nature chemical biology

Network pharmacology: the next paradigm in drug discovery

Andrew L Hopkins

The dominant paradigm in drug discovery is the concept of designing maximally selective ligands to act on individual drug targets. However, many effective drugs act via modulation of multiple proteins rather than single targets. Advances in systems biology are revealing a phenotypic robustness and a network structure that strongly suggests that exquisitely selective compounds, compared with multitarget drugs, may exhibit lower than desired clinical efficacy. This new appreciation of the role of polypharmacology has significant implications for tackling the two major sources of attrition in drug development—efficacy and toxicity. Integrating network biology and polypharmacology holds the promise of expanding the current opportunity space for druggable targets. However, the rational design of polypharmacology faces considerable challenges in the need for new methods to validate target combinations and optimize multiple structure-activity relationships while maintaining drug-like properties. Advances in these areas are creating the foundation of the next paradigm in drug discovery: network pharmacology.

Over the past decade, there has been a significant decrease in the rate that new drug candidates are being translated into effective therapies in the clinic. In particular, there has been a worrying rise in late-stage attrition in phase 2 and phase 3 (ref. 1). Currently, the two single most important reasons for attrition in clinical development are (i) lack of efficacy and (ii) clinical safety or toxicology, which each account for 30% of failures[1]. These late-stage attrition rates are at the heart of much of the relative decline in productivity of the pharmaceutical industry. Moreover, the decline in productivity is creating a major financial shock to the pharmaceutical industry. Owing to patents expiring on the current generation of marketed drugs, from 2010 onward, pharmaceutical companies will face the first fall in revenue in four decades.

Many reasons have been proposed for this decline in pharmaceutical research and development productivity. However, the fundamental problem may not be technological, environmental or even scientific but philosophical—there may be issues with the core assumptions that frame our approach to drug discovery. The increase in the rate of drugs failing in late-stage clinical development over the past decade has been concurrent with the dominance of the assumption that the goal of drug discovery is to design exquisitely selective ligands that act on a single disease target. This philosophy of rational drug design, or more specifically, the 'one gene, one drug, one disease' paradigm, arose from a congruence between genetic reductionism and new molecular biology technologies that enabled the isolation and characterization of individual 'disease-causing' genes[2], thereby enabling the full realization of Ehrlich's philosophy of 'magic bullets' targeting individual chemoreceptors[3]. The underlying assumption of the current approach is that safer, more effective drugs will result from designing very selective ligands where undesirable and potentially toxic side activities have been removed. However, after nearly two decades of focusing on developing highly selective ligands, the clinical attrition figures challenge this hypothesis.

Division of Biological Chemistry and Drug Discovery, College of Life Science, University of Dundee, Dow Street, Dundee DD1 3DF, UK. Correspondence should be addressed to A.L.H. (a.hopkins@dundee.ac.uk).

Published online 20 October 2008; doi:10.1038/nchembio.118

Need for a one-two punch

Clinical attrition rates are not the only data to challenge the current paradigm in drug discovery. Large-scale functional genomics studies in a variety of model organisms have revealed that under laboratory conditions, many single-gene knockouts by themselves exhibit little or no effect on phenotype, with approximately 19% of genes being essential across a number of model organisms[4–6]. In addition to the 19% lethality rate, systematic genome-wide homozygous gene deletion experiments in yeast reveal that only 15% of knockouts result in a fitness defect in ideal conditions[7]. A project to delete each of the druggable genes[8] in the mouse genome and profile each knockout across a battery of phenotypic assays has revealed that as few as 10% of knockouts demonstrate phenotypes that may be of value for drug target validation[4,9–11].

This robustness of phenotype can be understood in terms of redundant functions and alternative compensatory signaling routes[12]. Network analysis of biological pathways and interactions has revealed that much of the robustness of biological systems can derive from the structure of the network[13,14]. The scale-free nature of many biological networks results in systems that are resilient against random deletion of any one node but that are also critically dependent on a few highly connected hubs. The inherent robustness of interaction networks, as an underlying property, has profound implications for drug discovery; instead of searching for the 'disease-causing' genes, network biology suggests that the strategy should be to identify the perturbations in the disease-causing network[15].

Network biology analysis predicts that if, in most cases, deletion of individual nodes has little effect on disease networks, modulating

multiple proteins may be required to perturb robust phenotypes[13,16,17]. The emergent phenotype that occurs with the perturbation of multiple nodes is experimentally demonstrated by synthetic behaviors: synthetic lethality, synthetic sickness and synthetic rescue. Dual knockouts in a number of model systems have shown that although the isolated deletion of two individual genes may demonstrate no effect, the simultaneous deletion of the two genes can be lethal (synthetic lethality) or deleterious (synthetic sickness)[18]. A recent large-scale study by Hillenmeyer *et al.* demonstrates the extent of synthetic lethality when gene deletions are augmented by chemical interventions[19]. Under ideal conditions, only 34% of single-gene deletions in yeast resulted in lethality or sickness. However, when the whole genome panel of yeast single-gene knockouts was screened against a diverse small-molecule library and assayed against a wide range of environmental conditions, an additional 63% of gene knockouts showed a growth phenotype[19], resulting in 97% of genes demonstrating a fitness defect when challenged with a small molecule under at least one environmental condition. Thus, although the majority of genes may be redundant under any one environment, there seems to be little redundancy across a spectrum of conditions when a genetic perturbation is combined with a chemical insult.

As increased understanding of the role of networks in the robustness and redundancy of biological systems challenges the dominant assumption of single target drug discovery[16,20–23], a new approach to drug discovery—that of polypharmacology[16,17,20,21,24–29]—is emerging. Polypharmacology is not to be confused with the behavior of promiscuous aggregators (as identified by McGovern *et al.*) that arise from certain small molecules self-associating into colloids at high concentrations in biological buffers[30]. Instead, polypharmacology is the specific binding of a compound to two or more molecular targets. In an analogy to Paracelsus' axiom that the difference between a drug and a poison is the dose, the utility or toxicity of synthetically lethal and synthetically sick combinations is found in the biological context to which they are applied. Therefore, understanding the polypharmacology of a drug and its effect on biological networks and phenotype is essential if we wish to improve efficacy but also understand toxicity.

Synthetic lethality in cancer

The fundamental challenge of anticancer therapy is the need for agents that eliminate cancer cells with a therapeutic index that is safely tolerated by the patient. Most current anticancer drugs inhibit essential functions that are present in both normal and cancerous cells. Although these differentially impact rapidly dividing cancer cells, the essential nature of the targets of most cytotoxic anticancer drugs results in narrow therapeutic indices. In recent years, a new generation of drugs have targeted protein kinases, such as ABL, EGRF and ERBB2, that are differentially expressed in different cancers. These new drugs, which target non-essential proteins, have more manageable side effect profiles than cytotoxics; however, clinical efficacy is, in general, limited. An ideal cancer therapy, therefore, would be one that targeted proteins or interactions that are essential in cancer cells but non-essential in normal cells.

Cancer-specific molecular targets are rare. Most mutated oncogenic proteins are also present in normal cells, and selective inhibition of the mutant form can be a challenge. For example, the chronic myeloid leukemia (CML)-specific BCR-ABL fusion protein is inhibited by imatinib. However, imatinib also inhibits the non-oncogenic C-Abl kinase in normal cells, and long-term administration of the drug can lead to cardiotoxicity[31]. To overcome the difficulty of identifying and targeting differential features in a cancer, synthetic lethality has been proposed as a possible strategy for therapeutic intervention[32]. In the context of oncology, genetic and epigenetic changes in a cancer cell may change not only the relative expression levels but also the stoichiometry of the interaction network, and thus change the relative dependence on specific proteins relative to normal cells. Thus, two proteins that are non-essential in a normal cell may be essential in the context of a re-wired cancer cell network. In short, though the majority of the protein inventory in a cancer cell is the same as a normal cell, the differences in the topology of the biological networks could be targeted to produce an improved therapeutic index. Indeed, subtle differences in network stability and structure between cancer cells may explain the wide variance in cell fate that has been observed in individual cells of the same genetic lineage[33].

