The effects of feedback loops on disease comorbidity in human signaling networks

Le Duc Hau and Yung-Keun Kwon*

School of Computer Science and Information Technology, University of Ulsan, San 29, Mugeo 2-dong, Nam-gu, Ulsan 680-749.

Associate Editor: Dr. Olga Troyanskaya

ABSTRACT

Motivation: In general, diseases are more likely to be comorbid if they share associated genes or molecular interactions in a cellular process. However, there are still a number of pairs of diseases which show relatively high comorbidity but do not share any associated genes or interactions. This observation raises the need for a novel factor which can explain the underlying mechanism of comorbidity. We here consider a feedback loop structure ubiquitously found in the human cell signaling network as a key motif to explain the comorbidity phenomenon, since it is well known to have effects on network dynamics.

Results: For every pair of diseases, we examined its comorbidity and length of all feedback loops involved by the disease-associated genes in the human cell-signaling network. We found that there is a negative relationship between comorbidity and length of involved feedback loops. This indicates that a disease pair is more likely to comorbid if they are connected with feedback loops of shorter length. We additionally showed that such a negative relationship is more obvious when the number of positive involved feedback loops is larger than that of negative involved feedback loops. Moreover, we observed that the negative relationship between comorbidity and length of involved feedback loops holds especially for disease pairs that do not share any disease-associated genes. Finally, we proved all these results through intensive simulations, based on a Boolean network model.

Contact: kwonyk@ulsan.ac.kr

Supplementary information: Supplementary Information is available at BioInformatics online.

1 INTRODUCTION

When two disorders or illnesses occur in the same person, simultaneously or one after the other, they are called comorbid. It is reported that 80% of the elderly population has three or more chronic conditions (Caughey et al., 2008), which explains the importance of understanding disease comorbidity. There have been many studies observing the comorbidity phenomenon. Some studies have found high comorbidity rates of certain pairs of diseases (Gabriel et al., 2009; Ye et al., 2005; Raja et al., 2008) and addressed statistical results of prevalence and mortality of specified regions (Caughey et al., 2008; Tetsche et al., 2008). Other studies attempted to quantify the effect of a disease on other diseases by introducing comorbidity measures (Tang et al., 2008; Kelli et al., 2005). In a recent study of illness progression, a database was also constructed to summarize statistical correlations between phenotypic diseases from histories of more than 30 million patients in a phenotypic disease network (Hidalgo et al., 2009).

Another class of previous studies investigated factors to explain the cause of comorbidity. For example, Goh et al. constructed a human disease network in which a pair of diseases are linked when they share common disease-causing genes, and showed a common genetic origin of many diseases from this constructed network (Goh et al., 2007). In another study, Park et al. found statistically significant correlations between an underlying structure of cellular networks and comorbidity patterns in the human population by combining information on protein interactions, disease-gene associations, and population-level disease patterns extracted from Medicare data (Park et al., 2009). In that paper, the authors show positive correlations between the degree of comorbidity, the number of shared genes, and the number of shared protein interactions. A bipartite graph was also constructed in which nodes represent diseases and two diseases are linked if mutated enzymes associated with them catalyze adjacent metabolic reactions (Lee et al., 2008). Based on this graph, it was shown that two connected metabolic diseases sharing some pathways tend to show significant comorbidity. This result is consistent with a general proteomic notion that diseases may be related if they share protein interactions (Park et al., 2009) or proteins acting on the same pathway (Rual et al., 2005; Stelzl et al., 2005; Lim et al., 2006; Goehler et al., 2004; Calvano et al., 2005; Pujana et al., 2007; Oldham et al., 2006). Taken together, it can be generally accepted that diseases which share associated genes or molecular interactions in a cellular process, are more likely to be comorbid. However, it is interesting that many pairs of diseases show high comorbidity even though they do not share any associated genes or interactions (Park et al., 2009). Therefore, there is still a pressing need to find other factors which is related to comorbidity.

