BIOINFORMATICS ORIGINAL PAPER

Vol. 29 no. 5 2013, pages 630–637
doi:10.1093/bioinformatics/btt026

Systems biology

Advance Access publication January 17, 2013

A coherent feedforward loop design principle to sustain robustness of biological networks

Duc-Hau Le[1,2] and Yung-Keun Kwon[1,*]

[1]School of Electrical Engineering, University of Ulsan, 93 Daehak-ro, Nam-gu, Ulsan 680-749, Republic of Korea and [2]School of Computer Science and Engineering, Water Resources University, 175 Tay Son, Dong Da, Hanoi, Vietnam

Associate Editor: Jonathan Wren

ABSTRACT

Motivation: Many studies have investigated the relationship between structural properties and dynamic behaviors in biological networks. In particular, feedback loop (FBL) and feedforward loop (FFL) structures have received a great deal of attention. One interesting and common property of FBL and FFL structures is their coherency of coupling. However, the role of coherent FFLs in relation to network robustness is not fully known, whereas that of coherent FBLs has been well established.

Results: To establish that coherent FFLs are abundant in biological networks, we examined gene regulatory and signaling networks and found that FFLs are ubiquitous, and are in a coherently coupled form. This result was also observed in the species-based signaling networks that are integrated from KEGG database. By using a random Boolean network model, we demonstrated that these coherent FFLs can improve network robustness against update-rule perturbations. In particular, we found that coherent FFLs increase robustness because these structures induce downstream nodes to be robust against update-rule perturbations. Therefore, coherent FFLs can be considered as a design principle of human signaling networks that improve network robustness against update-rule perturbations.

Contact: kwonyk@ulsan.ac.kr

Supplementary information: Supplementary data are available at *Bioinformatics* online.

Received on October 17, 2012; revised on December 26, 2012; accepted on January 12, 2013

1 INTRODUCTION

Cellular dynamics are largely controlled by various topological motifs included in large-scale biological networks, which are represented by a huge number of interactions among biological components, such as genes, proteins, miRNAs and other small molecules (Klein *et al*., 2012). One of these topological motifs is a feedback loop (FBL) structure, and its effects on dynamics have been widely investigated. For example, it was shown that more robust networks tend to have larger numbers of positive FBLs and smaller numbers of negative FBLs (Kwon and Cho, 2008b), and the number of FBLs involving a node is positively correlated with the functional importance of the node (Kwon *et al*., 2007). It was also shown that two diseases are more likely to be comorbid if the genes associated with each disease are connected with FBLs of a relatively short length in a human signaling network (Le and Kwon, 2011a). These observations were validated as a design principle of biological networks through extensive simulations based on computational models. In addition, the role of FBLs was also investigated in some specific biological processes. For example, interlinked fast and slow positive FBLs drive reliable cell decisions (Brandman *et al*., 2005), interlinked positive and negative FBLs sustain the robustness of biological oscillation (Tsai *et al*., 2008) and FBLs play important role in modulating p53 pathway in individual cells (Harris and Levine, 2005; Lahav *et al*., 2004). Another well-known motif is the feedforward loop (FFL) structure that has also been investigated for its effects on network dynamics. For example, specific FFLs are needed for robust carbohydrate uptake in *Escherichia coli* (Kremling *et al*., 2008), as well as for adaptation to variations in the critical morphogen level where cell fate is switched (Rodrigo and Elena, 2011). In addition, a simple FFL motif composed of three genes was capable of processing external signals robustly (Macia *et al*., 2009), and the degree to which an FFL consisting of three positive transcriptional regulators was sensitive to primary level perturbation was related to the robustness (Hayot and Jayaprakash, 2005).

One interesting structural property common to FBLs and FFLs is coherent coupling. When a pair of FBLs or FFLs of the same sign is coupled, this is called a coherent coupling (see Fig. 1 for an illustrative example). Some previous studies have shown that an individual coherent structure can play an important role in the dynamic behavior of biological networks. For example, Mangan *et al*. (2003) showed that a coherent FFL serves as a sign-sensitive delay element in transcription networks, and Kalir *et al*. (2005) showed that a coherent FFL prolongs flagella expression in *Escherichia coli*. On the other hand, Kim *et al*. (2008) showed that a coupled positive FBL enhances signal amplification and bi-stable characteristics, whereas a coupled negative FBL realize enhanced homeostasis. In particular, we note a previous study that showed that coherently coupled FBLs can be a design principle in a human signaling network, as they make the network more robust (Kwon and Cho, 2008a). This result led us to question about whether the coherent FFL is also an important design principle for robustness-related dynamics, like the coherent FBLs in biological networks. Thus, in this study, we investigate the relationship between robustness-related dynamics and the coherent FFL structure through extensive simulations. A biological network such as a gene regulatory or signaling network can be represented by computational models. Among them, the Boolean network model (Glass and Kauffman, 1973; Kauffman, 1993) can be used to analyze the

*To whom correspondence should be addressed.

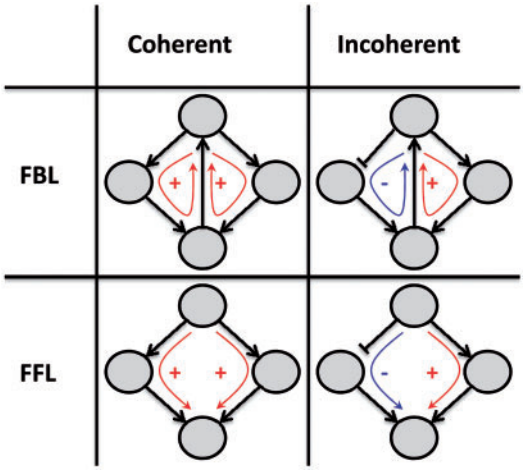

Fig. 1. An illustrative example of coherent/incoherent FBL and FFL structures. When a pair of FBLs or FFLs of the same (different) sign is coupled, this is called a coherent (resp., incoherent) coupling

complex dynamics of biological networks because it provides the most simplified representation for the network state (Bornholdt, 2005; Kauffman *et al*., 2003, 2004; Kwon and Cho, 2008b; Le and Kwon, 2011a; Shmulevich *et al*., 2005). Thus, we used the Boolean network model for the dynamics analysis.