Whitehurst *et al.* recently conducted a whole-genome synthetic lethality screen in combination with paclitaxel, resulting in the discovery of new drug-drug combinations[34]. From the whole-genome RNA interference screening, 87 initial genes were identified that sensitized a human non-small-cell lung cancer line to paclitaxel, including the gene encoding vacuolar ATPase, the target of salicylihalamide A. Subsequent testing of salicylihalamide A and paclitaxel in combination was shown to reduce cancer cell viability. Sensitization synthetic lethality screens can also be used to discover potential synergistic combinations that can enhance the effectiveness of therapies. For example, breast cancer cells with deficiencies in BRAC1 and BRAC2 show differential synthetic lethality to inhibition of poly(ADP-ribose)-polymerase-1 (PARP). Screening a PARP inhibitor for additional synthetic lethality with an RNAi library identified a set of kinases, including CDK5, whose knockdown resulted in increased sensitization to the PARP inhibitor[35]. In addition to whole-genome screening, hypotheses for new drug combinations can be discovered by analysis of gene expression signatures. For example, analysis of breast cancer gene expression data revealed that the "gang of four" (COX2, MMP1, MMP2 and epiregulin) is essential for lung metastasis[36,37]. Genetic and pharmacological inhibition of these four genes, in combination, resulted in the halting of metastatic progression in a mouse model[36]. Previous advances in cancer combination therapies, such as those against childhood leukemia, were developed empirically over three decades. Synthetic lethality chemical sensitization screens offer a promising method to help systematically explore candidate cancer drug combinations efficiently in the laboratory.

The relationship between kinetics and systemic responses to perturbations offers an intriguing additional dimension in which network pharmacology strategies can be applied; it also provides a framework for understanding systems responses[38–40]. The sequence in which a combination is dosed may create different perturbations to the network that may have a dramatic effect on efficacy[38,41]. The systemic effects of network perturbations suggest that further studies on dosing sequence should follow the discovery of even modest effects of combination therapies.

Antibacterial polypharmacology

The single-target approach has been a major assumption behind genomics-based drug discovery strategies so far, including the search for new antibacterial targets[42–45]. However, the biologically led strategy for new antibacterial drugs, usually consisting of the search for single proteins that are essential when deleted, is flawed for two fundamental reasons: the downstream difficulty of discovering small-molecule compounds has often been considered only after significant investment in biology and a single amino acid mutation in the target protein is often enough to confer drug resistance. Many effective antibiotics

act by targeting multiple proteins simultaneously rather than individual proteins[46]. For example, the antibacterial action of β-lactams is dependent on the inhibition of at least two of the multiple penicillin-binding proteins (PBPs). Indeed, because multiple PBPs can be deleted with no effect on phenotype[47], the strategy of single target essentiality would not have discovered this important class of antibacterial drug targets. Similarly, fluoroquinolone antibiotics are dual-targeted inhibitors of the proteins ParC and GyrA (ref. 48). D-Cycloserine inhibits four targets, both pairs of alanine racemases and D-Ala-D-Ala ligases. Likewise, fosfomycin overcomes the redundancy of UDP N-acetylglucosamine enolpyruvyl transferases by inhibiting them both. Therefore, if we wish to design single drugs that limit drug resistance, we could consider the development of methods to search and prioritize which combinations of targets can be inhibited by the same drug and are essential, either individually ('dual essentials') or in combination ('synthetic lethals').

A strategy of antibacterial polypharmacology can challenge the current approach to genome-based drug discovery of anti-infectives in four important ways. First, the druggability of a target is prioritized over single target essentiality. Second, the target does not need to be unique to the organism or absent in the host. Although many essential housekeeping enzymes may be common between the host and infectious agent, drug selectivity between the host and infectious agent can be determined at the binding site level. Third, targets are sought that are lethal in combination but may have been overlooked as non-essential in individual gene knockout studies, and fourth, groups of targets that are predicted to potentially bind the same compound are prioritized over singleton druggable targets. By targeting two or more essential genes with a single chemical agent, the ability to delay drug resistance is designed into the target discovery strategy from the start. Given the failure of current genome-based strategies for discovering new antibacterial drugs[45], learning the lessons of the previous successful generation of antibacterial drugs may encourage the development of antibacterial polypharmacology discovery strategies.

Topology of targets

The two key challenges facing the development of network pharmacology are identifying a node or combination of nodes in a biological network whose perturbation results in a desired therapeutic outcome[49,50], and discovering agents with the desired polypharmacology profile to perturb those nodes. Three complementary methods for the comparative analysis of disease networks are systematic screening, knowledge-based combinations and network analysis. Presently, combination screening of mixtures of drugs, chemical tools and RNAi in cell-based disease models is the most efficient means of systematically discovering new drug-drug combinations and synthetically lethal gene pairs. However, as with all preclinical models, active combinations discovered in the laboratory do not necessarily translate into the clinic[51]. Moreover, idealized synergist screens between drugs or RNAi in preclinical assays do not account for the complications of dosing, scheduling pharmacokinetics and metabolism necessary to optimize a therapeutic drug cocktail in the clinic. Owing to the size of the search matrix, RNAi–chemical sensitization screens are usually performed with a single drug against a whole-genome RNAi array[34]. Likewise, systematic combination screening of approximately 1,000 US Food and Drug Administration (FDA)-approved drugs required the use of high-throughput screening methods to assay the massive data matrix required for the factorial isobolgram analysis of each combination through the full dose range[52,53]. Therefore, if we are to effectively search for new drug combinations, there is a pressing need for computational methods that could reduce the global search space for target combinations[49,51].

Owing to the vast number of possible drug and target combinations and the ethical considerations and resource constraints on using *in vivo* models and conducting clinical trials, most new combinations have been selected for empirical testing based on a knowledge of the underlying disease biology. Such knowledge-based approaches are often incremental but can make a dramatic impact on disease outcomes, as demonstrated by the success of multidrug highly active antiretroviral therapy in decreasing human immunodeficiency virus (HIV) mortality rates in the developed world. Advances in pathway analysis[54] and text mining of the biomedical literature[55–58] can potentially be used to enable the large-scale text mining of disease knowledge to postulate new combination hypotheses by associative techniques of inductive and abductive inference[59,60]. However, though advances in informatics methods can aid the generation of new hypotheses from connecting concepts in the literature, the drawback of the knowledge-based methods is that they do not provide a robust modeling analysis of the emergent properties of a network and system when perturbed in new ways. Thus, counterintuitive, paradoxical and unexpected system responses cannot necessarily be predicted by these associative methods[39].

An intriguing possibility for systematic target identification is that the structures of biological networks themselves may provide valuable information in assessing targets and their combinations[14,17,38,61]. Early network analysis indicated the possibility of a direct correlation between lethality and the degree of connectivity of nodes, where highly connected hubs in protein interaction networks are more likely to be essential[62]. Subsequent re-analysis of the data challenged the relationship between the number of interactions of a protein and its essentiality[63]. However, the hypothesis that protein function relates to network topology has been strengthened by recent work that has refined the relationship between network topology and system function by focusing on 'betweenness centrality' (the number of nonredundant shortest paths traveling through a node[64,65]) and 'bridging centrality' (nodes between and connecting subgraph clusters defined by the ratio of the number of interactions of a neighboring node in a subgraph over the number of remaining edges in the subgraph[66]), in addition to the metric of the 'degree centrality' (the number of direct interactions intersecting a node[14]). Bottlenecks with high betweenness values tend to be better correlated with gene expression dynamics and essentiality than highly connected hubs[64,67]. These findings complement the analysis on 'party' and 'dates' hubs, which suggests that hubs with high betweenness values have pleiotropic functions across the network[68]. Unexpectedly, non-hub bottlenecks with transient interactions[65] and bridging proteins are less likely to be lethal than average and tend to be independently regulated[66]. Thus, given their position in communication between network clusters and their low lethality, bridging nodes have been suggested as potential drug targets, although modulation of the bridging targets themselves may still be indirect[66]. An initial network analysis on the current drug targets of approved drugs indicated that drug targets are commonly highly connected but not essential[69,70].

