PAPER

Co-controllability of drug-disease-gene network

Peng Gang Sun
School of Computer Science and Technology, Xidian University, Xi’an, 710071, People’s Republic of China
Institute of Computational Bioinformatics, Xidian University, Xi’an, 710071, People’s Republic of China
E-mail: psun@mail.xidian.edu.cn

Keywords: minimum dominating set, co-controllability, drug-disease-gene network

Abstract

Controllability of a single network often focuses on the determination of the network’s minimum dominating set, which aims to elaborate how to control the whole network with minimum driver nodes. This paper proposes a new framework, co-controllability of multiple networks, which stresses the control of one network by another network as well as the mutual control characteristics of multiple networks based on minimum dominating sets. We take a drug–disease–gene network that consists of a drug–drug network, a disease–disease network and a gene–gene network as an example to study co-controllability of multiple networks. The results show that driver nodes tend to be conserved, e.g. diseases highly associated with driver nodes of the drug–drug network tend to be driver nodes in the disease–disease network compared with random networks. In addition, co-controllability of multiple networks is probably associated with the networks’ node degree, which is more stringent than controllability of a single network that is mainly determined by the network’s degree distribution. We also find that diseases and drugs tend to be mapped as two different subnetworks of human protein–protein interaction (PPI) network, drugs are inclined to dominate diseases by controlling the PPI network, and the coded proteins of disease-related genes exhibit a low tendency to be drug targets for the control of diseases. The results in this paper not only play an important role in understanding co-controllability of multiple networks, but also are helpful for understanding the mechanisms of drug–disease–gene, disease treatments and drug design in a network-based framework.

1. Introduction

The study of complex networks and systems promotes the development of systems biology [2–4, 8, 10, 11, 13, 21, 25, 27–32, 40, 45]. Currently, biological network analysis as an efficient way to study the mechanisms of complex biological systems has become more and more important for systems biology [4, 5, 7, 8, 12, 14, 15, 17–22, 33–36, 47]. In general, we use biological data to build a biological network for a biological system, and through the study of the network’s characteristics, we try to elaborate how molecules, interactions and structures of the network determine the functions of the biological system [4, 5, 7, 8, 12, 14, 15, 17–22, 33–36, 47]. Biological network analysis not only can help us understand cellular organizations, processes and functions, but also is very helpful for disease diagnosis, treatments and drug design [1, 2, 4, 8, 11, 13, 21, 25, 27–30, 32, 40, 45, 48]. Systems biology strongly focuses on the vertical integration from microscopic processes to macroscopic biological phenomena [2–4, 8, 10, 11, 13, 21, 25, 27–30, 32, 40, 45, 48], while network physiology and network medicine as newly and rapidly developed research fields focus on the horizontal integration of different hierarchies of biological systems [49, 50]. In contrast, network physiology tries to understand physiologic functions as behaviors that are out of the collective coordination and communication of different dynamical components [49], and network medicine addresses the correlations among genes, drugs and diseases [50]. Therefore, the new fields of network physiology and network medicine are totally different from systems biology.
In the past several years, the study of characteristics of biological networks [2, 4–9, 14–20, 24, 33–39, 41, 47, 48] provides us with a new insight into understanding complex biological systems, e.g. Goh et al [12] built a human disease network and found that genes causing similar diseases exhibit an increased tendency for their protein products to interact with one another and tend to be coexpressed in specific human tissues based on the analysis of characteristics of the disease network. Furthermore, more works [18, 34, 36] have elaborated that biological network analysis plays an important role in understanding human diseases. Notably, Kim et al [48] developed an algorithm to identify the kernel of a network based on human signaling pathways and showed that kernel genes are more likely to be essential and highly associated with diseases and drugs. In addition, genes within the kernel not only maintain a low evolutionary rate, but also are ubiquitously expressed in various tissues and well conserved between species, which indicates that resultant kernel preserves the network’s dynamics [48]. Recently, a drug–disease–gene network that consists of three subnetworks, a drug–drug network (DrDrN), a disease–disease network (DiDiN) and a gene–gene network (GeGeN) was developed [33]. By the study of the network’s characteristics, this work tried to uncover the correlations between the network structures and the functions of biological systems as well as the mechanisms of drug–disease–gene in a network-based framework [33].

In recent years, the focus of complex networks has shifted towards the study of the networks’ controllability, which addresses the determination of minimum driver nodes in the dominating sets to achieve full control of the whole networks [23, 26, 41–44]. In particular, Liu et al [23] developed a framework to study the controllability of directed networks and achieve full control of the networks by identifying minimum driver nodes. They proved that the problem of minimum number of driver nodes can be determined by the maximum matching problem [23]. The results also showed that the number of driver nodes is mainly determined by the networks’ degree distribution [23]. Furthermore, Nacher and Akutsu [26] developed an equivalent optimization procedure to determine the minimum dominating sets in order to take full control of undirected networks. The results showed that it is easier to control a network with a more heterogeneous degree distribution [26]. Yuan et al [43] studied the minimum driver nodes on the networks with arbitrary network structures and link weight distributions. The results showed that it is hard to control dense networks with same link weights [43].

The controllability of complex networks motivates us to study the controllability of biological networks. Notably, Wuchty [41] studied the minimum driver nodes in the dominating set of protein–protein interaction (PPI) networks. The minimum dominating set of proteins is defined as an optimized set, and each non-driver protein interacts directly with a driver protein at least [41]. The results showed that driver proteins tend to be essential, disease-related and virus-targeted genes [41]. Although the controllability of biological networks provides us with a new insight into understanding the mechanisms of complex biological systems, the above still focuses on the controllability of a single network. A big biological system is actually composed of plenty of small biological networks that cooperate with one another to complete the functions of the whole biological system, which motivates us to study the co-controllability of multiple biological networks to elaborate the mechanisms of biological systems.

