Genome Biology

Minireview

Predicting drug side-effects by chemical systems biology

Nicholas P Tatonetti*, Tianyun Liu† and Russ B Altman†‡

Addresses: *Training Program in Biomedical Informatics, †Department of Bioengineering, Stanford University, ‡Department of Genetics, Stanford University, Stanford, CA 94305, USA.

Correspondence: Russ B Altman. Email: russ.altman@stanford.edu

Abstract

New approaches to predicting ligand similarity and protein interactions can explain unexpected observations of drug inefficacy or side-effects.

Drug-related adverse events affect approximately 2 million patients in the United States each year, resulting in about 100,000 deaths [1]. For example, highly publicized cases of severe adverse reactions recently resulted in a US Food and Drug Administration advisory panel suggesting that the popular pain relievers Percocet and Vicodin be banned [2]. Some adverse events are predictable consequences of the known mechanism of a drug, but others are not predicted and seem to result from 'off-target' pathways.

When developing novel chemical entities (NCEs) for a therapeutic application, knowledge of binding partners and affected biological pathways is useful for predicting both efficacy and side-effects. Traditional drug design has relied heavily on the one drug-one target paradigm [3], but this may overlook system-wide effects that cause the drug to be unsuccessful. Adverse side-effects and lack of efficacy are the two most important reasons a drug will fail clinical trials, each accounting for around 30% of failures [3]. The development of tools that can predict adverse events and system-wide effects might thus reduce the attrition rate. Such tools will most certainly include emerging information about protein-protein interactions, signaling pathways, and pathways of drug action and metabolism. A systems view of the body's responses to a drug threatens the simplicity of the one drug-one target paradigm, but could provide a framework for considering all effects, and not just those that are targeted.

The laboratory assays currently used to evaluate potential adverse drug effects can be costly and time-consuming. For example, an expensive two-year rodent bioassay is the current gold standard for determining the carcinogenicity of a NCE [4]. Some assays are also of doubtful utility - only around 15% of gene knockouts in the standard pharmaceutical model organisms show any fitness defect [3]. Therefore, drugs designed with a single target in mind may prove ineffective, not because they do not interact with the target in the expected way, but because of natural redundancies in pharmacological networks. To compound the problem, protein-ligand studies have found that a single drug can bind targets with vastly different pharmacology and that about 35% of known drugs have two or more targets [5]. It is not surprising that evolutionary relationships might lead to shared drug-binding capabilities in protein paralogs found across a wide range of cell types and biological pathways. These complexities, however, create new opportunities for therapeutic strategies involving the concerted use of drugs with multiple targets to achieve an increased specificity in effect. A recent review by Giordano and Petrelli, for example, describes their approach to developing multi-target drugs for cancer therapy while avoiding drug resistance by targeting multiple tyrosine kinase receptors [6].

Chemical systems biology, or the application of system-wide tools to the analysis of pharmacological responses, can help address the lack of efficacy and undesired off-target effects [3]. Understanding each of these requires the ability to characterize off-target side-effects *in silico*. In a recent study, Philip Bourne and colleagues (Xie *et al.* [7]) have used a chemical systems biology approach to explain the serious side-effects of a drug that was being trialed for prevention of cardiovascular disease.

Systems biology meets chemical biology

For our purposes here, systems biology means an approach to biology that looks at networks of molecular interactions (including gene products, endogenous small molecules and drugs) and processes these using qualitative graphical models or quantitative mathematical modeling [8]. Examples of implementations of quantitative methods include Flux Balance Analysis [9], differential equations [10], and Petri Nets [11]. Implementations of qualitative methods include Cytoscape [12], a graphical network representation, and Genoscape [13], a network-based knowledge integration extension tool. When the principles of systems biology are extended to medications, we get a network of interactions between drugs and the naturally occurring metabolic and signaling networks (Figure 1). These drugs may