Despite initial challenges, several compelling theoretical and experimental studies support the hypothesis that network topology is an essential feature in the emergent system function of the protein when it is perturbed; thus, this gives hope that systematic network analyses may be a useful basis for developing methods to prioritize drug targets and combinations of targets[17,61,71–77]. However, if we are to successfully exploit network analysis for target identification, then we must recognize the fundamentally different dynamics and kinetics in the

Box 1 Polypharmacology playbook

Three strategies are available to the designers of multitarget therapies. The first strategy, which is the most conventional, is to prescribe multiple individual medications. Multidrug combination cocktails are the mainstay of highly active antiretroviral therapy for HIV and a large number of anticancer protocols. The drawback of prescribing multiple medications is patient compliance and the danger of drug-drug interactions. To overcome these issues, a second strategy is the development of multicomponent drugs that contain two or more active ingredients formulated in the same delivery device, such as a single pill, capsule or inhaler[22,52]. Several successful drug combinations have now been reformulated in single multicomponent medicines, such as Atripla, Advair, Caduet, Combivir, Epzicom, Rebetron and Truvada. Advances in formulation technologies are expanding the number of drug combinations that can be effectively combined into a single delivery mechanism. However, given the significant differences in pharmacokinetics, metabolisms and bioavailability, reformulation of drug combinations is not a trivial problem. Further, two drugs that are generally safe when dosed individually cannot be assumed to be safe in combination. In addition to the possibility of adverse drug-drug interactions, if the theory of network pharmacology indicates that an effect on phenotype may derive from hitting multiple targets, then that combined phenotypic perturbation may be efficacious or deleterious. The major challenge to both drug combination strategies is the regulatory requirement for each individual drug to be shown to be safe as an individual agent and in combination. Therefore, most multicomponent drug development has focused on exploiting combinations of approved drugs, with the notable exception of the torcetrapib-atorvastatin combination that was being developed by Pfizer. A drawback of this approach to drug combinations is that the target universe of the current pharmacopeia is very limited. Current estimates are that the entire formulary of approximately 1,200 FDA-approved drugs only acts on about 320 molecular targets[115] (excluding the incredibly promiscuous kinase drug sunitinib (Sutent, SU11248), which itself binds to 79 protein kinases with $K_d < 10\ \mu M$)[116].

The third strategy for multitarget therapy is to design a single compound with selective polypharmacology[25,27,81]. The advent of wide-ligand profiling has revealed the extent of polypharmacology across the pharmacopeia, and it has also shown that many approved drugs act on multiple targets[25,115]. Dosing with a single compound may have advantages over a drug combination in terms of equitable pharmacokinetics and biodistribution. Indeed, troughs in drug exposure due to incompatible pharmacokinetics between components of a combination therapy may create a low-dose window of opportunity where a reduced selection pressure can lead to drug resistance. In terms of drug registration, approval of a single compound acting on multiple targets faces significantly lower regulatory barriers than approval of a combination of new drugs.

way drugs perturb networks compared to genetic deletions. Genetic deletions completely remove all the interactions and functions of the node from a network, whereas a drug may only partially ablate some interactions[78] (such as decreasing the concentration of a metabolite) but leave other links fully intact (such as protein-protein interactions). In contrast, agonists may strengthen particular links in a network. Thus, modeling attacks on links rather than nodes should provide a closer model of drug action[71,78]. Furthermore, an essential point is that only about 15% of any protein nodes in a network may be chemically tractable with small-molecule drugs[8]; therefore, it is necessary to map druggability and polypharmacology interactions on integrated biological networks in order to identify the optimal points of interaction for drug discovery.

Designer polypharmacology

Network analysis does not preclude the identification of individual targets that by virtue of their position in a disease network could be modulated to achieve a beneficial clinical outcome. Indeed, the theoretical work on synthetic rescue suggests that inactivation of one node, for example, by mutation or environmental factors, could be compensated for by the therapeutic inactivation of a second node[15,79]. Indeed, in terms of synthetic lethality, oncogenic mutations may themselves form half of a synthetically lethal pair, thereby resulting in the need to pharmacologically inhibit only one target[34,35].

However, if drug hunters are to embrace the wider opportunity posed by network pharmacology, then what is required is a 'playbook' of polypharmacology strategies to design therapies that act on several nodes in a disease network (**Box 1**). For good reason, medicinal chemists have tried to decrease the off-target effects of drug candidates to try to decrease the chances of off-target toxicities. Analysis of the Bioprint database of the complete screening matrix of FDA-approved drugs against approximately 200 assays reveals a strong relationship between calculated lipophilicity (*clogP*) and low-affinity off-target promiscuity[80]. The number of potential off-target activities appears to double above the value *clogP* = 3.75 (ref. 80). The goal of polypharmacology is not to lazily introduce nonspecific promiscuity into a compound by increasing the lipophilicity but to identify a compound with a desired biological profile across multiple targets whose combined modulation will perturb a disease state. Thus, understanding the broader polypharmacology profile of a compound and rationally modifying its profiles should equally benefit safety pharmacology as well as disease efficacy.

Specific design strategies are required to balance a set of biological activities in one compound. Morphy *et al.* have described the continuum of design strategies medicinal chemists have traditionally used to design drugs with multiple activities[81]. At one end of the spectrum are conjugated ligands, which contain separate pharmacopoeia entities connected by a linker. Ligands designed by conjugating two distinct pharmacophores are more likely to have high molecular weight and less likely to have oral drug-like physicochemical properties[82]. At the other end of the spectrum are ligands where multiple pharmacophores overlap or are highly integrated. Compounds with overlapping or integrated pharmacophores are likely to have lower molecular weight and potentially more drug-like physicochemical properties. Ligands where polypharmacology has been deliberately designed in, by conjugation or overlapping pharmacophores, tend to have lower ligand efficiency than general preclinical compounds. This finding is not surprising given that the compounds are not optimized for one single target.

Designers of polypharmacology can also take lessons from the work on drug resistance, particularly in the field of HIV-1 therapies. Mutations of single amino acid residues are often sufficient to confer drug resistance to many anti-HIV drugs, such as the non-nucleoside reverse transcriptase inhibitors (NNRTIs). Efforts to discover second-generation NNRTIs required compounds to be active against both the wild type and the common drug-resistant mutations. Thus, second-generation NNRTIs can be considered as multitarget drugs as they are required to bind to several structurally distinct binding sites on a spectrum of mutated reverse transcriptases. Crystallographic analyses of the NNRTIs revealed that one design strategy is to identify inhibitors that make strong molecular interactions with conserved regions of the binding site, such as structurally important residues and main chain atoms; this has the added advantage of reducing dependence on interactions with mutable residues[83]. Several of these design strategies have also been derived a priori by theoretical analysis of drug-resistant target ensembles[53].

