



Contents lists available at SciVerse ScienceDirect

## Pharmacology &amp; Therapeutics

journal homepage: [www.elsevier.com/locate/pharmthera](http://www.elsevier.com/locate/pharmthera)

## Structure and dynamics of molecular networks: A novel paradigm of drug discovery A comprehensive review

Peter Csermely<sup>a,\*</sup>, Tamás Korcsmáros<sup>a,b</sup>, Huba J.M. Kiss<sup>a,c</sup>, Gábor London<sup>d</sup>, Ruth Nussinov<sup>e,f</sup>

<sup>a</sup> Department of Medical Chemistry, Semmelweis University, P.O. Box 260, H-1444 Budapest 8, Hungary

<sup>b</sup> Department of Genetics, Eötvös University, Pázmány P. s. 1C, H-1117 Budapest, Hungary

<sup>c</sup> Department of Ophthalmology, Semmelweis University, Tömő str. 25-29, H-1083 Budapest, Hungary

<sup>d</sup> Department of Chemistry and Applied Biosciences, Swiss Federal Institute of Technology (ETH), Zurich, Switzerland

<sup>e</sup> Center for Cancer Research Nanobiology Program, SAIC-Frederick, Inc., National Cancer Institute, Frederick National laboratory for Cancer Research, Frederick, MD 21702, USA

<sup>f</sup> Sackler Institute of Molecular Medicine, Department of Human Genetics and Molecular Medicine, Sackler School of Medicine, Tel Aviv University, Tel Aviv 69978, Israel

## ARTICLE INFO

## Keywords:

Cancer  
Diabetes  
Drug target  
Network  
Side-effects  
Toxicity

## ABSTRACT

Despite considerable progress in genome- and proteome-based high-throughput screening methods and in rational drug design, the increase in approved drugs in the past decade did not match the increase of drug development costs. Network description and analysis not only give a systems-level understanding of drug action and disease complexity, but can also help to improve the efficiency of drug design. We give a comprehensive assessment of the analytical tools of network topology and dynamics. The state-of-the-art use of chemical similarity, protein structure, protein–protein interaction, signaling, genetic interaction and metabolic networks in the discovery of drug targets is summarized. We propose that network targeting follows two basic strategies. The “central hit strategy” selectively targets central nodes/edges of the flexible networks of infectious agents or cancer cells to kill them. The “network influence strategy” works against other diseases, where an efficient reconfiguration of rigid networks needs to be achieved by targeting the neighbors of central nodes/edges. It is shown how network techniques can help in the identification of single-target, edgetic, multi-target and allo-network drug target candidates. We review the recent boom in network methods helping hit identification, lead selection optimizing drug efficacy, as well as minimizing side-effects and drug toxicity. Successful network-based drug development strategies are shown through the examples of infections, cancer, metabolic diseases, neurodegenerative diseases and aging. Summarizing >1200 references we suggest an optimized protocol of network-aided drug development, and provide a list of systems-level hallmarks of drug quality. Finally, we highlight network-related drug development trends helping to achieve these hallmarks by a cohesive, global approach.

© 2013 Elsevier Inc. All rights reserved.

## Contents

1. Introduction . . . . .	0
2. An inventory of network analysis tools helping drug design . . . . .	0
3. The use of molecular networks in drug design . . . . .	0
4. Areas of drug design: an assessment of network-related added-value . . . . .	0
5. Four examples of network description and analysis in drug design . . . . .	0
6. Conclusions and perspectives . . . . .	0
Conflict of interest statement . . . . .	0
Acknowledgments . . . . .	0
References . . . . .	0

*Abbreviations:* ADME, absorption, distribution, metabolism and excretion; ADMET, absorption, distribution, metabolism, excretion and toxicity; FDA, USA Food and Drug Administration; GWAS, genome-wide association study; mTOR, mammalian target of rapamycin; NME, new molecular entity; QSAR, quantitative structure–activity relationship; QSPR, quantitative structure–property relationship; PPAR, peroxisome proliferator-activated receptor; SNP, single-nucleotide polymorphism.

\* Corresponding author. Tel.: +36 1 459 1500; fax: +36 1 266 3802.

E-mail address: [csermely.peter@med.semmelweis-univ.hu](mailto:csermely.peter@med.semmelweis-univ.hu) (P. Csermely).

0163-7258/\$ – see front matter © 2013 Elsevier Inc. All rights reserved.

