Send Orders for Reprints to reprints@benthamscience.ae

Mini-Reviews in Medicinal Chemistry, 2015, *15*, 000-000

Network based Approach to Drug Discovery: A Mini Review

Peng Li, Yingxue Fu and Yonghua Wang*

Center of Bioinformatics, College of Life Science, Northwest A and F University, Yang ling, Shaanxi, China

Please provide corresponding author(s) photograph

Abstract: With the rapid development of high-throughput genomic technologies and the accumulation of genome-wide datasets for human disease, it has been shown that using only reductionistic principles has been difficult to capture the complex biological networks and design rational medication. However, the emerging paradigm of "network based methodology" proposes to harness the power of networks to uncover relationships between various data types of interest for drug discovery. Recent advances include networks that encompass relationships between drugs, disease-related genes, therapeutic targets and diseases. It is shown how network techniques can help in the investigation of the mechanism of action of existing drugs, new molecules, or to identify novel disease genes and targets. We review how these different types of network analysis approaches facilitate drug discovery and their associated challenges. Some representative examples are reviewed to show that network analysis is a powerful, integrated, computational and experimental approach to improve the drug discovery process.

Keywords: Drug discovery, target identification, network analysis.

INTRODUCTION

In recent years, the prospect for complex disease pharmacotherapy seems to have "hit the wall", with multiple high-profile trial failures and declining industrial interest. Reasons for such predicament might include an intensive regulatory environment, a competitive market, the elevated bar of existing medicines for further innovation and the increasing cost of mega-trials. However, the most important and intrinsic reason comes up to the lack of mechanistic understanding of drug action and the complicated etiologies [1-3].

The strategy behind many modern pharmaceuticals is to restore the healthy state by inhibiting a molecular target that is central to the disease mechanism. Drug discovery efforts are, therefore, crucially dependent on identifying individual molecular targets and validating their relevance to a human disease. This target validation is followed by identification of specific chemical- or antibody-based modulators of the target. As an example, varieties of blockbusters for the therapy for cardiovascular diseases have been sprung up, such as statins, angiotensin-converting enzyme (ACE) inhibitors, anti-platelet agents and beta-blockers. However, many of these drugs play functional roles in biological processes outside the scope of the drug's intended effects [4, 5]. This often leads to unexpected situations at various stages during the drug discovery process. For example, torcetrapib (Pfizer, New York, NY, USA), an inhibitor of cholesteryl ester transfer protein (CETP), failed in the Investigation of Lipid Level Management to Understand Its Impact in Atherosclerotic Events (ILLUMINATE) trial for the increased risk of mortality and morbidity [6], due to the off-target effects of torcetrapib on hypertension [7]. On the contrary, the unpredictable off-target interaction may also give rise to safety effects on patients. For example, statins, originally designed to target elevated lipids for the treatment of atherosclerosis, might also confer cardiovascular benefit with their anti-inflammatory effects, independent of LDL-lowering effects [8].

Indeed, a growing body of post-genomic biology (as reflected for acquisition of high-throughput genomic, transcriptomic, proteomic, and metabolomic data) has been revealing a far more complex portrait of drug actions. It is appreciated that many drugs with a specific efficacy actually act on multiple protein targets [9, 10]. This so-called polypharmacology is an undesirable property in the conventional reductionist paradigm and it might be more suitable to view through the lens of systems-based approaches [10]. Common forms of complex diseases are caused by multiple genetic factors, each of which contributes modestly to the disease risk, and also environmental factors. Generally, it has become evident that many human diseases cannot be attributed to the malfunction of a single gene but to complex interactions among multiple genetic variants. Perturbations in several genes might only make subtle contributions to the susceptibility of a particular individual [11]. Their complexity resists traditional efforts which have been applied to identify a single gene or pathway to treat the disease [12]. Therefore, the disease causations should be studied on the basis of the entire body of knowledge including all genes that are associated with the clinical traits. Accordingly, a systems-based approach integrating all the potential related factors involved in the pathologies and a

*Address correspondence to this author at the Center of Bioinformatics, College of Life Science, Northwest A and F University, Yang ling, Shaanxi, 712100 China; Tel:/Fax: +86-029-87092262;
E-mail: yh_wang@nwsuaf.edu.cn

1389-5575/15 $58.00+.00

disease treatment in a network framework is required to address these complex issues.

Currently, network analysis has been a main branch of computational biology. A network offers a natural abstraction of a set of entities (nodes) and of the relationships (edges) occurring among them. These nodes and edges can have various attributes and annotations. Depending on the nature of the study, nodes can represent genes, proteins, small molecules or any other entities capable of interacting in the system being modeled. Edges connecting these nodes represent the physical interactions, genetic regulatory interactions and higher order relationships such as co-expression or some other shared properties linking the nodes. Edges can have directions, weights and other attributes that provide information about the hierarchy of effects. In this article we review some of the recent advances in the field of network pharmacology, starting with approaches relying on various network methodologies (Fig. **1**). Drugs targeting for diseases that involve various interactions between multiple systems and those with complex etiology will most likely benefit from a network-based approach. Based on the specific pathologies, two systems-based network approaches can be applied to identify novel drugs or targets for a given disease. The so-called "central hit" strategy for targeting disease (e.g., cancer) critical network nodes seeks to disrupt the network, whereas more rigid systems (e.g., cardiovascular disease) may need a "network influence" approach to identify and regulate nodes or edges of multi-tissue pathways for essentially redirecting or restoring biological networks. We will show examples of applications of these methodologies both in explaining how the network analysis can facilitate drug discovery and target identification.

DRUG-DRUG NETWORKS

In the drug-drug network, nodes are drugs, and two drugs are connected to each other based on their similarity profiles including chemical and structural properties, and also molecular and phenotypic profiles. Drug-drug network can be used to understand the relationships between different drug compounds and connect them to potential targets and diseases based on their similarities to other compounds in the network. The rational basis for drug-drug network approaches is rooted in that the known quantitative or qualitative profiles model the effect of drug compounds in a biological system. Then, the degrees of similarity that exist can be exploited using computational approaches for drug targeting and repositioning.

