Biophysical Journal
Article

Biophysical Society

Crosstalk and the Dynamical Modularity of Feed-Forward Loops in Transcriptional Regulatory Networks

Michael A. Rowland,[1,2] Ahmed Abdelzaher,[3] Preetam Ghosh,[3] and Michael L. Mayo[2,*]
[1]Oak Ridge Institute for Science and Education, Oak Ridge, Tennessee; [2]Environmental Laboratory, U.S. Army Engineer Research and Development Center, Vicksburg, Mississippi; and [3]Department of Computer Science, Virginia Commonwealth University, Richmond, Virginia

ABSTRACT Network motifs, such as the feed-forward loop (FFL), introduce a range of complex behaviors to transcriptional regulatory networks, yet such properties are typically determined from their isolated study. We characterize the effects of crosstalk on FFL dynamics by modeling the cross regulation between two different FFLs and evaluate the extent to which these patterns occur in vivo. Analytical modeling suggests that crosstalk should overwhelmingly affect individual protein-expression dynamics. Counter to this expectation we find that entire FFLs are more likely than expected to resist the effects of crosstalk (≈20% for one crosstalk interaction) and remain dynamically modular. The likelihood that cross-linked FFLs are dynamically correlated increases monotonically with additional crosstalk, but is independent of the specific regulation type or connectivity of the interactions. Just one additional regulatory interaction is sufficient to drive the FFL dynamics to a statistically different state. Despite the potential for modularity between sparsely connected network motifs, *Escherichia coli* (*E. coli*) appears to favor crosstalk wherein at least one of the cross-linked FFLs remains modular. A gene ontology analysis reveals that stress response processes are significantly overrepresented in the cross-linked motifs found within *E. coli*. Although the daunting complexity of biological networks affects the dynamical properties of individual network motifs, some resist and remain modular, seemingly insulated from extrinsic perturbations—an intriguing possibility for nature to consistently and reliably provide certain network functionalities wherever the need arise.

INTRODUCTION

Despite observations that many network motifs found within transcriptional-regulatory networks reside within larger modules of higher connectivity, the notion that certain dynamical functions (such as fold-change detection, response acceleration, or noise mitigation) can be attributed to individual motifs, has gained increasing acceptance, irrespective of whether they experience crosstalk or other extraneous regulatory interactions (1–4). The idea of a structure-function relationship for molecular networks has its roots in the notion of modularity—that despite their apparent complexity, parts of a transcriptional regulatory network operate without much influence from or on expression of the surrounding genes. These dynamically independent network modules have evolved, at least in part, to partition the network into smaller sections that can then be rewired to respond to a changing environment (5). As such, transcription factors found within such modules often participate in similar cellular processes (6,7). In just one example, the Hap family of transcription factors belongs to a module regulating cellular respiration (8). Although the overall network may exhibit a degree of functional modularity, whether the smaller transcriptional network motifs are similarly modular remains a challenging open question, one in which protein-expression dynamics should play a significant role (9).

One prominent example of a network motif is the feed-forward loop (FFL; Fig. 1 *A*), a hierarchical network module in which a gene is regulated by two transcription factors, one of which regulates expression of the other (10). FFLs are found to appear in much higher (or lower) quantities than expected in the transcriptional-regulatory networks of organisms, such as the bacterium *Escherichia coli* (*E. coli*) or the budding yeast *Saccharomyces cerevisiae*, when compared against counts obtained from degree-preserved randomizations of the original regulatory networks (10). Studies of FFL dynamics often draw functional conclusions from FFLs studied

Submitted September 13, 2016, and accepted for publication February 16, 2017.

*Correspondence: michael.l.mayo@usace.army.mil

Editor: Andrew Spakowitz.
http://dx.doi.org/10.1016/j.bpj.2017.02.044

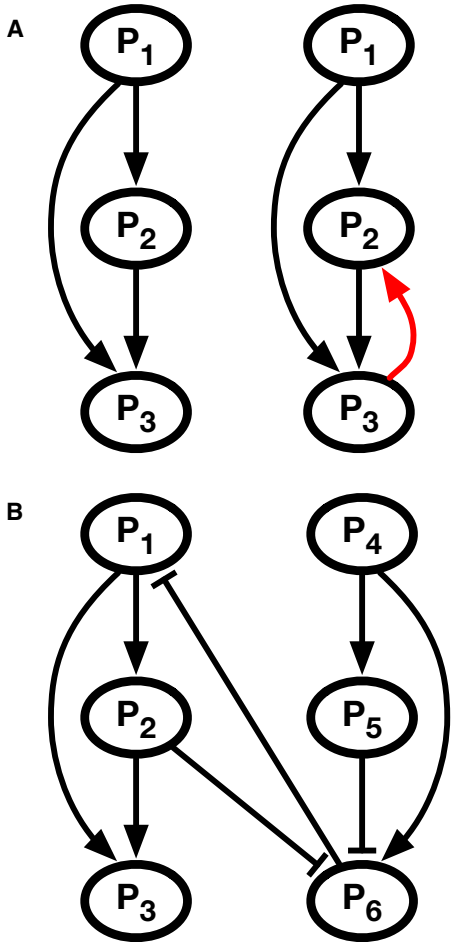

FIGURE 1 Diagrams of FFLs and inter-FFL crosstalk. (*A*) The left motif represents a canonical FFL. It has two regulators, P_1 and P_2. P_1 regulates P_2 and both regulate the target gene, P_3. The right motif represents an embedded FFL. While this motif contains the regulatory interactions from the canonical FFL, P_3 can also regulate P_2. Hence the three-interaction canonical motif is embedded within a four-interaction motif. In this work we are interested in crosstalk between canonical motifs. (*B*) An example of crosstalk between two canonical FFLs is shown. P_1-P_2-P_3 represents one FFL and P_4-P_5-P_6 represents the other. This pattern includes two crosstalk interactions: P_2 inhibiting P_6 expression and P_6 inhibiting P_1 expression. Note that this pattern also includes a feedback loop within the crosstalk among P_1, P_2, and P_6. To see this figure in color, go online.

in isolation outside of the embedded network environment (7,10–13). Although beneficial because it maintains parsimony by minimizing outside influence, conclusions drawn from such studies may provide neither a realistic nor a fair assessment of FFL functionality in vivo. Given that FFLs are generally not topologically isolated within the transcriptional regulatory network, they may be susceptible to the transient dynamics of other regulatory modules, network motifs, or expressed proteins (7,14).

Here we investigate the extent to which FFL dynamics may be coupled through crosstalk interactions to evaluate whether they can be considered dynamically and functionally modular in their native operating conditions within the *E. coli* transcriptional-regulatory network. Although FFLs could potentially be nontrivially coupled to any number of smaller regulatory network modules (e.g., bifans, diamonds), we focus here on FFL-FFL crosstalk for a number of reasons. First, very little is generally known about how any type of network motif connects to other regulatory modules within a network, the dynamical implications of those connections, or whether any biological functions can be associated with the interactions. From these many possibilities, dyadic FFLs are a parsimonious and symmetrical starting point upon which to build a knowledge base. Second, our choice of inter-FFL crosstalk drastically reduces topological complexity by limiting the configuration space of crosstalk interactions regulating each network motif. Lastly, many behaviors and functions of FFLs have been well documented, which provides clear expectations for dynamically effective parameter regimes. We leveraged an existing, well-studied dynamical FFL model that incorporates the experimental phenomenology of *cis*-regulatory logic gates positioned at the transcription-initiation sequence to deterministically predict qualitative trends in mean protein concentrations (10,15,16). This schema was adapted so that the dynamical regulation and logic-gate combination was used to model additional crosstalk interactions between individual FFLs. To study the effect of crosstalk on the FFLs' dynamics, we simulated time-series data for the regulated gene of each FFL in response to the activation of the top-level transcription factors, both with and without crosstalk. A bootstrapped statistical analysis of the difference in responses for each time point resulted in a binary decision as to whether the crosstalk significantly altered the dynamical response of either FFL. Using these models as a basis, we developed a metric termed the "modularity index", which measures the fraction of FFL pairs coupled through crosstalk that exhibit time-series profiles dissimilar to isolated FFLs. Our results demonstrate that the despite intuition suggesting that coupled FFLs should be dynamically dependent on each other, there are examples in which crosstalk has no significant impact on the dynamics of at least one of the FFLS. The likelihood of finding such dynamic insulation, however, decreases with additional crosstalk interactions.