Before examining how the dynamics are related to the coherent FFLs, we first showed that FFLs are ubiquitously found in gene regulatory and signaling networks and further, that they are more likely to form a coherent structure than randomly expected. In addition, we examined a large set of signaling networks of 1008 species integrated from the KEGG pathway database (Kanehisa and Goto, 2000) and showed that there is a larger number of coherent FFLs than randomly expected in each of 1001 species. In other words, coherent FFLs are frequently observed motifs in biological networks. To unravel the role of these coherent FFLs, we examined the robustness against initial-state or update-rule perturbations through the Boolean network model. To this end, two groups of random networks were generated and classified according to the degree of coherency of the FFLs. Through extensive simulations, we observed that the networks with relatively high coherency of FFLs are more robust against update-rule perturbations than those with relatively low coherency of FFLs, whereas there is little difference between them in terms of robustness against initial-state perturbations. In addition, we found that coherent FFLs can make downstream nodes robust against update-rule perturbations, and this induces the difference in robustness against update-rule perturbations. In summary, coherent FFLs seem to be a design principle to efficiently sustain robust behavior against update-rule perturbations in biological networks.

2 MATERIALS AND METHODS

2.1 Biological networks dataset

To investigate the distributions of coherent FFLs, we used four relatively small-scale gene regulatory networks and two large-scale signaling networks. In addition, we used the species-based signaling pathways in the KEGG database.

Gene regulatory networks: We collected four gene regulatory networks based on the literature: (i) Arabidopsis morphogenesis regulatory network (AMRN) with 10 nodes and 22 links (*Mendoza et al*., 1999); (ii) a yeast cell cycle network (YCCN) with 11 nodes and 34 links (Li *et al*., 2004); (iii) a mammalian cell cycle network (MCCN) with 10 nodes and 34 links (Faure *et al*., 2006); and (iv) *Arabidopsis* floral organ fate determination gene network (AFDN) with 15 nodes and 37 links (Espinosa-Soto *et al*., 2004). All of these networks were shown to be robust against some external or internal perturbations. For instance, AMRN was shown to robustly control the process of flower development (*Mendoza et al*., 1999). In addition, it was shown that AFDN has a robust developmental module (Espinosa-Soto *et al*., 2004), and YCCN is robustly designed against mutations (Li *et al*., 2004).

Signaling networks: Two large-scale signaling networks were collected based on the literature: (i) a human signaling network (HSN) including 1539 nodes and 4754 interactions (Cui *et al*., 2009) and (ii) a canonical signaling network (CSN) including 818 nodes and 1801 links (Kwon and Cho, 2008a). To analyze a large set of signaling networks, we collected signaling pathways from the KEGG (Kanehisa and Goto, 2000), containing pathways of 1574 species. After extracting only the activation and inhibition interactions from the pathways and integrating them for each species, we ultimately found 1008 species containing at least one FFL of length 2 (see Supplementary Table S1).

2.2 A random Boolean network model

A Boolean network is represented by a directed graph $G(V, A)$, where V is a set of Boolean variables, and A is a set of ordered pairs of the Boolean variables called directed links. Each $v_i \in V$ has a value of 1 ('on') or 0 ('off'), which represents the possible states of the corresponding elements. For example, the values 1 and 0 represent the 'turn-on' and 'turn-off' statuses, respectively. A directed link (v_i, v_j, τ) represents a positive (activating) or negative (inhibiting) relationship from v_i to v_j by setting τ to 1 or -1, respectively. The value of each variable v_i at time $t+1$ is determined by the values of k_i other variables $v_{i_1}, v_{i_2}, \ldots, v_{i_{k_i}}$ with a link to v_i at time t by the Boolean update function $f_i : \{0,1\}^{k_i} \rightarrow \{0,1\}$. All the variables are synchronously updated. Hence, the update rule can be written as $v_i(t+1) = f_i(v_{i_1}(t), v_{i_2}(t), \ldots, v_{i_{k_i}}(t))$ where we randomly select either a logical conjunction or disjunction in f_i with a uniform probability distribution. For example, if a Boolean variable v has a positive relationship with v_1, a negative relationship with v_2, and a positive relationship with v_3, then the conjunction and disjunction update rules are $v(t+1) = v_1(t) \wedge \bar{v}_2(t) \wedge v_3(t)$ and $v(t+1) = v_1(t) \vee \bar{v}_2(t) \vee v_3(t)$, respectively. In the case of a conjunction, the value of v at time $t+1$ is 1 only if the values of v_1, v_2 and v_3 at time t are 1, 0 and 1, respectively. Meanwhile, in the case of a disjunction, the value of v at time $t+1$ is 1 if at least one of the states of the clauses, $v_1(t)$, $\bar{v}_2(t)$ and $v_3(t)$ is 1. In previous studies, biological networks such as signaling and gene regulatory networks were successfully described with Boolean network models using only conjunction or disjunction update-rules (Albert, 2004; Faure *et al*., 2006; Huang and Ingber, 2000; Kwon and Cho, 2007).

Many studies have been performed to elucidate the dynamic behaviors of biological networks. These dynamic behaviors include noise propagation and signaling sensitivity (Hornung and Barkai, 2008), robustness and fragility (Kwon and Cho, 2008b) and cross-talk (Siso-Nadal *et al*., 2009). In this study, we focus on converging dynamics. In particular, we address robustness against perturbations in terms of Boolean dynamics. Given a Boolean network $G(V, A)$ with $V = \{v_1, v_2, \ldots, v_N\}$, a state of G is defined as a vector of values v_1 through v_N. A state trajectory starts from an initial state and eventually converges to either a fixed-point or limit-cycle attractor. These attractors can represent diverse biological network behaviors such as multi-stability, homeostasis and oscillation (Bhalla *et al*., 2002; Ferrell and Machleder, 1998; Pomerening *et al*., 2003). Based on the definition of the attractor, we can introduce the notion of robustness in terms of converging dynamics. If a network sustains the converged

attractor against a perturbation, it is called robust against the perturbation. This concept has been widely used (Ciliberti *et al.*, 2007; Huang *et al.*, 2005; Kitano, 2004a, b; Li *et al.*, 2004). Here, we considered two types of perturbations: an initial-state perturbation and an updating-rule perturbation. Given a Boolean network $G(V, A)$, a set of update-rules $F=\{f_1, f_2, \ldots, f_N\}$ and an initial state $s=[v_1(0), v_2(0), \ldots, v_N(0)]$ at $t=0$, an initial-state perturbation subject to a node $v_i \in V$ represents a situation in which s is mutated to $s'=[v_1(0), \ldots, 1\text{-}v_i(0), \ldots, v_N(0)]$, *i.e.* the corresponding initial value is switched to $\bar{v}_i(0)$ [the negation of $v_i(0)$]. An initial-state perturbation represents the abnormal (or malfunctioning) status of a protein or gene caused by a mutation. On the other hand, an update-rule perturbation subject to a node $v_i \in V$ involves a mutation where F is changed to $F'=\{f_1, \ldots, f_i', \ldots, f_N\}$, where f_i' is the disjunction rule if f_i is the conjunction rule and vice versa. The update-rule perturbation may represent a change of a relationship between nodes. For a set of considered initial states S, we now define the robustness of a node v_i against the initial-state perturbation and the update-rule perturbation, denoted by $\gamma_s(v_i)$ and $\gamma_r(v_i)$, respectively, as follows:

$$\gamma_s(v_i) = \frac{\sum_{s \in S} I(\alpha(s, G, F) = \alpha(s', G, F))}{|S|}$$

and

$$\gamma_r(v_i) = \frac{\sum_{s \in S} I(\alpha(s, G, F) = \alpha(s, G, F'))}{|S|},$$

where $\alpha(s, G, F)$ represents the attractor to which s will converge in a Boolean network G specified by the set of update-rules F, and $I(condition)$ denotes an indicator function that returns 1 if the *condition* is true and 0 otherwise. In other words, $\gamma_s(v_i)$ and $\gamma_r(v_i)$ represent the probability with which a network sustains the converging attractor against the initial-state and update-rule perturbations, respectively. In addition, a node v whose $\gamma_s(v)$ or $\gamma_r(v)$ is <1 is called a nonrobust node against the initial-state or the update-rule perturbation. Furthermore, the robustness of a network G against the initial-state perturbation and the update-rule perturbation, denoted by $\gamma_s(G)$ and $\gamma_r(G)$, respectively, are defined as follows:

$$\gamma_s(G) = \frac{\sum_{v \in V} \gamma_s(v)}{|V|} \text{ and } \gamma_r(G) = \frac{\sum_{v \in V} \gamma_r(v)}{|V|}.$$

In other words, the robustness of a network is computed by averaging the robustness of all nodes in the network.

To perform extensive simulations over a number of random Boolean networks, we used two random network generation models: a Barabási–Albert model (BA; Barabasi and Albert, 1999), and a variant model (ER) of an Erdős–Rényi model (Erdös and Rényi, 1959). These models have been widely used to discover design principles of biological networks (Kwon *et al.*, 2007; Kwon and Cho, 2008a, b; Le and Kwon, 2011a, b) [refer to our previous study (Le and Kwon, 2011b) for pseudo-codes to calculate $\gamma_s(v)$, $\gamma_r(v)$, $\gamma_s(G)$, $\gamma_r(G)$ and to generate random networks based on BA and ER models].

2.3 Feedforward loops

A feedforward loop is an important motif in network dynamics (Dongsan *et al.*, 2008; Yang *et al.*, 2009; Zecca and Struhl, 2010). Given a network $G(V, A)$, a sequence of nodes $w_0 \rightarrow w_1 \rightarrow \ldots \rightarrow w_{L\text{-}1} \rightarrow w_L$ is a simple path of length $L(\geq 1)$, if there exist links from $w_{i\text{-}1}$ to w_i [i.e. $(w_{i\text{-}1}, w_i) \in A$ $\forall i = 1, 2, \ldots, L$], with $w_j \neq w_k$ for $j, k \in \{0, 1, \ldots, L\}$. The sign of a simple path is determined by the parity of the number of negative links involved. If the parity number is even or zero, the path is positive, otherwise, it is negative. When there are two or more simple paths from a source node (v_i) and a target node (v_j) without any common intermediate node between those paths except v_i and v_j, the set of the simple paths is called an FFL starting from v_i and ending at v_j. Furthermore, an FFL is *coherent* if all simple paths involved have the same sign, otherwise, it is *incoherent*. In this work, we only consider FFLs of length two (i.e. the paths connecting a source and a target node have length less than or equal to two). We also define *NuFFL* of $G(V, E)$ as the number of FFLs found in the network. In other words, it represents the number of pairs of ($v \in G$, $w \in G$) such that there exists an FFL from v to w. Similarly, *NuCoFFL* and *NuInCoFFL* of $G(V, E)$ represent the number of coherent and incoherent FFLs found in the network, respectively. In addition, *NuCoFFL*(v) *and NuInCoFFL*(v) for a node $v \in V$ denote the number of coherent and incoherent FFLs involving the node v, respectively.

Let us consider an FFL consisting of M (≥ 2) different paths connecting a source (v_i) and a target node (v_j) of lengths $L_1, L_2, \ldots,$ and L_M. Assuming that a path is specified as being negative or positive independently and uniformly at random, the probability with which the FFL is coherent or incoherent is $2/2^M$ or $(2^M-2)/2^M$, respectively. Then, the *randomly expected ratio of coherent FFLs* in a network with N FFLs consisting of $M_1, M_2, \ldots, M_N$ paths can be defined as $\frac{1}{N}\sum_{i=1}^{N} \frac{2}{2^{M_i}}$.

3 RESULTS

3.1 Ubiquitousness of FFLs and their coherency in real biological networks

We first examined how frequently FFLs are found in the real biological networks. To this end, we compared *NuFFLs* of the real biological networks with those of random networks generated by shuffling the direction and the sign of every interaction from the real biological networks. More specifically, for a given real biological network represented by $G(V, A)$, a shuffled random network was obtained by replacing each interaction $(v_i, v_j, \tau) \in A$ with one of (v_i, v_j, τ), $(v_i, v_j, -\tau)$, (v_j, v_i, τ) and $(v_j, v_i, -\tau)$ uniformly at random. In this way, we generated 1000 shuffled random networks for each of the four gene-regulatory (AMRN, YCCN, MCCN and AFDN) and two signaling networks (HSN and CSN). Figure 2 shows that *NuFFLs* of HSN and AMRN are significantly larger than those of the shuffled random networks (using one-sample t-test, all P-values $< 10^{-7}$). Similar results were also observed for other signaling (CSN) and gene-regulatory (YCCN, MCCN and AFDN) networks (see Supplementary Fig. S1). This indicates that FFLs are ubiquitous in real biological networks. Second, we examined how predominantly FFLs are coherently organized. To this end, we compared the ratio of *NuCoFFL* to *NuFFL* in two real biological networks, HSN and AMRN, with that in the shuffled random networks (Fig. 3). The ratio of both HSN and AMRN is significantly greater than that of the shuffled networks (using one-sample t-test, all P-values $< 10^{-196}$). In addition, the ratio of *NuCoFFL* to *NuFFL* in the real biological networks is larger than the randomly expected ratio of coherent FFLs (see Materials and Methods for the definition of the randomly expected ratio of coherent FFLs). Similar results are observed for other signaling (CSN) and gene-regulatory networks (YCCN, MCCN and AFDN) (see Supplementary Fig. S2).

