Minireview

Drug-therapy networks and the prediction of novel drug targets
Zoltan Spiro, Istvan A Kovacs and Peter Csermely

Address: Department of Medical Chemistry, Semmelweis University, PO Box 260, H-1444 Budapest 8, Hungary.

Correspondence: Peter Csermely. Email: csermely@puskin.sote.hu

Abstract

A recent study in *BMC Pharmacology* presents a network of drugs and the therapies in which they are used. Network approaches open new ways of predicting novel drug targets and overcoming the cellular robustness that can prevent drugs from working.

Despite a significant and continuous increase in medical research spending, the number of new drugs approved and new drug targets identified each year has remained almost constant for the past 20-25 years, with about 20 new drugs and about five new targets per year. Lengthy development procedures and the high risk of unexpected side-effects in advanced-stage clinical trials reduce the ability of the drug development process to be innovative. At this rate it will take more than 300 years to double the number of available drugs [1]. However, there are several ways to overcome these burdens. Promising areas of drug design include: wide-range screens of existing drugs, seeking novel applications; combination therapy, that is, the use of several drugs or short DNA oligomers, called aptamers, together [2-4]; and the development of multi-target drugs [5].

The organization of our rapidly growing knowledge on diseases, disease-related genes, drug targets and their structures, and drugs and their chemical structures gives us another exciting way to discover novel areas of drug development. Several networks have recently been constructed to help drug discovery [1,6-9]. In the network concept a complex system is

perceived as a set of interacting elements bound together by links. Links usually have a weight, which characterizes their strength such as the affinity of binding between the two elements, or the propensity of one element to act on the other. Links can also be directional, when one of the elements has a larger influence on the other than vice versa [10,11].

In a recent study in *BMC Pharmacology*, Nacher and Schwartz [8] compiled a drug-therapy network in which all US-approved drugs and associated human therapies - that is, the therapeutic properties of the drugs involved according to the Anatomic Therapeutic Chemical (ATC) classification - were connected to each other. From a bipartite network of therapies and drugs (in which therapies were connected to drugs, but drugs were not connected to other drugs or therapies to other therapies) they constructed two other networks: a drug network, in which two drugs were connected if they were both involved in at least one common therapy; and a therapy network, in which two therapies were connected if a particular drug was implicated in both therapies. Their analysis [8] provides the first view of the relationships between therapies as defined by drug-

therapy interactions, and it highlights a few key drugs that connect distinct classes of therapy in a few steps.

Both the drug and the therapy networks [8] proved to be small worlds; that is, distant therapies were separated by an average of less than three chemicals [10,11]. Highly connected therapies, 'therapy hubs', are likely to be relevant in the therapy network, because this network behaved in a manner close to that of a tree-like network, in which the relative importance of hubs is high. Most drugs (79%) were grouped in clusters connected to a specific therapy. However, a minority of drugs (21%) formed bridges spanning different therapeutic classes; these drugs may have a particular significance.

Nacher and Schwartz [8] computed several measures of network 'centrality' characterizing the importance of the drugs in the network context of various therapies. They identified a subnetwork of the bridging drugs with high 'betweenness centrality' (drugs that participated in a large number of shortest paths connecting other drugs) in the drug-therapy network; these include scopolamine, morphine, tretionin and magnesium sulfate. For example, tolbutamide and magnesium sulfate defined a key shortest path of two steps between the distant classes of therapy 'insulins and analogs' and 'dermatological preparations'. Apparently unrelated disorders were thus separated by a much lower number of chemicals than might be expected. Most drugs act on one target, but a few drugs act on a large number of targets. Nacher and Schwartz [8] propose that drugs that have a high betweenness centrality and act on multiple targets may influence multiple metabolic pathways, and they especially highlighted hydroxocobalamin, vitamin B3, vitamin B12, atropine, ophenadrine and procaine as members of this category.

The network approach not only gives us a systematic way to organize our vast databases, but also provides a visual image that can help us to understand the daunting complexity of these systems. However, many networks, such as that of the 1,360 individual chemical substances studied by Nacher and Schwartz [8], are too big for easy visualization. The ATC classification system used by these authors [8] gave them the opportunity to construct a hierarchical representation of drug-therapy information, in which we can zoom in from the top layer of 15 anatomical main therapeutic groups, through the 66 therapeutic subgroups (second layer), the 123 pharmacological subgroups (third layer), the 448 chemical subgroups (fourth layer) until reaching the fifth layer of 1,360 individual chemical substances. Where the dataset is not as straightforwardly hierarchical as the ATC classification [8], network hierarchy can be explored by various other techniques [12-14].