MATERIALS AND METHODS

Dynamic modeling of FFL crosstalk

We define a canonical FFL as a three-gene pattern of regulatory interactions in which one apical transcription factor regulates expression of another transcription factor, both of which directly regulate a target gene (Fig. 1 *A*). Feed-forward loops embedded within three-node patterns that host regulatory interactions additional to the FFL substructure are not considered in this work (Fig. 1 *A*; *red interaction*). We refer to a crosstalk pattern as two FFLs coupled by at least one regulatory interaction in which a transcription factor in one FFL regulates the expression of a gene in the other. These concepts are illustrated in Fig. 1 *B*, as the distinct protein-coding genes P_1, P_2, and P_3 comprise one FFL, and those labeled

by P_4, P_5, and P_6 comprise another. Both FFLs are coupled by two regulatory interactions: P_6 represses the expression of P_1, and P_2 represses the expression of P_6.

We extend the mathematical models of FFL dynamics initially developed by Mangan and Alon (10) to include crosstalk interactions between two motifs. A dynamical change in the expression of a gene Y is modeled by reaction-limited chemical kinetics to result in the following ordinary differential equation:

$$\frac{dY}{dt} = B_y + \beta_y G\left(\left\{X_i^*\right\}, \left\{K_{X_i,Y}\right\}\right) - \alpha_y Y, \qquad (1)$$

in which B_y is the basal expression rate of Y, β_y is a weight for the expression rate due to regulation, and α_y is the loss of Y due to degradation and dilution. The logic function G is defined as:

$$G\left(\left\{X_i^*\right\}, \left\{K_{X_i,Y}\right\}\right) = \begin{cases} \prod_i f\left(X_i^*, K_{X_i,Y}\right) & \text{for an AND gate} \\ \sum_i f\left(X_i^*, K_{X_i,Y}\right) & \text{for an OR gate,} \end{cases}$$

in which $f(X_i^*, K_{Xi,Y}) = (X_i^*/K_{Xi,Y})^H/[1 + (X_i^*/K_{Xi,Y})^H]$ for an activator and $f(X_i^*, K_{Xi,Y}) = 1/[1 + (X_i^*/K_{Xi,Y})^H]$ for a repressor. Here, X_i^* is the concentration of active regulator X_i, set by multiplying the expressed concentration of X_i by stimulus S_{Xi}. The value $K_{Xi,Y}$ is a rate constant for the regulation of Y by X_i. All parameter values are scaled so they appear unitless in this model.

Following the work of Mangan and Alon (10), we set all $K_{Xi,Y} = 0.1$; S_{Xi}, β_y, and α_y to 1; and $B_y = 0$ for all dynamic simulations. For each two-FFL system, expression of the top-level regulators (P_1 and P_4 in the example given in Fig. 1 *B*) is initialized to 1 (activated), while the expression state of the remaining genes are initialized to 0 (deactivated). If the system has no regulators of the top-level regulators, then they remain fully expressed and active throughout the simulation. We chose AND gates to combine the action of multiple noncompetitive transcriptional regulators on protein production. Although the choice of regulatory logic (i.e., AND, OR) affects details of the protein production, such as the value of fixed points or the location of bifurcations, their qualitative dynamical features remain quite similar (17). We therefore do not lose much generality by choosing one AND over OR gate logic. Finally, dynamics of the targeted genes (P_3 and P_6 in the example given in Fig. 1 *B*) are recorded continuously throughout the simulation.

Parameter values were chosen because they put the time course of both FFLs within the dynamically changing domain to initially avoid a saturating signal and provide crosstalk with the best opportunity to alter FFL dynamics. This parameter set allows the regulation of the second regulator (e.g., P_2) by the primary regulator (e.g., P_1) to significantly affect the target gene (e.g., P_3) when studying FFLs in isolation (10) (see Supplemental Information). We treated all crosstalk links with identical parameters, which treats all links fairly by ensuring that regulation of expression by crosstalk is unbiased by kinetics.

Numerical methods

Our model consists of a system of ordinary differential equations which we integrate numerically using a dense linear solver with the backward differentiation formula, and Newton iteration methodology from CVODE from the SUNDIALS solver suite (http://computation.llnl.gov/projects/sundials/cvode) (18). In the simulations we start out with only the top-level transcription factors expressed, both of which are initialized to 1 (maximal expression). Cutoffs for time series were chosen by finding the longest elapsed time needed for the expression of the targeted gene, within any individually simulated FFL, to within 5% of its steady state (10 arbitrary time units). Statistics were performed using either custom scripts or with existing functions provided by R (19).

Evaluating affected FFL dynamics with bootstrapping

For each pairwise FFL crosstalk pattern, we obtained two time-series responses, one for each FFL; one simulation was first performed without the crosstalk interactions (i.e., the nominal isolated FFL response), and then another one performed including crosstalk. Our goal is to determine if the inclusion of crosstalk interactions significantly impacts the dynamics of either FFL; if this statistical challenge fails, then we term the state "affected" (see Supporting Materials and Methods). For each FFL, we compute the difference, in a logarithmic scale, between the dynamics with ($Z_{\text{crosstalk}}$) and without crosstalk ($Z_{\text{isolation}}$), on a timepoint-by-timepoint basis:

$$\Delta Z(t) = \log_{10} Z_{\text{crosstalk}}(t) - \log_{10} Z_{\text{isolation}}.$$

These differences constitute a set of cardinality 1000 (the number of time-points), from which we construct a probability distribution. We bootstrap this distribution 1000 times to obtain 95% confidence intervals for the median time-series difference. If zero was outside of the 95% confidence interval, then we conclude that crosstalk significantly affects the FFL dynamical state. Otherwise we conclude that crosstalk interactions had a statistically insignificant effect on FFL dynamics. We did not consider the obviously unaffected cases in which an FFL experienced no incident crosstalk, meaning that none of the nodes of the FFL are targets of crosstalk regulatory interactions from the other FFL (e.g., the P_4-P_5-P_6 FFL in Fig. S1 *A* experiences no incident interactions). For such cases, dynamics will remain unaffected due to a lack of external regulatory stimulus. Instead, we only evaluate FFL dynamics containing genes regulated in part by other FFLs.

This approach remains agnostic to the degree to which the states reflect dynamical effects: for example, systems in which crosstalk has a 20% change in dynamics and systems in which crosstalk has a 90% change in dynamics, are both referred to simply as "affected". This was done to avoid potential bias in setting arbitrary cutoffs that determine whether a differential response at individual time points should be considered significant. One drawback is that our algorithm allows for smaller fluctuations in the response of crosstalk-included topologies about the response for isolated topologies (in which two curves may seem qualitatively different), to be considered "unaffected", because it is insensitive to the magnitude of such fluctuations. If the median difference between two such responses is sufficiently close to 0, then there is potential for the algorithm to provide false-positive results. However, given the relative rarity of modular dynamics exhibited by cross-linked FFLs (see below), false positives are likely infrequent and therefore unlikely to significantly affect our results.

Largest connected component of the *E. coli* transcriptional regulatory network

We considered the transcriptional regulatory network from *E. coli* to investigate the abundance and dynamics of inter-FFL crosstalk patterns in a biological network. The network was rendered using GeneNetWeaver (http://gnw.sourceforge.net/)—a bioinformatics tool originally designed to assess the accuracy of network inference algorithms, which annotates transcriptional-regulatory interactions between protein-coding genes for the *E. coli* bacterium and the baker's yeast *Saccharomyces cerevisiae* (20,21). All such interactions have been experimentally verified, and interactions are labeled as stimulatory, inhibitory, dual (both interaction types), or unspecified.