When drug compounds are integrated into a network of relationships based on chemical and structural similarities, drug action can be inferred by simple chemical characteristics. This approach typically incorporates quantitative structure activity relationship (QSAR) data with physiochemical and structural properties of biological targets. The computational basis of chemical similarity approaches is to extract a set of chemical fingerprints as features for each drug, and then to relate the drugs directly to each other by clustering or constructing networks based on the extracted features [13]. Machine learning classifiers such as self-organizing maps, an unsupervised-machine learning technique, have been used to compare candidate ligands with the known drugs of a target protein to find new compounds [14, 15]. Recently, a notable advance which used network analysis to predict polypharmacology was the development of the similarity ensemble approach (SEA) [16]. In this approach, drug targets were represented by a set of ligands which were known to bind to the targets. To evaluate a novel query compound for a specific drug target, using the Tanimoto coefficient to calculate the pair-wise similarity score to represent the chemical similarity between two compounds, a raw score was derived by calculating the overall chemical similarity between the query compounds and each member of the set of ligands binding to the target. A statistical model based on the extreme value distribution was then used to normalize the sum of similarity scores to define a significant score, and scores surpassing the significance threshold indicated the query drug is a candidate ligand of the target. Many predictions by the SEA have been experimentally confirmed [16-18]. In addition, considerable efforts have been made to develop 3D molecular similarity methods [19, 20].

Another promising approach is to use phenotypic information to relate drugs to each other. Compared to the chemical property based network approach described above, phenotype based drug-drug network provides a representation of resultant physiological consequences of drug compound's biological activity. Phenotypes were divided into two types based on genome-wide screening: a high-dimensional intermediate phenotype including a gene expression profile, and a low-dimensional endpoint phenotype such as a side effect or cell-growth rate. These phenotypes can be treated as molecular signature of drugs or as a feature vector in the language of machine learning to classify different drugs [21].

Recently, Campillos *et al.* proposed a systematic method to relate drugs to each other based on the assumption that drugs with similar side-effects are likely to interact with similar target proteins [22]. For each drug, they extracted side effects from the drug package labels and translated them into standard vocabularies using the Unified Medical Language System (UMLS) Metathesaurus [23]. The side-effect information was weighted by using a scheme that integrated their frequency and correlation across all drugs in the set, and similarity scores of pairwise drugs were computed based on the sum of the respective weights of their overlapping side-effects. A randomization approach was used to establish the significance of the side effect similarity scores, which were further incorporated with a measure of the structural similarity between drugs to increase predictive power. The authors used this approach to calculate drug–drug similarity coefficients based on their shared side-effect profiles and used these as edges to construct a network. These phenotypic side-effect similarities helped the authors infer cases where two drugs were likely to act through common targets. The resulting drug–drug relationships were shown to recapitulate many shared target relationships between drugs, and several predicted novel drug–target relationships were experimentally confirmed.

Moreover, a new way developed to inferring drug-drug network is through the extent of their similarity in inducing cellular gene expression. In this drug-drug network, each

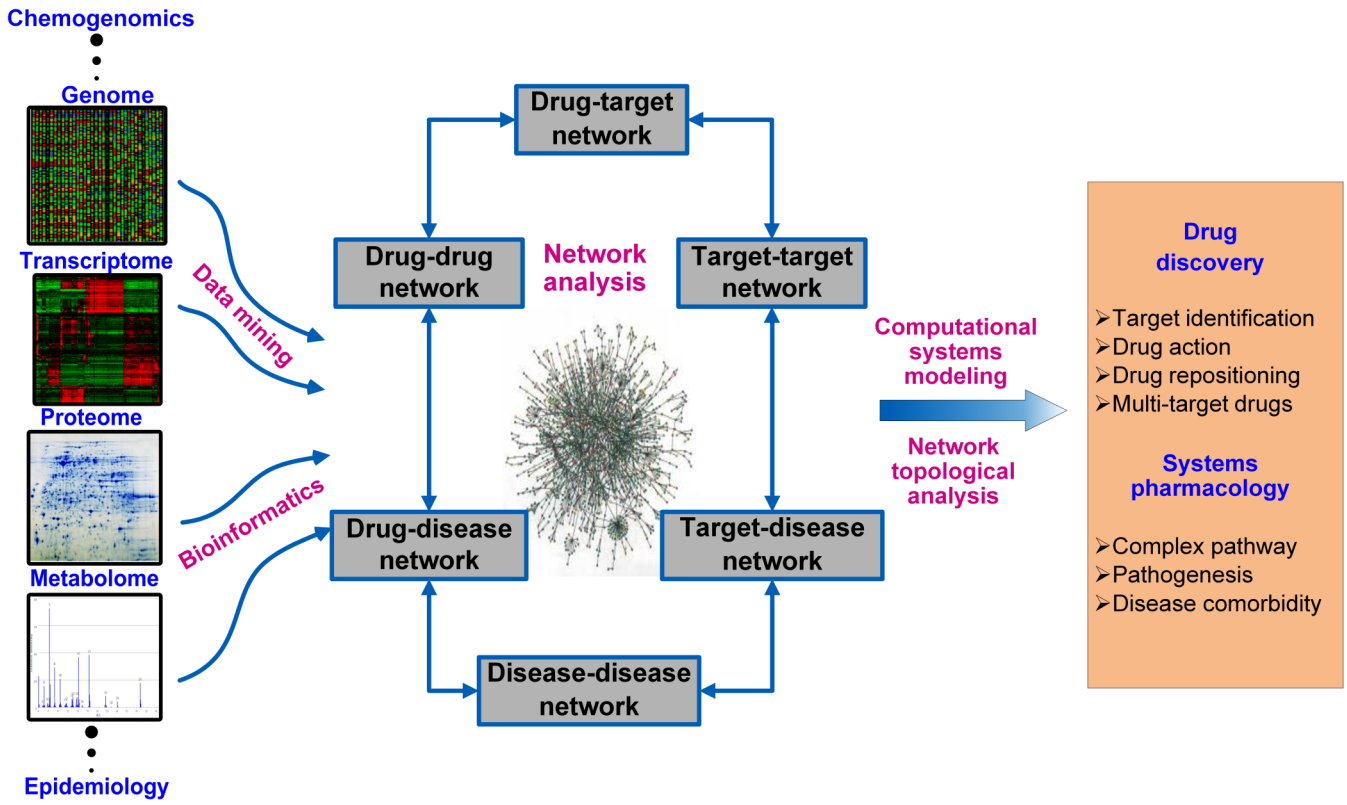

Fig. (1). Network analysis serves an integral role in systems approaches to drug discovery. High-throughput "OMIC" data acquisition from multiple levels of chemical and biological complexity can be integrated by network analysis. Interaction networks provide a global template for computational and mathematical systems modeling, simulation, and prediction. Nodes represent genes, proteins, small molecules or any other entities capable of interacting in the system being modeled. Edges connecting these nodes represent the physical interactions, genetic regulatory interactions and higher order relationships such as co-expression or some other shared properties linking the nodes. These networks can be asscіated with each other. For example, Starting from this graph, it generates two biologically relevant network projections: the target–target network and the drug-drug network. Network topological parameters also provide foci and targets for hypothesis generation and experimental testing. Together, network-based approaches facilitate efforts in drug discovery and systems pharmacology. Some specific applications of these various network-based systems approaches are outlined in the text.