Drug-target and related networks

To show the rich context through which the results of Nacher and Schwartz [8] can be interpreted, we show the power of network approaches for constructing various drug-target and related networks, for predicting new drug targets, and to get around unwanted resistance and side effects of drugs. These tools promise to increase the number of novel drug targets and improve the approval rate of new drugs.

When thinking about the possible network representations of diseases, drugs and drug targets, the elements of the network first have to be defined. For a list of the available databases of drugs and related information, see Table 1. The next step is to find a general rule determining the elements that are linked in the particular network and the nature (such as weight or directionality) of the links connecting them. Besides the drug-therapy network already mentioned [8], several other, recently published network-building rule-sets [1,6,7,9] give additional exciting and novel information on the vast datasets of diseases, drugs and drug targets. A summary of these representations is shown in Figure 1.

In some of these approaches [1,6,7,9], the network can be constructed by either linking two drug-target proteins if both bind one or more compounds, or by linking two compounds (drugs) if both have at least one protein as a common target. One result from this analysis is that the average molecular weight of compounds becomes smaller and smaller as we go from preclinical drug candidates to Phase I, II, III and approved drugs. Other physicochemical properties, such as hydrophobicity and the ability to form hydrogen bonds, reduce further the number of drug candidates that can be given orally, which is the method normally desired [7].

The analysis of the drug-target network [1,6,7,9] also reveals further elements of the low-risk behavior of the pharmaceutical industry. The network is particularly enriched in highly targeted proteins, and elements with many neighbors (called hubs) are preferentially connected to each other, forming a so-called 'rich club'. This is a result of the tendency to target an already validated target protein with alternative or follow-up compounds. Experimental drugs act on a greater diversity of target proteins, and show a more diverse localization of the targets than the plasma membrane, which is usually the preferred site of action. So far, however, these efforts have not led to a significant expansion of novel targets, that is, novel classes of protein or cellular compartments [6].

An additional approach to deciphering meaningful information for drug development efforts is to link human

Table 1

Useful links to therapy, disease, drug and drug-target network data

Name	Description	References
DrugBank	A bioinformatics-cheminformatics resource combining detailed drug data with comprehensive drug target information with over 4,900 drug entries (about 3,500 experimental) and about 1,500 non-redundant protein entries	[26]
Drug-target Network	Network data of 890 drugs and 394 target human proteins	[6]
Drug-Therapy Network	Three layers of drug-therapy networks according to the ATC classification	[8]
Online Mendelian Inheritance in Man (OMIM)	A knowledgebase of human genes and genetic disorders	[27]
Potential Drug Target Database (PDTD)	A three-dimensional drug target structure database with a target identification option	[28]
Predicted drug targets	A set of 1,383 predicted drug targets	[17]
Protein ligand network	A network of 4,208 ligands and about 15,000 binding sites	[9]
Therapeutic Target Database	Lists over 1,500 therapeutic targets, disease conditions and corresponding drugs	[29]
FDA Orange Book	Approved drug products with therapeutic equivalence evaluations	[32]
Investigational Drugs database (IDdb)	Thomson Investigational drugs database including information on 107,000 patents, 25,000 investigational drugs and 80,000 chemical structures	[33]
TDR Targets Database	Identification and ranking targets against neglected tropical diseases	[34]

diseases that have in common at least one gene involved in the development of the disease. This human disease network has also been converted to the other possible network of disease genes, in which two genes are connected if they are associated with the same disorder [15]. Among human diseases, several types of cancer, such as colon and breast cancer, are hubs that are genetically connected to more than 30 distinct disorders. Disease genes that contribute to a common disorder often have protein products that form larger complexes, are often co-expressed and have similar major functions [15]. Interestingly, those inheritable disease genes that are not essential occupy a peripheral position in the cellular network. This is in stark contrast to essential genes, which are more central [16]. By contrast with inheritable disease genes, disease genes associated with somatic mutations, such as somatic cancer genes, have a central position in cellular networks [15]. When comparing drug-target networks with the related diseases, an ongoing shift of drug development can be observed towards 'novel diseases' with associated genes that were not previous drug targets [6].