The *E. coli* transcriptional-regulatory network hosts 1565 nodes and 3758 interactions forming 23 disjoint components. The largest connected component (LCC) we obtained comprises 1477 genes and 3671 interactions. We used a depth-first search to identify and separate the LCC from the whole network. For simplicity, the self, dual, and unverified interactions

are pruned from the network, which comprise 3% of all annotated interactions. The resulting LCC includes 3582 interactions and only stimulatory or inhibiting regulatory action. The form of the LCC is a set of interactions of the type $A_i\ B_i\ C_i$ where, for interaction i, node A_i regulates node B_i with sign C_i, where the sign is either 1 (stimulatory) or -1 (inhibitory). The LCC network file has been included as Tables S1–S3 in the Supplemental Information.

Enumerating FFL crosstalk patterns

We used two approaches to first identify and then count the number of pairwise FFL crosstalk patterns in the LCC of the *E. coli* transcriptional regulatory network. In the first method we employ a brute force search, requiring neither previous knowledge nor context of the *E. coli* network. This depth-first search begins from every parent node and first enumerates all canonical FFLs in the network. Next, the algorithm loops over all pairs of identified FFLs to determine if they are topologically independent (i.e., the protein-coding genes of one FFL do not also compose the other FFL) and whether there is regulatory crosstalk between them; if crosstalk is found between topologically independent FFLs, then the regulatory structure is recorded. Finally, occurrences of each unique pattern are counted, and this number defines the recorded abundance of the pattern.

The second method involves identifying suitable regulatory subnetworks before searching the actual network for their instances. We start by identifying all possible topological patterns between two FFLs coupled by up to 18 links—chosen arbitrarily considering the complexity of the task, and without regard to the regulatory action of each edge. Next, the isomorphic subnetworks at this topological level are identified and removed; the remaining topological subnetworks are searched for in the network and each instance recorded. The resulting list is pruned of topological patterns exhibiting zero abundance in the gene-regulatory network. The resulting list is far smaller than the ensemble of all possible networks with up to eight interactions. Finally, we then looped over all pairs of canonical FFLs in the LCC, and compared their coupling patterns to those in the predefined list, enumerating matches in a separate array.

Randomizing transcriptional-regulatory networks

We produced a series of degree-preserved randomized networks with a Markov chain Monte Carlo (MCMC) randomization scheme to evaluate the statistical significance of each identified inter-FFL crosstalk pattern (see previous section). Each randomization was instantiated from the *E. coli* LCC. Network interactions that are not held constant (G) are randomized with an algorithm presented as pseudocode in the Supplemental Information. During each randomization step, two different interactions within the network are randomly selected with uniform probability. Nodes at the receiving ends are exchanged, which switches the interactions. For example, if nodes A regulate B and nodes C regulate D, then after a swap, A regulates D and C regulates B. Swaps are rejected if either of these new interactions already exists within the network; otherwise, the swap is accepted. Upon rejection, interactions are returned to their original, pre-swap states. This algorithm is similar to the simpler, but widely used configuration model, except that pairwise link randomization may be carried out indefinitely (22). The configuration model, however, is known to produce degree-correlated networks (23).

It has been shown that the minimum number of steps necessary in such a MCMC randomization process for the graph to forget the original network is the following (24):

$$N = E \ln(1/\varepsilon),$$

in which E is the number of links in the network, and ε is a positive weight that sets the desired randomness. Given 3564 links in the LCC, and putting $\varepsilon = 10^{-6}$, we find that just under 50,000 MCMC steps are required to randomize the LCC, which includes steps that lead to rejection of the swap.

We developed three randomization schemes based on this MCMC process. In the first, we allow all links to be randomized, subject to the restriction that no swap duplicates existing links. In the second scheme, we first identify the set of links composing all FFLs within the network, and then remove them from the selection pool. This method maintains continuity of the constituent FFLs. In the third scheme, we first identify the set of all constituent FFL links, but only for FFLs that experience crosstalk with at least one other FFL, and before this set of links is removed from the selection pool. The goal of these latter two schemes is to control for the presence of FFLs in the system by randomizing crosstalk between them.

RESULTS

Modularity is rare among FFLs coupled by crosstalk

We developed a crosstalk library of 106,935 sets of crosstalk interactions spanning from one to eight directed regulatory interactions between two FFLs, such that each interaction originates from a gene within one and regulates a gene within the other. To build this library, we first reconstructed all unique sets with $l = 1$–4 crosstalk interactions. For the remaining $l = 5$–8 interactions, we randomly sampled 20,000 sets from the relevant configuration space. The maximum number of crosstalk interaction sets is $\binom{18}{l}$, assuming three possible crosstalk edges starting from any of the six transcription factors present in the two FFLs. We applied each of the crosstalk interaction sets to each pairing of the eight unique FFL types described by Mangan and Alon (10). Given $\binom{8}{2} = 28$ unique FFL pairs, this results in $28 \times 106{,}935 = 2{,}994{,}180$ unique inter-FFL crosstalk patterns and evaluated the effects of crosstalk on the dynamics of either FFL using the mathematical model previously described by Mangan and Alon (10) (Eq. 1; see Supplemental Information for details).

We defined a metric termed the "modularity index" to determine whether modularity is commonplace among feed-forward loops. The modularity index is the ratio of the number of individual FFLs dynamically affected by crosstalk to the total number of evaluated FFLs (i.e., the number of FFLs that participate in crosstalk that experience at least one regulatory interaction from the other FFL; see Evaluating Affected FFL Dynamics with Bootstrapping). The complement of the modularity index gives the fraction of dynamically modular network motifs.

The trend in the modularity index is to increase monotonically across individual feed-forward loops that experience at least one incident interaction, but are embedded within patterns with an increasing number of total regulatory interactions (Fig. 2 *A*, *red dots*). To explain this dependence, we modeled the modularity index as the probability, $p(s|l)$, that a feed-forward loop's dynamical state is affected by the presence of l-many incident regulatory interactions. Of