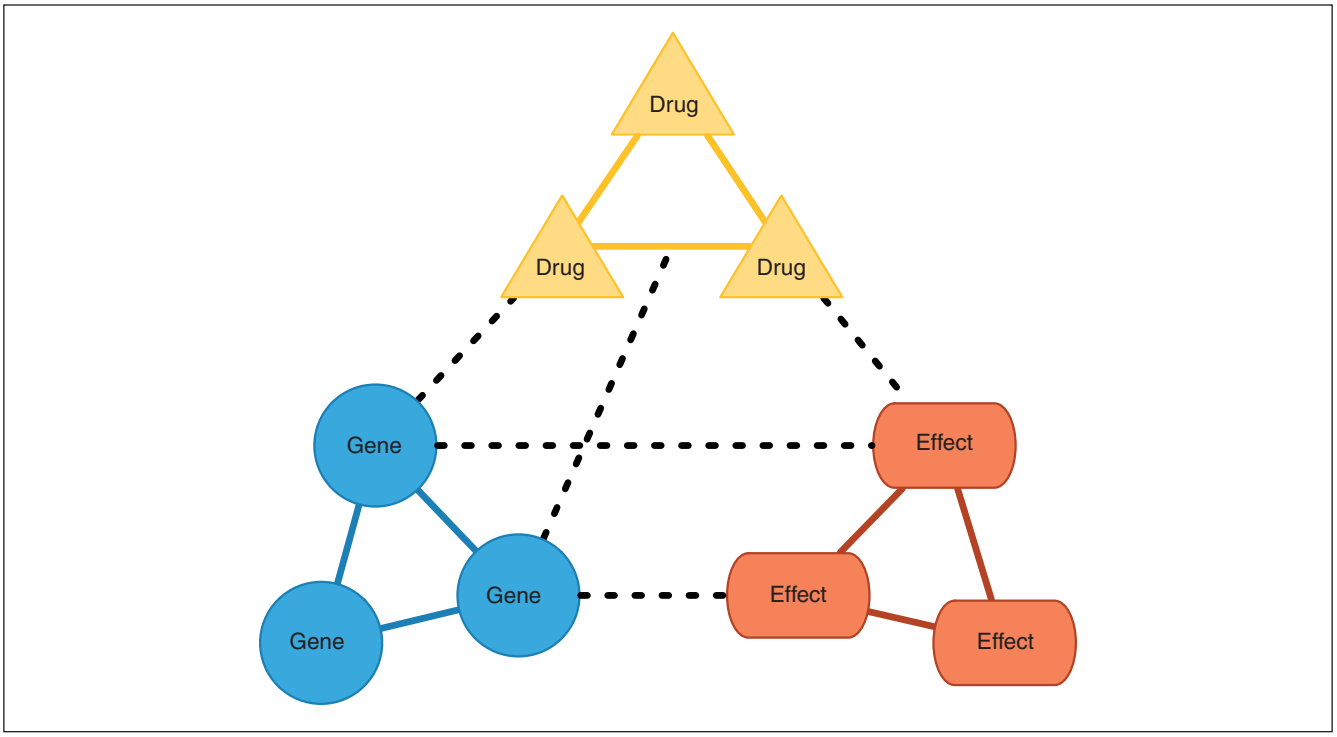

Figure 1

Meta-networks allow novel inferences. Systems approaches allow the generation of networks of genes based on common pathways or common evolutionary history, networks of drugs based on chemical similarity or similarity in biological effects, and networks of effects based on similar biological pathways and cellular compartments. The ability to link these three networks allows novel inferences.

connect otherwise disconnected and independent sub-networks, and this may cause both expected and unexpected effects. Pharmacological systems biology must combine the biological and chemical characteristics of small and large molecules to develop an understanding of drug action. These protein-drug joint networks provide two opportunities. First, they can provide more detailed descriptions (even signatures) of drug effects, and second, they can provide a framework for the design of novel therapeutic strategies [4].

The intersection of systems biology and chemical biology opens new avenues of research. In particular, there are opportunities to combine data from genomics, three-dimensional structure, large chemical screens, protein-protein interactions, protein-drug binding interactions, and cellular imaging and localization to assemble a high-fidelity model of how and where small molecules interact with cellular components. A harbinger of the opportunities that exist is the work by Apsel *et al.* [14], who have integrated chemical biology and systems biology techniques to design drugs that act as dual inhibitors of two families of oncogenes.

The recent work of Xie *et al.* [7] is another excellent example of the use of networks combining proteins and drugs. They investigated the reasons for the serious side-effects of torcetrapib, an inhibitor of cholesteryl ester transfer protein (CETP) that was in clinical trials as a preventive treatment for cardiovascular disease. The aim of torcetrapib was to raise the levels of the desirable high-density lipoprotein cholesterol (HDL-C), but torcetrapib turned out to have the side-effect of raising blood pressure, with potentially fatal effects in high-risk patients, and was withdrawn from development in 2006.

