused on oscillations within a single cell over time (Figure 1B). Activation is assumed to be constant throughout the simulation. Initial parameter estimates are in Table S2. In order to account for differences between the posterior and anterior PSM, we modeled two distinct GRNs, with the posterior network active during the first half of the time course and the anterior network active during the second half of the time course.
(Figure 1C). Though no morphological boundary marks the transition from posterior to anterior PSM, regulation of her genes is distinct in the two regions (GAJEWISKI et al. 2003; HOLLEY et al. 2000). For simplification, the transition between the two networks is discrete, with immediate cessation of transcription of posteriorly expressed genes at $t = \frac{1}{2} t_f$. Importantly, we allowed regulation of each her/hes gene to be unique.

The non-oscillatory gene hes6 responds to negative feedback: We performed perturbation experiments by knocking down the her/hes genes singly and in all pairwise combinations using morpholino injection at the 1-cell stage. We calculated the efficiency of each morpholino using a modified luciferase assay as previously described (TROFKÁ et al. 2012). We found that most morpholinos reduce protein levels by approximately 98%, with the exception of her11 morpholino (MO), which reduced luciferase reporter expression by $90.5 \pm 5.3\%$ (Figure S1). Using intronic primers, we measured the change in transcription rate for all six her/hes genes using quantitative Real-Time PCR (qPCR). In addition, we examined her1 expression pattern and somite polarity using in situ hybridization (Figure S2-S3). As expected, transcription rate of the oscillatory her genes increased for many of the knockdown conditions (File S1). Surprisingly, we also noticed a greater than two-fold increase in hes6 transcription for seven of the 21 knockdown conditions, despite hes6 being expressed in a posterior to anterior gradient in the posterior tailbud (Figure 2A-B). In particular, simultaneous knockdown of hes6 and a second her gene have
the greatest effects on *hes6* transcription. Since we did not observe significant DNA-binding by the Hes6/Hes6 homodimer in vitro, this suggests the Hes6 heterodimers may act to negatively autoregulate transcription of *hes6*. The response of *hes6* to *her/hes* knockdown suggests that transcription of *hes6* is regulated, at least to some extent, by the negative feedback loop. Therefore, we allowed for regulation of *hes6* by the Her/Hes dimers in our computational model.

We noted that the relative response of each *her/hes* gene to a given knockdown condition was more consistent across experimental replicates than the absolute response of each transcriptional target. Therefore, we normalized the qPCR data in order to focus on the relative response between *her/hes* genes (see Methods). We found that *her12* responds more strongly than the other *her/hes* genes across almost all knockdown conditions studied. In agreement with the limited response of *hes6* for most knockdown treatments, *hes6* typically responds less than the oscillatory *her* genes (Figure 3A). The few exceptions correspond to those knockdowns where *hes6* production rate is increased by more than three-fold.

**Simulated annealing can be used to fit GRN models to data from perturbation experiments:** Within the zebrafish segmentation network, there are 36 possible dimer-DNA interactions. If we allow each repressive interaction pair to be either weak or strong, this means there are $10^{17}$ possible network topologies. In order to efficiently search such a large parameter space, we utilized simulated annealing, a global optimization method, to compare in silico
models of the her/hes network with in vivo qPCR data. The experimentally measured changes in transcription rate of six her/hes genes under 21 different morphant conditions provide 126 data points with which to fit models of her/hes regulatory interactions (Figure 3A). To further constrain the system, we also specified that oscillations of her1 and her7 must persist under wild-type conditions.

Simulated annealing (SA) works by minimizing the difference between experimental qPCR data and in silico qPCR values calculated using a dynamic model of the segmentation clock. To identify the best fit, SA methodically tests many different GRN configurations. As each regulatory network is tested, SA asks whether the latest network performs better than the previous. Depending on the improvement in performance, the new network will be accepted with some probability and further small changes to the network will be made in order to evolve performance. Simulated annealing is a Monte Carlo method, so we ran 100 trials for each test, with approximately 90 networks completing successfully, and ranked network predictions two different ways, by energy score and Jensen-Shannon divergence. Energy score is a measure of the total difference between experiment and simulation across all qPCR measurements, where a network that perfectly matches the qPCR data would have an energy score of zero. Jensen-Shannon divergence values quantify how similar two probability distributions are, and range from zero (perfectly similar) to one (highly divergent). To calculate Jensen-Shannon divergence, we binned qPCR measurements so as to reduce
the complexity of the data, and created probability distributions based on the
response state of all six genes within a given knockdown.