We further investigated degree of coherency of FFLs in 1008 species-based signaling networks integrated from the KEGG signaling pathway database (see Materials and Methods section). For each signaling network, we calculated the ratio of *NuCoFFL* to *NuFFL* and the randomly expected ratio of coherent FFLs (Fig. 4). Intriguingly, the ratio of *NuCoFFL* to *NuFFL* is greater

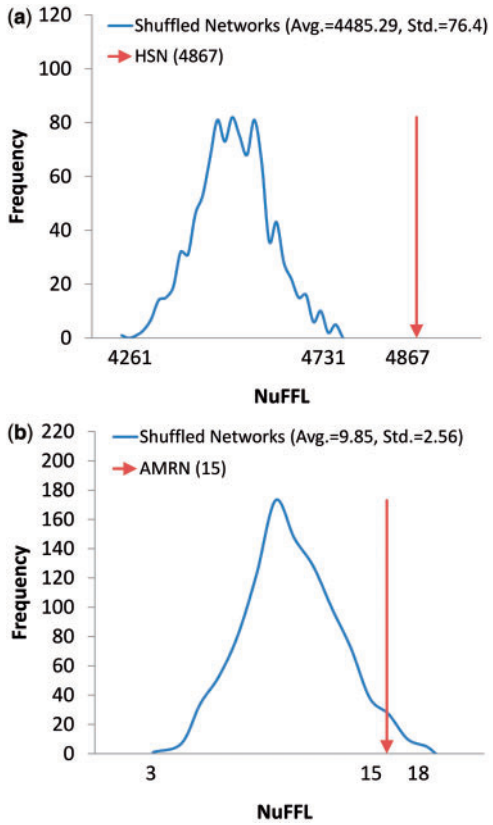

Fig. 2. Comparison of *NuFFL* between the real biological networks and the shuffled networks. **(a)** Results of HSN. *NuFFL* of HSN (4867) is significantly larger than that of the shuffled networks. (The average and standard deviation of *NuFFL* are 4485.29 and 76.4, respectively.) **(b)** Results of AMRN. *NuFFL* of AMRN (15) is significantly larger than that of the shuffled networks. (The average and standard deviation of *NuFFL* are 9.85 and 2.56, respectively.) In each subfigure, 1000 shuffled networks were generated

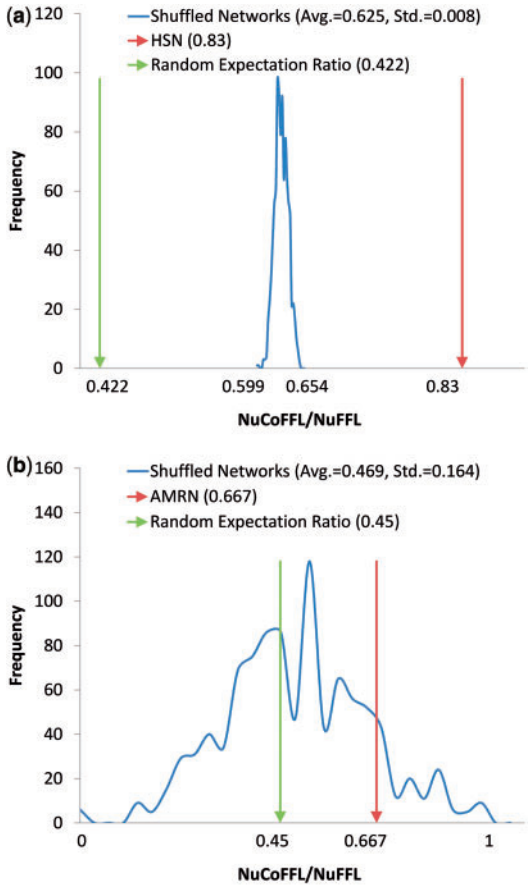

Fig. 3. The ratio of *NuCoFFL* to *NuFFL* in the real biological networks and its distribution in the shuffled networks. **(a)** Results of HSN. The ratio of HSN (0.83) is significantly larger than the randomly expected ratio (0.422) and that of the shuffled networks. (The average and standard deviation of *NuCoFFL*/*NuFFL* are 0.625 and 0.008, respectively.) **(b)** Results of AMRN. The ratio of AMRN (0.667) is significantly larger than the randomly expected ratio (0.45) and that of the shuffled networks. (Average and standard deviation of *NuCoFFL*/*NuFFL* are 0.469 and 0.164, respectively.) In each subfigure, 1000 shuffled networks were generated

than the randomly expected ratio in 1001 cases of signaling networks (see Supplementary Table S1). Similar results were observed for individual pathways. For example, there were 100 pathways of Homo sapiens having at least one FFL of length 2, and the ratio of *NuCoFFL* to *NuFFL* was greater than the randomly expected ratio in all the pathways (see Supplementary Table S2). Figure 5 shows the case of the Type II diabetes mellitus pathway in Homo sapiens, where a total of 116 different FFLs are found, and all of them are coherent. Therefore, coherent FFLs are more frequently found in real biological networks than randomly expected, and this raises the question of why coherent FFLs are so abundant in real biological networks.

3.2 Effect of an individual coherent FFL on robustness

Before investigating the effect of multiple coherent FFLs on network robustness, we examine the effect of an individual coherent FFL. Thus, we investigate the robustness against the initial-state and the update-rule perturbations in the Boolean networks consisting of only a single three-node FFL (Fig. 6). There can be eight different networks that are classified into two groups of networks containing only a single coherent (Fig. 6a) or incoherent FFL (Fig. 6b), respectively. Then, the following two lemmas with respect to $\gamma_s(G)$ and $\gamma_r(G)$, respectively, can be proved. In both lemmas, all possible initial states are considered for robustness computation, and v_S, v_M and v_T denote a source, a mediate and a target node in Figure 6, respectively. In addition, $F=\{f_S, f_M, f_T\}$ represents the set of update rules for v_S, v_M and v_T, respectively.

LEMMA 1. Given a Boolean network $G(V=\{v_S, v_M, v_T\}, A)$ consisting of a three-node FFL as shown in Figure 6, $\gamma_s(G)$ is 2/3 for both coherent and incoherent FFL cases. In particular, v_S is the unique nonrobust node.