edge represents a significant similar gene expression profile between two drug nodes. Each drug or disease may be considered as an indicator inducing an array of specific gene expression changes in a cell [24]. The usefulness of a drug-drug network is based on the hypothesis that although the precise mechanism of action is not well-understood for many approved compounds, high-throughput molecular expression changes can be used to represent a signature of drug effect in a biological system.

One of the most comprehensive and systematic data resource that enables network analysis is the Connectivity Map (CMap) project [25], the first large installment of gene expression profiles following drug treatment on five human cancer cell lines with 309 small molecules. The molecular activity profile of each drug in the set was ordered into a ranked list according to expression changes after exposure to the drug compound. These profiles can be used as the basis of comparison to connect drugs based on shared gene expression profiles in the CMap. Using the CMap data, Iorio *et al.* first generated a drug-drug network using statistically significant transcriptional pairwise similarities between drugs as the edges [26]. The pairwise similarity between the gene expression profiles following drug treatment computed using a novel, "drug distance metric" based on Gene Set Enrichmen Analysis (GSEA) [27]. Drugs were then organized into a network using the resulting similarity scores, and analyzed with graph-theoretic tools to identify coherent "communities" of drugs consisting of groups of densely interconnected drug nodes. The resulting drug communities were indeed enriched for drugs sharing with similar mechanism of actions and therapeutic application. The authors used this network to classify both known and novel HSP90 inhibitors and CKD2 inhibitors, and also predict and experimentally validate previously unknown cellular autophagy activity for the rho-kinase inhibitor Fasudil, a known ROCK inhibitor approved in Japan against blood vessel obstruction.

Taken together, these results show the ability of drug networks in identifying novel applications for existing drugs, as well as to characterize novel molecules by looking at the known properties of their connected neighboring compounds.

DRUG-TARGET INTERACTION NETWORKS

In the framework of network-based drug discovery, at the most basic level is the drug-target network, in which a drug and a protein are connected to each other if the protein is a known target of the drug. Such a network has been shown to form a bipartite graph with some natural characteristic features of its network topology. Based on the drug-target network, one can generate two biologically relevant network projections: the drug-drug network is typically constructed by linking two drugs if they share a number of targets or target-target network which comprises links between targets that are targeted by the same set of drugs. Yildirim *et al.* applied network analysis to FDA approved drugs and drug targets in the drug–target network revealing a rich network of polypharmacology interactions between drugs and their targets [9].

Detection of compound–protein interactions with drug-drug network is based on the hypothesis that similar ligands are likely to interact with similar proteins. However, in the drug-target network, the prediction can be performed based on drug space, target space and the topology of drug–target interactions. A straightforward statistical approach is to use binary classification methods where compound–protein pairs are trained as an input for machine learning classifiers including support vector machine (SVM) and neural network or the other statistical approach such as the distance learning in the framework of supervised bipartite graph inference [28-30].

As one example of this utility, Yamanishi *et al.* [30] predicted drug–target interactions from the integration of chemical structure information, genomic sequence data and known drug–target network topology simultaneously on a large scale. The authors constructed the drug–target interaction network for each protein class using a bipartite graph representation and further investigated the correlations between the drug structure similarity, the target sequence similarity and the drug–target interaction network topology. Then the network was analyzed using a new supervised bipartite graph learning method to infer unknown drug–target interactions by integrating chemical space and genomic space into a unified "pharmacological space" space. In the method, chemical space represents the chemical structure similarity space of chemical compounds, genomic space represents the protein sequence similarity space of different proteins and pharmacological space represents the interaction space reflecting the drug–target interaction network, where interacting drugs and target proteins are close to each other. The authors predicted novel drug–target interactions for four protein classes involving enzymes, ion channels, GPCRs and nuclear receptors. Moreover, the authors have included pharmacological effect of drug compounds into the computational model [31]. In this study, not only chemical and genomic data but also pharmacological data were used to predict unknown drug–target interactions on a large scale. Pharmacological effects of given compounds were firstly predicted by their chemical structures. Then, unknown drug–target interactions were identified based on the pharmacological effect similarity in the framework of supervised bipartite graph inference. The authors found that drug–target interactions are more correlated with pharmacological effect similarity than with chemical structure similarity.

Taken together, more integrative methods developed taking into account not only drug chemical structures and target protein sequences but also the available known drug–target network information in the drug-target network. Moreover, during the past few years, the community developed many algorithms to specifically compare drug binding regions, which are now widely used in system-level applications in pharmacology [32].

TARGET-TARGET NETWORKS

Here, target-target networks, also named biological networks or interactome mainly define the networks that can combine genome scale datasets with information about specific genes and proteins. In recent years, most attention has been directed towards protein interaction networks, where nodes are proteins that are linked to each other by physical (binding) interactions [33, 34]; metabolic networks, in which nodes are metabolites that are linked if they participate in the same biochemical reactions [35-37]; protein signaling networks, in which links represent the impact of transcription factors and other signaling molecules on downstream events including changes in gene expression and post-translational states of proteins [38]. They have provided insight into the origins of overall cellular behaviors and evolutionary design principles, as well as more focused fields of study concerning specific cell biological processes or diseases. With advances in network analysis tools and the advent of systems biology approaches, one can devise experimentally testable hypotheses ranging from prediction of novel functions of specific genes to genome scale properties of human cellular networks. In a similar manner, analysis of networks for pharmacologic studies provides better decision making for therapeutic interventions.