To verify that SA is competent to solve a problem as complex as the
her/hes GRN, we validated the method using synthetic in silico networks. In
brief, we designed five synthetic negative feedback networks that have the same
number of nodes as the in vivo network. We then calculated qPCR data for these
networks in silico, and added noise to the qPCR values similar in magnitude to
the error observed experimentally. Given an initial qPCR data set, we then used
simulated annealing to solve for each synthetic network. Since the network is
known a priori, we can calculate the accuracy of each prediction, where accuracy
is defined as the proportion of network connections (out of 36 total) that match
the synthetic input GRN. Using energy score and Jensen-Shannon divergence,
we find that we can sort the networks according to accuracy, and that across the
five synthetic networks, the most accurate predictions range from 64-97% identity
(Figure 3B). The high level of accuracy we can achieve for the synthetic networks
suggests that our data set is comprehensive enough to make reasonable
predictions of the gene regulatory network, given appropriate assumptions and
parameter values.

We also found that adding constraints to the model regarding under which
conditions oscillations should be maintained greatly impacted the performance of
simulated annealing, and further constrained the possible network predictions.
Therefore, we incorporated two additional constraints for the experimental data
set. Loss of both her1 and her7, or both her1 and hes6, have been shown to
abolish tissue level oscillations in the PSM and perturb somite borders along the entire anterior-posterior axis (Figure S4 and File S2) (HENRY et al. 2002; SCHROTER et al. 2012; SCHROTER and OATES 2010; SIEGER et al. 2006). Therefore, when solving for the in vivo network, we stipulated that oscillations of her1 and her7 must be lost in the anterior and posterior PSM under these two knockdown conditions.

In order to fit a computational model of the her/hes network, we needed to account for two features of our experimental data. (1) Since we measured transcriptional response in pooled PSMs, both the peak and trough of her oscillations would be represented in the RNA sample. Therefore, we averaged the response of each her/hes gene over time within the simulation. (2) Previous studies have found that higher-order interactions can be predicted based on pairwise data (CHATTERJEE et al. 2010; GEVA-ZATORSKY et al. 2010). In order to generate the large number of perturbation conditions required for an exhaustive pairwise analysis, we used morpholino injection rather than mutants. Because some morpholinos have been shown to stabilize mRNA transcripts (GAJEWSKI et al. 2003), we used intronic primers to detect only nascent transcripts via qPCR. The use of intronic primers meant we measured transcription rate rather than mRNA level in vivo, so we also calculated the average production rate in silico (see Methods).

One network is repeatedly identified using simulated annealing: We ran several tests of the simulated annealing algorithm using a number of different
assumptions regarding regulatory logic and parameter values (see Methods). Using additive logic and the parameters in Table S2, we ran 78 simulated annealing trials. While several GRNs were identified twice, one network was identified five times, suggesting the GRN structure is highly constrained under these conditions. This network fits 55.6% of the training set qPCR values within two standard deviations, and the average difference between simulated and experimental qPCR data is 0.102 (Figure S5). Again, a set of 100 randomly generated networks not fit using simulated annealing have an average difference between simulated and qPCR values ranging from 0.13 to 0.44. As expected, given the small number of repressors, the network is sparsely connected in the anterior PSM, with the Her1 homodimer repressing her1, her7, and her11, while Her11/Her11 represses only her11 (Figure 4A). The posterior PSM is more densely connected, with her1 repressed by all five dimers; hes6 and her15 are regulated by the same four dimers, but with different strengths; and her7 is regulated by three dimer species, including Her1/Her1, Her7/Hes6 and Her12/Her12. In contrast, her12 is repressed by a single dimer species, Her1/Her1 (Figure 4A). The limited regulation on her12 may explain why the production rate of her12 changes so much more than the other her/hes genes in response to perturbation of the negative feedback loop (Figure 3A). As the concentration of Her1 changes, the regulation of her12 is strongly affected. In contrast, regulation of her1 is buffered against concentration changes in a single dimer, given its more complex regulation. We also note that hes6 is strongly
repressed by all three Hes6 heterodimers, in agreement with qPCR results suggesting negative autoregulation by Hes6-containing dimers (Figure 2B).