Based on the discussions above, we break through the limitation of controllability of a single network that only focuses on the control of a network with minimum driver nodes and propose a new framework, co-controllability of multiple networks, which stresses the control of one network by another network as well as the mutual control characteristics of multiple networks based on minimum dominating sets. We take a drug–disease–gene network that consists of a drug–drug network, a disease–disease network and a gene–gene network as an example and try to elaborate the co-controllability of multiple networks as well as the mechanisms of drug–disease–gene in a network-based framework.

The rest of the paper is organized as follows. In section 2, we describe the drug–disease–gene network. Section 3 discusses controllability of the drug–disease–gene network. Section 4 studies co-controllability of the drug–disease–gene network. The conclusion is provided in section 5.

2. Drug–disease–gene network

Recently, a drug–disease–gene network that consists of a drug–drug network (DrDrN), a disease–disease network (DiDiN) and a gene–gene network (GeGeN) was built to elaborate the mechanism of drug–disease–gene in a network-based framework [33]. The network is constructed based on 5018 drug–disease associations [9], 3702 drug–target (drug–gene) associations [37] and 877 disease–gene associations [16, 24].

In the network, the DrDrN, the DiDiN and the GeGeN are also composed of two parts respectively, e.g. the DrDrN consists of a DrDrN\textsuperscript{disease} and a DrDrN\textsuperscript{gene}. In the DrDrN\textsuperscript{disease} (21 429 links, 288 drugs), two drugs are connected if they are associated with a same disease [12, 33]. In the DrDrN\textsuperscript{gene} (14 665 links, 704 drugs), two drugs are connected if they are associated with a same drug target gene [12, 33].
Here, we take the drug–disease–gene network as an example to elaborate the controllability and the co-controllability of the DrDr\textsubscript{disease}, the DrDr\textsubscript{gene}, the DiDi\textsubscript{drug}, the DiDi\textsubscript{gene}, the GeGe\textsubscript{disease} and the GeGe\textsubscript{drug}.

3. Controllability of drug–disease–gene Network

We often study the controllability of a single network by determining the network’s minimum dominating set. In this paper, \( G(V, E) \) is used to denote an undirected network, where \( V \) and \( E \) correspond to the set of nodes and the set of links respectively, and \( |V| = n, |E| = l \). \((a_{ij})_{n \times n}\) denotes the adjacency matrix of \( G \), and \( a_{ij} = 1 \) indicates that node \( i \) and node \( j \) interact directly, and 0 otherwise.

A minimum dominating set of \( G \) is defined as a subset of \( V' \subseteq V \) that every node in \( V \) is either adjacent to an element of \( V' \) at least or is an element of \( V' \), i.e. \( \forall i \in V, \text{if } i \notin V', \text{then } \exists j \in V', a_{ij} = 1 \) \((i \neq j)\), and \( i \in V' \) otherwise (see figures 1(b) and (c)) \([23, 26, 41–44]\).

The minimum dominating set problem can be solved by ‘0-1’ integer programming, which is mentioned in \([23, 26, 41–44]\).

\[
\min \sum_{i \in V} x_i \tag{1}
\]

subject to

\[
x_i + \sum_{j \in V} a_{ij}x_j \geq 1 \tag{2}
\]

where \( x_i = 1 \) indicates \( i \in V' \), and 0 otherwise.

Table 1 shows the characteristics of minimum dominating sets of the DrDr\textsubscript{disease}, the DrDr\textsubscript{gene}, the DiDi\textsubscript{drug}, the DiDi\textsubscript{gene}, the GeGe\textsubscript{disease} and the GeGe\textsubscript{drug} compared with three randomization procedures based on degree distribution conservation, node degree conservation and Erdős–Rényi (E–R) model \([10]\) respectively. In the table, \( n, l \) correspond to the number of nodes and the number of links in the networks; \( n_{mds} \) indicates the number of nodes or the size of minimum dominating sets; \( \%_{mds} \) indicates the fraction of nodes that belong to minimum dominating sets; \( \langle k \rangle \) indicates the average node degree of the networks; and \( \langle k_{mds} \rangle \) indicates the average node degree of minimum dominating sets.

From the table, we can see that the controllability of a single network is mainly determined by the network’s degree distribution \([23]\), e.g. the DrDr\textsubscript{gene} and the randomized DrDr\textsubscript{gene} based on degree distribution conservation share the same \( n_{mds} \) and almost the same \( \langle k_{mds} \rangle \) compared with the randomized DrDr\textsubscript{gene} based on node degree conservation and that of the E–R model. We also can see that \( \langle k_{mds} \rangle \) is not only associated with the network’s degree distribution, but also the network’s density, e.g. in densely connected networks such as the DrDr\textsubscript{disease} and the DiDi\textsubscript{drug}, the average degree of driver nodes tends to be greater than the networks’ average degree.

The genes in the GeGe\textsubscript{drug} (17 248 links, 1036 genes) contain 19 524 links in human protein–protein interaction (PPI) network, called PPI (GeGe\textsubscript{drug}) and the genes in the GeGe\textsubscript{disease} (2173 links, 638 genes) contains 3423 links in human PPI network, called PPI (GeGe\textsubscript{disease}). In figure 2(a), we can see that the
GeGeN\text{drug} and the GeGeN\text{disease} share 123 genes and 85 links, while the PPI (GeGeN\text{drug}) and the PPI (GeGeN\text{disease}) share 123 genes and 461 links. In addition, the GeGeN\text{drug} is just a spanning subnetwork of the PPI (GeGeN\text{drug}), and the GeGeN\text{disease} is also a spanning subnetwork of the PPI (GeGeN\text{disease}).