In the analysis and visual representation of drug-therapy and drug-target networks, the weight of the links (such as the number of drugs binding to both of two linked targets in the drug-target network) is seldom assessed. In addition, these networks have not been thoroughly analyzed by defining their groups, or modules [10-14]. Both additions

will certainly provide more detailed information on these exciting datasets. Important messages could be drawn from the additional networks shown in Figure 1. Not only drugs, but also their respective drug targets, can be linked to the various therapies. As an additional, rich source of data, patient records can be analyzed for the diseases diagnosed as well as the drugs prescribed. Patient medication records can be transcribed to a patient-drug target network, which may reveal novel aspects of the phenotype variability of diseases. Yet another set of data lies in the symptoms of patients, which can serve as a basis to construct symptom-disease, symptom-therapy or symptom-drug networks (Figure 1).

Drugs may also form a structural network, where two drugs are linked if they contain the same, signature-like chemical segment or feature. Drugs can also be assembled to form a side-effect network, or toxicity network, which may give an overall view of these two key maladies of drug development. As more and more data will be available in the future, patient symptoms can be extended by appropriately selected patient transcriptome, proteome, metabolome, oral microbiome and gut microbiome data. This 'inflation' of drug and drug-related networks is unlikely to solve the current problems of drug design; rather, it may be that the more networks we add, the less clarity and focus we will enjoy. Drug- and disease-related network representations will certainly have their own evolution, however, and it is not
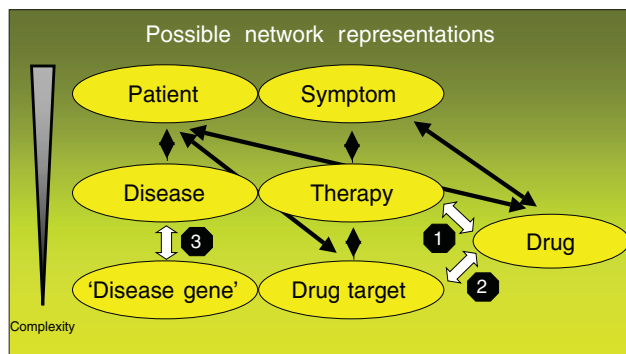
Figure 1
Overview and possible extensions of therapy, disease, drug and drug-target networks. The ovals represent data types (datasets) that have already been used for network construction (connected with white arrows) or that could be used to construct similar networks in the future (connected with black arrows). Datasets are positioned from top to bottom in the approximate order of their decreasing complexity; datasets in the same row overlap each other. Numbers indicate: (1) the drug-therapy network of Nacher and Schwartz [8]; (2) drug-target networks [1,6,7,9]; (3) a disease-gene network [15].

yet clear which of them will give the most straightforward, non-obvious visual and analytical information.

Prediction of novel drug targets using network analysis

Existing segments of drug-target networks may have hidden information on additional drug targets that are not yet included in the network. Extension of existing networks by predicting links and elements is a recent, exciting field of network studies [13,17,18]. The identification of novel drug-target candidates can be accomplished by finding missing links in all networks in which drug targets serve as links, for example, in drug [1,6,7,9], therapy [8] or patient networks joined by common drug targets. Methods for discovering new links can identify new nodes in all bipartite networks, in which any of the nodes of one type can be converted to links joining nodes to the other type. This may give us novel methods to predict, discover, test and extend gene regulatory, metabolic, opinion (recommendation), collaboration (co-authorship), sexual and any other affiliation-type networks.

Robustness is an intrinsic property of cellular networks that enables them to maintain their functions despite various perturbations [19]. Networks of different topology vary by orders of magnitude in their robustness to mutations and noise. Enhanced robustness is a property of only a very small number of all possible network topologies [20]. Cellular networks in both health and disease belong to this extreme minority and show this robust behavior.

Many times when a drug fails or produces side effects, cellular robustness provides most of the explanation. A drug can be ineffective when the robustness of cellular networks of disease-affected cells or parasites compensates for the changes caused by it. By contrast, drug side effects can be the result of hitting an unexpected point of fragility in the affected networks [21]. Robustness analysis is already being used to reveal primary drug targets [22], and the first methods have also been established to give a quantitative measure of changes in robustness during drug action [23].