(PROOF) Let us consider an arbitrary initial state $s=\{v_S(0), v_M(0), v_T(0)\}\in\{0,1\}^3$. For convenience, $v(s, t)$ represents the value of a node v at time t when the network is initialized with a state s.

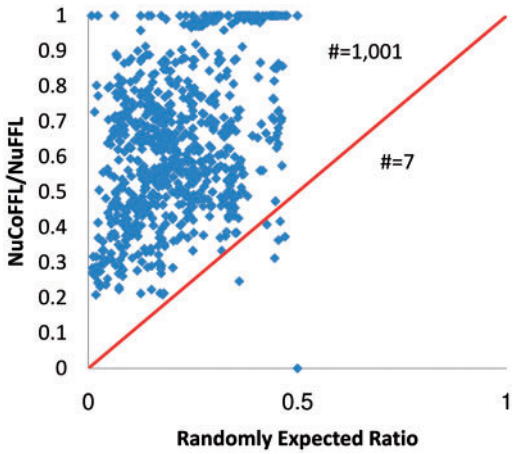

Fig. 4. Comparison of the ratio of *NuCoFFL* to *NuFFL* observed in the signaling networks integrated from KEGG and the randomly expected ratio of coherent FFLs. A point represents a result of a signaling network of one species, and the values of the *y*-axis and *x*-axis represent the ratio of *NuCoFFL* to *NuFFL* observed in the signaling networks and the randomly expected ratio of coherent FFLs, respectively. The former is larger than the latter in 1001 cases out of 1008 species

(i) When v_S is perturbed, the perturbed initial state s' is set to $\{1-v_S(0), v_M(0), v_T(0)\}$. Then, $v_S(s, t)\neq v_S(s', t)$ for $t\geq 0$ and thus $\gamma_s(v_S)=0$.
(ii) When v_M is perturbed, s' is set to $\{v_S(0), 1-v_M(0), v_T(0)\}$. Then, $v_S(s, t)=v_S(s', t)$ for $t\geq 0$ and thus $v_M(s', 1)=f_M(v_S(s', 0))=f_M(v_S(s, 0))=v_M(s, 1)$. That is, $v_M(s, t)=v_M(s', t)$ for $t\geq 1$. Similarly, $v_T(s', 2)=f_T(v_S(s', 1), v_M(s', 1))=f_T(v_S(s, 1), v_M(s, 1))=v_T(s, 2)$. That is, $v_T(s, t)=v_T(s', t)$ for $t\geq 2$. Therefore, $\gamma_s(v_M)=1$.
(iii) When v_T is perturbed, s' is set to $\{v_S(0), v_M(0), 1-v_T(0)\}$. It is trivial that $v_S(s, t)=v_S(s', t)$ and $v_M(s, t)=v_M(s', t)$ for $t\geq 0$. Then, $v_T(s', 1)=f_T(v_S(s', 0), v_M(s', 0))=f_T(v_S(s, 0), v_M(s, 0))=v_T(s, 1)$. Thus, $v_T(s, t)=v_T(s', t)$ for $t\geq 1$ and then $\gamma_s(v_T)=1$.

Note that (*i*), (*ii*), and (*iii*) hold for all types of networks shown in Figure 6. Therefore, $\gamma_s(G)$ is 2/3 and, in particular, v_S is the unique nonrobust node.

LEMMA 2. Given a Boolean network $G(V=\{v_S, v_M, v_T\}, A)$ consisting of a three-node FFL as shown in Figure 6, $\gamma_r(G)$ is 1 for coherent FFL cases whereas it is 2/3 for incoherent FFL cases. In particular, T is the unique nonrobust node for incoherent FFL cases.

(PROOF) For convenience, $v(F, t)$ represents the value of node v at time t when the network is given with the set of update-rules F.

(i) When v_S or v_M is perturbed, the perturbed update rule F' is set to $\{f_S', f_M, f_T\}$ or $\{f_S, f_M', f_T\}$, respectively. We note that $f_S=f_S'$ and $f_M=f_M'$ since f_S and f_M are constant and identity functions, respectively. Therefore, $\gamma_r(v_S)=1$ and $\gamma_r(v_M)=1$.
(ii) When v_T is perturbed, the perturbed update rule F' is set to $\{f_S, f_M, f_T'\}$. Then, $v_T(F', t)=f_T'(v_S(f_S, t-1), v_M(f_M, t-1))=A\odot B$ where A is $v_S(f_S, t-1)$ or $1-v_S(f_S, t-1)$, $\odot$ is the conjunction or disjunction function, and B is $v_M(f_M, t-1)$ or $1-v_M(f_M, t-1)$, respectively, for all $t\geq 1$. Then, it is true that $A=B$ for coherent FFL cases but $A\neq B$ for incoherent cases. In addition, we note that the outputs of the conjunction and the disjunction functions are same when the two inputs (A and B) are equivalent whereas they are different otherwise. Therefore, $\gamma_r(v_T)=1$ for coherent FFL case whereas $\gamma_r(v_T)=0$ for incoherent FFL case.

From (*i*) and (*ii*), $\gamma_r(G)$ is 1 for coherent FFL cases whereas 2/3 for incoherent FFL cases. In particular, v_T is the unique nonrobust node for incoherent FFL cases.

These lemmas say that the robustness against the initial-state perturbation is 2/3 in all networks. In other words, there is no difference between the networks composed of a single coherent or incoherent FFL with respect to $\gamma_s(G)$. On the other hand, the robustness against the update-rule perturbation was 1 for the networks of a single coherent FFL, whereas it was 2/3 for the networks of a single incoherent FFL. Therefore, coherency of FFLs improved robustness against update-rule perturbations, but not against the initial-state perturbations. Moreover, it was found that the source node denoted by S in Figure 6 was always a unique nonrobust node subject to initial-state perturbation in both the single coherent and incoherent FFL networks [i.e. $\gamma_s(v_S)=0$ and $\gamma_s(v_M)=\gamma_s(v_T)=1$ in all networks]. On the other hand, the target node denoted by T was a nonrobust node subject to update-rule perturbation in only the networks composed of a single incoherent FFL [i.e. $\gamma_r(v_S)=\gamma_r(v_M)=1$ and $\gamma_r(v_T)=0$], whereas there was no nonrobust node in the networks of coherent FFLs [i.e. $\gamma_r(v_S)=\gamma_r(v_M)=\gamma_r(v_T)=1$]. This interesting result can be similarly observed in the case of four-node FFLs (see Supplementary Fig. S3). More specifically, a network G consisting of only a single four-node FFL shows $\gamma_s(G)=3/4$, since the source node (S) is always not robust. On the other hand, for the target node (T), $\gamma_r(v_T)=1$ for every coherent FFL network (Supplementary Fig. S3a), but $\gamma_r(v_T)=0$ for every incoherent FFL network (Supplementary Fig. S3b). From this result, we can infer that the network robustness against the perturbation subject to the upstream area is not controlled by the coherency of FFLs. However, the coherent FFLs can better sustain network robustness against the update-rule perturbation subject to the downstream area.