In this study, we consider a feedback loop (FBL) structure as a novel factor to explain the comorbidity phenomenon. FBLs are a
well-known critical motif to affect dynamics in biological networks (Snoussi, 1998, Mendoza et al., 1999, Milo et al., 2002, Yeger-Lotem et al., 2004, Prill et al., 2005). In particular, FBLs were shown to play an important role in robustly sustaining steady state of networks against perturbations (Kwon et al., 2008, Kwon et al., 2007b). In this regard, we investigated the relationship between comorbidity and FBLs involved with disease-associated genes in a human signaling network. Using integrated data from a disease-gene association database and a human cell-signaling network, we show that there is a negative relationship between comorbidity and the length of involved FBLs. In other words, a pair of diseases is highly comorbid if their associated genes are connected with feedback loops of relatively short length. It is interesting to note that this relationship is valid especially for disease pairs that do not share any associated genes. Moreover, such a negative relationship is more clearly observed when positive feedback loops are more abundant than negative feedback loops between a disease pair. We also show the negative relationship between comorbidity and the length of feedback loops through intensive simulations, based on random Boolean network models.

2 METHODS

2.1 Datasets

To obtain comorbidity information between pairs of diseases, we used the published dataset (Park et al., 2009) containing two quantified comorbidity values, RR (relative risk) and PHI (φ-correlation coefficient), for a total of 83,924 disease pairs. The two comorbidity measures were defined as RR = C1/NC0 and PHI = (NC0 − I/N) / √(I/N)(N − I/N)(N − I) , respectively, where N is the number of patients, I denotes incidence of disease i, C0 denotes the number of patients who were simultaneously diagnosed with diseases i and j, respectively, and C1 is I/N. In that study, the Medicare database which includes the clinical history of 13,039,018 patients was used for comorbidity evaluation. When two diseases co-occur more frequently than expected by chance, we have RR > 1 and PHI > 0. Each measure was reported to carry unique biases that are complementary (Hidalgo et al., 2009). Additionally, we defined the morbidity of a disease i as I/N (i.e., the prevalence of a disease).

In this paper, we analyze a large-scale human signaling network to find a topological characteristic related to the comorbidity phenomenon. To this end, we first obtained the human signaling network (Cui et al., 2007) which had been built up by integrating one cancer-related signaling network obtained from Cancer Cell Map (http://cancer.cellmap.org/cellmap.php), and three general signaling networks, which are not cancer-specific, obtained from BioCarta (http://www.biocarta.com) and two previous results of Ma'ayan et al., 2005 and Awan et al., 2007, respectively. We removed 62 small molecules such as Ca++, glutocorticoid, and H2O2, and their interactions from the network. As a result, the network consisting of 1,517 nodes and 4,761 links was constructed for this investigation (see Table S1 in Supplementary Information). In addition, we also obtained the entire list of disease-gene associations from the previous study (Park et al., 2009). After matching these databases (See Figure S1 in Supplementary Information for more details of database relationship diagram), we finally constructed a list of 334 diseases, their associated genes (see Table S2 in Supplementary Information), and their corresponding names (see Table S3 in Supplementary Information). We also collected comorbidity values of 18,896 pairs of those diseases (see Table S4 in Supplementary Information).

2.2 Definitions of topological properties in a network

We defined some topological properties with respect to either a single gene or a set of genes. In this paper, we consider a network represented by a directed graph G = (V, A), where V is a set of nodes and A is the set of ordered pairs of the nodes called directed links. A directed link (v, vj) is assigned with either a positive ("activating") or negative ("inhibiting") relationship from v ∈ V to vj ∈ V. When the human signaling network is represented by G(V, A), we denote the set of genes associated with a disease D by V(D)⊆V and then |V(D)| represents the number of genes associated with D. For a gene v, we consider the connectivity of v, which is defined as the number of links involving v. Moreover, connectivity of D is defined as the average connectivity over the set of genes in V(D).