In addition to the relationship between lipophilicity and promiscuity, a strong relationship has also been reported between molecular weight[25,82] and molecular complexity[84] and promiscuity. In an

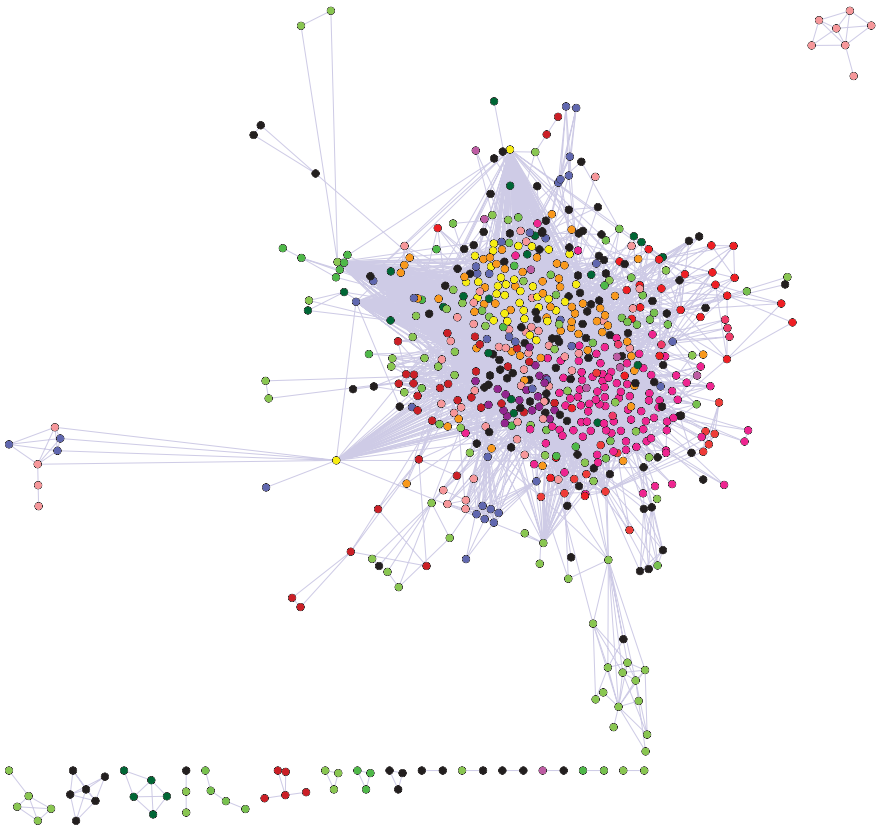

Figure 1 Human polypharmacology interactions network at ten-fold selectivity. Network representations of the integrated chemogenomics space defined by Paolini *et al.*[90]. Two proteins (nodes) are defined as interacting in chemical space (edges) if they bind at least 10% of shared screened compounds with a difference in potency of only ten-fold below an activity cutoff of 10 μM (that is, if a compounds exhibits an IC_{50} = 10 nM against target A, it must show an activity below IC_{50} = 100 nM against target B to be considered interacting in this network). At these thresholds, 675 proteins are connected by 10,016 interactions in the total network by at least one compound. Nodes are color-coded by gene family: aminergic G protein–coupled receptors (GPCRs), yellow; peptide GPCRs, orange; other GPCRs, light pink; ion channels, light blue; nuclear hormone receptors, brown; phosphodiesterases, purple; protein kinases, pink; enzymes, green; proteases, red; others, black.

analysis of a corporate screening database of 70,000 biologically active small molecules screened over 200 molecular assays, Hopkins *et al.* observed a tight relationship between molecular weight and the polypharmacology of a compound, in IC_{50} measurements below 10 μM[25]. These empirical observations are complemented by theoretical work by Hann *et al.*, in which using a simple model of ligand-receptor pharmacophoric interaction revealed that the probability of a randomly chosen ligand binding decreases precipitously as the ligand becomes more complex[84]. Assuming that molecular weight is a proxy measure of molecular complexity, these observations from experiments and models provide the foundation for understanding the success of fragment-based lead discovery[85], in which a small library (500–2,000) of low-molecular-weight (100–250 Da) compounds can be successfully used to find ligand-efficient[86] 'hits' against a large number of molecular targets. Therefore, it has been proposed that multitarget fragment screens could be a promising approach for the discovery of drug-like promiscuous ligands[87]. In particular, surface plasmon resonance[88] may be well suited to multitarget fragment screening owing to the increased sensitivity of the new generation of instruments, the low protein consumption required compared with NMR- and X-ray–based fragment screening, and the multiple channels that can be screened simultaneously[89].

Alternatively, advances in cheminformatics and the availability of large-scale structure-activity relationship (SAR) databases provide the tools necessary to develop computational methods of multitarget design. Large-scale, integrated chemogenomics knowledge bases, such as that described by Paolini *et al.*, enable the systematic search across large datasets of integrated structure-activity data for compounds that are observed or predicted to bind to multiple targets[90]. For example, 35% of biologically active compounds in the data warehouse built by Paolini *et al.* are observed to bind to more than one target, and the majority are active against targets within the same gene family. However, as we observed from the structure of a polypharmacology interaction network that maps 12,000 interactions of 700 human proteins in chemical space, there is a surprising level of interaction between gene families (**Fig. 1**). Though the predicted number of druggable targets may be a relatively small fraction of the proteome, the observed number of chemically tractable combinations (with integrated pharmacophores) is over an order of magnitude larger[90].

Wermuth expands this logic into a design strategy that can be used for multitarget drug discovery called selective optimization of side activities (SOSA)[91,92]. The SOSA idea provides a pragmatic approach to designing polypharmacology: rather than attempting the difficult problem of merging and integrating pharmacophores, the starting point is an integrated pharmacophore that already provides some of the nascent activity profile. An opportunistic strategy may be to investigate, given the polypharmacology profile of a particular compound, what phenotypic behavior it exhibits. For example, an analysis of the polypharmacology interaction network between drug targets associated with asthma is shown in **Figure 2**. An SOSA analysis of this graph provides a wealth of opportunities for identifying lead series that already exhibit interesting mixtures of pharmacology (for example, the edges between the nodes) that were identified by integrating a number of SAR data sources. As in all drug discovery, the choice of lead matter is crucial in determining the success of a project. Traditionally, medicinal chemists have attempted to combine activities to create multitarget drugs. A more productive approach may be to focus efforts on the discovery of lead compounds with interesting multitarget profiles from systematic data mining or multitarget screening. Thus, one of the most pragmatic applications of network pharmacology thinking could be to systematically reassess drug candidates and drug discovery programs in the new light of understanding their wider biological activity profiles[93].

Multitarget design

The key to identifying multitarget drugs is in appreciating their limits. Although the opportunity space for compounds with specific polypharmacology profiles may be significantly larger than the

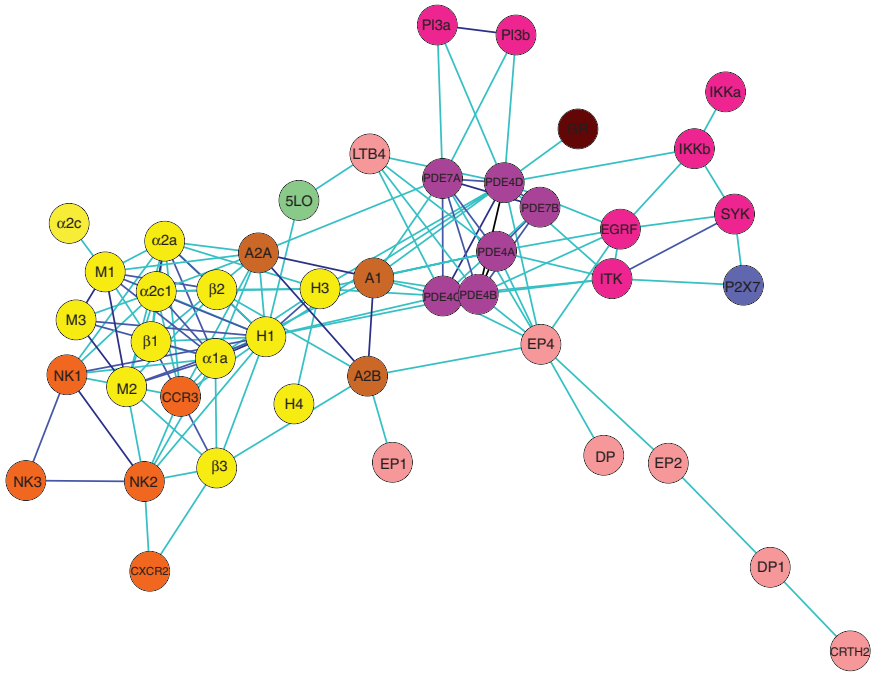

Figure 2 Expanding opportunity for drug discovery space with polypharmacology. A subset of the network data shown in **Figure 1** for literature targets associated with asthma. Drug targets are represented as nodes, and chemical matter that binds to two or more nodes is represented as edges. Targets are colored by gene family. The color of the edges represents the strength of the chemical network between two targets as defined by the number of shared compounds that are active against both targets below an affinity of 1 μM: light blue (1 to 10 compounds) to black (>1,000 compounds). Of the 44 targets described in the literature as potential drug targets for the treatment of asthma, 44 share polypharmacology of existing chemical matter with another potential target. These 44 targets are identified for 137 target combinations across >10,000 compounds in this portfolio. Thus, by considering both single-target and dual pharmacology approaches, at least 181 potential profile opportunities can be examined. As this is an analysis of known chemical matter and biological activities, many of these profiles could be tested immediately in appropriate disease models. The network is represented in Cytoscape[117]. Drug targets are color-coded by gene family: aminergic GPCRs, yellow; peptide GPCRs, orange; lipophilic GPCRs, light pink; ion channels, light blue; nuclear hormone receptors, brown; phosphodiesterases, purple; protein kinases, pink; enzymes, green.