Cellular robustness can be caused by strong links forming negative or positive feedbacks that help the cell to return to the original state or jump to another, respectively; by weak links that provide alternative, redundant pathways; or by a range of other mechanisms [11,19,21]. But achieving robustness always has a price. Robust cells have their fragile point, their 'Achilles heel', and cannot be optimized for all other aspects of cellular life, such as proliferation. This gives us chances to conquer or redirect cellular robustness by the application of drugs. We can develop drugs that, for example, find the Achilles heel of the cellular robustness of disease-affected cells or parasites, or that decrease robustness, for example, by inhibiting the effect of weak links [11,19,21]. Note that a decrease in robustness makes the cellular network noisier and less predictable, which means that robustness-decreasing drugs will be more difficult to find than conventional drugs using currently available analytical methods that assume an 'equilibrium' network. The development of 'fuzzy', stochastic network analysis [14,18] and the comparison of network time series may help to overcome this difficulty.

Our current knowledge on cellular networks and their analytical methods has arrived at a time when testing the effects of drug candidates with known cellular targets or target-sets on the robustness of cellular networks is becoming possible. A robustness test, revealing both resistance-related failures and side effects, should, in our opinion, be a mandatory element of standard drug-development protocols. The more we know about tissue- and disease-specific changes in cellular networks and about variations in these changes between individuals, the better we will be able to predict the efficiency of drugs in *in silico* experiments.

In summary, the recently published drug-therapy [8] and drug-target networks [1,6,7,9] as well as their potential extensions (Figure 1) provide a powerful and exciting tool for the organization of the expanding drug-development data and give us a global view on major trends and limitations. The advent of combinatorial therapies [2-4] and multi-target drugs [5] may greatly help us to break or re-direct the robust behavior of the cell. For the

knowledge-based design of appropriate drug combinations and multi-target drugs, however, we need novel approaches and techniques to explore the dynamic complexity of cellular networks after multiple perturbations [24,25].

Note added in proof

During the processing of this manuscript Campillos *et al.* [30] published a network of 502 drugs and their side effects using these data to predict novel drug targets based on the similarity of side-effect of two chemically dissimilar drugs. From the same data a side-effect network could also be constructed, where two side effects are linked, if a drug exists, which has both. This side-effect network in combination with the link-prediction methods outlined above [13,17,18] opens the possibility to predict additional side effects of existing drugs and drug candidates. Network-based side-effect prediction would greatly help the development of better clinical trial protocols, and would uncover additional possible dangers before the large-scale use of novel drugs.