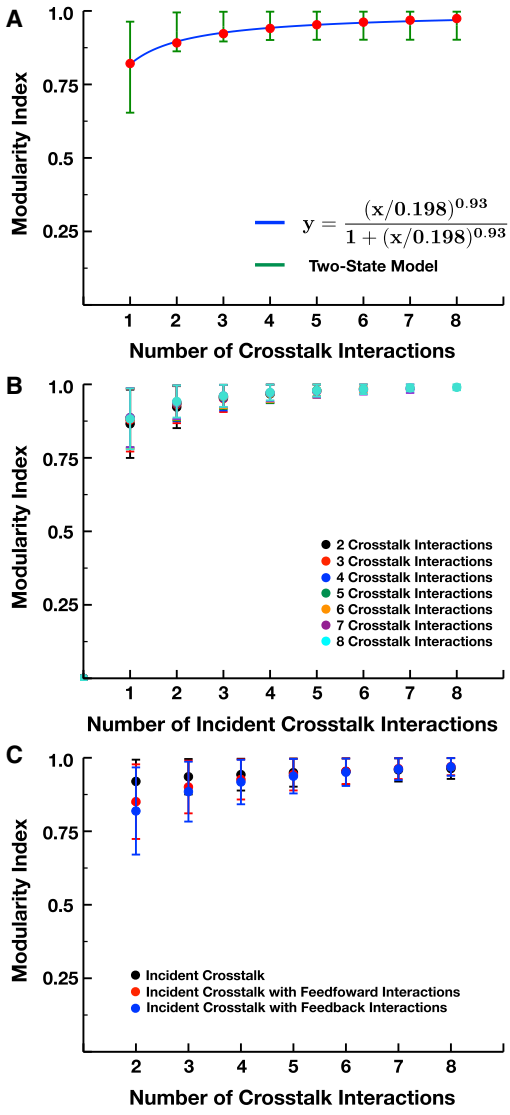

FIGURE 2 The modularity indices of inter-FFL crosstalk patterns increase with additional crosstalk. (*A*) Here, we show the modularity index as a function of the number of crosstalk interactions. The modularity index for a set of patterns is the ratio of the number of FFLs whose dynamics are significantly altered due to crosstalk to the total number of FFLs that contain at least one gene whose expression is regulated by the other FFL. The increase in the modularity indices with larger numbers of crosstalk interactions is modeled with a power-law expression (*blue line*) and the results of the two-state model (*green*). (*B*) The modularity index as a function of the number of incident interactions per FFL. Within patterns with $N = 2$–8 crosstalk interactions, we obtained the modularity index for FFLs with $M = 1$ to N incident crosstalk interactions. The error bars represent the mean $\pm$ SD in the expected value (modularity index) from the Bernoulli distribution of FFLs with significantly/not significantly impacted dynamics. Note that the expected values demonstrate very little variance between the different numbers of total crosstalk interactions in the pattern for any particular number of incident crosstalk interactions. (*C*) Here, we show the modularity index as a function of the number of crosstalk interactions. In this case we split up the FFLs into three subgroups: those that contain feed-forward interactions among the incident crosstalk (*red*), those that contain feedback interactions among the incident crosstalk (*blue*), and those that have incident crosstalk without feed-forward or feedback interactions (*black*). To see this figure in color, go online.

interest is the transition from the dynamically unaffected state, $\bar{s}$ to the affected one, s, as a function of the number of regulatory interactions incident to a feed-forward loop. For simplicity, we assume that dynamical states of patterns with l interactions depend only on the dynamical states associated with $l - 1$ crosstalk interactions (i.e., adding individual regulatory interactions to a pattern of crosstalk constitutes a Markov chain). Taken together, these considerations lead to a difference equation:

$$p(s\,|\,l) - p(s\,|\,l-1) \;=\; \omega p(\bar{s}\,|\,l-1) - \nu p(s\,|\,l-1),$$

wherein $\omega = p(s|\bar{s})$ is the probability that a feed-forward loop transitions from the unaffected to the affected state, and $\nu = p(\bar{s}|s)$ is interpreted similarly. This difference equation can be solved exactly by method of the generating function (see Supplemental Information):

$$p(s\,|\,l) \;=\; \frac{1-(1-\omega-\nu)^{l}}{1+\nu/\omega}. \qquad (2)$$

This model naturally follows a Bernoulli distribution, so the mean is simply given by Eq. 2.

Fig. 2 *A* illustrates the modularity index for FFLs coupled through crosstalk and which experience at least one incident interaction (ordinate), plotted against the total number of crosstalk interactions present in the pattern (abscissa). This ensures that only FFLs potentially affected by crosstalk remain in the sample space used to estimate the index. The best-fit parameter values of the two-state model to these data, $\omega = 0.808$ [0.753, 0.864] and $\nu = 0.0428$ [0.0215, 0.0641] (bracketed are 95% confidence intervals), were found by curve-fitting Eq. 2 to the data of Fig. 2 *A* using a maximum likelihood estimation method. In Fig. 2 *A*, we only analyzed the dynamics of FFLs that experience incident interactions. Fitting the two-state model to these data, we find that $\omega = 0.875$ [0.808, 0.943] and $\nu = 0.0207$ [0.00795, 0.0334]. Therefore, the total number of crosstalk interactions between two FFLs is predictive of their modularity, as long as a considered FFL experiences at least one incident interaction.

Despite its rather simple assumptions, the two-state model (Eq. 2) mostly captures the variability observed in the modularity index data, evidenced by the overlap of the model (Fig. 2 *A*; mean $\pm$ variance, *green bars*) with results from the crosstalk library (Fig. 2 *A*, *red circles*). While the probability that the presence of any single crosstalk interaction in a pattern alters FFL dynamics is quite high ($\approx$81%), there remains a small but substantial chance ($\approx$4.3%) that additional interactions induce modularity in patterns affected by crosstalk. It follows that we can expect to find modular FFLs in patterns with larger l, but they appear less abundantly. This prevalence of modular FFLs is unexpected, given that just a single up- or down-regulating incident interaction should be enough to alter

protein-expression dynamics for a regulated gene (see Supplemental Information).

Modular FFLs, especially those with a larger number of crosstalk interactions, likely arise from a combination of the direction of crosstalk regulation, the activity level of the regulators, and the magnitude of the impact of the crosstalk regulation on its target (e.g., a stimulatory interaction to a node that, without crosstalk, is already highly active would have little effect on the target node). Changing the parameter values used for the crosstalk interactions (i.e., strengthening the magnitude of the impact of the crosstalk regulation on the target node) may increase modularity indices, while reducing the magnitude of the crosstalk interactions should decrease modularity indices. However, if parameter values were varied uniformly across all of the tested patterns, then we should not expect the relative performance of the modularity indices of Fig. 2, and their monotonically increasing trend, to change.

In our two-state model we did not specify the type of regulatory interactions, nor did we consider their specific connectivity of the targeted motif, nor did we account for the contribution of any topological features such as feedback on the protein dynamics. This begs the question: do more complex crosstalk features contribute substantially to the modularity index? Individual model fits to the data of Fig. 2, *A* and *B* are similar enough to suggest that only the number of total crosstalk interactions within a pattern is sufficient to predict the modularity of the regulated FFLs. However, we also investigated the role of more complex crosstalk topology on the modularity of FFLs. Fig. 2 *C* demonstrates that FFLs do not appear significantly affected by the inclusion of feedback and feed-forward interactions within the crosstalk. An example of such a topology is given by Fig. 1 *B*, wherein interactions among P_1, P_2, and P_6 exhibit a feedback loop. We observe only small differences in the modularity indices between patterns with the same total number of crosstalk interactions that include FFL interactions (Fig. 2 *C*, *red*), patterns that include feedback interactions (Fig. 2 *C*, *blue*), and patterns that include neither (Fig. 2 *C*, *black*). These results reinforce the notion that only the number of crosstalk interactions between FFLs constitutes the primary driver of dynamical FFL response.

Statistical bias toward inter-FFL crosstalk patterns in *E. coli* transcriptional networks

Up to now we have characterized the effects of inter-FFL crosstalk patterns on FFL dynamics for a large randomly generated library of inter-FFL crosstalk patterns. However, this library did not account for the biological context, nor did it consider any potential consequences of inter-FFL crosstalk on dynamical activity of the greater transcriptional regulatory network. Biological networks are the result of selected evolutionary mutations, and are therefore nonrandom. Thus, changes to the network often provide some appreciable phenotypic benefit or admit no considerable negative effects on fitness (7,25,26). As such, it would be reasonable to hypothesize that if an inter-FFL crosstalk pattern is found within a transcriptional-regulatory network, then it elicits little to no effect on the organism or may otherwise improve its fitness.

To progress toward testing this hypothesis, we first searched the *E. coli* transcriptional regulatory network's LCC to catalog all instances of inter-FFL crosstalk patterns. The LCC hosts 934 FFLs, providing at most 435,711 FFL pairs. In this context we identified 221,382 potential pairings based on the FFLs within the network—$\approx$51% of the maximum, including 2593 unique patterns. The remaining FFLs were found to be uncoupled; either they lacked crosstalk interactions or they shared one or more common transcription factors. We did not further investigate these individual network motifs.