opportunity universe for single-target drugs (which is outlined by the druggable genome[8]), not all target pair combinations will be accessible to a single agent with drug-like properties. Predicting chemically tractable combinations can increase the chances of finding a multi-target compound. In order to develop a drug with a desired polypharmacology profile, two problems need to be solved. First, a lead compound needs to be identified with the desired biological activity against multiple targets; then, this lead needs to be optimized into a clinical candidate that combines the desired polypharmacology with a safe, drug-like pharmaceutical profile.

The integration of *in silico* methods, combined with wide-ligand biological profiling against protein assays and gene expression arrays, can provide drug designers with a new toolbox with which to assess polypharmacology. First, proteins can be related by binding exact or similar endogenous ligands or proteins; this can be determined by exploiting data in gene ontologies and metabolic databases such as KEGG[94,95]. Second, proteins can be related by known observed polypharmacology of large sets of ligands such as those found in pharmaceutical company screening sets and the medicinal chemistry literature. A chemical network can be created that relates proteins that bind the same ligand, where the strength of each edge can be calculated using metrics of promiscuity[69,70,90,96–99].

If experimentally derived ligand data are not available, proteins can be related by predictions of polypharmacology[90,100–104]. Sets of ligands for each protein obtained from chemogenomic data sources can be used to train machine learning algorithms to predict pharmacology activity profiles. Bayesian approaches can be used to classify chemical structures based on chemical fingerprints[90,100,101]. Dynamically cross-comparing the chemical similarity of sets of ligands for each protein is a complementary and promising method of predicting polypharmacology profiles from chemical structure[100,101]. The structure of the observed polypharmacology network itself also provides information beyond chemical structure similarity that could also be utilized using Bayesian network methods, as has been the case in protein-protein interaction networks[105]. Though promising, chemical fingerprint–based similarity methods paradoxically are incompatible with the two simplest relationships—those of lipophilicity and molecular weight with promiscuity—as larger compounds tend to exhibit a greater number of fingerprints. The development of improved polypharmacology prediction methods will be an important topic in cheminformatics and toxicology. Complementary to chemogenomics approaches to predict drug-target relationships is the use of phenotypic data (such as clinical side effects)[104,106] or gene expression profiles[107] to cluster chemical structures by functional effects. Clustering compounds by phenotypic and gene signature profiles enables unknown mechanisms to be inferred.

Structural bioinformatics methods that map gene family sequence alignments onto the structural information of protein binding sites provide a valuable first-order assessment of likely selectivity and promiscuity within a gene family[95,108–110]. Comparing protein binding site similarity with the observed polypharmacology networks could provide insight into protein binding site parameters that map with the behavior of chemical networks. For example, analysis of the protein kinase superfamily in the human genome reveals discrete clusters of subfamilies when the full-length sequences are compared (**Fig. 3**). However, these subfamily clusters break down when the sequence similarity is measured at the level of the ATP binding site, where most kinase inhibitors bind. Hence, at a first approximation the difficulties drug designers have in fine tuning the selectivity profiles of competitive kinase inhibitors is not surprising[14,111,112]. Despite the inherent challenge of achieving finely tuned kinase activity profiles, deriving kinase SAR data[112] from the wealth of large-scale chemogenomics analyses and using subtle modification of existing kinase scaffolds to exploit 'de-hydrons'[111] are two promising strategies. Finally, if three-dimensional protein structures of the targets of interest are available, then parallel large-scale multitarget virtual screening is also a promising method[113].

a

>gil1170188|sp|P08631|HCK_HUMAN TYROSINE-PROTEIN KINASE HCK

MGGRSSCEDPGCPRDEERAPRMGSMKSKFLQVGGNTFSKTETSASPHCPVYVPDPTSTIKPGPNSHNSNTPGIREAGSEDIIVVALYDYEAIHHEDL
SFQKGDQMVVLEESGEWWKARSLATRKEGYIPSNYVARVDSLETEEWFFKGISRKDAERQLLAPGNMLGSFMIRDSETTKGSYSLSVRDYDPRQGDT
VKHYKIRTLDNGGFYISPRSTFSTLQELVDHYKKGNDGLCQKLSVPCMSSKPQKPWEKDAWEIPRESLKLEKKLGAGQFGEVWMATYNKHTKVAVKT
MKPGSMSVEAFLAEANVMKTLQHDKLVKLHAVVTKEPIYIITEFMAKGSLLDFLKSDEGSKQPLPKLIDFSAQIAEGMAFIEQRNYIHRDLRAANIL
VSASLVCKIADFGLARVIEDNEYTAREGAKFPIKWTAPEAINFGSFTIKSDVWSFGILLMEIVTYGRIPYPGMSNPEVIRALERGYRMPRPENCPEE
LYNIMMRCWKNRPEERPTFEYIQSVLDDFYTATESQYQQQP

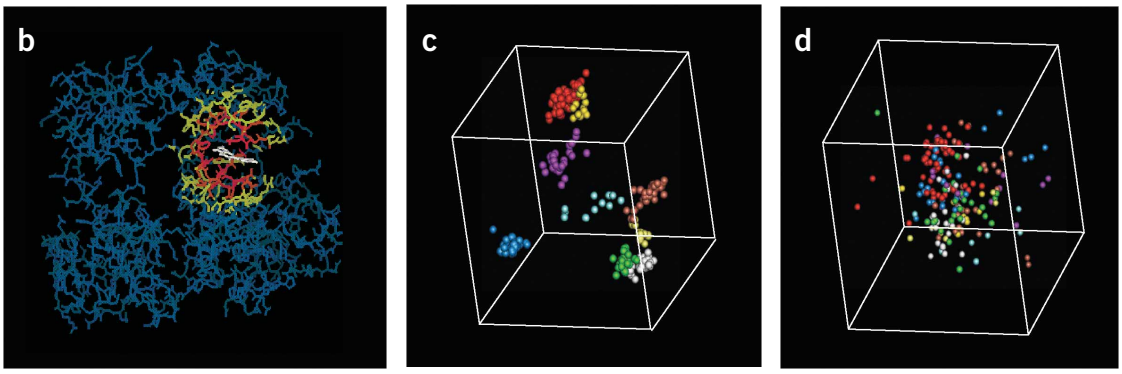

Figure 3 Protein kinase inhibitor promiscuity as a function of binding site sequence similarity. (**a**,**b**) The full-length sequence of human protein tyrosine kinase HCK (**a**), where the amino acids surrounding the ATP binding site are color-coded by their distance from the binding site surface when mapped onto the canonical protein kinase structure (**b**). (**c**) Multidimensional scaling of the human kinome to cluster kinases using full sequences reveals that the kinases cluster into discrete families. (**d**) Multidimensional scaling of the same kinases using the binding site–weighted sequences as defined in **a** and **b** reveals the breakdown of the subfamily clustering and the similarity at the binding site level of many diverse protein kinases. Graphic courtesy of Colin Groom (Cambridge Crystallographic Data Centre, Cambridge, UK).