### Table 1. Characteristics of minimum dominating sets.

| DrDrNgene | DrDrNdisease | DiDiN\text{drug} | DiDiNgene | GeGeN\text{drug} | GeGeN\text{disease} |
|-----------|--------------|------------------|-----------|-----------------|---------------------|
| n         | 704          | 288              | 267       | 193             | 1036                |
| l         | 14,665       | 21,429           | 31,253    | 300             | 17,248              |
| $n_{\text{mds}}$ | 66          | 2                | 1         | 58              | 111                 |
| $p_{\text{mds}}$ | 9.38%      | 0.69%            | 0.37%     | 30.05%          | 10.71%              |
| $\langle k \rangle$ | 41.66      | 148.81           | 234.10    | 3.11            | 33.30               |
| $\langle k_{\text{mds}} \rangle$ | 23.89       | 281              | 266       | 3.66            | 21.11               |
| Degree distribution conservation (randomized) |
| DrDrNgene | DrDrNdisease | DiDiN\text{drug} | DiDiNgene | GeGeN\text{drug} | GeGeN\text{disease} |
| n         | 66           | 2                | 1         | 58              | 111                 |
| $n_{\text{mds}}$ | 66          | 2                | 1         | 58              | 111                 |
| $p_{\text{mds}}$ | 9.38%      | 0.69%            | 0.37%     | 30.05%          | 10.71%              |
| $\langle k \rangle$ | 23.89       | 281              | 266       | 3.66            | 21.11               |
| $\langle k_{\text{mds}} \rangle$ | 23.89       | 281              | 266       | 3.66            | 21.11               |
| Node degree conservation (randomized) |
| DrDrNgene | DrDrNdisease | DiDiN\text{drug} | DiDiNgene | GeGeN\text{drug} | GeGeN\text{disease} |
| n         | 223          | 6                | 4         | 54              | 111                 |
| $n_{\text{mds}}$ | 31.68%      | 2.08%            | 1.50%     | 27.98%          | 10.71%              |
| $\langle k \rangle$ | 10.92        | 211.83           | 222.25    | 4.91            | 95.64               |
| $\langle k_{\text{mds}} \rangle$ | 20.39        | 281              | 266       | 3.66            | 21.11               |
| E–R (randomized) |
| DrDrNgene | DrDrNdisease | DiDiN\text{drug} | DiDiNgene | GeGeN\text{drug} | GeGeN\text{disease} |
| n         | 54           | 5                | 2         | 74              | 101                 |
| $n_{\text{mds}}$ | 7.67%      | 1.74%            | 0.75%     | 38.34%          | 9.75%               |
| $\langle k \rangle$ | 40.59        | 148.8            | 237.5     | 2.74            | 31.41               |
| $\langle k_{\text{mds}} \rangle$ | 40.59        | 148.8            | 237.5     | 2.74            | 31.41               |

Figure 2. Characteristics of the GeGeN\text{drug}, the GeGeN\text{disease}, the PPI (GeGeN\text{drug}) and the PPI (GeGeN\text{disease}). (a) The GeGeN\text{drug} and the GeGeN\text{disease} tend to be considered as two different subnetworks of human PPI network. (b)–(e) correspond to $n_{\text{mds}}$, $p_{\text{mds}}$, $\langle k \rangle$ and $\langle k_{\text{mds}} \rangle$ of the GeGeN\text{drug}, the GeGeN\text{disease}, the PPI (GeGeN\text{drug}) and the PPI (GeGeN\text{disease}), respectively.
results may indicate that diseases and drugs can be mapped as a subnetwork of human PPI network respectively, and two subnetworks tend to be different.

Figures 2(b)–(e) shows the characteristics of minimum dominating sets of the GeGeNdrug, the GeGeNdisease, the PPI (GeGeNdrug) and the PPI (GeGeNdisease). Since the PPI (GeGeNdrug) and the PPI (GeGeNdisease) are densely connected compared with the GeGeNdrug and the GeGeNdisease respectively, they need fewer driver nodes to control and have a greater $\langle k_{\text{mds}} \rangle$.

In the DiDiNdrug, only one driver node (OMIM ID: 166710) can control the whole network, and we need two driver nodes (DB00316, DB00783) to control the DrDrNdisease. We also can see that in both the GeGeNdrug and the GeGeNdisease driver nodes are not housekeeping genes (maintenance genes) [38], but 26.13% (29/111) and 41.22% (61/148) are essential genes respectively [46], which may indicate that disease-related genes tend to be essential compared with drug target genes.

4. Co-controllability of drug–disease–gene network

Controllability of a single network often focuses on the determination of the network’s minimum dominating set, which aims to elaborate how to control the whole network with minimum driver nodes [23, 26, 41–44]. A large biological system is actually composed of plenty of small biological networks that cooperate with one another to complete the functions of the biological system, which motivates us to study co-controllability of multiple biological networks to elaborate the mechanisms of biological systems. Here, we propose a new framework, co-controllability of multiple networks, which stresses how a network can control another network and the mutual control of multiple networks. Figures 3 and 4 illustrate the co-controllability of multiple networks.

Here, we first define some measures to evaluate the co-controllability of multiple networks, which are described as follows:

Figure 3. Illustration of indirect co-controllability of multiple networks. (a)–(b) show 4 networks, A, B, C and D. (c) shows the minimum dominating sets of the 4 networks. (d) A indirectly controls C via B, and C indirectly controls A via B. (e)–(f) correspond to Contr\text{ind}(A \rightarrow C) and Contr\text{ind}(C \rightarrow A) respectively. (g)–(h) correspond to Contr\text{net}(A \rightarrow C) and Contr\text{net}(C \rightarrow A), respectively.
where \( f_{mds}(A \rightarrow B) \) denotes a mapping from the minimum dominating set of the network \( A \) to that of the network \( B \) (see figures 3(e), (f)).