We determined whether the presence of each type of crosstalk pattern was statistically significant by comparing the number of times it appears within the LCC against its counts in three sets of 960 degree-preserved randomizations of the LCC network using the edge-switching method from Kashtan and Alon (7), Capra et al. (25), and Rowland and Deeds (26). These three sets were constructed by 1) randomizing the nodes to which each edge connects, but rejecting any randomization that leads to double edges or self-loops; 2) randomizing edges according to scheme 1, except for those edges composing an identified FFL; and 3) randomizing edges according to scheme 2, except for crosstalk edges that connect any two nodes of different FFLs. We find that $\approx$63% of the inter-FFL crosstalk patterns identified in the LCC are overrepresented, which suggests an evolutionary preference to develop and maintain these patterns (Fig. 3, *black*). Approximately 7% of the identified patterns are underrepresented in the LCC, suggesting a preference against developing and maintaining these patterns (Fig. 3, *red*). The LCC shows no bias toward the remaining 30% of the patterns; they may be considered evolutionarily neutral (Fig. 3 *B*, *blue*). Despite their conceptual differences, there is no evidence that the choice of randomization scheme has any impact on the bias (or lack thereof) toward any of the patterns (Tables S1 and S2; $\chi^2 > 2000$). This seems counterintuitive, as one would assume that the randomization schemes that maintain FFL structures (i.e., schemes 2 and 3) would exhibit a different bias toward inter-FFL crosstalk, because they naturally favor representation of FFLs in the network.

The presence of FFL crosstalk was robust across all networks created from the randomization schemes. The number of intra-FFL interactions and interactions between genes that did not participate in FFLs is lower for random networks versus the LCC. Inter-FFL crosstalk interactions are more prevalent in the random networks (Fig. S3). This trend continues for interactions between genes that

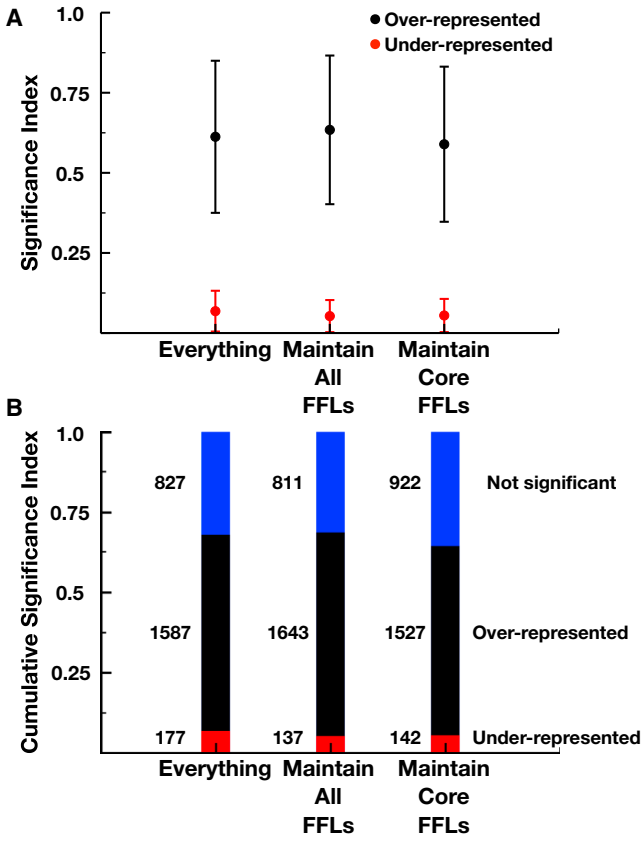

FIGURE 3 The bias toward inter-FFL crosstalk patterns in the *E. coli* LCC. (*A*) Here, we show the significance indices (the fraction of patterns that are over- (*black*) or underrepresented (*red*) in *E. coli*) as determined by comparing against degree-preserved randomizations of the LCC network. We used three randomization schemes, as follows: 1) allow all interactions to be randomized, 2) randomize all interactions except those involved in FFL motifs in the LCC, and 3) randomize all interactions except those involved in FFL motifs that participate in inter-FFL crosstalk patterns in the LCC. (*B*) The cumulative significance indices for over- (*black*) and under-represented patterns (*red*) and patterns that *E. coli* demonstrates no bias toward (*blue*), is given. Also provided are the numbers of patterns in each category. To see this figure in color, go online.

participate in FFLs, and for those that do not. These results confirm that there is a bias for the evolution of FFLs within the LCC (5,10,27). As the total number of inter- and intra-FFL interactions is higher in the LCC than the random networks, there also seems to be a general biological bias toward inter-FFL crosstalk (Fig. S3). These evolutionary preferences suggest that topology matters, and that the communication between FFLs interacting through crosstalk is potentially beneficial to the fitness of the organism.

Modularity of *E. coli* crosstalk patterns

Are FFLs found within crosstalk patterns of *E. coli*'s transcriptional regulatory network dynamically modular? To answer this question we employed the same protocol described above in our analysis of the crosstalk pattern library. For each of the 2593 unique patterns found within LCC, we determined whether the dynamics of each FFL could be considered to be significantly or insignificantly affected by crosstalk, and computed modularity indices based on the number of crosstalk interactions between FFLs of a pattern.

In Fig. 4, we show how modularity indices differ between those found within *E. coli* and those of the pattern library. For the overrepresented patterns of *E. coli*, those with fewer crosstalk interactions (1–4 edges) demonstrated significantly different modularity indices from patterns of similar size in the crosstalk library (Fig. 4 *A*; *red*, $p < 10^{-4}$, Fisher's exact test). Specifically, patterns found within *E. coli* that support only one or two crosstalk interactions are more likely to exhibit dynamically modular FFLs, while patterns with three or four interactions exhibit more nonmodular FFLs. We found a similar trend in *E. coli*'s underrepresented patterns, wherein those with one crosstalk interaction were more likely to demonstrate modularity, while FFLs experiencing three crosstalk interactions were more likely to be nonmodular (Fig. 4 *B*, *red*; $p < 10^{-4}$, Fisher's exact test). Finally, among patterns found without a bias toward abundance (neither significantly nor insignificantly represented), those with only one or two crosstalk interactions showed significantly lower modularity indices. In contrast, patterns from this same group with three or more interactions showed substantially higher modularity indices when compared to those of the pattern library (Fig. 4 *C*, *red*; $p < 10^{-4}$, Fisher's exact test). This suggests, in *E. coli*, either that well-connected FFLs are selected to minimize the dynamic modularity of the FFLs or that well-connected FFLs naturally reduce modularity. At the same time, there is an evolutionary pressure to maintain individual FFL dynamics in instances in which two FFLs have only one or two crosstalk interactions.

Overrepresented patterns exhibited a stronger preference toward more crosstalk than either underrepresented or insignificantly abundant patterns. The average number of crosstalk interactions among overrepresented patterns is ≈ 3.76 interactions (Fig. 5 *A*), while the averages for underrepresented and insignificant patterns were both significantly smaller, ≈ 2.35 interactions (Fig. 5, *B* and *C*; $p < 10^{-5}$, two-sample permutation test). This result, combined with the observation that *E. coli* patterns with three or more crosstalk interactions demonstrate modularity indices equal to or larger than expected, suggests that *E. coli* either selects against modular FFL crosstalk, or that such crosstalk does not negatively affect fitness. Put another way, *E. coli* appears to favor crosstalk with a larger number of interactions between FFLs that makes modularity a less likely evolutionary outcome.