Recently, a series of increasingly sophisticated network-based tools have been developed to predict potential disease genes, including linkage methods, disease module-based methods and diffusion-based methods [39].

Linkage methods assume that the direct interaction partners of a disease protein are likely to be associated with the same disease phenotype. Oti *et al.* used this approach in the protein-protein network to predict disease genes for genetically heterogeneous hereditary diseases [40]. The authors hypothesized if disease proteins had interaction partners which were located within other loci associated with that same disease; such interaction partners were considered to be candidate disease genes. The results showed that exploiting protein–protein interactions can greatly increase the likelihood of finding positional candidate disease genes.

Disease module-based methods assume that the cellular components that belong to the same topological, functional or disease module have a high likelihood of being involved in the same disease [12, 41]. These methods start with identifying the disease modules and inspecting their members as potential disease genes. Lage *et al.* integrated quality-controlled interactions of human proteins with a validated, computationally derived phenotype similarity

score, to identify previously unknown complexes likely to be associated with diseases [12]. In the first step there is a virtual pull down of all the candidate proteins involved in a particular disorder. Each complex is named as the candidate complex. Second, proteins involved in disorders are identified from the candidate complex, and the pairwise similarity is determined by text mining. In the concluding step, scoring and ranking the candidates were carried out by the Bayesian predictor based on phenotypes associated with the proteins in the candidate complex.

Diffusion-based methods find the paths that are closest to the known disease genes. In these algorithms, "random walkers" are "released" and are allowed to diffuse along the links of the interactome, moving to any neighbouring node with equal probability. In this way, one can identify the nodes and links that are closest to the known disease genes, as they will be those that are most often visited by the random walkers. Proteins that interact with several disease proteins will gain a high probabilistic weight, as well as those that may not directly interact with any disease proteins but are in close network proximity to them. Vanunu *et al.* proposed a network-based method for prioritizing disease genes and inferring protein complex associations based on formulating constraints on the prioritization function that relate to its smoothness over the network and usage of prior information [42]. This method was applied to study prostate cancer, Alzheimer and type 2 diabetes mellitus and found several novel causal genes and protein complexes for further investigation. On the same data set, diffusion-based methods have the best predictive performance [43].

In addition to the above methods that utilize the information that is encoded in the network topology as well as the placement of the known disease genes, recent studies have also included gene transcription changes and 3D structural interactions in the context of PPI networks [44, 45]. The study of signaling networks also have potential to enhance our understanding of drug's mode of action. There are multiple approaches to construct a mathematical model of a signaling pathway with different level including logic-based models, Bayesian network and mass action model to study the effects of drug treatments and complex behaviors such as the synergistic combinations of drugs [46, 47].

Genome-scale metabolic networks are designed to model metabolism at a cellular level. Flux Balance Analysis (FBA) is a constraint-based modeling approach that is suitable for modeling cellular metabolism [48]. FBA uses stoichiometric constraints and rates of extracellular metabolite uptake and production as input. In FBA, the metabolic network is modeled as a system in a pseudo-steady state as it assumes that the growth rate of a cell is constant. FBA does not require prior knowledge of enzyme kinetics and concentrations of intracellular metabolites. These models provide a framework for computationally interpreting rates of intracellular reactions. Recently, Li *et al.* have utilized the FBA method for the prediction of novel targets in the host pathogen network [49]. Folger *et al.* further used a human genome-scale metabolic network to model cancer metabolism [50]. This model was used to predict new cytostatic drug targets that inhibit cancerous cell growth. The authors found that a total of 40% of the proteins, identified by this approach, were known anticancer drug targets thus demonstrating a potential application for drug discovery. In addition, the model has been successfully used for discovering combinations of synthetic lethal drug targets.

DRUG-DISEASE NETWORKS

There are mainly two types of drug-disease networks: one network is to utilize knowledge of drug indications for disease to compose a directed graph connecting drugs to their indications, and the other is the network in which each node represents a drug (or disease) and each edge represents a significant similarity or "anti-similarity" in cellular expression profiles between drug and disease nodes.

For example, Gottlieb *et al.* proposed a computational approach PREDICT that utilizes multiple drug–drug and disease–disease similarity measures to directly predict novel drug-disease associations for both FDA approved drugs and experimental molecules on a large scale [51]. Their approach was designed in three phases. Firstly, authors constructed five drug-drug similarity measures and two disease-disease similarity measures. Then, they analyzed these similarity measures to build classification features and subsequent learn classification rule that can distinguish between true and false drug-disease associations by using these similarity measures. And finally authors applied a logistic regression classifier to predict any new possible drug-disease associations. Thus for a given drug-disease association from the gold standard (experimentally curated list of drug-disease interactions), the authors computed an association score by considering all the other known drug-disease association. In addition, Nacher *et al.* constructed a bipartite graph in which drug nodes are connected to their therapeutic indications [43]. The authors generated two biologically relevant network projections: a network of drugs connected if they are used for similar indications and a network of diseases connected if they are treated with the same drugs. Based on these networks, the authors found diseases clustered based on their treatments as well as treatments clustered based on the diseases. Authors calculated the topological importance of the drugs in the network and found that certain drugs, with multiple targets, are used to treat distinct diseases in different parts of the network.

Based on the cellular expression profiles, drug-disease networks rely on direct comparison between the molecular activity signatures of drugs with those of a disease state. If the gene expression profile of a drug is significantly anti-similar with that characterizing a disease state, it can be hypothesized that the drug could "revert" the disease state, hence the disease phenotype [52]. Hu and Agarwal [53] created a drug-disease network using publicly available gene expression profiles, and integrated this network with molecular profiles and knowledge of drugs and drug targets to infer novel drugs for diseases, whose relationships were not been previously known. They began by extracting disease-relevant expression data sets from the NCBI Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/) and computed differential gene expression profiles to find drugs and diseases that clustered together by their gene expression profiles. Similar efforts using the text mining and microarray

data involved in drugs and diseases have also been successful predict drug-disease associations in a more contextualized view that is provided by network biology [54].