In this paper, we consider feedback loops as an important topological property. Feedback loops are ubiquitously found and play an important role in dynamical behaviors of cellular signaling networks (Milo et al., 2002, Yeger-Lotem et al., 2004; Prill et al., 2005). A feedback loop is defined as a circular chain of relationships. For example, given a network G(V, A), v0 → v1 → v2 → ... → vL → v0 is a feedback loop of length L if there are links from v0 to v1, v1 to v2 and so on, up to vL to v0. In a similar way, when G(V, A) represents the human signaling network and a pair of diseases, D and D’, are given, D and D’ are called connected with a feedback loop of a maximal length L if there exists at least one feedback loop of length L involving both D and D’. In a similar way, when G(V, A) represents the human signaling network and a pair of diseases, D and D’, are given, D and D’ are called connected with a feedback loop of a maximal length L if there exists at least one feedback loop of length L involving both D and D’. Such that v and v’ are connected with a feedback loop of a maximal length L. Additionally, we denote the number of feedback loops between a pair of nodes or diseases by NuFBL(v, v’), and NuFBL(D, D’), respectively.

To analyze topological properties between a pair of nodes or diseases, we extend the definitions with respect to feedback loops, as follows. Given a network G(V, A) and a pair of nodes v ∈ V and v’ ∈ V, we call v and v’ connected with a feedback loop of a maximal length L if there exists at least one feedback loop of length L involving both v and v’. In a similar way, when G(V, A) represents the human signaling network and a pair of diseases, D and D’, are given, D and D’ are called connected with a feedback loop of a maximal length L if there exists at least one feedback loop of length L involving both D and D’. Such that v and v’ are connected with a feedback loop of a maximal length L. Additionally, we denote the number of feedback loops between a pair of nodes or diseases by NuFBL(v, v’), and NuFBL(D, D’), respectively.

2.3 Definitions of dynamical properties in a network

To prove our hypothesis, we employed a Boolean network model, which has been widely used to represent biological networks and successfully captured some biological characteristics (Kaufman et al., 2003; Shmulevich et al., 2003, 2004; Shmulevich et al., 2005; Kwon et al., 2007a). In particular, it has been also frequently used in simulating the dynamics of various signaling networks such as a guard cell abscisic acid signaling (Saadatpour et al., 2010), a central intrinsic and extrinsic apoptosis pathway (Maï et al., 2009; Schlatter et al., 2009), a mammalian ERBB signaling pathway (Sahin et al., 2009), a T-cell receptor signaling (Saez-Rodriguez et al., 2007), a neurotransmitter signaling pathway (Saez-Rodriguez et al., 2007), and so on.

2.3.1 A random Boolean network

When a Boolean network is represented by a directed graph G(V, A), each v ∈ V has a value of 1 ("on") or 0 ("off"), which represents the possible states of the corresponding elements. The value of each variable v, at time t+1 is determined by the values of ki other variables v1, v2, ..., vn, with a link to v, at time t by the Boolean function f: {0, 1}n → {0, 1}. Hence, we can write the update rule as v(t+1) = fi(v(t), v1(t), ..., vn(t)) where we randomly select either a logical conjunction or disjunction for all signed relationships in fi with a uniform probability distribution. For example, if a Boolean variable v has a positive relationship from v1, a negative relationship from v2, and a positive relationship from v3, then the conjunction and disjunction update rules are v(t+1) = vi(t) ∨ vj(t) and v(t+1) = vi(t) ∧ vj(t), respectively. In the case of a conjunction, the
value of $v$ at time $t + 1$ is 1 if only if the values of $v_1$, $v_2$, and $v_3$ at time $t$ are 1, 0, and 1, respectively whereas, 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, $v_1(t)$, $v_2(t)$, and $v_3(t)$ is 1. Although there can be many other logical functions in addition to conjunction and disjunction functions, biological networks were successfully described by Boolean models using only those two functions in many previous studies (Helikar et al., 2008; Kwon et al., 2007a; Albert, 2004; Faure et al., 2006; Huang et al., 2000). In addition, the sign of each link is determined between positive and negative ones uniformly at random.

To generate a large number of random Boolean networks, we considered three models. The first model generates random Boolean networks in a way that the connectivity of every node is $\geq 2$. On the other hand, the second model generates random Boolean networks in a way that every node has at least one incoming link and at least one outgoing link. These two models are denoted as Model-A and Model-B, respectively. The difference between the two models is that the first model permits presence of input and output nodes, while the second one does not (An input or output node means one which has no incoming or outgoing link, respectively, and there is no employed constraint on the number of input and output nodes in Model-A). Thus, the reason why we considered those different models is to show that our simulation results are not dependent on the presence of input and output nodes in the Boolean networks. Two visualization examples of random Boolean networks generated by Model-A and Model-B are shown in Figure S2 in Supplementary Information. In addition to the two models, we considered the other model proposed by Barabasi and Albert (Barabasi et al., 1999) which can generate random networks with a scale-free property, namely a power-law degree distribution (see Figure S3 in Supplementary Information for a more detailed generation process). We employed this model, which is denoted by Model-C, to examine whether our simulation results also hold for scale-free networks.