To further characterize the network, we used this model of negative feedback to predict novel perturbations of the zebrafish segmentation clock. We predicted the qPCR data for three triple her/hes knockdowns that were not used in the training dataset. We predicted the change in transcription rate of all six her/hes genes after morpholino injection of her1+her12+her15 MO, her1+hes6+her15 MO, and her7+her11+her15 MO. Injecting three different morpholinos required us to reduce the concentration of her1 and her15 morpholino. Therefore, we recalculated knockdown efficiency so as to properly model these morphant conditions (Figure S1). We then measured her/hes response to the triple knockdowns as before. We found that the network model predicted the results of these novel perturbations with an average difference between each simulated and experimental qPCR data point of 0.134 (Figure 4B). As we stipulated in the simulated annealing algorithm, the GRN identified here oscillates under wild-type conditions (Figure 4C) and, unlike many other networks returned by SA, fails to oscillate under her1+her7 or her1+hes6 knockdown conditions (Figure 4D). However, a number of discrepancies exist between the behavior of this GRN and the in vivo network, including specific expression patterns in different knockdown conditions and the more than 40% of qPCR data that are poorly fit. The highly accurate fits possible with the synthetic networks suggests that some of our assumptions and parameter values are not in line with
the biological system. Furthermore, hes6 oscillates in this model given our initial biochemical parameters (Figure 4C).

**Tuning production and protein degradation rates minimizes hes6 oscillations:** For the initial simulation, we set the biochemical parameters describing maximum transcription, translation, degradation, and dimerization as equal between the six her/hes genes. Given that hes6 responds to negative feedback and is clearly influenced by the Her/Hes dimers, we altered biochemical parameters, rather than the number and strength of repressors, to reduce the amplitude or damp the oscillations of hes6 observed in our simulation. Keeping the delay times, rates describing production and decay of the her genes, and regulatory network interactions constant, we again used SA to scan for parameter sets that fit the single and double knockdown qPCR data. We allowed parameters describing transcription, translation, and degradation of hes6 to vary independently, as well as the parameters describing total strength of repression ($S_{\text{weak}}$) and the relative difference in strength between weak and strong dimers ($n$). We did not add any additional constraints to the model beyond those used when predicting the GRN. We searched the output parameter sets to find those that minimized hes6 oscillations relative to oscillations of her1. Overall, 12% (11/92) of the optimized parameter sets oscillated under wild-type conditions but not after her1+her7 MO or her1+hes6 MO knockdown, and had hes6 oscillations with amplitude at least less than half that of her1. Indeed, the median amplitude of hes6 oscillations from parameter sets with minimized hes6 oscillations was
1/10th that found with the remaining parameter sets. A representative example of the output from one such parameter set is shown in Figure 5A. We ran the SA algorithm 100 times, but the parameter sets predicted from eight of these runs had high energy scores and failed to exhibit appropriate *her1* and/or *her7* oscillations under wild-type conditions. We excluded these parameter sets from analysis due to their failure to meet even the minimal criteria of wild-type oscillations.