Equation (4) denotes the indirect control from the driver nodes of \( A \) to that of \( C \) via the driver nodes of \( B \).

\[
\text{Contr}_{mds}^{\text{net}}\left(\frac{\overline{AB} \rightarrow C}{A \rightarrow B} \right) = \left| \frac{f_{mds}(A \rightarrow B) \rightarrow C}{\left| C_{mds}^{\text{net}} \right|} \right|
\]

where \( |X| \) is the size of \( X \), \( C_{mds}^{\text{net}} \) is the minimum dominating set of the network \( C \), and \( f_{mds}(A \rightarrow C) \subseteq C_{mds}^{\text{net}} \). Equation (4) denotes the indirect control from the driver nodes of \( A \) to that of \( C \) via the driver nodes of \( B \).

\[
\text{Contr}_{mds}^{\text{net}}\left(\frac{\overline{AB} \rightarrow C}{A \rightarrow B} \right) = \left| \frac{g_{\text{expand}}(f_{mds}(A \rightarrow B) \rightarrow C)}{|C|} \right|
\]

where \( g_{\text{expand}}(X) = \{i, j | i \in X \text{ or } a_{ij} = 1\} \), \( |C| \) is the number of nodes of \( C \), \( g_{\text{expand}}(f_{mds}(A \rightarrow C)) = g_{\text{expand}}(f_{mds}(f_{mds}(A \rightarrow B) \rightarrow C)) \). Equation (5) denotes the indirect control from the network \( A \) to the network \( C \) via the network \( B \) (see figures 3(g), (h)).

\[
\text{CoContr}_{mds}^{\text{net}}\left(\frac{\overline{AB} \leftrightarrow C}{A \leftrightarrow B} \right) = \left| \frac{f_{mds}(A \rightarrow B) \rightarrow C}{\left| A_{mds} \right| + \left| C_{mds}^{\text{net}} \right|} \right|
\]

Equation (6) denotes the indirect co-control between the driver nodes of \( A \) and that of \( C \) via the driver nodes of \( B \).
CoContr\textsuperscript{net}\left(\frac{AB}{A \leftrightarrow C}\right) = \frac{\mathcal{S}_{\text{expand}}\left(f_{\text{mds}}\left(\frac{AB}{A \rightarrow C}\right)\right)}{|A| + |C|} + \frac{\mathcal{S}_{\text{expand}}\left(f_{\text{mds}}\left(\frac{AB}{C \rightarrow A}\right)\right)}{|A| + |C|}

(7)

Equation (7) denotes the indirect co-control between the network A and the network C via the network B. Similarly, it is easy to define direct control and direct co-control (see figure 4):

\text{Contr}_{\text{mds}}\left(\frac{AB}{A \rightarrow B}\right) = \left|\frac{f_{\text{mds}}\left(\frac{AB}{A \rightarrow B}\right)}{|B|}\right|

(8)

Equation (8) denotes the direct control from the driver nodes of A to that of B, which is defined as the fraction of driver nodes of B that are associated with the driver nodes of A (see figures 4(b), (c)). The greater Contr\textsuperscript{mds}\left(\frac{AB}{A \rightarrow B}\right) indicates that the driver nodes of B tend to be conserved with respect to that of A.

\text{Contr}\textsuperscript{net}\left(\frac{AB}{A \rightarrow B}\right) = \left|\frac{\mathcal{S}_{\text{expand}}\left(f_{\text{mds}}\left(\frac{AB}{A \rightarrow B}\right)\right)}{|B|}\right|

(9)

Equation (9) denotes the direct control from the network A to the network B. The greater Contr\textsuperscript{net}\left(\frac{AB}{A \rightarrow B}\right) means the stronger control ability from the network A to the network B (see figures 4(d), (e)).

\text{CoContr}_{\text{mds}}\left(\frac{AB}{A \leftrightarrow B}\right) = \left|\frac{f_{\text{mds}}\left(\frac{AB}{A \rightarrow B}\right) + f_{\text{mds}}\left(\frac{AB}{B \rightarrow A}\right)}{|A| + |B|}\right|

(10)

Equation (10) denotes the direct co-control between the driver nodes of A and that of B.

\text{CoContr}\textsuperscript{net}\left(\frac{AB}{A \leftrightarrow B}\right) = \left|\frac{\mathcal{S}_{\text{expand}}\left(f_{\text{mds}}\left(\frac{AB}{A \rightarrow B}\right)\right) + \mathcal{S}_{\text{expand}}\left(f_{\text{mds}}\left(\frac{AB}{B \rightarrow A}\right)\right)}{|A| + |B|}\right|

(11)

Equation (11) denotes the direct co-control between the network A and the network B.

Figure 5 shows the results of Contr\textsuperscript{mds}\left(\frac{AB}{A \rightarrow B}\right) for the DrDrN\textsuperscript{disease}, the DrDrN\textsuperscript{gene}, the DiDiN\textsuperscript{drug}, the DiDiN\textsuperscript{gene}, the GeGeN\textsuperscript{disease}, the GeGeN\textsuperscript{drug}, and the randomized DrDrN\textsuperscript{disease}, the randomized DrDrN\textsuperscript{gene}, the randomized DiDiN\textsuperscript{drug}, the randomized DiDiN\textsuperscript{gene}, the randomized GeGeN\textsuperscript{disease}, the randomized GeGeN\textsuperscript{drug} based on degree distribution conservation, node degree conservation and E–R model respectively. From figures 5(a)–(c), we can find that driver nodes in real networks tend to be conserved compared with random networks, e.g. driver nodes in the DiDiN\textsuperscript{drug} are highly associated with that of the DrDrN\textsuperscript{disease} and the DrDrN\textsuperscript{gene}. Similarly, in figure 5(b), driver nodes in the DrDrN\textsuperscript{disease} are closely related to that of the DiDiN\textsuperscript{gene}. In figure 5(c), driver nodes in the DiDiN\textsuperscript{drug} are highly associated with that of the GeGeN\textsuperscript{disease}, which may indicate that the gene–gene network can be considered as a bridge linking diseases and drugs.