Fusing multiple prediction methods is likely to improve the overall success rates of *in silico* lead identification. Once a lead is identified, the second computational design problem lies in the constraints of optimizing in multiple dimensions. If we assume, as the data suggest, that very few drugs are truly selective, then most biologically active small molecules have a degree of promiscuity by their nature. The challenge for medicinal chemists is to understand the profile of each compound, to fine tune the profile, and then to select for clinical development those compounds whose profiles maximally modulate a disease network with the minimum level of toxicity and side effects.

Conclusion

Network pharmacology is an approach to drug design that encompasses systems biology, network analysis, connectivity, redundancy and pleiotropy. Network pharmacology offers a way of thinking about drug discovery that simultaneously embraces efforts to improve clinical efficacy and understand side effects and toxicity—two of the most important reasons for failure. A variety of studies have shown the power of network analysis in understanding biological systems. Furthermore, emergent phenotypes beyond those seen in single-gene deletion experiments have been observed through synthetic behaviors, combinations and chemical biology probes. The biological rationale for considering multitarget strategies over single-target approaches is compelling, yet such strategies are at present a minority activity in the pharmaceutical industry. The reason is that optimizing multiple activities, while trying to balance drug-like properties and control unwanted off-target effects, is a difficult task. We do not yet have a robust set of design tools with which to apply this approach routinely. Structure-based drug design took nearly two decades of multiple, parallel technological improvements to arrive at its current mainstream position in medicinal chemistry. Developments in computer graphics, high-power radiation sources, computational processing power, refinement protocols, virtual screening and cryocrystallography were all necessary to create the environment for rapid, iterative structure-based drug discovery. To make network pharmacology commonplace, a different set of tools, concerned with combinatorial and network search algorithms and methods for predicting the biological profiles, will need to be refined. Network pharmacology re-introduces the old idea that understanding the biological and kinetic profile of the drug is more important than individual validation of targets or combinations of targets. In many ways, the network pharmacology strategy outlined here is a modern reinvention of Paul Janssen's original and incredibly successful methods of drug discovery, where side activities of compounds are explored through a broad spectrum of structure-activity relationships[114]. Given the crisis in translation facing the pharmaceutical industry, network pharmacology offers a new framework for thinking about how to innovate drug discovery, and thus it is an idea whose time has come.

Published online at http://www.nature.com/naturechemicalbiology/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Kola, I. & Landis, J. Can the pharmaceutical industry reduce attrition rates? *Nat. Rev. Drug Discov.* **3**, 711–716 (2004).
2. Sams-Dodd, F. Target-based drug discovery: is something wrong? *Drug Discov. Today* **10**, 139–147 (2005).
3. Kaufmann, S.H.E. Paul Ehrlich: founder of chemotherapy. *Nat. Rev. Drug Discov.* **7**, 373 (2008).
4. Zambrowicz, B.P. & Sands, A.T. Modeling drug action in the mouse with knockouts and RNA interference. *Drug Discov. Today Targets* **3**, 198–207 (2004).
5. Winzeler, E.A. *et al.* Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* **285**, 901–906 (1999).
6. Giaever, G. *et al.* Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* **418**, 387–391 (2002).
7. Deutschbauer, A.M. *et al.* Mechanisms of haploinsufficiency revealed by genome-wide profiling in yeast. *Genetics* **169**, 1915–1925 (2005).
8. Hopkins, A.L. & Groom, C.R. The druggable genome. *Nat. Rev. Drug Discov.* **1**, 727–730 (2002).