Given a Boolean network with $N$ Boolean variables, $v_1, v_2, ..., v_N$, we define a network state as a vector consisting of values of the Boolean variables: there are $2^N$ states in total. Each state transits to another state through a set of $N$ Boolean update functions, $f_1, f_2, ..., f_N$. We can construct a state transition diagram that represents the transition of each state. A state trajectory starts from an initial state and eventually converges to either a fixed-point or a limit-cycle attractor. Attractors can represent diverse behaviors of biological networks, such as multi-stability, homeostasis, and oscillation (Ferrell et al., 1998; Bhalla et al., 2002; Pomerening et al., 2003). In addition, we define a transient sequence of values of a node $v$ as follows: When a Boolean network $G(V, A)$ was initialized with $v_i(0)$, $v_j(0)$, ..., and $v_N(0)$ at the starting time 0, $v_i(t_1, t_2)$ represents a sequence of the transient values of a node $v_i$ during the time interval from $t_1$ to $t_2$.

### 2.3.2. Mutual-effectiveness in a random Boolean network

In Boolean networks, we propose a novel measure, mutual-effectiveness, to quantify the mutual influence between a pair of nodes or node groups in terms of the network dynamics. To define it, we first introduce two types of perturbations, an initial-state perturbation and an updating-rule perturbation. Given a Boolean network initialized with $v_i(0)$, $v_j(0)$, ..., and $v_N(0)$, the initial-state perturbation at a node $v_i \in V$ means flipping $v_i(0)$ to $T_i(0)$. On the other hand, the updating-rule perturbation at a node $v_i \in V$ means switching the updating-rule at $v_i$ from a conjunction function to a disjunction function or vice versa, depending on the current function type. Assuming a perturbation at $v_i$, we define the effectiveness from $v_i$ to another node $v_j$, $\mu(v_i, v_j)$, as follows:

1. Let $\tau_i$ the valid constringent time of $v_i$, defined as $\tau_i = \max(T_i, T_j)$, where $T_i$ or $T_j$ represent the time steps for the network to converge to an attractor when $v_i$ was subject to the perturbation or not, respectively.
2. We obtain two different transient sequences of $v_i$, $v_i(0, \tau_i)$ and $v_i'(0, \tau_i)$, when $v_i$ was subject to the perturbation or not, respectively.
3. Then, we compute $\mu(v_i, v_j) = \frac{d(v_i, v_j)}{\rho(v_i, v_j)} / \tau_i$ where $d(\cdot)$ means the Hamming distance (i.e., the number of bits having different values) between two sequences. Thus, $\mu(v_i, v_j)$ represents how largely the trajectory with respect to $v_i$ was affected by the perturbation at $v_j$.

Since $\mu$ is not commutative, we derive the mutual-effectiveness for a pair of nodes $v_i$ and $v_j$, $\rho(v_i, v_j)$, as follows: $\rho(v_i, v_j) = (\mu(v_i, v_j) + \mu(v_j, v_i)) / 2$.

Therefore, mutual-effectiveness is a measure about how largely each node is mutually affected by perturbation at the other node in terms of dynamics. In this regard, mutual-effectiveness in Boolean networks can be used to represent the comorbidity phenomenon in signaling networks. Figure 1 shows an example of the calculation of mutual-effectiveness of a node pair, $v_i$ and $v_j$. To compute $\mu(v_i, v_j)$, we get two transient sequences of $v_i$, $v_i(0, \tau_i)$ and $v_i'(0, \tau_i)$, when $v_i$ was subject to a perturbation or not, respectively. In the same way, $\mu(v_i, v_j)$ is computed and finally $\rho(v_i, v_j)$ are obtained by averaging $\mu(v_i, v_j)$ and $\mu(v_j, v_i)$.