TARGET-DISEASE NETWORKS

The target-disease network is a bipartite graph that connects diseases and their therapeutic targets. This network helps not only understand the similarity and difference in treating different diseases but also explore the potential therapeutic effects for drug compounds of the certain targets. Our group has built the connection between the ingredients of various Chinese medicines and their related diseases via the target–disease network [55-58]. For example, Li *et al.* [57] collected potential targets of compounds in Compound Danshen Formula (CDF), a widely used Traditional Chinese Medicine (TCM) applied in clinical treatment of cardiovascular diseases, from the PharmGkb, TTD and DrugBank databases, and the obtained disease-target interactions were used to build the target-disease Network. Based on the assumption that certain drugs acting on same protein associated with different diseases in a network may cause different diseases, the authors further applied the network to find some novel therapeutic effects of CDF.

With similar process, we can also construct gene-disease network consisting of two disjoint sets of nodes: one set corresponds to disorders and the other set corresponds to all disease-related genes in the human genome. The most complete and best-curated list of known phenotype-gene associations is maintained in the Morbid Map (MM) of the Online Mendelian Inheritance in Man (OMIM) [59]. Each entry of the MM is composed of four fields, the name of the disorder, the associated gene symbols, its corresponding OMIM id, and the chromosomal location. Each entry of the MM is composed of four fields, the name of the disorder, the associated gene symbols, its corresponding OMIM id, and the chromosomal location. Goh et al analyzed the complete data set available in the OMIM database and used the phenotypic information and disease associated genes to construct a gene-disease network [60]. Starting from the diseasome bipartite graph, authors generated two biologically relevant network projections the "human disease network" nodes represent disorders, and two disorders are connected to each other if they share gene in which mutations are associated with both disorders. In the "disease gene network" nodes represent disease genes, and two genes are connected if they are associated with the same disorder. In this study, disease genes were found to show significant functional clustering in the studied network suggesting the existence of disorder-specific functional modules. Furthermore, Anna *et al.*, included mendelian, complex and environmental diseases in an integrated gene-disease network and showed that the concept of modularity applies for all of these diseases [61].

In addition, gene–disease network can also be applied for drug repositioning based on genome-wide association studies (GWAS; http://www.genome.gov/gwastudies) [62].Sanseau et al compiled GWAS gene–disease associations and found out the "druggable" targets by small molecules in these genes, thereby connecting drugs with GWAS diseases. The authors hypothesized that if drugs have the same or closely related indication to the GWAS traits; it increases confidence in the pursued indication. Conversely, drug with indications different from the GWAS traits represent drug repositioning opportunities.

DISEASE-DISEASE NETWORKS

During these years, huge efforts have been devoted to the use of networks (disease network) to integrate different genetic, proteomic, metabolic and phenotypic datasets to elucidate the entangled origins of many diseases [53, 63, 64]. The disease-disease network is constructed by connecting two disorders if they have similar treatment or molecular profiles. The systematic identification of such network-based dependencies among cardiovascular disorders offers a sufficient resolution and specificity for etiologic heterogeneity and clinical treatment of diseases. Indeed, huge efforts have been devoted to the use of disease networks (diseasome) to integrate different genetic, proteomic, metabolic and phenotypic datasets to elucidate the entangled links of diseases [65, 66]. Uncovering such links between diseases could help understand how and why different disorders are linked at the molecular level and also aid drug discovery, in particular when it comes to the use of approved drugs to treat molecularly linked diseases.

Chiang *et al.* performed a network-based, guilt by association approach to discover novel drug indications based on the shared treatment profile from any disease pairs The authors built 5,549 disease-disease associations if two diseases shared at least one FDA approved drug in common. For each disease pair, those drugs that were used against only one of the two may also be therapeutic for the other disease. Novel drug-indication associations could be inferred by associating drugs with novel indications by expanding from simple pairs into network clusters. Finally, by using a guilt-by-association approach, the authors generated approximately 57,000 robust novel drug uses, of which, a number of novel drug uses were found to be highly enriched in clinical trials, indicating external validation of these predictions.

Goh *et al.* generated disease network by connecting two disorders if they are associated with the same genes and demonstrated the existence of these widespread molecular connections between linked diseases. Furthermore, Park *et al.* found statistically significant correlations between the underlying structure of cellular networks and disease comorbidity patterns in the human population [66]. The authors showed that two disease pairs in which the network-based information offers a plausible mechanism for statistically significant comorbidity patterns. In addition, Lee *et al.* demonstrated that metabolic diseases can be also organized in a metabolic disease network if the enzymes and their associated diseases are linked through metabolic pathways [64]. This study found that metabolic diseases connected through shared pathways tend to show significant comorbidity.

Similarly, disease networks can also be built based on the cellular expression profiles. Gene coexpression analyses have identified shared candidate genes that associate normal

myocardial development to myocardial hypertrophy and failure [68], and revealed novel molecular connections between Alzheimer disease and cardiovascular diseases [69]. More recently, Rende *et al.* constructed a cardiovascular disease functional linkage network that carries significant interconnections among modules representing cardiovascular diseases with other complex disorders such as infection by Listeria monocytogenes, myasthenia gravis, hemorrhagic diatheses, and protein S deficiency [70].

CONCLUSION

Network approaches allow biomedical researchers to rapidly organize current knowledge by integrating different types of large datasets to systematically descript drug action, identify novel medications and understand complex diseases. In this article, we have highlighted a variety of fundamental network analysis approaches, which are being used to facilitate drug discovery. One point should be emphasized is that these types of networks are not independent but are interrelated to each other. The integrated network analyses are critical to systems biology and pharmacology to uncover previously unknown relationships. For example, Hu *et al.* has reported a integrative network platform named VisANT that allows users to construct 11 different types of networks based on the disease and therapy hierarchy, disease–gene and therapy–drug associations to analyze the correlations between disease, therapy, genes and drugs systematically [71]. In addition, there are still large gaps in our knowledge that prevent us from reconstructing the ultimate drug discovery network. The fidelity of network is limited by the lack of sufficient and high-quality data. There are still many other factors such as environmental stress, epigenetic modifications and invasion of pathogens also contribute to diseases. Incorporating these factors will further improve the coverage and significance of the networks [72]. In addition to the static network analysis, the dynamic networks are more important, should be better integrated in the follow-up studies [73]. As methodologies evolve, the systems network methodology is believed to provide a complete picture that allows us to appreciate the networked nature of human diseases, to design new pharmacological models and then to guide the experiments to new drug discovery and disease treatment.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflicts of interest.