In order to determine which parameters are most important for minimizing hes6 oscillations, we calculated the coefficient of variation of each parameter across all 92 optimized parameter sets, and compared this value to the coefficient of variation for only those eleven parameter sets that minimized hes6 oscillations. Degradation rates for Hes6 protein and Hes6 homodimers, as well as hes6 maximum transcription and Hes6 translation rates, are more constrained when hes6 oscillations are reduced (Figure 5B). In addition, the average value for hes6 maximum transcription, translation, and Hes6/Hes6 dissociation rates are significantly different between parameter sets with weak versus strong hes6 oscillations (Figure 5B, asterisks). Across parameter sets with minimized hes6 oscillations, the average maximum transcription rate for hes6 was much lower (approximately 15-fold) compared to the hes6 transcription rate averaged over all parameter sets, while average Hes6 translation rate was 3-fold higher (Table S2). Though Hes6 protein and Hes6/Hes6 dimer degradation rates are more constrained when hes6 oscillations are minimized, the average value of these parameters is not significantly different from the entire dataset. In addition, the
average Hes6/Hes6 dissociation rate differs significantly between the two groups, but the dissociation rate is less constrained across parameter sets with minimized hes6 oscillations, suggesting this term is less important for reducing hes6 oscillation amplitude. We also note that two parameters are highly constrained across all SA runs, regardless of hes6 oscillation strength. Both $S_{\text{weak}}$ and $n$ have coefficients of variation less than 0.2 across all 92 parameter sets, suggesting these values play a vital role in maintaining her gene oscillations in our model. Indeed, the eight excluded parameter sets that failed to maintain her1 or her7 oscillations had highly aberrant $S_{\text{weak}}$ and $n$ values, with $S_{\text{weak}}$ values up to 16-fold higher, and $n$ values up to 8-fold higher than the average from 92 successful runs (data not shown).

We wondered whether hes6 was unique in showing a response to negative feedback while still having graded expression. In both chick and mouse, components of the Fgf and Wnt pathways oscillate during somitogenesis. However, the identity of the oscillating components is poorly conserved between vertebrate species (KROL et al. 2011). Therefore, we analyzed the response of Fgf and Wnt target genes that show graded expression in the zebrafish. We found that the log fold change response of sef, an Fgf target, was significantly correlated with the six her/hes genes across all twenty-one knockdown treatments. In contrast, the Wnt target gene axin2 was significantly correlated only with hes6, while hes6 itself had no significant correlation with the other her genes (Figure 5C). The oscillating genes her1, her7, her11, her12 and her15
were all significantly correlated in their response to single and double her/hes knockdown (data not shown).

DISCUSSION

Despite the value in defining precise gene connectivity networks, creating a detailed GRN in vertebrates remains a difficult task, as enhancers sometimes lie far from the gene they control (KLEINJAN et al. 2006; RUF et al. 2011), can span up to 50 kilobases (kb) (WHYTE et al. 2013) and may be highly complex. Within the 3.7 kb region upstream of her7, for example, there are 11 putative Her binding sites, all of which bind a subset of the Her/Hes dimers to varying degrees (TROFKÁ et al. 2012). Here, we developed a method to determine the basic structure of the her/hes module within the zebrafish segmentation clock GRN. Rather than relying on biochemical data detailing transcription factor binding, which does not always result in changes in transcription (BIGGIN 2011), we used perturbation experiments and quantitative measurements of gene transcription rate to infer possible regulatory interactions within the her/hes network. Although we lack enough data to confidently predict the in vivo network, we were able to use our predicted GRN to explore which parameters may govern the evolutionary transition from oscillatory to non-oscillatory expression.

We used simulated annealing (SA) to explore a broad parameter space and identify a GRN designed to fit single and pairwise her/hes knockdown data. Simulated annealing has been used previously to infer GRNs from microarray
datasets and fit multiple parameter values (Chen et al. 2010; Tomshine and Kaznessis 2006). However, these studies focused only on in silico or synthetic networks where the underlying topology was already known. Here, we verified our simulated annealing method using in silico networks with known topologies, before going on to analyze an in vivo GRN. Even with the inclusion of noise, four out of five in silico networks could be predicted with greater than 75% accuracy. Further, these network predictions could be separately sorted by two metrics: energy score and Jensen-Shannon divergence, in order to find the more accurate predictions. Once certain that SA was able to solve a GRN using only qPCR data from single and pairwise perturbation experiments, we extended our method to predict the in vivo her/hes regulatory network.