In a Boolean network, a node is called a functional important node if a perturbation at the node makes the network converge to another attractor, which is different from the original attractor to which the network converged when the node was not subject to the perturbation. Therefore, mutual-effectiveness in biological networks can be used to consider the cellular dynamical behavior. In all simulations of this study, we generated random Boolean networks, such that the ratio of functionally important nodes over the total number of nodes is larger than or equal to 0.05.

### 3 RESULTS

#### 3.1 Analysis of morbidity in terms of the number of disease-associated genes, connectivity, and number of feedback loops

Before we investigated comorbidity, we first addressed how well morbidity could be explained in terms of some topological characteristics in the human signaling network. Morbidity of a disease was defined as the prevalence of a disease (see Materials and Me-
thod for the definition) and three topological properties were considered for analysis: the number of associated genes of a disease, the connectivity of a disease, and the number of feedback loops involved with a disease (see Materials and Method for the definitions). We plotted the relation of disease prevalence to each topological property (Fig. 2). This result shows that the correlation between disease prevalence and the three topological properties is very small. In other words, morbidity of a disease is not easy to simply understand in terms of topological properties in the human signaling network.

3.2 Analysis of comorbidity in terms of length of feedback loops in the human signaling network

We investigated the relationship of feedback loops to comorbidity in the human signaling network as follows: When a length $L$ was specified, we collected a set of pairs of diseases which are connected to a feedback loop of length $\leq L$ (see Materials and Methods for the definition) and computed the average comorbidity over the set of collected pairs of diseases (Fig. 3). Varying $L$ from 2 to 7, we examined two kinds of comorbidity values, i.e., $RR$ (Fig. 3(a)) and $PHI$ (Fig. 3(b)). $RR$ is maximal when $L$ is 2 while $PHI$ is so when $L$ is 3. Although there is such a difference between the peak points of $RR$ and $PHI$, the relationship trend between each comorbidity value and the maximal feedback loop length is interestingly negative (p-value=0.00324 in Fig. 3(a) and p-value=0.00824 in Fig. 3(b)). In other words, the comorbidity value of a pair of diseases is more likely to be high, as their associated genes are involved with feedback loops of shorter lengths. Considering it was not easy to find any topological property correlated to morbidity in Fig. 2, this finding is intriguing. Moreover, other topological properties such as the length of the shortest path between a disease pair and the number of feedback loops connecting a pair of diseases did not show any obvious relationship to comorbidity values (see Figure S4(a) and (b) in Supplementary Information).

3.3 The effect of feedback loop length on mutual-effectiveness in random Boolean networks

To understand why the comorbidity trend is negatively related to feedback loop length, we performed extensive simulations based on random Boolean networks models. We generated 100 random Boolean networks with $|V|=50$ and $|A|=75$, collected a group of functional important node pairs, which are connected with a feedback loop of length $\leq L$ by varying $L$ from 2 to 10, and examined the mutual effectiveness of each group (Fig. 4). We used three kinds of random Boolean networks models: one that permits the presence of input/output nodes (Model-A; Fig. 4(a)), another does not (Model-B; Fig. 4(b)), and the other generates scale-free networks (Model-C; Fig. 4(c)) (see Materials and Methods for the definitions). In addition, we considered two types of perturbations: an initial-state perturbation and an update-rule perturbation (see Materials and Method for the definitions). We observed a strong negative relationship between the maximal FBL length and mutual effectiveness (all p-values < 0.001), irrespective of the generation models and types of perturbations. Moreover, we also performed the same simulations with random Boolean networks of different network sizes (IV) and different network densities (ratio of $|A|$ over $|V|$). We found that the relationship between the maximal feedback

---

**Fig. 2.** Correlations between disease prevalence and topological properties over 334 diseases. Considered topological properties are (a) the number of genes associated with a disease, (b) the connectivity of a disease and (c) the number of feedback loops involved with a disease. Pearson correlation coefficients are 0.260, -0.065, and -0.041, respectively. All axes are logarithmic in scale.