ACKNOWLEDGEMENTS

This work is supported by grants from Northwest A & F University, National Natural Science foundation of China (11201049 and 31170796). It was partially supported by China Academy of Chinese Medical Sciences (ZZ0608), and National 973 Program of China (2013CB531805).

REFERENCES

[1] Csermely, P.; Agoston, V.; Pongor, S. The efficiency of multi-target drugs: the network approach might help drug design. *Trends Pharmacol. Sci.*, **2005**, *26*(4), 178-182.

[2] Korcsmáros, T.; Szalay, M.S.; Böde, C.; Kovács, I.A.; Csermely, P. How to design multi-target drugs: target search options in cellular networks. *Expert Opin. Drug Discov.*, **2007**, *2*(6), 799-808.

[3] Sams-Dodd, F. Target-based drug discovery: is something wrong? *Drug Discov. Today*, **2005**, *10*(2), 139-147.

[4] Zhao, S.; Iyengar, R. Systems pharmacology: network analysis to identify multiscale mechanisms of drug action. *Annu. Rev. Pharmacol. Toxicol.*, **2012**, *52*, 505-521.

[5] Xie, L.; Xie, L.; Kinnings, S.L.; Bourne, P.E. Novel computational approaches to polypharmacology as a means to define responses to individual drugs. *Annu. Rev. Pharmacol. Toxicol.*, **2012**, *52*, 361-379.

[6] Barter, P.J.; Caulfield, M.; Eriksson, M.; Grundy, S.M.; Kastelein, J.J.; Komajda, M.; Lopez-Sendon, J.; Mosca, L.; Tardif, J.-C.; Waters, D.D. Effects of torcetrapib in patients at high risk for coronary events. *N. Engl. J. Med.*, **2007**, *357*(21), 2109-2122.

[7] Forrest, M.J.; Bloomfield, D.; Briscoe, R.J.; Brown, P.; Cumiskey, A.M.; Ehrhart, J.; Hershey, J.; Keller, W.; Ma, X.; McPherson, H. Torcetrapib-induced blood pressure elevation is independent of CETP inhibition and is accompanied by increased circulating levels of aldosterone. *Br. J. Pharmacol.*, **2008**, *154*(7), 1465-1473.

[8] Schönbeck, U.; Libby, P. Inflammation, immunity, and HMG-CoA reductase inhibitors statins as antiinflammatory agents? *Circulation*, **2004**, *109*(21 Suppl 1), II18-II26.

[9] Yıldırım, M.A.; Goh, K.I.; Cusick, M.E.; Barabási, A.L.; Vidal, M. Drug-target network. *Nat Biotechnol.*, **2007**, *25*(10), 1119-1126.

[10] Chan, S.Y.; Loscalzo, J. The emerging paradigm of network medicine in the study of human disease. *Circ. Res.*, **2012**, *111*(3), 359-374.

[11] Ideker, T.; Galitski, T.; Hood, L. A new approach to decoding life: systems biology. Annu. Rev. *Genomics Hum. Genet.*, **2001**, *2*(1), 343-372.

[12] Lage, K.; Karlberg, E.O.; Størling, Z.M.; Olason, P.I.; Pedersen, A.G.; Rigina, O.; Hinsby, A.M.; Tümer, Z.; Pociot, F.; Tommerup, N. A human phenome-interactome network of protein complexes implicated in genetic disorders. *Nat Biotechnol*., **2007**, *25*(3), 309-316.

[13] Eckert, H.; Bajorath, J. Molecular similarity analysis in virtual screening: foundations, limitations and novel approaches. *Drug Discov. Today*, **2007**, *12*(5), 225-233.

[14] Schneider, P.; Tanrikulu, Y.; Schneider, G. Self-organizing maps in drug discovery: compound library design, scaffold-hopping, repurposing. *Curr. Med. Chem.*, **2009**, *16*(3), 258-266.

[15] Byvatov, E.; Fechner, U.; Sadowski, J.; Schneider, G. Comparison of support vector machine and artificial neural network systems for drug/nondrug classification. *J. Chem. Inf. Comput. Sci.*, **2003**, *43*(6), 1882-1889.

[16] Keiser, M.J.; Roth, B.L.; Armbruster, B.N.; Ernsberger, P.; Irwin, J.J.; Shoichet, B.K. Relating protein pharmacology by ligand chemistry. *Nat Biotechnol.*, **2007**, *25*(2), 197-206.

[17] Keiser, M.J.; Setola, V.; Irwin, J.J.; Laggner, C.; Abbas, A.I.; Hufeisen, S.J.; Jensen, N.H.; Kuijer, M.B.; Matos, R.C.; Tran, T.B. Predicting new molecular targets for known drugs. *Nature*, **2009**, *462*(7270), 175-181.

[18] Lounkine, E.; Keiser, M.J.; Whitebread, S.; Mikhailov, D.; Hamon, J.; Jenkins, J.L.; Lavan, P.; Weber, E.; Doak, A.K.; Côté, S. Large-scale prediction and testing of drug activity on side-effect targets. *Nature*, **2012**, *486*(7403), 361-367.

[19] Jenkins, J.L.; Glick, M.; Davies, J.W. A 3D similarity method for scaffold hopping from known drugs or natural ligands to new chemotypes. *J Med Chem.*, **2004**, *47*(25), 6144-6159.

[20] Jain, A.N. Morphological similarity: a 3D molecular similarity method correlated with protein-ligand recognition. *J. Comput. Aided Mol. Des.*, **2000**, *14*(2), 199-213.

[21] Li, Y.; Tesson, B.M.; Churchill, G.A.; Jansen, R.C. Critical reasoning on causal inference in genome-wide linkage and association studies. *Trends Genet.*, **2010**, *26*(12), 493-498.

[22] Campillos, M.; Kuhn, M.; Gavin, A.-C.; Jensen, L.J.; Bork, P. Drug target identification using side-effect similarity. *Science*, **2008**, *321*(5886), 263-266.

[23] Bodenreider, O. The unified medical language system (UMLS): integrating biomedical terminology. *Nucleic Acids Res.*, **2004**, *32*(suppl 1), D267-D270.

[24] Iorio, F.; Rittman, T.; Ge, H.; Menden, M.; Saez-Rodriguez, J. Transcriptional data: a new gateway to drug repositioning? *Drug Discov. Today*, **2013**, *18*(7), 350-357.