References
1. Ma'ayan A, Jenkins SL, Goldfarb J, Iyengar R: **Network analysis of FDA approved drugs and their targets.** *Mt Sinai J Med* 2007, **74:**27-32.
2. Borisy AA, Elliott PJ, Hurst NW, Lee MS, Lehar J, Price ER, Serbedzija G, Zimmermann GR, Foley MA, Stockwell BR, Keith CT: **Systematic discovery of multicomponent therapeutics.** *Proc Natl Acad Sci USA* 2003, **100:**7977-7982.
3. Wong PK, Yu F, Shahangian A, Cheng G, Sun R, Ho CM: **Closed-loop control of cellular functions using combinatory drugs guided by a stochastic search algorithm.** *Proc Natl Acad Sci USA* 2008, **105:**5105-5110.
4. Fitter S, James R: **Deconvolution of a complex target using DNA aptamers.** *J Biol Chem* 2005, **280:**34193-34201.
5. Csermely P, Agoston V, Pongor S: **The efficiency of multi-target drugs: the network approach might help drug design.** *Trends Pharmacol Sci* 2005, **26:**178-182.
6. Yildirim MA, Goh KI, Cusick ME, Barabasi AL, Vidal M: **Drug-target network.** *Nat Biotechnol* 2007, **25:**1119-1126.
7. Paolini GV, Shapland RHB, van Hoorn WP, Mason JS, Hopkins AL: **Global mapping of pharmacological space.** *Nat Biotechnol* 2006, **24:**805-815.
8. Nacher JC, Schwartz JM: **A global view of drug-therapy interactions.** *BMC Pharmacol* 2008, **8:**5.
9. Park K, Kim D: **Binding similarity of network of ligand.** *Proteins* 2008, **71:**960-971.
10. Barabási AL, Oltvai ZN: **Network biology: understanding the cell's functional organization.** *Nat Rev Genet* 2004, **5:**101-113.
11. Csermely P: *Weak Links: a Universal Key for Network Diversity and Stability*. Heidelberg: Springer Verlag; 2006.
12. Song C, Havlin S, Makse HA: **Self-similarity of complex networks.** *Nature* 2005, **433:**392-395.
13. Clauset A, Moore C, Newman MEJ: **Hierarchical structure and the prediction of missing links in networks.** *Nature* 2008, **453:**98-101.
14. Kovács IA, Szalay MS, Csermely P, Korcsmáros T: **Method for analyzing the fine structure of networks.** *Patent application number: WO2007093960.* 2006. (http://www.wipo.int/pctdb/en/wo.jsp?wo=2007093960&IA=WO2007093960&DISPLAY=STATUS).
15. Goh KI, Cusick ME, Valle D, Childs B, Vidal M, Barabási AL: **The human disease network.** *Proc Natl Acad Sci USA* 2007, **104:**8685-8690.
16. Feldman I, Rzhetsky A, Vitkup D: **Network properties of genes harbouring inherited disease mutations.** *Proc Natl Acad Sci USA* 2008, **105:**4323-4328.
17. Li Q, Lai L: **Prediction of potential drug targets based on simple sequence properties.** *BMC Bioinformatics* 2007, **8:**353.
18. Nepusz T, Petróczi A, Négyessy L, Bazsó F: **Fuzzy communities and the concept of bridgeness in complex networks.** *Phys Rev E Stat Nonlin Soft Matter Phys* 2008, **77:**016107.
19. Kitano HH: **Biological robustness.** *Nat Rev Genet* 2004, **5:**826-837.
20. Ciliberti S, Martin OC, Wagner A: **Robustness can evolve gradually in complex regulatory gene networks with varying topology.** *PLoS Comput Biol* 2007, **3:**e15.
21. Kitano HH: **A robustness-based approach to systems-oriented drug design.** *Nat Rev Drug Discov* 2007, **6:**202-210.
22. Hallén K, Björkegren J, Tegnér J: **Detection of compound mode of action by computational integration of whole-genome measurements and genetic perturbations.** *BMC Bioinformatics* 2006, **7:**51.
23. Moriya H, Shimizu-Yoshida Y, Kitano H: **In vivo robustness analysis of cell division cycle genes in *Saccharomyces cerevisiae*.** *PLoS Genet* 2006, **2:**e111.
24. Korcsmáros T, Szalay MS, Böde C, Kovács IA, Csermely P: **How to design multi-target drugs: Target-search options in cellular networks.** *Expert Opin Drug Discov* 2007, **2:**1-10.
25. Antal MA, Bode C, Csermely P: **Perturbation waves in proteins and protein networks: Applications of percolation and game theories in signaling and drug design.** *Curr Prot Pept Sci* 2008, in press, arXiv:0802.2330v1.
26. Wishart DS, Knox C, Guo AC, Cheng D, Shrivastava S, Tzur D, Gautam B, Hassanali M: **DrugBank: a knowledgebase of drugs, drug actions and drug targets.** *Nucleic Acids Res* 2008, **36(Database issue):**D901-D906.
27. Hamosh A, Scott AF, Amberger JS, Bocchini CA, McKusick VA: **Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders.** *Nucleic Acids Res* 2005, **33(Database issue):**D514-D517.
28. Gao Z, Li H, Zhang H, Liu X, Kang L, Luo X, Zhu W, Chen K, Wang X, Jiang H: **PDTD: a web-accessible protein database for drug target identification.** *BMC Bioinformatics* 2008, **9:**104.
29. Chen X, Ji ZL, Chen YZ: **TTD: Therapeutic Target Database.** *Nucleic Acids Res* 2002, **30:**412-415.
30. Campillos M, Kuhn M, Gavin A-C, Jensen LJ, Bork P: **Drug target identification using side-effect similarity.** *Science* 2008, **321:**263-266.
31 **LINK-Group** [www.linkgroup.hu]
32. **FDA Orange Book** [http://www.fda.gov/cder/ob/]
33. **Investigational Drugs database (IDdb)** [http://scientific.thomson.com/products/iddb/]
34. **The TDR Targets Database** [http://tdrtargets.org/]