9. Austin, C.P. *et al.* The knockout mouse project. *Nat. Genet.* **36**, 921–924 (2004).
10. Zambrowicz, B.P., Turner, C.A. & Sands, A.T. Predicting drug efficacy: knockouts model pipeline drugs of the pharmaceutical industry. *Curr. Opin. Pharmacol.* **3**, 563–570 (2003).
11. Zambrowicz, B.P. & Sands, A.T. Knockouts model the 100 best-selling drugs–will they model the next 100? *Nat. Rev. Drug Discov.* **2**, 38–51 (2003).
12. Kitano, H. Towards a theory of biological robustness. *Mol. Syst. Biol.* **3**, 137 (2007).
13. Barabási, A.L. & Oltvai, Z.N. Network biology: understanding the cell's functional organization. *Nat. Rev. Genet.* **5**, 101–113 (2004).
14. Albert, R., Jeong, H. & Barabasi, A.L. Error and attack tolerance of complex networks. *Nature* **406**, 378–382 (2000).
15. Chen, Y. *et al.* Variations in DNA elucidate molecular networks that cause disease. *Nature* **452**, 429–435 (2008).
16. Csermely, P., Agoston, V. & Pongor, S. The efficiency of multi-target drugs: the network approach might help drug design. *Trends Pharmacol. Sci.* **26**, 178–182 (2005).
17. Korcsmáros, T., Szalay, M.S., Böde, C., Kovács, I. & Csermely, P. How to design multi-target drugs: target search options in cellular networks. *Expert Opin. Drug Discov.* **2**, 1–10 (2007).
18. Ooi, S.L. *et al.* Global synthetic-lethality analysis and yeast functional profiling. *Trends Genet.* **22**, 56–63 (2006).
19. Hillenmeyer, M.E. *et al.* The chemical genomic portrait of yeast: uncovering a phenotype for all genes. *Science* **320**, 362–365 (2008).
20. Roth, B.L., Sheffler, D.J. & Kroeze, W.K. Magic shotguns versus magic bullets: selectively non-selective drugs for mood disorders and schizophrenia. *Nat. Rev. Drug Discov.* **3**, 353–359 (2004).
21. Wermuth, C.G. Multitarget drugs: the end of the 'one-target-one-disease' philosophy? *Drug Discov. Today* **9**, 826–827 (2004).
22. Keith, C.T., Borisy, A.A. & Stockwel, B.R. Multicomponent therapeutics for networked systems. *Nat. Rev. Drug Discov.* **4**, 71–78 (2005).
23. Petrelli, A. & Giordano, S. From single- to multi-target drugs in cancer therapy: when aspecificity becomes an advantage. *Curr. Med. Chem.* **15**, 422–432 (2008).
24. Mencher, S.K. & Wang, L.G. Promiscuous drugs compared to selective drugs (promiscuity can be a virtue). *BMC Clin. Pharmacol.* **5**, 3 (2005).
25. Hopkins, A.L., Mason, J.S. & Overington, J.P. Can we rationally design promiscuous drugs? *Curr. Opin. Struct. Biol.* **16**, 127–136 (2006).
26. Flordellis, C.S., Manolis, A.S., Paris, H. & Karabinis, A. Rethinking target discovery in polygenic diseases. *Curr. Top. Med. Chem.* **6**, 1791–1798 (2006).
27. Dessalew, N. & Workalemahu, M. On the paradigm shift towards multitarget selective drug design. *Curr. Comput. Aided Drug Des.* **4**, 76–90 (2008).
28. Keskin, O., Gursoy, A., Ma, B. & Nussinov, R. Towards drugs targeting multiple proteins in a systems biology approach. *Curr. Top. Med. Chem.* **7**, 943–951 (2007).
29. Zimmermann, G.R., Lehár, J. & Keith, C.T. Multi-target therapeutics: when the whole is greater than the sum of the parts. *Drug Discov. Today* **12**, 34–42 (2007).
30. McGovern, S.L., Caselli, E., Grigorieff, N. & Shoichet, B.K. A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening. *J. Med. Chem.* **45**, 1712–1722 (2002).
31. Kerkelä, R. *et al.* Cardiotoxicity of the cancer therapeutic agent imatinib mesylate. *Nat. Med.* **12**, 908–916 (2006).
32. Kaelin, W.G., Jr. The concept of synthetic lethality in the context of anticancer therapy. *Nat. Rev. Cancer* **5**, 689–698 (2005).
33. Gascoigne, K.E. & Taylor, S.S. Cancer cells display profound intra- and interline variation following prolonged exposure to antimitotic drugs. *Cancer Cell* **14**, 111–122 (2008).
34. Whitehurst, A.W. *et al.* Synthetic lethal screen identification of chemosensitizer loci in cancer cells. *Nature* **446**, 815–819 (2007).
35. Turner, N.C. *et al.* A synthetic lethal siRNA screen identifying genes mediating sensitivity to a PARP inhibitor. *EMBO J.* **27**, 1368–1377 (2008).
36. Gupta, G.P. *et al.* Mediators of vascular remodelling co-opted for sequential steps in lung metastasis. *Nature* **446**, 765–770 (2007).
37. Eltarhouny, S.A. *et al.* Genes controlling spread of breast cancer to lung "gang of 4". *Exp. Oncol.* **30**, 91–95 (2008).
38. Kitano, H. A robustness-based approach to systems-oriented drug design. *Nat. Rev. Drug Discov.* **6**, 202–210 (2007).
39. Bond, R.A. Can intellectualism stifle scientific discovery? *Nat. Rev. Drug Discov.* **1**, 825–829 (2002).
40. Dudekula, N., Arora, V., Callaerts-Vegh, Z. & Bond, R.A. The temporal hormesis of drug therapies. *Dose Response* **3**, 414–424 (2006).
41. Schwartz, G.K. & Shah, M.A. Targeting the cell cycle: a new approach to cancer therapy. *J. Clin. Oncol.* **23**, 9408–9421 (2005).
42. Mills, S.D. When will the genomics investment pay off for antibacterial discovery? *Biochem. Pharmacol.* **71**, 1096–1102 (2006).
43. Pucci, M.J. Use of genomics to select antibacterial targets. *Biochem. Pharmacol.* **71**, 1066–1072 (2006).
44. Payne, D.J., Gwynn, M.N., Holmes, D.J. & Rosenberg, M. Genomic approaches to antibacterial discovery. *Methods Mol. Biol.* **266**, 231–259 (2004).
45. Payne, D.J., Gwynn, M.N., Holmes, D.J. & Pompliano, D.L. Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nat. Rev. Drug Discov.* **6**, 29–40 (2007).
46. Lange, R.P., Locher, H.H., Wyss, P.C. & Then, R.L. The targets of currently used antibacterial agents: lessons for drug discovery. *Curr. Pharm. Des.* **13**, 3140–3154 (2007).
47. Denome, S.A., Elf, P.K., Henderson, T.A., Nelson, D.E. & Young, K.D. *Escherichia coli* mutants lacking all possible combinations of eight penicillin binding proteins: viability, characteristics, and implications for peptidoglycan synthesis. *J. Bacteriol.* **181**, 3981–3993 (1999).
48. Janoir, C., Zeller, V., Kitzis, M.D., Moreau, N.J. & Gutmann, L. High-level fluoroquinolone resistance in *Streptococcus pneumoniae* requires mutations in parC and gyrA. *Antimicrob. Agents Chemother.* **40**, 2760–2764 (1996).
49. Ramaswamy, S. Rational design of cancer-drug combinations. *N. Engl. J. Med.* **357**, 299–300 (2007).
50. Mayer, L.D. & Janoff, A.S. Optimizing combination chemotherapy by controlling drug ratios. *Mol. Interv.* **7**, 216–223 (2007).
51. Dancey, J.E. & Chen, H.X. Strategies for optimizing combinations of molecularly targeted anticancer agents. *Nat. Rev. Drug Discov.* **5**, 649–659 (2006).
52. Borisy, A.A. *et al.* Systematic discovery of multicomponent therapeutics. *Proc. Natl. Acad. Sci. USA* **100**, 7977–7982 (2003).
53. Radhakrishnan, M.L. & Tidor, B. Optimal drug cocktail design: methods for targeting molecular ensembles and insights from theoretical model systems. *J. Chem. Inf. Model.* **48**, 1055–1073 (2008).
54. Tsui, I.F., Chari, R., Buys, T.P. & Lam, W.L. Public databases and software for the pathway analysis of cancer genomes. *Cancer Inform.* **3**, 389–407 (2007).
55. Swanson, D.R. Medical literature as a potential source of new knowledge. *Bull. Med. Libr. Assoc.* **78**, 29–37 (1990).
56. Wren, J.D., Bekeredjian, R., Stewart, J.A., Shohet, R.V. & Garner, H.R. Knowledge discovery by automated identification and ranking of implicit relationships. *Bioinformatics* **20**, 389–398 (2004).
57. Bekhuis, T. Conceptual biology, hypothesis discovery, and text mining: Swanson's legacy. *Biomed. Digit. Libr.* **3**, 2 (2006).
58. Loging, W., Harland, L. & Williams-Jones, B. High-throughput electronic biology: mining information for drug discovery. *Nat. Rev. Drug Discov.* **6**, 220–230 (2007).
59. Peirce, C.S. *The Essential Peirce* Vol. 2 (Indiana University Press, Bloomington, Indiana, USA, 1998).
60. Kell, D.B. & Oliver, S.G. Here is the evidence, now what is the hypothesis? The complementary roles of inductive and hypothesis-driven science in the post-genomic era. *Bioessays* **26**, 99–105 (2004).
61. Latora, V. & Marchiori, M. Vulnerability and protection of infrastructure networks. *Phys. Rev. E* **71**, 015103R (2005).
62. Jeong, H., Mason, S.P., Barabási, A.L. & Oltvai, Z.N. Lethality and centrality in protein networks. *Nature* **411**, 41–42 (2001).
63. Coulomb, S., Bauer, M., Bernard, D. & Marsolier-Kergoat, M.C. Gene essentiality and the topology of protein interaction networks. *Proc. Biol. Sci.* **272**, 1721–1725 (2005).
64. Joy, M.P., Brock, A., Ingber, D.E. & Huang, S. High-betweenness proteins in the yeast protein interaction network. *J. Biomed. Biotechnol.* **2**, 96–103 (2005).
65. Yu, H., Kim, P.M., Sprecher, E., Trifonov, V. & Gerstein, M. The importance of bottlenecks in protein networks: correlation with gene essentiality and expression dynamics. *PLoS Comput. Biol.* **3**, e59 (2007).
66. Hwang, W.C., Zhang, A. & Ramanathan, M. Identification of information flow-modulating drug targets: a novel bridging paradigm for drug discovery. *Clin. Pharmacol. Ther.* published online, doi:10.1038/clpt.2008.129 (9 July 2008).
67. Hahn, M.W. & Kern, A.D. Comparative genomics of centrality and essentiality in three eukaryotic protein-interaction networks. *Mol. Biol. Evol.* **22**, 803–806 (2005).
68. Han, J.D. *et al.* Evidence for dynamically organized modularity in the yeast protein-protein interaction network. *Nature* **430**, 88–93 (2004).
69. Yildirim, M.A., Goh, K.I., Cusick, M.E., Barabási, A.L. & Vidal, M. Drug-target network. *Nat. Biotechnol.* **25**, 1119–1126 (2007).
70. Nacher, J.C. & Schwartz, J.M. A global view of drug-therapy interactions. *BMC Pharmacol.* **8**, 5 (2008).
71. Motter, A.E., Nishikawa, T. & Lai, Y. Range-based attack on links in scale-free networks: are long-range links responsible for the small-world phenomenon? *Phys. Rev. E* **66**, 065103 (2002).
72. Moriya, H., Shimizu-Yoshida, Y. & Kitano, H. *In vivo* robustness analysis of cell division cycle genes in *Saccharomyces cerevisiae*. *PLoS Genet.* **2**, e111 (2006).
73. Wunderlich, Z. & Mirny, L.A. Using the topology of metabolic networks to predict viability of mutant strains. *Biophys. J.* **91**, 2304–2311 (2006).
74. Kovacs, I., Csermely, P., Korcsmaros, T. & Szalay, M. WO patent application WO 2007-IB50471 2007 0213 (2007).
75. Watterson, S., Marshall, S. & Ghazal, P. Logic models of pathway biology. *Drug Discov. Today* **13**, 447–456 (2008).
76. Gerber, S., Assmus, H., Bakker, B. & Klipp, E. Drug-efficacy depends on the inhibitor type and the target position in a metabolic network—a systematic study. *J. Theor. Biol.* **252**, 442–455 (2008).
77. Potapov, A.P., Goemann, B. & Wingender, E. The pairwise disconnectivity index as a new metric for the topological analysis of regulatory networks. *BMC Bioinformatics* **9**, 227 (2008).
78. Agoston, V., Csermely, P. & Pongor, S. Multiple weak hits confuse complex systems: a transcriptional regulatory network as an example. *Phys. Rev. E* **71**, 051909 (2005).
79. Motter, A.E., Gulbahce, N., Almaas, E. & Barabási, A.L. Predicting synthetic rescues in metabolic networks. *Mol. Syst. Biol.* **4**, 168 (2008).
80. Leeson, P.D. & Springthorpe, B. The influence of drug-like concepts on decision-making in medicinal chemistry. *Nat. Rev. Drug Discov.* **6**, 881–890 (2007).