<http://dx.doi.org/10.1016/j.pharmthera.2013.01.016>

Please cite this article as: Csermely, P., et al., Structure and dynamics of molecular networks: A novel paradigm of drug discovery, *Pharmacol. Ther.* (2013), <http://dx.doi.org/10.1016/j.pharmthera.2013.01.016>

## 1. Introduction

'Business as usual' is no longer an option in drug industry (Begley & Ellis, 2012). There is a growing recognition that systems-level thinking is needed for the renewal of drug development efforts. However, interrelated data have grown to such an unforeseen complexity, which argues for novel concepts and strategies. The Introduction aims to convey to the Reader that the network description and analysis can be a suitable method to describe the complexity of human diseases and help the development of new drugs.

### 1.1. Drug design as an area requiring a complex approach

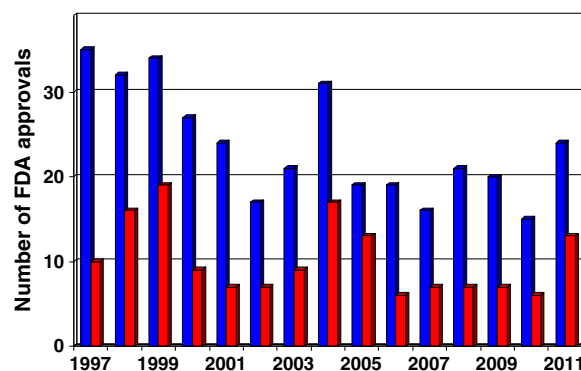
The population of Earth is growing and aging. Some of the major health challenges, such as many types of cancers and infectious diseases, diabetes and neurodegenerative diseases are in desperate need of innovative medicines. Despite of this challenge, fast and affordable drug development is a vision that contrasts sharply with the current state of drug discovery. It takes an average of 12 to 15 years and (depending on the therapeutic area) as much as 1 billion USD to bring a single drug into market. In the USA, pharmaceutical industry was the most R&D-intensive industry (defined as the ratio of R&D spending compared to total sales revenue) until 2003, when it was overtaken by communications equipment industry (Austin, 2006; Chong & Sullivan, 2007; Bunnage, 2011).

The increasingly high costs of drug development are partly associated

- with the high percentage of projects that fail in clinical trials,
  - with the recent focus on chronic diseases requiring longer and more expensive clinical trials,
  - with the increased safety concerns caused by catastrophic failures in the market and
  - with more expensive research technologies.
- Moreover, direct costs are doubled, where the second half comes from the 'opportunity cost', i.e. the financial costs of tying up investment capital in multiyear drug development projects (Austin, 2006; Chong & Sullivan, 2007; Bunnage, 2011).

We have a few hundreds of targets of approved drugs from the >20,000 non-redundant proteins of the human proteome. Despite the considerably higher R&D investment after the millennium, the number of new molecular entities (NMEs) approved by the USA Food and Drug Administration (FDA) remained constant at an annual 20 to 30 compounds. The number of NMEs potentially offering a substantial advance over conventional therapies is an even more sobering number of 6 to 17 per year in the last decade (Fig. 1). However, it is worth to note that looking only at the number of new drugs without considering their therapeutic value omits an important factor in the analysis (Austin, 2006; Overington et al., 2006; Chong & Sullivan, 2007; Bunnage, 2011; Edwards et al., 2011; Scannell et al., 2012).

Part of the slow progress is related to the high risks of investments. The development of an NME-drug costs approximately four times more than that of a non-NME. Moreover, the 'curse of attrition' steadily remained the biggest issue of the pharmaceutical industry in the last decades (Fig. 2). Each NME launched to the market needs about 24 development candidates to enter the development pipeline. Attrition of phase II studies is the key challenge, where only 25% of the drug-candidates survive. The 25% survival includes new agents against known targets (the 'me-too' or 'me-better' drugs), and therefore may be a significant overestimate of the survival rate of drug-candidates directed towards new targets. The low survival rate is exacerbated further by the very high costs of a failing compound at this late development stage (Brown & Superti-Furga, 2003; Austin, 2006; Bunnage, 2011; Ledford, 2012). These high risks made the drug industry cautious, and sometimes perhaps over-cautious. As the pharmacologist and Nobel Laureate James Black said: "the most fruitful basis for the discovery of a new drug is to start with an old drug" (Chong & Sullivan,