Biological significance of inter-FFL crosstalk in *E. coli*

Transcription factors that belong to a single module tend to work toward similar cellular processes (6,7). Because

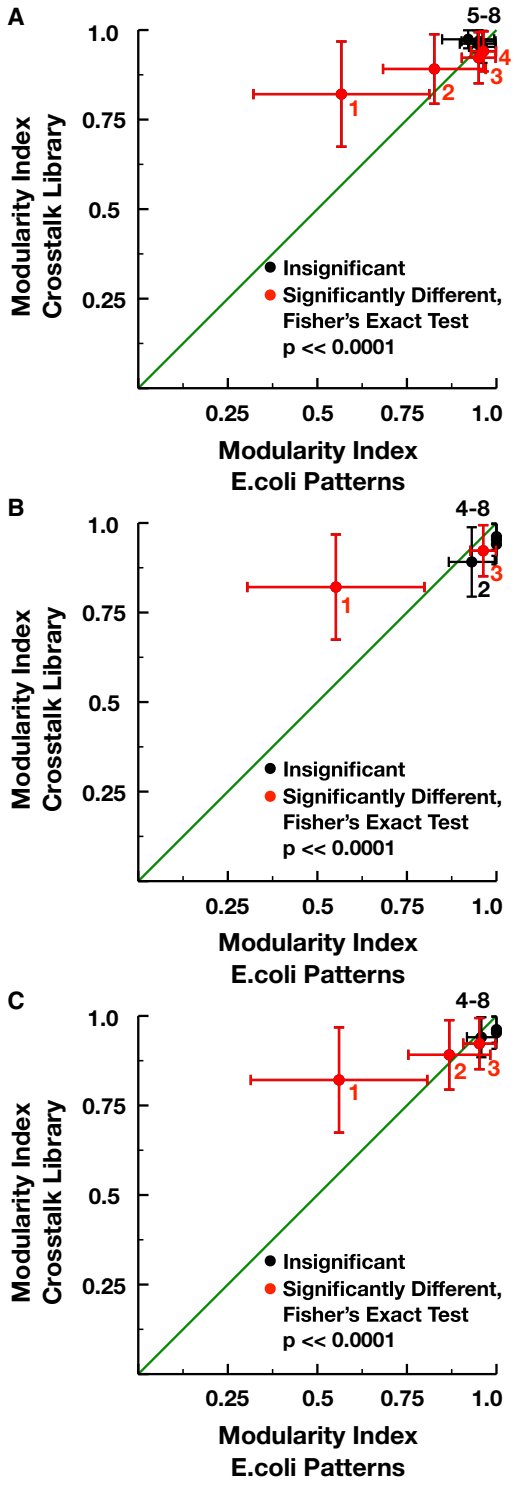

FIGURE 4 Modularity indices for cross-linked FFL "patterns" with $N \geq 1$ crosstalk interactions. Smaller values of the modularity index are associated with more dynamically modular patterns. Modularity indices are shown for patterns found within the *E. coli* LCC that are either over-represented (*A*), under-represented (*B*), or exhibit no statistically significant bias (*C*) when compared against the abundance of decoupled FFLs from the crosstalk pattern library. Modularity indices from the LCC patterns that are either significantly greater or less than the modularity indices from the crosstalk library are colored red in (*A*)–(*C*). Statistical significance is established via Fisher's exact test, $p << 10^{-4}$. Shown points represent a comparison between modularity indices for patterns with $n = 1, \ldots, 8$ crosstalk interactions found in the *E. coli* LCC and those found in the crosstalk library. The number of crosstalk interactions associated with points is labeled so as to match the significance of their points. To see this figure in color, go online.

overrepresented inter-FFL crosstalk patterns in *E. coli* tend to be dynamically coupled, we hypothesize that if the involved transcription factors are associated with similar biological processes, then they belong to a common network module.

As a test of this hypothesis, we obtained the gene ontology (GO) annotations for each of the genes that participate in the overrepresented inter-FFL crosstalk patterns. The GO annotations for each FFL included all annotations for its component genes. We can then describe the set of similar annotations for two FFLs connected by crosstalk as the intersection of GO annotations for the paired FFLs. For example, if one FFL of a pattern, labeled FFL_1, included the annotations "oxidoreductase activity" and "metal ion binding", whereas the other FFL, labeled FFL_2, included "oxidoreductase activity" and "membrane protein complex", then "oxidoreductase activity" can be attributed to the FFL_1-FFL_2 pairing. We enumerated GO annotations for all FFL pairings following this schema.

How does crosstalk impact the functions of genes of FFLs? For each pair of crosstalking FFLs, we identified a fraction of GO annotation overlap (i.e., the number of GO annotations for any of the three genes in one FFL that are also associated with any of the three genes in the other divided by the total number of GO annotations associated with all six genes). The functional similarity of FFLs apparently increases with the fraction of overlap. We found that the average fraction of overlap is $\approx 0.22 \pm 0.07$ (Fig. 6, *black*). To put this in context, we performed a similar analyses on $n = 10^7$ pairs of sets of three randomly chosen genes from the *E. coli* LCC to obtain an average overlap of 0.064 ± 0.049 (Fig. 6, *red*; $p < 2.2 \times 10^{-16}$, Student's *t*-test). Genes associated with cross-linked FFLs are therefore more functionally similar than expected from chance alone, given that the fraction of overlap is much more substantial in the *E. coli* network when compared against genes paired randomly.

To determine whether certain GO annotations were significantly overrepresented among *E. coli* crosstalk patterns, annotation counts for the overrepresented patterns of the LCC were compared for statistical significance against a distribution of counts derived from analyzing sets of randomly sampled genes of the network. More specifically, two distinct sets of three individual genes were randomly sampled from each network a total of 221,382 times, and this three-gene pairing procedure was repeated 960 times. To the randomly chosen genes of each run, we applied a GO analysis similar to the one described above for the

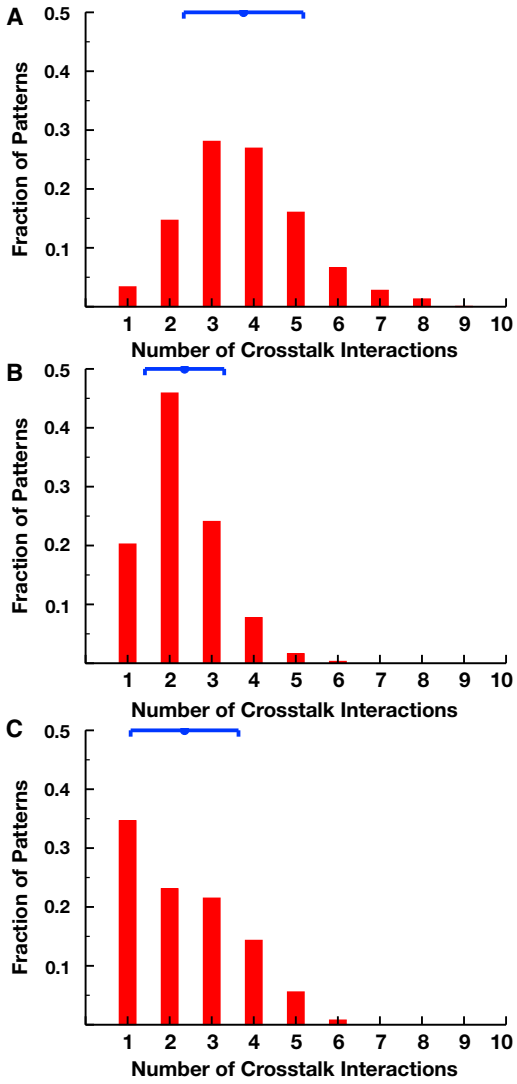

FIGURE 5 The distribution of patterns with $N = 1$–10 crosstalk interactions in the LCC. (*A*) Here we show the histogram of the fraction of overrepresented patterns with N crosstalk interactions. The blue dot and horizontal error bars represent the mean ± SD of the distribution (mean = 3.76, SD = 1.42). (*B*) Here, we show a histogram of the fraction of patterns with no significant bias with N crosstalk interactions (mean = 2.35, SD = 0.94). (*C*) Here, we show a histogram of the fraction of underrepresented patterns with N crosstalk interactions (mean = 2.35, SD = 1.28). Note that the mean number of crosstalk interactions in the overrepresented patterns is significantly greater than the mean number of crosstalk interactions in underrepresented patterns or patterns with no significant bias ($p < 10^{-5}$, two-sample permutation test). To see this figure in color, go online.