Here, we describe a GRN recovered in over 6% of SA runs. This network matches training data with an average difference of 0.102 between simulated and experimental qPCR data, oscillates under wild-type conditions, fails to oscillate after knockdown of her1+her7 or her1+hes6, and predicts qPCR data from triple knockdown experiments not used in the training dataset with an average error of 0.134. We ran multiple tests using different model assumptions (see Methods) and several networks were identified that could fit the data equally well, if not slightly better. However, no network met all biological criteria, including fitting all qPCR data, the her1+her7 morpholino and her1+hes6 morpholino synergies and in vivo expression patterns. We chose to focus further analyses on the GRN that was selected five of seventy eight times.
We do not suggest that the GRN we have identified perfectly matches the in vivo regulatory network. Indeed, our model does not recapitulate all known experimental observations. For example, we do not see the her7+hes6 morpholino rescue described previously (SCHROËTER et al. 2012; TROFKA et al. 2012) using the initial parameter set. However, using the parameter set shown in Figure 5A, we observe that single knockdown of her7 results in damped oscillations in the posterior PSM, which are restored in the her7+hes6 MO knockdown (data not shown). There are a number of reasons why the GRN we have identified may miss certain aspects of the in vivo network. First, our model is greatly simplified. We assume only two levels of repression exist and use equal affinities for all Her/Hes dimer species. Second, the GRN is not fully optimized. Even when running simulations in parallel for 168 hours, we are still only able to fit less than 60% of the data points. We currently lack the computing power to substantially improve the initial fit. Third, we did not vary the biochemical parameters and network configuration simultaneously. In order to reduce the size of the optimization problem, we fixed the biochemical parameters using previous estimates (CINQUIN 2007), since few of these rates have been directly measured (AY et al. 2013). Therefore, the GRN we identified may be specific to this parameter set, and the biochemical parameters in vivo may be quite different. Indeed, the high level of constraint seen on values for S\text{weak} and n when we varied hes6 biochemical parameters suggests that our GRN may be tightly tuned to the initial parameter set. Fourth, when fitting multiple parameters at once, individual parameters may exhibit “sloppy sensitivity spectra”
(Gutenkunst et al. 2007). This means that the value of specific parameters may be poorly constrained, even when the model as a whole provides accurate predictions. Finally, a remaining caveat for our analysis is that regulation of segmentation may change over time. Defects due to loss of her1 occur early in somitogenesis, while defects in her7/− mutants present later (Chooraopikayil et al. 2012). The timing of defects in her1 and her7 expression are also uncoupled in the tortuga mutant (Dill and Amacher 2005). Our analysis was conducted between the eight and ten somite stage, and therefore may only apply to formation of the posterior trunk somites.

Though we cannot be confident regarding any one particular network connection, our GRN model does provide insight into the zebrafish segmentation clock. We saw that the inclusion of hes6 as a target of negative feedback resulted in oscillations of hes6 mRNA, contradictory to the graded expression seen in vivo. This phenomenon was not limited to the GRN described here, as all networks we recovered that provided reasonable fits of the qPCR data and fit the her1+her7 morpholino and her1+hes6 morpholino synergies also showed hes6 oscillations. In order to understand how a gene regulated by negative feedback could fail to oscillate, we fixed the negative feedback network in our model and fit biochemical parameters for hes6 using simulated annealing. Without adding any additional constraints, we then looked for parameter sets that resulted in minimized oscillations of hes6 compared to her1. Again, due to sloppy sensitivities, individual parameter values may reveal less information about the network than the identity of highly constrained parameters. We find that when
selecting for reduced \textit{hes6} oscillation amplitude, the transcription and translation rates of \textit{hes6} are highly constrained and their average value is significantly different from that of parameter sets that do not have minimized \textit{hes6} oscillations. In particular, parameter sets that give minimized \textit{hes6} oscillations have an average \textit{hes6} maximum transcription rate \~60\% that of the oscillating \textit{her} genes, and Hes6 translation rate approximately 3-fold greater than the \textit{her} genes. This is in contrast to recent work that proposed production rates for \textit{hes6} that are 9-fold higher than that of \textit{her1} or \textit{her7} (SCHRÖTER \textit{et al.} 2012). There are several possible reasons for this discrepancy. First, in Schröter \textit{et al.}, the relative rates of \textit{hes6} and \textit{her1} production were chosen to recapitulate the change in clock period in \textit{hes6}\textsuperscript{-/-} mutants, which we do not address. Second, Schröter \textit{et al.} collapse the processes of transcription and translation into a single parameter, which inhibits our ability to compare the two values. For example, though our model suggests the maximum \textit{hes6} transcription is lower than that of \textit{her1}, our average Hes6 translation rate is three times as high as that of the oscillating Her proteins. Finally, we implicitly selected against high levels of \textit{hes6}, since that tended to increase the amplitude of oscillations.