Table 2 shows the results of direct control between the drug networks and the disease networks. We can see in the table that driver nodes of the disease networks as targets of control are highly associated with that of the drug networks. For example, the driver node (OMIM ID: 166710) as a target of control from the DrDrN\textsuperscript{disease} to the DiDiN\textsuperscript{drug} is associated with osteoporosis and bone mineral density variation QTL [16, 24]. The driver node (DB00873) as a target of control from the DiDiN\textsuperscript{drug} to the DrDrN\textsuperscript{disease} is associated with loteprednol [16, 24]. We can use loteprednol to treat allergic conjunctivitis, uveitis, acne rosacea, selected infective conjunctivitides, herpes zoster keratitis, iritis, cyclitis, and superficial punctate keratitis [16, 24, 37]. It is used for the treatment and management of seasonal allergic rhinitis [16, 24, 37]. The driver node (DB00316) as a target of control from the DrDrN\textsuperscript{disease} to the DiDiN\textsuperscript{gene} is associated with acacetaminophen [16, 24, 37]. Acacetaminophen (DB00316) is also known as paracetamol, and is widely used due to its analgesic and antipyretic effects. Acacetaminophen’s therapeutic effects are similar to salicylates, but it lacks gastric ulcerative, antiplatelet and anti-inflammatory effects [16, 24, 37]. Acacetaminophen as an analgesic and antipyretic drug is widely used for the relief of fever, headaches [16, 24, 37]. Acacetaminophen is also a major ingredient in numerous cold and flu medications and many prescription analgesics [16, 24, 37]. Acacetaminophen is often used separately or in combination with pseudophedrine, dextromethorphan, diphenhydramine, doyaxilamine, codeine, chlorpheniramine, hydrococode, or oxycodeone [16, 24, 37].

Driver nodes as targets of control from the DrDrN\textsuperscript{gene} to the DiDiN\textsuperscript{gene} are associated with rheumatoid arthritis (OMIM ID: 180300), osteosarcoma (OMIM ID: 259500), bladder cancer (OMIM ID: 109800), Ehlers-
Danlos syndrome (OMIM ID: 130060), colorectal cancer (OMIM ID: 114500), diabetes (OMIM ID: 125853), schizophrenia (OMIM ID: 181500), mycobacterium tuberculosis (OMIM ID: 607948), Crohn disease-associated growth failure and inflammatory bowel disease (OMIM ID: 266600), pulmonary fibrosis, idiopathic (OMIM ID: 178500), lung cancer (OMIM ID: 211980), pulmonary disease, chronic obstructive, severe early-onset (OMIM ID: 606963), gastric cancer (OMIM ID: 137215), meningioma (OMIM ID: 607174), mitochondrial DNA depletion syndrome (OMIM ID: 251880), glaucoma (OMIM ID: 137760), tuberous sclerosis-1 (OMIM ID: 191100), asthma (OMIM ID: 600807), Alzheimer disease (OMIM ID: 104300), hypercholesterolemia (OMIM ID: 143890), amyloidosis, familial visceral (OMIM ID: 105200), migraine with or without aura (OMIM ID: 157300), hypertension (OMIM ID: 145500), and bronchiectasis with or without elevated sweat chloride 1 (OMIM ID: 211400). In addition, driver nodes as targets of control from the DiDiNgene to the DrDrNgene are associated with...
Table 2. Direct control between the drug networks and the disease networks.

| Drug (OMIM ID) | Gene (OMIM ID) | Drug (OMIM ID) | Gene (OMIM ID) | Drug (OMIM ID) | Gene (OMIM ID) | Drug (OMIM ID) | Gene (OMIM ID) |
|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| ADRB1 (DB00475) | DB00171, DB00170, DB00559 | DB00125, DB01159, DB00227 | DB00121, DB00848, DB01221, DB00630, DB00363, DB00171, DB00157, DB00367 | DB00570, DB00193 | DB00170, DB00367 | DB00570, DB00193 | DB00170, DB00367 |
| ADRA1A (DB00227) | DB00121, DB00227, DB00559 | DB00125, DB01159, DB00227 | DB00121, DB00848, DB01221, DB00630, DB00363, DB00171, DB00157, DB00367 | DB00570, DB00193 | DB00170, DB00367 | DB00570, DB00193 | DB00170, DB00367 |
| MTR (DB00212) | DB00121, DB00227, DB00559 | DB00125, DB01159, DB00227 | DB00121, DB00848, DB01221, DB00630, DB00363, DB00171, DB00157, DB00367 | DB00570, DB00193 | DB00170, DB00367 | DB00570, DB00193 | DB00170, DB00367 |
| ADORA2A (DB00384) | DB00121, DB00227, DB00559 | DB00125, DB01159, DB00227 | DB00121, DB00848, DB01221, DB00630, DB00363, DB00171, DB00157, DB00367 | DB00570, DB00193 | DB00170, DB00367 | DB00570, DB00193 | DB00170, DB00367 |

Table 3. Direct control between the drug networks and the gene networks.

| Drug (OMIM ID) | Gene (OMIM ID) | Drug (OMIM ID) | Gene (OMIM ID) | Drug (OMIM ID) | Gene (OMIM ID) | Drug (OMIM ID) | Gene (OMIM ID) |
|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| ADRB1 (DB00475) | DB00171, DB00170, DB00559 | DB00125, DB01159, DB00227 | DB00121, DB00848, DB01221, DB00630, DB00363, DB00171, DB00157, DB00367 | DB00570, DB00193 | DB00170, DB00367 | DB00570, DB00193 | DB00170, DB00367 |
| ADRA1A (DB00227) | DB00121, DB00227, DB00559 | DB00125, DB01159, DB00227 | DB00121, DB00848, DB01221, DB00630, DB00363, DB00171, DB00157, DB00367 | DB00570, DB00193 | DB00170, DB00367 | DB00570, DB00193 | DB00170, DB00367 |
| MTR (DB00212) | DB00121, DB00227, DB00559 | DB00125, DB01159, DB00227 | DB00121, DB00848, DB01221, DB00630, DB00363, DB00171, DB00157, DB00367 | DB00570, DB00193 | DB00170, DB00367 | DB00570, DB00193 | DB00170, DB00367 |
| ADORA2A (DB00384) | DB00121, DB00227, DB00559 | DB00125, DB01159, DB00227 | DB00121, DB00848, DB01221, DB00630, DB00363, DB00171, DB00157, DB00367 | DB00570, DB00193 | DB00170, DB00367 | DB00570, DB00193 | DB00170, DB00367 |