3.3 Effect of multiple coherent FFLs on robustness in random Boolean networks

Inspired by the observation of the effect of a simple motif composed of a single coherent FFL on network robustness, we hypothesized that multiple coherent FFLs in a network play an important role in sustaining robustness, in particular against the update rule perturbation. To confirm the hypothesis, we performed extensive simulations over many random Boolean networks. To this end, we generated two sets of 10 000 random Boolean networks with 30 nodes and 75 links (i.e. $|V|=30$ and $|A|=75$) using the ER and BA models, respectively (see Materials and Methods), and $|S|$ is set to 100. Then, they were classified into two groups of networks according to the difference between the ratio of *NuCoFFL* to *NuFFL* (α) and the randomly expected ratio of coherent FFLs (β), '$\alpha>\beta$' and

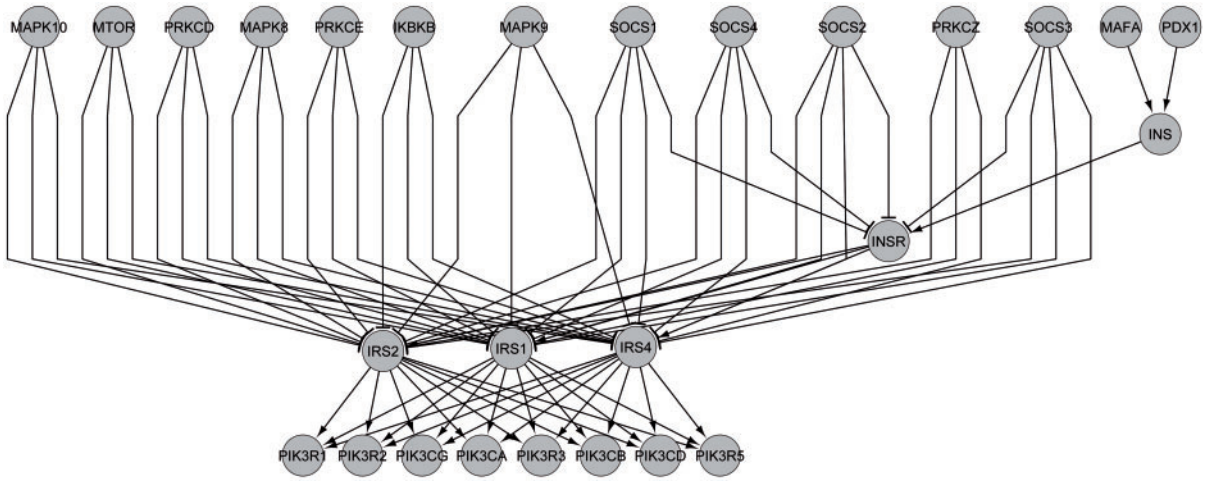

Fig. 5. Type II diabetes mellitus pathway in KEGG. In this pathway, there are a total of 116 FFLs of length 2, and every one is a coherent FFL. Arrow and bar links denote the activation and the inhibition interactions, respectively

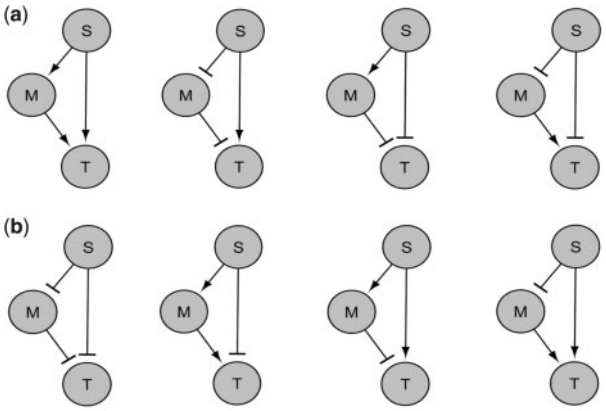

Fig. 6. Robustness of a Boolean network consisting of only a single three-node FFL. **(a)** Four different networks consisting of a single coherent FFL. **(b)** Four different networks consisting of a single incoherent FFL. In the figure, *S*, *M* and *T* represent a source, a mediate and a target node, respectively. With respect to the initial-state perturbation, $\gamma_s(G)$ is 2/3 for every network G in both (a) and (b). On the other hand, with respect to the update-rule perturbation, $\gamma_r(G)$ is 1 for every network G in (a), whereas $\gamma_r(G)$ is 2/3 for every network G in (b)

'$\alpha < \beta$' groups. We compared the average $\gamma_s(G)$ and $\gamma_r(G)$ between those two groups (Table 1). As shown in the table, the robustness against the update-rule perturbation of '$\alpha > \beta$' is significantly larger than that of '$\alpha < \beta$' (using two-sample *t*-test, *P*-value < 0.01), whereas the difference between the two groups with respect to robustness against initial-state perturbations is not obvious (using two-sample *t*-test, *P*-value > 0.1). Similar results were observed for the random Boolean networks with different size and density (see Supplementary Table S3). Therefore, more abundant coherent FFLs than randomly expected help a network to sustain robustness against update-rule perturbations, whereas they do not affect the robustness against the initial-state perturbations. This result is consistent with the observation in Section 3.2 about the simple network consisting of a single FFL.

In addition, we proved that a single coherent FFL better sustains robustness against update-rule perturbations subject to a target node than a single incoherent FFL, and this causes a robustness difference between the single coherent and incoherent FFLs. Therefore, we additionally investigated whether the robustness difference between '$\alpha > \beta$' and '$\alpha < \beta$' network groups, as shown in Table 1, can be explained with the same reasoning. To this end, we classified every source or target node (v) included in the 10 000 random Boolean networks into '*Group1*' if $NuCoFFL(v) > NuInCoFFL(v)$ and into '*Group2*' if $NuCoFFL(v) < NuInCoFFL(v)$. Next, we compare average $\gamma_s(v)$ and $\gamma_r(v)$ between the two groups. As shown in Table 2, there was no difference between the two node groups of the source nodes with respect to average $\gamma_s(v)$ and $\gamma_r(v)$. For the target nodes, whereas there was no difference between the two node groups with respect to average $\gamma_s(v)$, the average $\gamma_r(v)$ of *Group1* is significantly greater than that of *Group2* (using two-sample *t*-test, all *P*-values $< 10^{-48}$). Similar results were observed for the random networks with different size and density (See Supplementary Table S4). This result indicates that the coherent FFLs make a network more robust against the update-rule perturbation subject to target nodes.