Xie *et al.* [7] generated off-target binding networks by comparing the structure of ligand-binding sites in all known protein structures. The proteins identified as having similar binding domains were ranked by a normalized docking score and clustered by their structural and functional characteristics into a gene network that includes metabolic and regulation pathways. Using this analysis, the authors identified possible off-targets for torcetrapib even though the binding site of CETP itself is not fully described. Perhaps most interestingly, they incorporated biological pathways into their off-target networks and found a potential explanation for the poorly understood effects of torcetrapib on blood pressure. By combining a simple gene regulation model with the predicted binding affinities to activators and inhibitors of

the renin-angiotensin-aldosterone system (RAAS), they showed that torcetrapib caused more severe effects since it has a higher affinity for more RAAS activators.

To validate this approach, the investigators compared the off-target networks for drugs with different side-effect profiles, and show that the networks are quite different and consistent with the different effects of the drugs on blood pressure [7]. Their method can, however, only use proteins with known structures - a small fraction of the human proteome. As a result, pharmacologists may become fans of high-throughput structural biology!

An alternative approach to discovering off-target effects relies on identifying common chemical features among drugs with the same set of adverse reactions [15]. This approach links chemical sub-structures to specific toxicities and can be used to determine the potential side-effects of a drug with a novel chemical structure. An implementation of this technique is described by Scheiber *et al.* [15] to relate chemical substructures to side-effects and by Campillos *et al.* [16] to combine drug chemical similarity to side effect similarity to predict shared drug targets. Recent work by Shoichet and colleagues (Hert *et al.* [17]) in this field uses the similarity ensemble approach with a Bayesian method to build chemoinformatics networks based on chemical similarities between drugs, instead of on structural or sequence similarities between drug targets. Comparisons between the ligand-based network of Hert *et al.* [17] and the target-based network of Xie *et al.* [7] might provide interesting insights. If the networks' information content is complementary, as opposed to redundant, then a method that utilized both network may outperform either one alone.

Other investigators have taken a complementary approach. Instead of looking for common chemical sub-structures, they focus on common adverse reactions. Scheiber *et al.* [1] have incorporated data from a variety of databases and identified drugs with shared toxicities. They then apply an understanding of the molecular pathways underlying these toxicities to predict drug targets. In this way, they can develop data-driven hypotheses about the mechanisms of a particular side-effect. This approach is particularly useful when chemicals with very different structures (not likely to be recognized using measures of chemical similarity) interact with the same biological pathway. The toxicities are effectively used as a proxy for the biological pathways that the drug is involved with.

The success of network-based methods relies heavily on the development and curation of high-quality biological and pharmacological databases. The new high-throughput technologies have provided a huge amount of data on protein-protein and gene-gene interaction networks. The meta-database pathguide.org [18] currently lists more than 70 such databases that are freely available. However, as Blow points out in a recent review [19], no one database is complete, and combining datasets will yield more information. The study by Xie *et al.* [7], for example, incorporates data from eight different sources. The availability of these databases will fuel the next generation of chemical systems biology tools and lead to major advances in drug discovery and repositioning. Databases that attempt to integrate these different sources of data are becoming available. One such, STITCH, tries to consolidate knowledge about interactions between proteins and small molecules [20]. Although undoubtedly useful, these huge databases do raise the issue of false discovery. Incorporating domain knowledge to rank genes by their propensity to cause a modulated drug response may be one way of addressing this issue [21].

The ability to predict and even design the effects of new drugs is critical for the future pharmaceutical industry. By integrating biological and chemical knowledge, the pharmacological effects of drugs can be more completely understood and used to create predictive models. Recent work has focused on relating drugs to targets by chemical similarity, target structural similarity and even side-effect similarity. In each case, the results have illustrated the power of thinking about drug responses in the context of a network of interactions, and from a systems perspective.

Acknowledgements

NPT is supported by training grant NIH LM007033. TL is supported by LM05652. RBA is supported by LM05652 and the NIH/NIGMS Pharmacogenetics Research Network and Database and the PharmGKB resource (NIH U01GM61374).