melatonin (DB01065), icosapent (DB00159), lovastatin (DB00227), lidocaine (DB00281), gemcitabine (DB00441), dimethyl sulfoxide (DB01093), vitamin E (DB0163), theophylline (DB00277), hydroxychloroquine (DB01611), colchicine (DB01394), telmisartan (DB00966), halothane (DB01159), biotin (DB00121), levamisole (DB000848), ketamine (DB01221), alendronate (DB00630), clozapine (DB00363), adenosine triphosphate (DB00171), menadione (DB00170), levonorgestrel (DB00367), vinblastine (DB00570), and tramadol (DB00193) [16, 24]. Table 3 shows the results of direct control between the drug networks and the gene networks, and we also can find the results of direct control between the disease networks and the gene networks in table 4. From the tables, we can see that driver nodes of the gene networks as drug target genes/disease-causing genes are closely associated with that of the drug networks/the disease networks respectively. For example, driver genes as targets of control from the Drug gene to the Drug network are associated with cellular response to chemical stimulus of biological process with p-value = 4.254 × 10⁻¹⁰, and proteinaceous extracellular matrix of cellular component with p-value = 2.481 × 10⁻³ [1]. Table 5 shows the gene ontology (GO) enrichment of genes as targets of control from the Drug gene to the Drug network, and the genes are enriched in GO terms [1]. For instance, driver genes as targets of control from the Drug gene to the Drug network are related to behavioral response to cocaine of biological process with p-value = 1.493 × 10⁻⁴, drug binding of molecular function with p-value = 2.495 × 10⁻², and cell body of cellular component with p-value = 2.495 × 10⁻² [1]. Driver genes as targets of control from the Drug gene to the Drug network are associated with cellular response to chemical stimulus of biological process with p-value = 7.341 × 10⁻², small
Table 4. Direct control between the disease networks and the gene networks.

| A               | B               | \( f_{sub}(A \rightarrow B) \) (Genes) | \( f_{sub}(B \rightarrow A) \) (OMIM ID) |
|----------------|----------------|-------------------------------------|----------------------------------------|
| DiDiNgene      | GeGeNdrug      | —                                   | —                                      |
| DiDiNgene      | GeGeNdisease   | —                                   | 166710                                 |
| DiDiNgene      | GeGeNdrug      | MTR, KCNJ11, NOS2, ERBB2, ENPP1    | 157300, 600807, 104300, 143890          |
| DiDiNgene      | GeGeNdisease   | CIITA, JAK2, PAX6, MAPT, NRAS, RET  | 180300, 252150, 607208, 251880          |
| DiDiNgene      | GeGeNdrug      | PIK3CA, NDA1, HBA1, HNF1A, TERT    | 256000, 130860, 145001, 219100          |
| DiDiNgene      | GeGeNdisease   | DRD4, CCL2, TP53, ITGB4, IFNGR1, CYP2A6, TSC1, EPO, COL6A2, MTR | 273300, 125833, 160150, 101600 |
| DiDiNgene      | GeGeNdrug      | FGFR3, FGA, OPTN, BRCA2, GDF5, HTR2A, CDH23, MZP, HFE, SELE | 607174, 137760, 60807, 601626 |
| DiDiNgene      | GeGeNdisease   | EDARADD, SDHB, MTHFR, MLH1, NOD2, KRAS, KRIT1, ARHGAP26 | 104300, 143890, 605074, 608930 |

Table 5. GO enrichment of genes as targets of control from the drug/disease network to the gene networks.

| A               | B               | \( f_{sub}(A \rightarrow B) \) / GO enrichment | \( p\)-value       |
|----------------|----------------|-----------------------------------------------|--------------------|
| DrDrNgene      | GeGeNdrug      | biological process response to chemical       | 4.254 \times 10^{-9} |
| DrDrNgene      | GeGeNdrug      | molecular function ion binding                | 4.138 \times 10^{-4} |
| DrDrNgene      | GeGeNdrug      | cellular component proteinaceous extracellular matrix | 2.481 \times 10^{-3} |
| DrDrNgene      | GeGeNdrug      | biological process behavioral response to cocaine | 1.493 \times 10^{-4} |
| DiDiNgene      | GeGeNdrug      | molecular function drug binding               | 1.838 \times 10^{-4} |
| DiDiNgene      | GeGeNdrug      | cellular component cell body                  | 2.495 \times 10^{-2} |
| DiDiNgene      | GeGeNdrug      | biological process cellular response to chemical stimulus | 7.341 \times 10^{-2} |
| DiDiNgene      | GeGeNdrug      | molecular function small molecule binding     | 5.084 \times 10^{-5} |
| DiDiNgene      | GeGeNdrug      | cellular component ATP-sensitive potassium channel complex | 7.655 \times 10^{-3} |
| DiDiNgene      | GeGeNdrug      | biological process response to stimulus       | 2.379 \times 10^{-9} |
| DiDiNgene      | GeGeNdrug      | molecular function protein binding             | 1.407 \times 10^{-4} |
| DiDiNgene      | GeGeNdrug      | cellular component cell part                  | 7.600 \times 10^{-4} |

Table 6. Essential genes as targets of control from the drug/disease network to the gene networks.

| A               | B               | \( f_{sub}(A \rightarrow B) \) / Essential genes | Percentage       |
|----------------|----------------|-----------------------------------------------|------------------|
| DrDrNgene      | GeGeNdrug      | PDGFRB, NR3C1, ACE, PTG2, MTR, CACNA1B, GSS, HMGCR, FKBP1A, ACHE, SCNSA, SERPIN1 | 37.50% (12/32)  |
| DrDrNgene      | GeGeNdrug      | MTR, CURN, PCCA, ACE                           | 40.00% (4/10)    |
| DiDiNgene      | GeGeNdrug      | MTR, ERBB2                                     | 40.00% (2/5)     |
| DiDiNgene      | GeGeNdrug      | EDARADD, BRCA2, TP53, ITGB4, PAX6, MTR, PIK3CA, MLH1, KRT1, HNF1A, JAK2, RET, NDA1, KRAS, EP300, FGFR3, TSC1 | 43.59% (17/39)  |

molecule binding of molecular function with \( p\)-value = 5.084 \times 10^{-2}, and ATP-sensitive potassium channel complex of cellular component with \( p\)-value = 7.655 \times 10^{-3} [1]. Driver genes as targets of control from the DiDiNgene directly to the GeGeNdisease involve in response to stimulation of cellular process with \( p\)-value = 2.379 \times 10^{-9}, protein binding of molecular function with \( p\)-value = 1.407 \times 10^{-4}, and cell part of cellular component with \( p\)-value = 7.600 \times 10^{-4} [1].