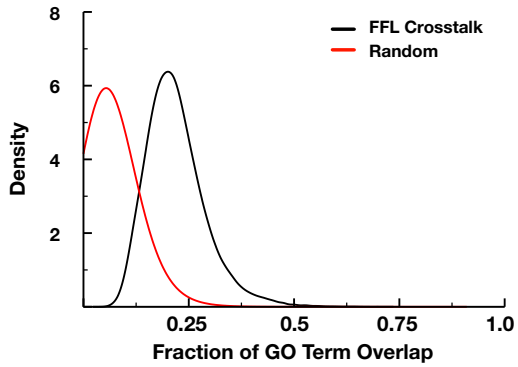

FIGURE 6 The fraction of GO annotation overlap between crosstalking FFLs. The fraction of overlap is defined as the number of unique GO annotations associated with genes in both FFLs divided by the total number of unique GO annotations associated with genes in either FFL. The crosstalking FFLs in the *E. coli* LCC demonstrate an average fraction of overlap of 0.22 ± 0.07 (*black*), whereas a sample of $n = 10^7$ pairs of sets of three randomly selected genes from the LCC demonstrate an average fraction of overlap of 0.064 ± 0.049 (*red*). To see this figure in color, go online.

FFL crosstalk patterns. We compiled annotations for each of the two gene triplets of each sample and determined their common annotations.

These analyses resulted in 57 significantly overrepresented GO annotations (Table S3). Many of these annotations are included within energy and metabolism, signal transduction, and electron transport processes. Metabolic networks adapt to many protein-level perturbations, such as genetic mutations or deletions, by self-regulating metabolic fluxes (28–31). Across many species, signaling networks exhibit varying degrees of evolvability and robustness. In bacteria, for example, two-component signaling demonstrates a surprisingly high sensitivity to crosstalk, which may reduce its signaling efficacy and therefore result in lowered overall fitness. In contrast, metazoan networks are very robust, and yet easily adapt to both subcellular and environmental changes (25,26,29,32–35). That coupled feed-forward loops are both prevalent and involved in many vital metabolic processes adds a rich layer of dynamical complexity to the static picture informed by methods such as flux-balance analysis. A more comprehensive understanding regarding the role of dynamics in functional modularity of network motifs may help to clarify the role of, and distinguish between, similar notions of adaptability and robustness intrinsic to many signaling networks.

DISCUSSION

Although FFL transcriptional motifs have been extensively studied theoretically and experimentally, conclusions about functionality reached by ignoring the influence of extraneous regulatory interactions may yield an incomplete picture of in vivo network motif behavior (7,10–13). In this work we used modeling and simulation to address this question. Specifically, we investigated how the dynamics of two FFLs experiencing crosstalk differed from that without crosstalk; the modularity of FFLs can be so tested under a wide range of conditions that include changing crosstalk connectivity or regulation type. We found that while the magnitude of the dynamical impacts from crosstalk generally depend on the functional context of the incoming signals, it does not depend on the specific type of connectivity made between the protein-coding genes, nor does it

generally depend on the specific type of regulation assigned to each interaction. Only the number of crosstalk interactions between two FFLs is sufficient to predict whether their dynamics will be affected; the probability that an FFL exposed to crosstalk remains dynamically modular decreases monotonically with only the number of such interactions.

Given the crosstalk library data (Fig. 2), we might expect modularity between coupled FFLs to be somewhat rare in the transcriptional regulatory network for *E. coli*. We enumerated over 2000 unique FFL crosstalk patterns in the largest connected component of the *E. coli* transcriptional regulatory network. We found that $\approx 63\%$ of these patterns are overrepresented, and these patterns demonstrate a bias toward higher numbers of crosstalk interactions between the FFLs. Consistent with the crosstalk library, these overrepresented patterns exhibited a higher inclination toward being nonmodular than did the underrepresented or statistically neutral patterns. Our results are consistent with an increasing realization in the field of systems biology that dynamics are required to understand whether network motifs behave in a modular manner (9). Network motifs are not often topologically isolated, due to being embedded within a tangle of regulatory interactions with the potential to experience input signals with differing amplitude and duration. There is, however, evidence that some regulatory modules may respond similarly despite exposure to multiple unique signals as a means to guard against spurious effects of crosstalk—an effect termed "kinetic insulation" (36). Our results suggest that FFL crosstalk patterns are unaffected by this or other mechanisms used by cells to maintain signaling modularity in other circumstances.

This tendency for overrepresentation of FFL crosstalk patterns could be explained from a signaling perspective. The presence of crosstalk decreases the gross cost of a perturbation—a measure of whether perturbation of a network interaction affects the cell's phenotype, because more crosstalk increases the number of perturbation targets and channels that convey the transient effects of other stimuli (37). If more interactions equate with a kind of dilution of perturbing signals, then crosstalk patterns with more interactions may be overrepresented in the biological network as a means to protect against spurious environmental fluctuations. Given that many of the genes within the *E. coli* network identify with FFLs ($\approx 36\%$ of genes in the LCC), more crosstalk between such genes could reasonably coincide with a biological preference for overrepresentation.

Bacterial genomes evolve rather quickly, either by duplication and divergence of existing coding regions or by lateral gene transfer. The wiring of the network is thus susceptible to change, by duplicating genes or through horizontal gene transfer followed by domain shuffling with subsequent finetuning of specificities with site mutations (25,26,38–41). Gene functions can diverge rapidly, allowing for the quick evolution of new interactions—or their quick removal (26,39,42). We hypothesize that fast evolution of bacterial genomes accounts for the higher degree of dynamical modularity observed for patterns of one and two crosstalk interactions between FFLs in *E. coli*. This mechanism allows for patterns that negatively impact bacterial fitness to be quickly removed from the genome while beneficial patterns are fixed in the genome. This allows for the possibility of evolving further crosstalk between the FFLs, accounting for the higher number of crosstalk interactions seen in the overrepresented patterns in *E. coli*. This is not to say, however, that other organisms, through drift and selection, might possess dissimilar transcriptional regulatory network structures; rather, we merely remark that only the rapid evolution of bacterial genomes in real time should allow bacteria to explore a configuration space more thoroughly in the same amount of real time than higher organisms with longer lifecycles. Dynamically neutral patterns, however, would be selected against to maintain network parsimony (37). Our results suggest that further investigation into the evolution of the *E. coli* transcriptional regulatory network is necessary to understand the implications of over- and underrepresented inter-FFL crosstalk patterns. Additionally, these results only apply to crosstalk between canonical FFLs. There exists a host of other three-node or larger motifs. These motifs have the potential to demonstrate a wide variety of dynamics in isolation, and provide more or less opportunity for extraneous crosstalk from the embedding network to alter their complex dynamics. Given the combinatorial complexity surrounding the number of motifs that could be generated with three or more nodes, the range of dynamical behaviors arising from intermotif crosstalk represents a great challenge ahead for systems biology.

Our results illuminate a link between network topology and dynamics relating changes in structure to changes in functionality and, possibly, to altered behavior. Understanding how individual protein-coding genes react to crosstalk, and how this response affects network motif dynamics, is a first step toward extrapolating effects on fitness to the whole cell. This is confounded by the fact that transcriptional regulatory networks are incredibly complex systems with a multitude of components and network substructures whose combined and interdependent states result in great variation of phenotypic responses to stimuli (43–45). Ultimately, overcoming these challenges is paramount for leveraging the unique properties attributed to network motifs to aid, for example, in the rational design of the ever-more complex synthetic circuits that promise to revolutionize systems biology.

SUPPORTING MATERIAL

Supporting Materials and Methods, three figures, three tables, and one data file are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(17)30294-1.

AUTHOR CONTRIBUTIONS

M.A.R., P.G., and M.L.M. designed the research. M.A.R. and A.A. performed the research. M.A.R., A.A., P.G., and M.L.M. wrote the manuscript.

ACKNOWLEDGMENTS

The authors thank Kevin Pilkiewicz for his helpful comments on the project.

Funding was provided by the U.S. Army's Environmental Quality and Installations 6.1 Basic Research program. Opinions, interpretations, conclusions, and recommendations are those of the author(s) and are not necessarily endorsed by the U.S. Army. This research was supported in part by an appointment to the Postgraduate Research Participation Program at the U.S. Engineer Research and Development Center – Environmental Laboratory administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and the Engineer Research and Development Center.