[25] Lamb, J.; Crawford, E.D.; Peck, D.; Modell, J.W.; Blat, I.C.; Wrobel, M.J.; Lerner, J.; Brunet, J.P.; Subramanian, A.; Ross, K.N. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science*, **2006**, *313*(5795), 1929-1935.

[26] Iorio, F.; Bosotti, R.; Scacheri, E.; Belcastro, V.; Mithbaokar, P.; Ferriero, R.; Murino, L.; Tagliaferri, R.; Brunetti-Pierri, N.; Isacchi, A. Discovery of drug mode of action and drug repositioning from transcriptional responses. *Proc. Natl. Acad. Sci. U S A*, **2010**, *107*(33), 14621-14626.

[27] Subramanian, A.; Tamayo, P.; Mootha, V.K.; Mukherjee, S.; Ebert, B.L.; Gillette, M.A.; Paulovich, A.; Pomeroy, S.L.; Golub, T.R.; Lander, E.S. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*, **2005**, *102*(43), 15545-15550.

[28] Bleakley, K.; Yamanishi, Y. Supervised prediction of drug–target interactions using bipartite local models. *Bioinformatics*, **2009**, *25*(18), 2397-2403.

[29] Jacob, L.; Vert, J.P. Protein-ligand interaction prediction: an improved chemogenomics approach. *Bioinformatics*, **2008**, *24*(19), 2149-2156.

[30] Yamanishi, Y.; Araki, M.; Gutteridge, A.; Honda, W.; Kanehisa, M. Prediction of drug–target interaction networks from the integration of chemical and genomic spaces. *Bioinformatics*, **2008**, *24*(13), i232-i240.

[31] Yamanishi, Y.; Kotera, M.; Kanehisa, M.; Goto, S. Drug-target interaction prediction from chemical, genomic and pharmacological data in an integrated framework. *Bioinformatics*, **2010**, *26*(12), i246-i254.

[32] Konc, J.; Janežič, D. Binding site comparison for function prediction and pharmaceutical discovery. *Curr. Opin. Struct. Biol.*, **2014**, *25*, 34-39.

[33] Rual, J.F.; Venkatesan, K.; Hao, T.; Hirozane-Kishikawa, T.; Dricot, A.; Li, N.; Berriz, G.F.; Gibbons, F.D.; Dreze, M.; Ayivi-Guedehoussou, N. Towards a proteome-scale map of the human protein–protein interaction network. *Nature*, **2005**, *437*(7062), 1173-1178.

[34] Stelzl, U.; Worm, U.; Lalowski, M.; Haenig, C.; Brembeck, F.H.; Goehler, H.; Stroedicke, M.; Zenkner, M.; Schoenherr, A.; Koeppen, S. A human protein-protein interaction network: a resource for annotating the proteome. *Cell*, **2005**, *122*(6), 957-968.

[35] Jeong, H.; Tombor, B.; Albert, R.; Oltvai, Z.N.; Barabási, A.L. The large-scale organization of metabolic networks. *Nature*, **2000**, *407*(6804), 651-654.

[36] Fell, D.A.; Wagner, A. The small world of metabolism. *Nat. Biotechnol.*, **2000**, *18*(11), 1121-1122.

[37] Duarte, N.C.; Becker, S.A.; Jamshidi, N.; Thiele, I.; Mo, M.L.; Vo, T.D.; Srivas, R.; Palsson, B.Ø. Global reconstruction of the human metabolic network based on genomic and bibliomic data. *Proc. Natl. Acad. Sci. U S A*., **2007**, *104*(6), 1777-1782.

[38] Pieroni, E.; de la Fuente van Bentem, S.; Mancosu, G.; Capobianco, E.; Hirt, H.; de la Fuente, A. Protein networking: insights into global functional organization of proteomes. *Proteomics*, **2008**, *8*(4), 799-816.

[39] Barabási, A.L.; Gulbahce, N.; Loscalzo, J. Network medicine: a network-based approach to human disease. *Nat. Rev. Genet.*, **2011**, *12*(1), 56-68.

[40] Oti, M.; Snel, B.; Huynen, M.A.; Brunner, H.G. Predicting disease genes using protein–protein interactions. *J. Med. Genet.*, **2006**, *43*(8), 691-698.

[41] Navlakha, S.; Kingsford, C. The power of protein interaction networks for associating genes with diseases. *Bioinformatics*, **2010**, *26*(8), 1057-1063.

[42] Vanunu, O.; Magger, O.; Ruppin, E.; Shlomi, T.; Sharan, R. Associating genes and protein complexes with disease via network propagation. *PLoS Comput. Biol.*, **2010**, *6*(1), e1000641.

[43] Nacher, J.C.; Schwartz, J.M. A global view of drug-therapy interactions. *BMC Pharmacol.*, **2008**, *8*(1), 5.

[44] Kotlyar, M.; Fortney, K.; Jurisica, I. Network-based characterization of drug-regulated genes, drug targets, and toxicity. *Methods*, **2012**, *57*(4), 499-507.

[45] Wang, X.; Wei, X.; Thijssen, B.; Das, J.; Lipkin, S.M.; Yu, H. Three-dimensional reconstruction of protein networks provides insight into human genetic disease. *Nat. Biotechnol.*, **2012**, *30*(2), 159-164.

[46] Al-Lazikani, B.; Banerji, U.; Workman, P. Combinatorial drug therapy for cancer in the post-genomic era. *Nat. Biotechnol.*, **2012**, *30*(7), 679-692.

[47] Iorio, F.; Saez-Rodriguez, J.; Di Bernardo, D. Network based elucidation of drug response: from modulators to targets. *BMC Syst. Biol.*, **2013**, *7*(1), 139.

[48] Price, N.D.; Papin, J.A.; Schilling, C.H.; Palsson, B.O. Genome-scale microbial in silico models: the constraints-based approach. *Trends Biotechnol.*, **2003**, *21*(4), 162-169.

[49] Li, L.; Zhou, X.; Ching, W.K.; Wang, P. Predicting enzyme targets for cancer drugs by profiling human metabolic reactions in NCI-60 cell lines. *BMC Bioinformatics*, **2010**, *11*(1), 1471-2105.