Table 6 shows the essential genes as targets of control from the drug/disease networks to the gene networks. From the table, we can see that driver genes tend to be essential in the control of the gene networks [46]. For instance, 43.59% (17/39) of genes such as EDARADD, BRCA2, TP53, ITGB4, PAX6, MTR, PIK3CA as targets of control from the DiDiNgene to the GeGeNdisease are essential [46]. In addition, we also can see that the driver genes are not housekeeping genes [38].

Figure 6 shows the results of \( \overrightarrow{AB} \) for the DrDrNdisease, the DrDrNgene, the DiDiNdisease, the DiDiNgene, the GeGeNdisease, and the randomized DrDrNdisease, the randomized DrDrNgene, the randomized DiDiNdisease, the randomized DiDiNgene, and the randomized GeGeNdisease based on degree distribution conservation, node degree conservation and E–R model respectively. Just as we have found in the above that the gene networks play an important role in the link of diseases and drugs, e.g. in figure 6(c), the GeGeNdisease highly controls the DrDrNdisease and the DiDiNdisease.
Figure 7 shows the results of $\text{CoContr}^{\text{md}}(A \overset{\text{AB}}{\rightarrow} B)$ and $\text{CoContr}^{\text{net}}(A \overset{\text{AB}}{\leftrightarrow} B)$ for the DrDrN\text{\textsuperscript{disease}}, the DrDrN\text{\textsuperscript{gene}}, the DiDiN\text{\textsuperscript{drug}}, the DiDiN\text{\textsuperscript{gene}}, the GeGeN\text{\textsuperscript{disease}}, and the GeGeN\text{\textsuperscript{drug}}. From the figure, we can find that the results of co-controllability of random networks based on node degree conservation are closer to those of the real networks compared with random networks based on degree distribution conservation and the E–R model (see figures 7(a), (c) and (e), (g)). That is to say, the co-controllability of multiple networks is highly associated with the networks' node degree, which is more stringent than the controllability of a single network that is mainly determined by the network's degree distribution.
Figure 8 shows the results of $\text{Contr}^{\text{mds}}(A \rightarrow B)$, $\text{CoContr}^{\text{mds}}(A \rightarrow B)$, $\text{Contr}^{\text{net}}(A \rightarrow B)$ and $\text{CoContr}^{\text{net}}(A \rightarrow B)$ for the DrDrNgene, the DiDiNgene, the GeGeNdisease and the GeGeNdrug. In figures 8 (a)–(c), we can see that driver nodes exhibit a low tendency to be conserved based on indirect control/co-control compared with direct control/co-control. In figure 8(d), we can find that the ability with which the DrDrNgene indirectly controls the DiDiNgene via the GeGeNdrug is not only stronger than that of the GeGeNdisease, but also is comparable with the direct control from the DrDrNgene to the DiDiNgene. In addition, just as we have found in the above, diseases and drugs can be considered as two different subnetworks of human PPI network. Therefore, we may conclude that drugs tend to dominate diseases by controlling the PPI network, and the coded proteins of disease-related genes exhibit a low tendency to be drug targets for the control of diseases. The results above not
only can help us understand the mechanism between disease phenotypes and herbal formulae, but also discover the effective compounds and their combinations and develop a rational drug design.

In general, network-based approaches give us a new way to understand complex biological systems, and co-controllability of multiple networks as a new powerful tool can help us study the correlation between the control of biological networks and the control of diseases as well as the mechanisms of drug–disease–gene in a network-based framework, which is also helpful for new drug design and disease treatments.

5. Conclusions

This paper proposes a new framework, co-controllability of multiple networks, which stresses the control of one network by another network as well as the mutual control characteristics of multiple networks based on minimum dominating sets. We take a drug–disease–gene network that consists of a drug–drug network, a disease–disease network and a gene–gene network as an example to study the co-controllability of multiple networks. The results show that driver nodes tend to be conserved, e.g. diseases highly associated with driver nodes of the drug–drug network tend to be driver nodes in the disease–disease network compared with random networks. In addition, the co-controllability of multiple networks is probably associated with the networks’ node degree, which is more stringent than the controllability of a single network, which is mainly determined by the network’s degree distribution. We also find that diseases and drugs tend to be mapped as two different subnetworks of human PPI network, the drug–drug network tends to dominate the disease–disease network by controlling the PPI network, and the coded proteins of disease-related genes are hard to be drug targets for the control of diseases. The results in this paper not only play an important role in understanding the co-controllability of multiple networks, but also are helpful for understanding the mechanisms of drug–disease–gene, disease treatments and drug design in a network-based framework. In the future work, we will still focus on the network model and study the co-controllability of drug–disease–gene in different species.

Acknowledgments

The author would like to thank the anonymous reviewers for their valuable comments and suggestions to improve the quality of the paper. This work is supported by China Scholarship Council (CSC) for Studying Abroad, the National Natural Science Foundation of China (grant no. 61202175), the Fundamental Research Funds for the Central Universities (grant no. BDF181417), and the Research Fund for the Doctoral Program of Higher Education of China (grant no. 20120203120015).