The question of why \textit{hes6} transcription does not oscillate in zebrafish has broader implications. In addition to the conserved \textit{her/nes} genes that oscillate as part of the segmentation clock, components of Wnt and Fgf signaling oscillate in the PSM in mouse and chick. However, the cadre of oscillating genes is not well conserved across species (KROL \textit{et al.} 2011). We propose that one mechanism whereby genes may transition from oscillatory to non-oscillatory expression is by
altering specific biochemical parameters, in addition to adding or deleting transcription factor binding sites. In the case of hes6, our model suggests that the discrepancy between a non-oscillatory gene that responds to negative feedback may be resolved by reducing Hes6 degradation rate, decreasing mRNA transcription rate and increasing protein translation rates (Table S2). We identify sef as another gene in zebrafish that responds to negative feedback despite graded expression (Figure S6 and File S1). In fact, the transcriptional response of sef to her/hes knockdown correlates more strongly with oscillating her genes than even hes6, suggesting that stability and production rates of mature sef mRNA, rather than its transcriptional regulation, are responsible for the in vivo graded expression pattern. Given that hes6 transcription is regulated by Fgf signaling, the Her negative feedback network could also indirectly regulate hes6 transcription via modulation of the Fgf pathway.

Could there be a benefit to regulating gene expression via an oscillatory input, if the target does not need to oscillate? Theoretical work suggests that, for proteins with a lifetime greater than the period of oscillation, an oscillatory input reduces the effects of noise on protein expression levels (TOSTEVIN et al. 2012). Thus, in zebrafish, where the clock period is considerably shorter than in mouse or chick, segmentation clock negative feedback may act on multiple target genes to produce consistent levels of expression rather than entrain them to clock oscillations. Minimizing noise may be particularly important during somitogenesis in the case of hes6, as levels of hes6 have been suggested to play a role in
regulating the changing period of oscillations as the they move from posterior to anterior (CINQUIN 2007).

Acknowledgements

We thank Thierry Emonet, Corey O'Hern, Carl Schreck, Gunter Wagner and Hongyu Zhao for helpful discussions. We thank Michael Sneddon and William Pontius for assistance writing Matlab codes, Michael Stulberg for providing sef DIG probe, and Patrick McMillen for critical comments on the manuscript. Funding provided by NIH predoctoral training grant T32 GM 007499 and National Science Foundation Graduate Research Fellowship to J.S.S. and NIH grants RO1 HD045738 and R21 HD076173 to S.A.H. This work was also supported by the facilities and staff of the Yale University Faculty of Arts and Sciences High Performance Computing Center.
REFERENCES


OATES, A. C., and R. K. Ho, 2002 Hairy/E(spl)-related (Her) genes are central components of the segmentation oscillator and display redundancy with the


Sieger, D., B. Ackermann, C. Winkler, D. Tautz and M. Gajewski, 2006 her1 and her13.2 are jointly required for somitic border specification along the entire axis of the fish embryo. Developmental Biology 293: 242-251.


Figure 1. Design of computational model for her/hes negative feedback network: (A) Schematic showing autoregulation of a generic her gene by Her dimers, including key biochemical parameters. The her genes are targets of Notch signaling, and the Notch intracellular domain (NICD) activates transcription along with the cofactor Su(H). Processes of transcription, translation, dimerization, and degradation are modeled explicitly. Maximum transcription rate is given by $\beta_m$ and repression by Her/Hes dimers is a function of the dimer concentration, $f(D)$. Simulations consider reactions within a single cell over time. (B) Example of simulation results showing oscillations of multiple mRNA species over time. For the first half of the simulation the posterior PSM network is used, and the anterior PSM network is used during the second half of the simulation. (C) Schematic showing known Her/Hes dimers and their respective expression domains. In the anterior PSM, her1, her7 and her11 are transcribed, and two DNA-binding dimers can form: Her1/Her1 and Her11/Her11. In the posterior PSM, her1, her7, her12, hes6 and her15 are expressed; six DNA-binding dimers and three non DNA-binding dimers can form. B and C use the same gene/protein color-coding.
**Figure 2. Derepression of hes6 transcription in response to her/hes knockdown:**