[50] Folger, O.; Jerby, L.; Frezza, C.; Gottlieb, E.; Ruppin, E.; Shlomi, T. Predicting selective drug targets in cancer through metabolic networks. *Mol. Syst. Biol.*, **2014**, *7*(1), 501.

[51] Gottlieb, A.; Stein, G.Y.; Ruppin, E.; Sharan, R. PREDICT: a method for inferring novel drug indications with application to personalized medicine. *Mol. Syst. Biol.*, **2011**, *7*(1), 496.

[52] Dudley, J.T.; Deshpande, T.; Butte, A.J. Exploiting drug–disease relationships for computational drug repositioning. *Brief. Bioinform.*, **2011**, *12*(4):303-311.

[53] Hu, G.; Agarwal, P. Human disease-drug network based on genomic expression profiles. *Plos One*, **2009**, *4*(8), e6536.

[54] Qabaja, A.; Alshalalfa, M.; Alanazi, E.; Alhajj, R. Prediction of novel drug indications using network driven biological data prioritization and integration. *J. Cheminformatics*, **2014**, *6*(1), 1.

[55] Li, P.; Chen, J.; Wang, J.; Zhou, W.; Wang, X.; Li, B.; Tao, W.; Wang, W.; Wang, Y.; Yang, L. Systems pharmacology strategies for drug discovery and combination with applications to cardiovascular diseases. *J. Ethnopharmacol.*, **2014**, *151*(1), 93-107.

[56] Tao, W.; Xu, X.; Wang, X.; Li, B.; Wang, Y.; Li, Y.; Yang, L. Network pharmacology-based prediction of the active ingredients and potential targets of Chinese herbal Radix Curcumae formula for application to cardiovascular disease. *J. Ethnopharmacol.*, **2013**, *145*(1), 1-10.

[57] Li, X.; Xu, X.; Wang, J.; Yu, H.; Wang, X.; Yang, H.; Xu, H.; Tang, S.; Li, Y.; Yang, L. A system-level investigation into the mechanisms of chinese traditional medicine: compound danshen formula for cardiovascular disease treatment. *Plos One*, **2012**, *7*(9), e43918.

[58] Wang, X.; Xu, X.; Li, Y.; Li, X.; Tao, W.; Li, B.; Wang, Y.; Yang, L. Systems pharmacology uncovers Janus functions of botanical drugs: activation of host defense system and inhibition of influenza virus replication. *Integr. Biol.*, **2013**, *5*(2), 351-371.

[59] Hamosh, A.; Scott, A.F.; Amberger, J.S.; Bocchini, C.A.; McKusick, V.A. Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic Acids Res.*, **2005**, *33*(suppl 1), D514-D517.

[60] Goh, K.I.; Cusick, M.E.; Valle, D.; Childs, B.; Vidal, M.; Barabási, A.L. The human disease network. *Proc. Natl. Acad. Sci. U S A.*, **2007**, *104*(21), 8685-8690.

[61] Bauer-Mehren, A.; Bundschus, M.; Rautschka, M.; Mayer, M.A.; Sanz, F.; Furlong, L.I. Gene-disease network analysis reveals functional modules in mendelian, complex and environmental diseases. *Plos One*, **2011**, *6*(6), e20284.

[62] Sanseau, P.; Agarwal, P.; Barnes, M.R.; Pastinen, T.; Richards, J.B.; Cardon, L.R.; Mooser, V. Use of genome-wide association studies for drug repositioning. *Nat. Biotechnol.*, **2012**, *30*(4), 317-320.

[63] Hidalgo, C.A.; Blumm, N.; Barabási, A.-L.; Christakis, N.A. A dynamic network approach for the study of human phenotypes. *PLoS Comput. Biol.*, **2009**, *5*(4), e1000353.

[64] Lee, D.S.; Park, J.; Kay, K.; Christakis, N.; Oltvai, Z.; Barabási, A.L. The implications of human metabolic network topology for disease comorbidity. *Proc Natl. Acad. Sci. U S A.*, **2008**, *105*(29), 9880-9885.

[65] Sieberts, S.K.; Schadt, E.E. Moving toward a system genetics view of disease. Mamm. Genome, 2007, *18*(6-7), 389-401.

[66] Park, J.; Lee, D.S.; Christakis, N.A.; Barabási, A.L. The impact of cellular networks on disease comorbidity. *Mol. Syst. Biol.*, **2009**, *5*, 262.

[67] Chiang, A.P.; Butte, A.J. Systematic evaluation of drug–disease relationships to identify leads for novel drug uses. *Clin. Pharmacol. Ther.*, **2009**, *86*(5), 507-510.

[68] Dewey, F.E.; Perez, M.V.; Wheeler, M.T.; Watt, C.; Spin, J.; Langfelder, P.; Horvath, S.; Hannenhalli, S.; Cappola, T.P.; Ashley, E.A. Gene coexpression network topology of cardiac development, hypertrophy, and failure. *Circ. Cardiovasc. Genet.*, **2011**, *4*(1), 26-35.

[69] Ray, M.; Ruan, J.; Zhang, W. Variations in the transcriptome of Alzheimer's disease reveal molecular networks involved in cardiovascular diseases. *Genome Biol.*, **2008**, *9*(10), R148.

[70] Rende, D.; Baysal, N.; Kirdar, B. A novel integrative network approach to understand the interplay between cardiovascular disease and other complex disorders. *Mol. BioSyst.*, **2011**, *7*(7), 2205-2219.

[71] Hu, Z.; Chang, Y. C.; Wang, Y.; Huang, C. L.; Liu, Y.; Tian, F.; Granger, B.; DeLisi, C. VisANT 4.0: Integrative network platform to connect genes, drugs, diseases and therapies. *Nucleic Acids Res.*, **2013**, *41*(W1), W225-W231.

[72] Csermely, P.; Korcsmáros, T.; Kiss, H.J.; London, G.; Nussinov, R. Structure and dynamics of molecular networks: A novel paradigm of drug discovery: A comprehensive review. *Pharmacol. Ther.*, **2013**, *138*(3), 333-408.

[73] Ideker, T.; Krogan, N.J. Differential network biology. *Mol. Syst. Biol.*, **2012**, *8*, 565.

Received: September 29, 2014 Revised: October 25, 2014 Accepted: November 06, 2014