4 CONCLUSIONS

In a recent study, Kwon and Cho (2008a) showed that FBLs have a tendency to be coherently coupled in human signaling networks, and these coherently coupled FBLs seem to be a design principle, as they enhance network robustness against initial-state perturbation. In this study, we extended this previous finding to the coherent FFL structure. We first observed that FFL is a ubiquitously found motif in human signaling networks, and it is likely to be coherently organized. Therefore, we investigated whether such a coherent FFL is correlated with robustness-related behavior of a network through simulations with random Boolean network models. We found that a single coherent FFL makes a network more robust against update-rule perturbation than a single incoherent FFL, whereas there is little

Table 1. Effect of the multiple coherent FFLs on robustness in random Boolean networks

Model	Network group	Number of networks	Average $\gamma_s(G)$	*P*-value	Average $\gamma_r(G)$	*P*-value
ER	$\alpha>\beta$	4804	0.91282	3.06×10^{-1}	0.43288	2.12×10^{-5}
	$\alpha<\beta$	4985	0.91389		0.42658	
BA	$\alpha> \beta$	4881	0.92394	6.69×10^{-1}	0.52034	5.63×10^{-3}
	$\alpha< \beta$	5085	0.92436		0.51613	

A total of 10 000 random Boolean networks with $|V| = 30$ and $|A| = 75$ are generated by using ER and BA models individually, and $|S|$ is set to 100. Here, '$\alpha > \beta$' or '$\alpha < \beta$' represents that the set of random Boolean networks where the ratio of *NuCoFFL* to *NuFFL* observed in a network is larger or smaller than the randomly expected ratio of coherent FFLs, respectively.

Table 2. Robustness comparisons between source/target node groups classified by the difference between the numbers of coherent and incoherent FFLs involved by the node

Model of nodes	Location	Node group	Number of nodes	Average $\gamma_s(v)$	*P*-value	Average $\gamma_r(v)$	*P*-value
ER	Source	*Group1*	4102	0	NA	1	NA
		Group2	4523	0		1	
	Target	*Group1*	4049	1	NA	0.25575	9.42×10^{-50}
		Group2	4419	1		0.14386	
BA	Source	*Group1*	6195	0	NA	1	NA
		Group2	7200	0		1	
	Target	*Group1*	6344	1	NA	0.2953	3.44×10^{-49}
		Group2	7170	1		0.19585	

A total of 10 000 random Boolean networks with $|V| = 30$ and $|A| = 75$ are generated by using ER and BA models individually, and all the source and target nodes are extracted. $|S|$ is set to 100. '*Group1*' and '*Group2*' represent that the set of nodes where a node involves a larger or smaller number of coherent FFLs than that of incoherent FFLs, respectively. NA, Not Available.

difference between these two types of FFLs with respect to initial-state perturbation. Interestingly, this difference is caused by difference in the robustness against update-rule perturbations that are subject to the target node of the FFLs. This phenomenon and reasoning held consistently for networks containing multiple FFLs. In summary, coherent FFLs can be a design principle of the human signaling networks in that they improve network robustness against update-rule perturbations.

The findings in this article can help deepen understanding with regard to the dynamical role of coherent FFLs. For example, the structure of multiple coherent FFLs can be considered as a kind of redundancy engineering for biological robustness as observed in alternative regular pathways that aid competition and survival in changing environment (Pasek *et al*., 2006; Zhu *et al*., 2012). In addition, the role of coherent FFLs found in this article is differentiated from those of incoherent FFLs exposed in the previous studies such as the fold-change detection in gene regulation (Goentoro *et al*., 2009), the stable biochemical adaption (Ma *et al*., 2009), the precise gene expression programming (Osella *et al*., 2011). Therefore, many other dynamics comparison between coherent and incoherent FFLs will be a good future study.

Funding: This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 102.03-2012.16. This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2010-0010533 and 2012R1A1A2040357).

Conflict of Interest: none declared.

REFERENCES

Albert,R. (2004) Boolean modeling of genetic regulatory networks. *Lect. Notes Phys*., **650**, 459–481.

Barabasi,A.L. and Albert,R. (1999) Emergence of scaling in random networks. *Science*, **286**, 509–512.

Bhalla,U.S. *et al*. (2002) MAP kinase phosphatase as a locus of flexibility in a mitogen-activated protein kinase signaling network. *Science*, **297**, 1018–1023.

Bornholdt,S. (2005) Systems biology: Less is more in modeling large genetic networks. *Science*, **310**, 449–451.

Brandman,O. *et al*. (2005) Interlinked fast and slow positive feedback loops drive reliable cell decisions. *Science*, **310**, 496–498.

Ciliberti,S. *et al*. (2007) Robustness can evolve gradually in complex regulatory gene networks with varying topology. *PLoS Comput. Biol*., **3**, e15.

Cui,Q. *et al*. (2009) Protein evolution on a human signaling network. *BMC Syst. Biol*., **3**, 21.

Dongsan,K. *et al*. (2008) The biphasic behavior of incoherent feed-forward loops in biomolecular regulatory networks. *BioEssays*, **30**, 1204–1211.

Erdös,P. and Rényi,A. (1959) On random graphs, I. *Publicationes Mathematicae (Debrecen)*, **6**, 290–297.