References

1. Scheiber J, Chen B, Milik M, Sukuru SC, Bender A, Mikhailov D, Whitebread S, Hamon J, Azzaoui K, Urban L, Glick M, Davies JW, Jenkins JL: **Gaining insight into off-target mediated effects of drug candidates with a comprehensive systems chemical biology analysis.** *J Chem Inf Model* 2009, **49:**308-317.
2. Fosnocht D, Taylor JR, Caravati EM: **Emergency department patient knowledge concerning acetaminophen (paracetamol) in over-the-counter and prescription analgesics.** *Emerg Med J* 2008, **25:**213-216.
3. Hopkins AL: **Network pharmacology: the next paradigm in drug discovery.** *Nat Chem Biol* 2008, **4:**682-690.
4. Plant N: **Can systems toxicology identify common biomarkers of non-genotoxic carcinogenesis?** *Toxicology* 2008, **254:** 164-169.
5. Keiser MJ, Roth BL, Armbruster BN, Ernsberger P, Irwin JJ, Shoichet BK: **Relating protein pharmacology by ligand chemistry.** *Nat Biotechnol* 2007, **25:**197-206.
6. Petrelli A, Giordano S: **From single- to multi-target drugs in cancer therapy: when aspecificity becomes an advantage.** *Curr Med Chem* 2008, **15:**422-432.
7. Xie L, Li J, Xie L, Bourne PE: **Drug discovery using chemical systems biology: identification of the protein-ligand binding network to explain the side effects of CETP inhibitors.** *PLoS Comput Biol* 2009, **5:**e1000387.
8. Peterson RT: **Chemical biology and the limits of reductionism.** *Nat Chem Biol* 2008, **4:**635-638.

9. Lee JM, Gianchandani EP, Papin JA: **Flux balance analysis in the era of metabolomics.** *Brief Bioinformatics* 2006, **7:**140-150.
10. McAdams HH, Shapiro L: **Circuit simulation of genetic networks.** *Science* 1995, **269:**650-656.
11. Peleg M, Rubin D, Altman RB: **Using Petri Net tools to study properties and dynamics of biological systems.** *J Am Med Inform Assoc* 2005, **12:**181-199.
12. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T: **Cytoscape: a software environment for integrated models of biomolecular interaction networks.** *Genome Res* 2003, **13:**2498-2504.
13. Clément-Ziza M, Malabat C, Weber C, Moszer I, Aittokallio T, Letondal C, Rousseau S: **Genoscape: a Cytoscape plug-in to automate the retrieval and integration of gene expression data and molecular networks.** *Bioinformatics* 2009. [Epub ahead of print]
14. Apsel B, Blair JA, Gonzalez B, Nazif TM, Feldman ME, Aizenstein B, Hoffman R, Williams RL, Shokat KM, Knight ZA: **Targeted polypharmacology: discovery of dual inhibitors of tyrosine and phosphoinositide kinases.** *Nat Chem Biol* 2008, **4:**691-699.
15. Scheiber J, Jenkins JL, Sukuru SC, Bender A, Mikhailov D, Milik M, Azzaoui K, Whitebread S, Hamon J, Urban L, Glick M, Davies JW: **Mapping adverse drug reactions in chemical space.** *J Med Chem* 2009, **52:**3103-3107.
16. Campillos M, Kuhn M, Gavin AC, Jensen LJ, Bork P: **Drug target identification using side-effect similarity.** *Science* 2008, **321:**263-266.
17. Hert J, Keiser MJ, Irwin JJ, Oprea TI, Shoichet BK: **Quantifying the relationships among drug classes.** *J Chem Inf Model* 2008, **48:**755-765.
18. **Pathguide: the pathway resource list** [http://www.pathguide.org]
19. Blow N: **Systems biology: untangling the protein web.** *Nature* 2009, **460:**415-418.
20. Kuhn M, von Mering C, Campillos M, Jensen LJ, Bork P: **STITCH: interaction networks of chemicals and proteins.** *Nucleic Acids Res* 2008, **36(Database issue):**D684-D688.
21. Hansen N, Brunak S, Altman R: **Generating genome-scale candidate gene lists for pharmacogenomics.** *Clin Pharmacol Ther* 2009, **86:**183-189.

Published: 02 September 2009
doi:10.1186/gb-2009-10-9-238