REFERENCES

1. Dobrin, R., Q. K. Beg, …, Z. N. Oltvai. 2004. Aggregation of topological motifs in the *Escherichia coli* transcriptional regulatory network. *BMC Bioinformatics.* 5:10.
2. Goentoro, L., O. Shoval, …, U. Alon. 2009. The incoherent feedforward loop can provide fold-change detection in gene regulation. *Mol. Cell.* 36:894–899.
3. Mangan, S., S. Itzkovitz, …, U. Alon. 2006. The incoherent feed-forward loop accelerates the response-time of the gal system of *Escherichia coli*. *J. Mol. Biol.* 356:1073–1081.
4. Osella, M., C. Bosia, …, M. Caselle. 2011. The role of incoherent microRNA-mediated feedforward loops in noise buffering. *PLoS Comput. Biol.* 7:e1001101.
5. Milo, R., S. Shen-Orr, …, U. Alon. 2002. Network motifs: simple building blocks of complex networks. *Science.* 298:824–827.
6. Qi, Y., and H. Ge. 2006. Modularity and dynamics of cellular networks. *PLoS Comput. Biol.* 2:e174.
7. Kashtan, N., and U. Alon. 2005. Spontaneous evolution of modularity and network motifs. *Proc. Natl. Acad. Sci. USA.* 102:13773–13778.
8. Bar-Joseph, Z., G. K. Gerber, …, D. K. Gifford. 2003. Computational discovery of gene modules and regulatory networks. *Nat. Biotechnol.* 21:1337–1342.
9. Alexander, R. P., P. M. Kim, …, M. B. Gerstein. 2009. Understanding modularity in molecular networks requires dynamics. *Sci. Signal.* 2:pe44.
10. Mangan, S., and U. Alon. 2003. Structure and function of the feed-forward loop network motif. *Proc. Natl. Acad. Sci. USA.* 100:11980–11985.
11. Alon, U. 2007. Network motifs: theory and experimental approaches. *Nat. Rev. Genet.* 8:450–461.
12. Kaplan, S., A. Bren, …, U. Alon. 2008. The incoherent feed-forward loop can generate non-monotonic input functions for genes. *Mol. Syst. Biol.* 4:203.
13. Murugan, R. 2012. Theory on the dynamics of feedforward loops in the transcription factor networks. *PLoS One.* 7:e41027.
14. Ishihara, S., K. Fujimoto, and T. Shibata. 2005. Cross talking of network motifs in gene regulation that generates temporal pulses and spatial stripes. *Genes Cells.* 10:1025–1038.
15. Buchler, N. E., U. Gerland, and T. Hwa. 2003. On schemes of combinatorial transcription logic. *Proc. Natl. Acad. Sci. USA.* 100:5136–5141.
16. Setty, Y., A. E. Mayo, …, U. Alon. 2003. Detailed map of a *cis*-regulatory input function. *Proc. Natl. Acad. Sci. USA.* 100:7702–7707.
17. Oyarzún, D. A., M. Chaves, and M. Hoff-Hoffmeyer-Zlotnik. 2012. Multistability and oscillations in genetic control of metabolism. *J. Theor. Biol.* 295:139–153.
18. Hindmarsh, A. C., P. N. Brown, …, C. S. Woodward. 2005. SUNDIALS: suite of nonlinear and differential/algebraic equation solvers. *ACM Trans. Math. Softw.* 31:393–396.
19. R Core Team. 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
20. Marbach, D., T. Schaffter, …, D. Floreano. 2009. Generating realistic in silico gene networks for performance assessment of reverse engineering methods. *J. Comput. Biol.* 16:229–239.
21. Schaffter, T., D. Marbach, and D. Floreano. 2011. GeneNetWeaver: in silico benchmark generation and performance profiling of network inference methods. *Bioinformatics.* 27:2263–2270.
22. Newman, M. 2005. Random graphs as models of networks. *In* Handbook of Graphs and Networks: From the Genome to the Internet. S. Bornholdt and H. G. Schuster, editors. Wiley, Weinheim, Germany, pp. 35–65.
23. Catanzaro, M., M. Boguñá, and R. Pastor-Satorras. 2005. Generation of uncorrelated random scale-free networks. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 71:027103.
24. Ray, J., A. Pinar, and C. Seshadhri. 2012. Are we there yet? When to stop a Markov chain while generating random graphs. *In* Algorithms and Models for the Web Graph: 9th International Workshop, WAW 2012, Halifax, NS, Canada, June 22–23, 2012, Proceedings. A. Bonato and J. Janssen, editors. Springer, New York, NY.
25. Capra, E. J., B. S. Perchuk, …, M. T. Laub. 2012. Adaptive mutations that prevent crosstalk enable the expansion of paralogous signaling protein families. *Cell.* 150:222–232.
26. Rowland, M. A., and E. J. Deeds. 2014. Crosstalk and the evolution of specificity in two-component signaling. *Proc. Natl. Acad. Sci. USA.* 111:5550–5555.
27. Shen-Orr, S. S., R. Milo, …, U. Alon. 2002. Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nat. Genet.* 31:64–68.
28. Becker, S. A., A. M. Feist, …, M. J. Herrgard. 2007. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. *Nat. Protoc.* 2:727–738.
29. Kotte, O., J. B. Zaugg, and M. Heinemann. 2010. Bacterial adaptation through distributed sensing of metabolic fluxes. *Mol. Syst. Biol.* 6:355.
30. Lewis, N. E., and A. M. Abdel-Haleem. 2013. The evolution of genome-scale models of cancer metabolism. *Front. Physiol.* 4:237.
31. Segrè, D., D. Vitkup, and G. M. Church. 2002. Analysis of optimality in natural and perturbed metabolic networks. *Proc. Natl. Acad. Sci. USA.* 99:15112–15117.
32. Jones, C. B. 2012. Robustness, Plasticity, and Evolvability in Mammals: A Thermal Niche Approach. Springer, New York, NY.
33. Kitano, H. 2004. Biological robustness. *Nat. Rev. Genet.* 5:826–837.
34. Rowland, M. A., W. Fontana, and E. J. Deeds. 2012. Crosstalk and competition in signaling networks. *Biophys. J.* 103:2389–2398.
35. Rowland, M. A., B. Harrison, and E. J. Deeds. 2015. Phosphatase specificity and pathway insulation in signaling networks. *Biophys. J.* 108:986–996.
36. Behar, M., H. G. Dohlman, and T. C. Elston. 2007. Kinetic insulation as an effective mechanism for achieving pathway specificity in intracellular signaling networks. *Proc. Natl. Acad. Sci. USA.* 104:16146–16151.
37. Leclerc, R. D. 2008. Survival of the sparsest: robust gene networks are parsimonious. *Mol. Syst. Biol.* 4:213.
38. Chen, K., and N. Rajewsky. 2007. The evolution of gene regulation by transcription factors and microRNAs. *Nat. Rev. Genet.* 8:93–103.
39. McAdams, H. H., B. Srinivasan, and A. P. Arkin. 2004. The evolution of genetic regulatory systems in bacteria. *Nat. Rev. Genet.* 5:169–178.
40. Serres, M. H., A. R. Kerr, …, M. Riley. 2009. Evolution by leaps: gene duplication in bacteria. *Biol. Direct.* 4:46.

41. Toft, C., and S. G. Andersson. 2010. Evolutionary microbial genomics: insights into bacterial host adaptation. *Nat. Rev. Genet.* 11:465–475.

42. Olson, M. V. 1999. When less is more: gene loss as an engine of evolutionary change. *Am. J. Hum. Genet.* 64:18–23.

43. Davidson, E., and M. Levin. 2005. Gene regulatory networks. *Proc. Natl. Acad. Sci. USA.* 102:4935.

44. Tsuchiya, M., K. Selvarajoo, …, A. Giuliani. 2009. Local and global responses in complex gene regulation networks. *Physica A.* 388:1738–1746.

45. Huang, S., and S. A. Kauffman. 2009. Complex gene regulatory networks—from structure to biological observables: cell fate determination. *In* Encyclopedia of Complexity and Systems Science. R. A. Meyers, editor. Springer, New York, NY, pp. 1180–1213.