Espinosa-Soto,C. *et al.* (2004) A gene regulatory network model for cell-fate determination during arabidopsis thaliana flower development that is robust and recovers experimental gene expression profiles. *Plant Cell*, **16**, 2923–2939.
Faure,A. *et al.* (2006) Dynamical analysis of a generic boolean model for the control of the mammalian cell cycle. *Bioinformatics*, **22**, e124–e131.
Ferrell,J.E., Jr and Machleder,E.M. (1998) The biochemical basis of an all-or-none cell fate switch in xenopus oocytes. *Science*, **280**, 895–898.
Glass,L. and Kauffman,S.A. (1973) The logical analysis of continuous, non-linear biochemical control networks. *J. Theor. Biol.*, **39**, 103–129.
Goentoro,L. *et al.* (2009) The incoherent feedforward loop can provide fold-change detection in gene regulation. *Mol. Cell*, **36**, 894–899.
Harris,S.L. and Levine,A.J. (2005) The p53 pathway: positive and negative feedback loops. *Oncogene*, **24**, 2899–2908.
Hayot,F. and Jayaprakash,C. (2005) A feedforward loop motif in transcriptional regulation: induction and repression. *J. Theor. Biol.*, **234**, 133–143.
Hornung,G. and Barkai,N. (2008) Noise propagation and signaling sensitivity in biological networks: A role for positive feedback. *PLoS Comput. Biol.*, **4**, e8.
Huang,S. *et al.* (2005) Cell fates as high-dimensional attractor states of a complex gene regulatory network. *Phys. Rev. Lett.*, **94**, 128701.
Huang,S. and Ingber,D.E. (2000) Shape-dependent control of cell growth, differentiation, and apoptosis: Switching between attractors in cell regulatory networks. *Exp. Cell Res.*, **261**, 91–103.
Kalir,S. *et al.* (2005) A coherent feed-forward loop with a SUM input function prolongs flagella expression in *Escherichia coli*. *Mol. Syst. Biol.*, **1**, 2005.0006.
Kanehisa,M. and Goto,S. (2000) KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.*, **28**, 27–30.
Kauffman,S. (1993) *The Origins of Order: Self-organization and Selection in Evolution*. Oxford University Press, USA.
Kauffman,S. *et al.* (2003) Random boolean network models and the yeast transcriptional network. *Proc. Natl. Acad. Sci. US A*, **100**, 14796–14799.
Kauffman,S. *et al.* (2004) Genetic networks with canalyzing Boolean rules are always stable. *Proc. Natl. Acad. Sci. US A*, **101**, 17102–17107.
Kim,J.R. *et al.* (2008) Coupled feedback loops form dynamic motifs of cellular networks. *Biophys. J.*, **94**, 359–365.
Kitano,H. (2004a) Biological robustness. *Nat. Rev. Genet.*, **5**, 826–837.
Kitano,H. (2004b) Cancer as a robust system: Implications for anticancer therapy. *Natl. Rev. Cancer*, **4**, 227–235.
Klein,C. *et al.* (2012) Structural and dynamical analysis of biological networks. *Brief. Funct. Genomics*, **11**, 420–433.
Kremling,A. *et al.* (2008) A feed-forward loop guarantees robust behavior in *Escherichia coli* carbohydrate uptake. *Bioinformatics*, **24**, 704–710.
Kwon,Y.K. and Cho,K.H. (2007) Boolean dynamics of biological networks with multiple coupled feedback loops. *Biophys. J.*, **92**, 2975–2981.
Kwon,Y.K. and Cho,K.H. (2008a) Coherent coupling of feedback loops: A design principle of cell signaling networks. *Bioinformatics*, **24**, 1926–1932.
Kwon,Y.K. and Cho,K.H. (2008b) Quantitative analysis of robustness and fragility in biological networks based on feedback dynamics. *Bioinformatics*, **24**, 987–994.
Kwon,Y.K. *et al.* (2007) Investigations into the relationship between feedback loops and functional importance of a signal transduction network based on Boolean network modeling. *BMC Bioinformatics*, **8**, 384.
Lahav,G. *et al.* (2004) Dynamics of the p53-Mdm2 feedback loop in individual cells. *Nat. Genet.*, **36**, 147–150.
Le,D.H. and Kwon,Y.K. (2011a) The effects of feedback loops on disease comorbidity in human signaling networks. *Bioinformatics*, **27**, 1113–1120.
Le,D.H. and Kwon,Y.K. (2011b) NetDS: a Cytoscape plugin to analyze the robustness of dynamics and feedforward/feedback loop structures of biological networks. *Bioinformatics*, **27**, 2767–2768.
Li,F. *et al.* (2004) The yeast cell-cycle network is robustly designed. *Proc. Natl Acad. Sci. USA*, **101**, 4781–4786.
Ma,W. *et al.* (2009) Defining network topologies that can achieve biochemical adaptation. *Cell*, **138**, 760–773.
Macia,J. *et al.* (2009) Specialized or flexible feed-forward loop motifs: A question of topology. *BMC Syst. Biol.*, **3**, 84.
Mangan,S. *et al.* (2003) The coherent feedforward loop serves as a sign-sensitive delay element in transcription networks. *J. Mol. Biol.*, **334**, 197–204.
Mendoza,L. *et al.* (1999) Genetic control of flower morphogenesis in Arabidopsis thaliana: A logical analysis. *Bioinformatics*, **15**, 593–606.
Osella,M. *et al.* (2011) The role of incoherent microrna-mediated feedforward loops in noise buffering. *PLoS Comput. Biol.*, **7**, e1001101.
Pasek,S. *et al.* (2006) The role of domain redundancy in genetic robustness against null mutations. *J. Mol. Biol.*, **362**, 184–191.
Pomerening,J.R. *et al.* (2003) Building a cell cycle oscillator: Hysteresis and bistability in the activation of Cdc2. *Nat. Cell Biol.*, **5**, 346–351.
Rodrigo,G. and Elena,S.F. (2011) Structural discrimination of robustness in transcriptional feedforward loops for pattern formation. *PLoS One*, **6**, e16904.
Shmulevich,I. *et al.* (2005) Eukaryotic cells are dynamically ordered or critical but not chaotic. *Proc. Natl. Acad. Sci. USA*, **102**, 13439–13444.
Siso-Nadal,F. *et al.* (2009) Cross-talk between signaling pathways can generate robust oscillations in calcium and cAMP. *PLoS One*, **4**, e7189.
Tsai,T.Y.C. *et al.* (2008) Robust, tunable biological oscillations from interlinked positive and negative feedback loops. *Science*, **321**, 126–129.
Yang,J. *et al.* (2009) Feed-forward signaling of TNF-{alpha} and NF-{kappa}B via IKK-{beta} pathway contributes to insulin resistance and coronary arteriolar dysfunction in type 2 diabetic mice. *Am. J. Phys. Heart Circ. Phys.*, **296**, H1850–H1858.
Zecca,M. and Struhl,G. (2010) A feed-forward circuit linking wingless, fat-dachsous signaling, and the Warts-hippo pathway to *Drosophila* wing growth. *PLoS Biol.*, **8**, e1000386.
Zhu,L. *et al.* (2012) Engineering the robustness of industrial microbes through synthetic biology. *Trends Microb.*, **20**, 94–101.