(A) Expression of her1 (green) and hes6 (red) in a wild type 8-10-somite-stage tailbud. Stripes of her1 are indicative of oscillating expression, while hes6 expression is graded.

(B) Response of hes6 transcription rate to all single and pairwise her/hes gene knockdowns. Knockdowns are shown along x-axis, and y-axis indicates average fold change of hes6 transcription rate relative to wild type with standard error. Highlighted in pink are the seven treatments resulting in greater than two-fold increase in hes6 transcription: her15 MO, her1+hes6 MO, her7+hes6 MO, her11+hes6 MO, her12+hes6 MO, her12+her15 MO, and hes6+her15 MO.
Figure 3. Transcription rate data and validation of simulated annealing: (A) Normalized qPCR data for all six her/hes genes. Each column shows the normalized log fold change of transcription relative to wild type across all single and pairwise knockdowns. Values are color-coded according to the scale on the right side. Red indicates greater than average response and blue squares indicate less than average response, where the average response refers to all her/hes genes within a given knockdown treatment. Note the abundance of red squares for her12 and blue squares for hes6. (B) Validation of simulated annealing using synthetic networks. Results for 50 trials each of five different in silico networks. Each network was of equal complexity to the in vivo network, and noise was added to in silico “qPCR” data. Network predictions are color-coded for accuracy, and sorted along the x-axis by Jensen-Shannon divergence and along the y-axis by energy score. Here, accuracy corresponds to the proportion of network connections that match the synthetic input GRN. Note the cluster of red and orange (high accuracy) networks near the origin. For the five different networks, the best predictions ranged from 64-97% accuracy.
Figure 4. Performance of the predicted GRN for her/ hes negative feedback: (A) BioTapestry (LONGABAUGH et al. 2005) diagram of the predicted negative feedback network showing regulatory connections between all DNA-binding dimers and their targets in either the posterior or anterior PSM. Thin lines indicate “weak” repression, and thick lines indicate “strong” repression. For clarity, activation of genes and non DNA-binding dimers are not shown. (B) Comparison of prediction and experiment for three novel triple knockdowns. Pink circles indicate prediction made using GRN model, and green bars are standard deviation for three experimental replicates. (C) Results of simulation under wild-type conditions showing oscillations of mRNA within a single cell over time. Each gene is color-coded as indicated in the key. Note the oscillations of hes6 (pink). (D) The predicted GRN also recapitulates synergy seen between her1+her7 and her1+hes6. mRNA oscillations cease or are highly damped under both knockdown conditions in the posterior and anterior PSM. The only remaining cyclic gene is her11 in the anterior PSM, though these oscillations are still damped compared to the wild-type simulation. mRNA traces are color coded as in C. Some gene traces may not be visible since they overlap completely with others.
Figure 5. Oscillations of hes6 can be minimized by fitting biochemical parameters: (A) Simulation results using one sample parameter set that results in minimized hes6 oscillations. Traces for each mRNA are color coded as indicated on right. (B) Bar graph showing the coefficient of variation for 8 biochemical parameters describing hes6 and repressor strength across all parameter sets (yellow) or only those with minimized hes6 oscillations (pink). Parameters are along x-axis. Parameters with average values significantly different between the two groups are marked with an asterisk below the x-axis (see also Table S2). (C) Bar graph showing the linear pairwise correlation of qPCR response across all 21 knockdowns between hes6 (pink), sef (green), and axin2 (blue) and the her/hes genes. Significant correlations are marked with an asterisk above the bar. Correlation was calculated using Pearson’s linear correlation on the log fold change qPCR response. All qPCR values are provided in File S1.