**Fig. 3.** The relationship between average comorbidity and maximal FBL length in the human signaling network. All y-axis values represent the average comorbidity with a 95% confidence level. Blue lines represent linear regression of the average comorbidity values. (a) Result of $RR$ comorbidity (slope of linear regression $\approx -0.40655$) (b) Result of $PHI$ comorbidity (slope of linear regression $\approx -0.0003$).
Moreover, other topological properties such as the number of other nodes are also involved in that transference. Effect at a point cannot be well transferred to other points since a feedback loop involves a large number of other nodes and, thus, a perturbation feedback loops is as follows. A feedback loop of longer length is more likely to induce limit-cycle attractors on the other hand, networks with a relatively large number of negative feedback loops are more likely to induce fixed-point attractors; on the other hand, networks with a relatively large number of positive feedback loops are more likely to induce limit-cycle attractors (Kwon et al., 2007a). Therefore, we further examined the effect of the sign of feedback loop on comorbidity in the random Boolean networks (Fig. 5) and mutual-effectiveness in random Boolean networks (Fig. 6). When a maximal feedback loop length \( L \) was specified, we collected a set of disease pairs, \( D \) and \( D' \), which are connected with a feedback loop of length \( \geq L \) and classified them into two categories, ‘\( \text{NuFBL}, \geq \text{NuFBL} \)’ and ‘\( \text{NuFBL}, < \text{NuFBL} \)’, according to difference of \( \text{NuFBL}(D, D') \) and \( \text{NuFBL}(D, D') \). We compared comorbidity between those two categories in the human signaling network (Fig. 5). For both \( RR \) and \( PHI \), we observed that disease pairs belonging to the ‘\( \text{NuFBL}, \geq \text{NuFBL} \)’ group showed a clear negative relationship between average comorbidity and maximal FBL length compared to those belonging to the ‘\( \text{NuFBL}, < \text{NuFBL} \)’ group (\( p\)-value=0.00229 in Fig. 5(a) and \( p\)-value=2.277\times10^{-5} \) in Fig. 5(b)). The slope difference between the two groups was larger in nodes did not show any obvious relationship to mutual-effectiveness values as in the case of comorbidity in Section 3.2 (see Figure S4(c) and (d) in Supplementary Information).

3.4 The effect of sign of feedback loop on comorbidity and mutual-effectiveness

It was also reported that the dynamic behavior of networks depends on the sign of feedback loops. In terms of converging dynamics, networks with a relatively large number of positive feedback loops are more likely to induce fixed-point attractors; on the other hand, networks with a relatively large number of negative feedback loops are more likely to induce limit-cycle attractors (Kwon et al., 2007a). Therefore, we further examined the effect of the sign of feedback loops on comorbidity in the signaling network (Fig. 5) and mutual-effectiveness in random Boolean networks (Fig. 6). When a maximal feedback loop length \( L \) was specified, we collected a set of disease pairs, \( D \) and \( D' \), which are connected with a feedback loop of length \( \geq L \) and classified them into two categories, ‘\( \text{NuFBL}, \geq \text{NuFBL} \)’ and ‘\( \text{NuFBL}, < \text{NuFBL} \)’, according to difference of \( \text{NuFBL}(D, D') \) and \( \text{NuFBL}(D, D') \). We compared comorbidity between those two categories in the human signaling network (Fig. 5). For both \( RR \) and \( PHI \), we observed that disease pairs belonging to the ‘\( \text{NuFBL}, \geq \text{NuFBL} \)’ group showed a clear negative relationship between average comorbidity and maximal FBL length compared to those belonging to the ‘\( \text{NuFBL}, < \text{NuFBL} \)’ group (\( p\)-value=0.00229 in Fig. 5(a) and \( p\)-value=2.277\times10^{-5} \) in Fig. 5(b)). The slope difference between the two groups was larger in
the case of PHI than RR. In a similar way, we examined mutual-effectiveness in random Boolean networks generated by three models, Model-A (Fig. 6(a)), Model-B (Fig. 6(b)) and Model-C (Fig. 6(c)). For all models, we also observed that pairs of nodes belonging to the ‘NuFBL ≥ NuFBL’ group showed a more outstanding negative relationship between average of mutual-effective and maximal FBL length than those belonging to ‘NuFBL < NuFBL’. (For cases of initial-state perturbations, p-values are 0.00535, 0.02164, and 0.01898 in Fig. 6(a), (b), and (c), respectively; For cases of updating-rule perturbations, p-values are 0.04647, 0.00028, and 0.00151 in Fig. 6(a), (b), and (c), respectively). As shown in Fig. 6, this observation was consistent irrespective of the type of perturbation and network generation model. Also, it was independent of network size and density (See Figure S6 in Supplementary Information). Taken together, the negative relationship between comorbidity/mutual-effectiveness and the length of the involved feedback is more apparent in the case in which the number of involved positive feedback loops is larger than that of the involved negative feedback loops. This may be because positive FBLs are mainly related to amplifying signals, while negative FBLs play a role in inhibiting signals (Mendoza et al., 1999, Claire, 2004).