81. Morphy, R., Kay, C. & Rankovic, Z. From magic bullets to designed multiple ligands. *Drug Discov. Today* **9**, 641–651 (2004).
82. Morphy, R. & Rankovic, Z. The physicochemical challenges of designing multiple ligands. *J. Med. Chem.* **49**, 4961–4970 (2006).
83. Hopkins, A.L. *et al.* Design of non-nucleoside inhibitors of HIV-1 reverse transcriptase with improved drug resistance properties. 1. *J. Med. Chem.* **47**, 5912–5922 (2004).
84. Hann, M.M., Leach, A.R. & Harper, G. Molecular complexity and its impact on the probability of finding leads for drug discovery. *J. Chem. Inf. Comput. Sci.* **41**, 856–864 (2001).
85. Leach, A.R., Hann, M.M., Burrows, J.N. & Griffen, E.J. Fragment screening: an introduction. *Mol. Biosyst.* **2**, 430–446 (2006).
86. Hopkins, A.L., Groom, C.R. & Alex, A. Ligand efficiency: a useful metric for lead selection. *Drug Discov. Today* **9**, 430–431 (2004).
87. Morphy, R. & Rankovic, Z. Fragments, network biology and designing multiple ligands. *Drug Discov. Today* **12**, 156–160 (2007).
88. Neumann, T., Junker, H.D., Schmidt, K. & Sekul, R. SPR-based fragment screening: advantages and applications. *Curr. Top. Med. Chem.* **7**, 1630–1642 (2007).
89. Hämäläinen, M.D. *et al.* Label-free primary screening and affinity ranking of fragment libraries using parallel analysis of protein panels. *J. Biomol. Screen.* **13**, 202–209 (2008).
90. Paolini, G.V., Shapland, R.H., van Hoorn, W.P., Mason, J.S. & Hopkins, A.L. Global mapping of pharmacological space. *Nat. Biotechnol.* **24**, 805–815 (2006).
91. Wermuth, C.G. Selective optimization of side activities: another way for drug discovery. *J. Med. Chem.* **47**, 1303–1314 (2004).
92. Wermuth, C.G. Selective optimization of side activities: the SOSA approach. *Drug Discov. Today* **11**, 160–164 (2006).
93. Millan, M. Multi-target strategies for the improved treatment of depressive states: conceptual foundations and neuronal substrates, drug discovery and therapeutic application. *Pharmacol. Ther.* **110**, 135–370 (2006).
94. Ji, H.F. *et al.* Distribution patterns of small-molecule ligands in the protein universe and implications for origin of life and drug discovery. *Genome Biol.* **8**, R176 (2007).
95. Park, K. & Kim, D. Binding similarity network of ligand. *Proteins* **71**, 960–971 (2008).
96. Kuhn, M., von Mering, C., Campillos, M., Jensen, L.J. & Bork, P. STITCH: interaction networks of chemicals and proteins. *Nucleic Acids Res.* **36**, D684–D688 (2008).
97. Günther, S. *et al.* SuperTarget and Matador: resources for exploring drug-target relationships. *Nucleic Acids Res.* **36**, D919–D922 (2008).
98. Spiro, Z., Kovacs, I.A. & Csermely, P. Drug-therapy networks and the prediction of novel drug targets. *J. Biol.* **7**, 20 (2008).
99. Bonchev, D. & Buck, G.A. From molecular to biological structure and back. *J. Chem. Inf. Model.* **47**, 909–917 (2007).
100. Keiser, M.J. *et al.* Relating protein pharmacology by ligand chemistry. *Nat. Biotechnol.* **25**, 197–206 (2007).
101. Hert, J., Keiser, M., Irwin, J., Oprea, T. & Shoichet, B. Quantifying the relationships among drug classes. *J. Chem. Inf. Model.* **48**, 755–765 (2008).
102. 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* **24**, i232–i240 (2008).
103. Kuhn, M., Campillos, M., González, P., Jensen, L.J. & Bork, P. Large-scale prediction of drug–target relationships. *FEBS Lett.* **582**, 1283–1289 (2008).
104. Campillos, M., Kuhn, M., Gavin, A.C., Jensen, L.J. & Bork, P. Drug target identification using side-effect similarity. *Science* **321**, 263–266 (2008).
105. Scott, M.S. & Barton, G.J. Probabilistic prediction and ranking of human protein-protein interactions. *BMC Bioinformatics* **8**, 239 (2007).
106. Fliri, A.F., Loging, W.T., Thadeio, P.F. & Volkmann, R.A. Biospectra analysis: model proteome characterizations for linking molecular structure and biological response. *J. Med. Chem.* **48**, 6918–6925 (2005).
107. Lamb, J. *et al.* The connectivity map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* **313**, 1929–1935 (2006).
108. Davies, J.R., Jackson, R.M., Mardia, K.V. & Taylor, C.C. The Poisson index: a new probabilistic model for protein ligand binding site similarity. *Bioinformatics* **23**, 3001–3008 (2007).
109. Baroni, M., Cruciani, G., Sciabola, S., Perruccio, F. & Mason, J.S. A common reference framework for analyzing/comparing proteins and ligands. Fingerprints for ligands and proteins (FLAP): theory and application. *J. Chem. Inf. Model.* **47**, 279–294 (2007).
110. Zhang, Z. & Grigorov, M.G. Similarity networks of protein binding sites. *Proteins* **62**, 470–478 (2006).
111. Zhang, X., Crespo, A. & Fernández, A. Turning promiscuous kinase inhibitors into safer drugs. *Trends Biotechnol.* **26**, 295–301 (2008).
112. Aronov, A.M., McClain, B., Moody, C.S. & Murcko, M.A. Kinase-likeness and kinase-privileged fragments: toward virtual polypharmacology. *J. Med. Chem.* **51**, 1214–1222 (2008).
113. Jenwitheesuk, E., Horst, J.A., Rivas, K.L., Van Voorhis, W.C. & Samudrala, R. Novel paradigms for drug discovery: computational multitarget screening. *Trends Pharmacol. Sci.* **29**, 62–71 (2008).
114. Van Gestel, S. & Schuermans, V. Thirty-three years of drug discovery and research with Dr. Paul Janssen. *Drug Dev. Res.* **8**, 1–13 (1986).
115. Overington, J.P., Al-Lazikani, B. & Hopkins, A.L. How many drug targets are there? *Nat. Rev. Drug Discov.* **5**, 993–996 (2006).
116. Fabian, M.A. *et al.* A small molecule–kinase interaction map for clinical kinase inhibitors. *Nat. Biotechnol.* **23**, 329–336 (2005).
117. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).