References

[1] Ashburner M et al 2000 The gene ontology consortium gene ontology: tool for the unification of biology Nat. Genet. 25 25–9
[2] Albert R and Barabási A-L 2002 Statistical mechanics of complex networks Rev. Mod. Phys. 74 47–97
[3] Barabási A-L and Albert R 1999 Emergence of scaling in random networks Science 286 509–12
[4] Barabási A-L and Oltvai Z N 2004 Network biology: Understanding the cell’s functional organization Nat. Rev. Genet. 5 101–13
[5] Bu D et al 2003 Topological structure analysis of the protein–protein interaction network in budding yeast Nucleic Acids Res. 31 2443–50
[6] Becker K G, Barnes K C, Bright T and Wang A 2004 The genetic association database Nat. Genet. 36 431–2
[7] Bader G D and Hogue C W 2003 An automated method for finding molecular complexes in large protein interaction networks BMC Bioinf. 4 1–17
[8] Chen L, Wang R and Zhang X S 2009 Biomolecular Networks Methods and Applications in Systems Biology (Hoboken, NJ: Wiley)
[9] Davis A et al 2009 Comparative toxicogenomics database: a knowledgebase and discovery tool for chemical–gene–disease networks Nucleic Acids Res. 37 D786–92
[10] Erdős P and Rényi A 1960 On the evolution of random graphs Publ. Math. Inst. Hung. Acad. Sci. 5 17–61
[11] Fortunato S 2010 Community detection in graphs Phys. Rep. 486 75–174
[12] Goh K-I et al 2007 The human disease network Proc. Natl Acad. Sci. USA 104 8685–90
[13] Girvan M and Newman M E J 2002 Community structure in social and biological networks Proc. Natl. Acad. Sci. 99 7821–6
[14] Hahn M W and Kern A D 2009 Comparative genomics of centrality and essentiality in thress eukaryotic protein–interaction networks Mol. Biol. Evol. 22 803–6
[15] Hartwell L H et al 1999 From molecular to modular cell biology Nature 402 C47–52
[16] Hamosh A et al 2005 Online mendelian inheritance in man (OMIM), a knowledgebase of human genes and genetic disorders Nucleic Acids Res. 33 D514–7
[17] Hwang S, Kim E, Yang S, Marcotte E M and Lee I 2014 MORPHIN: a web tool for human disease research by projecting model organism biology onto a human integrated gene network Nucleic Acids Res. 42 W147–53
[18] Idier T and Sharan R 2008 Protein networks in disease Genome Res. 18 644–52
[19] Jeong H et al 2001 Lethality and centrality in protein networks Nature 411 41–2
[20] Junker B H, Koschützki D and Schreiber F 2006 Exploration of biological network centralities with CentiBiN BMC Bioinf. 7 1–19
[21] Kitano H 2002 Computational systems biology Nature 420 206–10
[22] Kutalik Z, Beckmann J S and Bergmann S 2008 A modular approach for integrative analysis of large-scale gene-expression and drug-response data Nat. Biotechnol. 26 531–9
[23] Liu Y Y, Slotine J J and Barabási A L 2011 Controllability of complex networks Nature 473 167–73
[24] McKusick V 2007 Mendelian inheritance in Man and its online version, OMIM Am. J. Hum. Genet. 80 588–604
[25] Newman M E J 2006 Modularity and community structure in networks Proc. Natl Acad. Sci. USA 103 8577–82
[26] Nacher J C and Akutsu T 2012 Dominating scale-free networks with variable scaling exponent: Heterogeneous networks are not difficult to control New J. Phys. 14 073005
[27] Palla G, Derényi I, Farkas I and Vicsek T 2005 Uncovering the overlapping community structure of complex networks in nature and society Nature 435 814–8
[28] Rosvall M and Bergstrom C T 2007 An information-theoretic framework for resolving community structure in complex networks Proc. Natl Acad. Sci. USA 104 7327–31
[29] Radicchi F, Castellano C, Cecconi F, Loreto V and Parisi D 2004 Defining and identifying communities in networks Proc. Natl Acad. Sci. USA 101 2658–63
[30] Reichardt J and Bornholdt S 2006 Statistical mechanics of community detection Phys. Rev. E 74 016110
[31] Smoot M, Ono K, Ruscheinski J, Wang P-L and Ideker T 2011 Cytoscape 2.8: new features for data integration and network visualization Bioinformatics 27 431–2
[32] Strogatz S H 2001 Exploring complex networks Nature 410 268–76
[33] Sun P G 2015 The human drug–disease–gene network Inf. Sci. 306 70–80
[34] Vidal M, Cusick M E and Barabási A-L 2011 Interactome networks and human disease Cell 144 986–98
[35] Wishart D et al 2008 DrugBank: a knowledgebase for drugs, drug actions and drug targets Nucleic Acids Res. 36 D901–6
[36] Wu X, Jiang R, Zhang M Q and Li S 2008 Network-based global inference of human disease genes Mol. Syst. Biol. 189 1–11
[37] Wuchty S 2014 Controllability in protein interaction networks Proc. Natl Acad. Sci. USA 111 7156–60
[38] Zhang A 2009 Protein Interaction Networks Computational Analysis (Cambridge: Cambridge University Press)
[39] Zhang R, Ou H Y and Zhang C T 2004 DEG, a database of essential genes Nucleic Acids Res. 32 D271–2
[40] Zhang Y, Tao C, Jiang G, Nair A A, Su J, Chute C G and Liu H 2014 Network-based analysis reveals distinct association patterns in a semantic MEDLINE-based drug-disease–gene network J. Biomed. Semant. 5 1–13
[41] Bashan A, Bartsch R P, Kantelhardt J W, Havlin S and Ivanov P C 2012 Network physiology reveals relations between network topology and physiological function Nat. Commun. 3 702
[42] Wang R S, Oldham W M and Loscalzo J 2014 Network-based association of hypoxia-responsive genes with cardiovascular diseases New J. Phys. 16 105014