3.5 The effect of feedback loops on comorbidity when there is no common disease gene

A previous study showed that a disease pair becomes more comorbid as they share a larger number of disease genes (Park et al., 2009). Inspired by the result, we further investigated the relationship between comorbidity and feedback loop length for the group of disease pairs that do not share any disease genes (Fig. 7). We classified every disease pair into two groups: “Shared” (set of disease pairs having at least one gene or no genes, respectively. (a) Result of RR comorbidity (slopes of blue and green line are -0.14136 and -0.46398, respectively). (b) Result of PHI comorbidity, (slopes of blue and green line are 0.00025 and -0.00027, respectively). All y-axis values represent the average comorbidity with a 95% confidence level.

Fig. 6. Comparisons of the relationship between average mutual-effectiveness of functional important nodes and maximal FBL length according to the majority sign of the involved feedback loops in random Boolean networks with |V| = 50 and |A| = 75. Each pair of nodes is classified into two categories: NuFBL ≥ NuFBL, if the number of involved positive FBLs is larger than or equal to that of involved negative FBLs, and NuFBL < NuFBL, if otherwise. Two types of perturbations, initial-state perturbation (blue) and updating-rule perturbation (green), were considered. All the y-axis values represent the average mutual-effectiveness values with a 95% confidence level. Solid and dashed lines represent a linear regression of average mutual-effectiveness values of ‘NuFBL ≥ NuFBL’ and ‘NuFBL < NuFBL’ categories, respectively. (a) Results of random Boolean networks generated by Model-A (slopes of blue solid, blue dashed, green solid, and green dashed lines are -0.02846, -0.01325, -0.01295, and -0.00888, respectively.) (b) Results of random Boolean networks generated by Model-B (slopes of blue solid, blue dashed, green solid, and green dashed lines are -0.04067, -0.02371, -0.02687, and -0.00753, respectively.) (c) Results of random Boolean networks generated by Model-C (slopes of blue solid, blue dashed, green solid, and green dashed lines are -0.01978, -0.00692, -0.00581, and -0.00233, respectively.)
shared" group (p-value=0.04746 in Fig. 7(a) and p-value=0.0075878 in Fig. 7(b)). This means that the number of shared genes is an important indicator for comorbidity, as shown in the previous study (Park et al., 2009). In addition, we observed that there is a negative relationship between comorbidity and length of feedback loops in the “Not-shared” group (p-value=0.01352 in Fig. 7(a) and p-value=0.00353 in Fig. 7(b)), while the relationship is ambiguous in the “Shared” group (The relationship is even positive in the case of PHI). This implies that a pair of diseases having no shared disease gene can be highly comorbid if they are connected with feedback loops of relatively short length. Considering most of disease pairs connected with feedback loops of relatively short length can be highly comorbid especially when they do not share any gene. Through intensive simulations based on Boolean networks, we proved the relationship between feedback loops and comorbidity is a fundamental property with respect to network dynamics.

In synthetic biology, the findings of this paper can be useful in order to control mutual effectiveness between components of an artificial cell. Also, the findings of this study can be considered evidence showing the usefulness of a Boolean network model in studying the dynamics of biological networks. In addition, it will be a significant future study to find out a more realistic measure than mutual-effectiveness to simulate comorbidity phenomenon in a Boolean network model.

**ACKNOWLEDGEMENTS**

**Funding:** This work was supported by National Research Foundation of Korea Grant funded by the Korean Government (KRF-2008-313-D00995).

**REFERENCES**