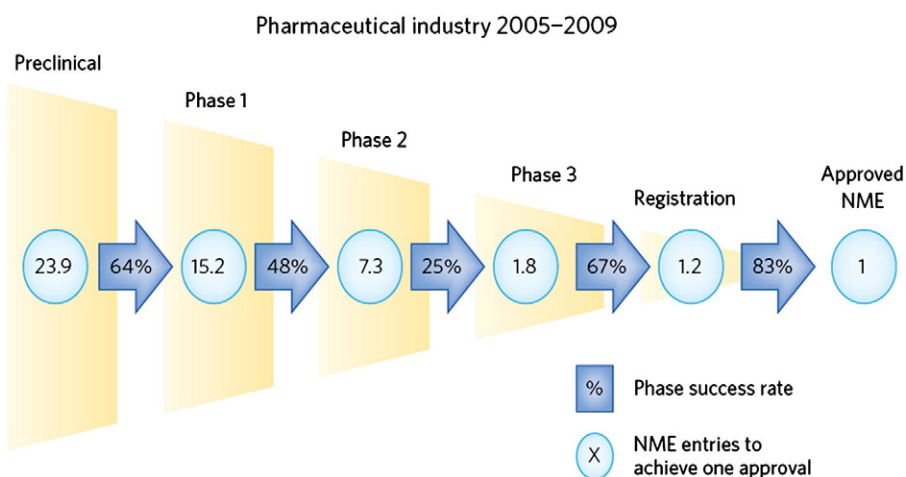
**Fig. 1.** Number of new molecular entities (NME, a drug containing an active ingredient that has not been previously approved by the US FDA) approved by the US Food and Drug Administration (FDA). Blue bars represent the total number of NMEs, whereas red bars represent "priority" NMEs that potentially offer a substantial advance over conventional therapies. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)  
Source: <http://www.fda.gov/Drugs/default.htm>.

2007). In fact, analysis of structure–activity relationship (SAR) pattern evolution, drug–target network topology and literature mining studies all showed the same behavior trend indicating that more than 80% of the new drugs tend to bind targets, which are connected to the network of previous drug targets (Cokol et al., 2005; Yildirim et al., 2007; Iyer et al., 2011a).

Improving the quality of target selection is widely considered as the single most important factor to improve the productivity of the pharmaceutical industry. From the 1970s target selection was increasingly separated from lead identification. Drug development process often fell to the 'druggability trap', where the attraction of working on a chemically approachable target encouraged development teams to push forward projects having a poor target quality. Additionally, chemical leads were often discovered to have unwanted side-effects and/or be toxic at later development phases (Brown & Superti-Furga, 2003; Hopkins, 2008; Bunnage, 2011).

The decline in the productivity of the pharmacological industry may stem partly from the underestimation of the complexity of cells, organisms and human disease (Lowe et al., 2010). We will illustrate the high level of this complexity by three examples.

- Under ideal conditions only 34% of single-gene deletions in yeast resulted in decrease in proliferation. However, when knockouts were screened against a diverse small-molecule library and a wide range of environmental conditions, 97% of the gene-deletions demonstrated a fitness defect (Hillenmeyer et al., 2008).
- Many of the most prevalent diseases, such as cancer, diabetes and coronary artery disease have a genetic background including a large number of genes (see Section 5 and Brown & Superti-Furga, 2003; Hopkins, 2008; Fliri et al., 2010). Following a treatment with a chemotherapeutic agent almost all of 1000 tagged proteins of cancer cells showed a dynamic response, when their temporal expression levels and localization were tracked (Cohen et al., 2008).
- As Loscalzo and Barabasi (2011) summarized in their excellent review, diseases are typically recognized and defined by their late-appearing manifestations in a partially dysfunctional organ-system. As a part of this, therapeutic strategies often do not focus on truly unique, targeted disease determinants, but (rightfully) address the patho-phenotypes of the already advanced disease stage. These advanced patho-phenotypes have a large number of symptoms, which are not primarily disease-specific (such as inflammation). This definition of disease may obscure subtle, but potentially important differences among patients with clinical presentations, and may also neglect pathobiological mechanisms extending the disease-defining organ system. Loscalzo and Barabasi (2011) argue that the complexity of



**Fig. 2.** Success rate of new molecular entities (NMEs) by R&D development phases. The figure shows the combined R&D survival by development phase for 14 large pharmaceutical companies. Note that attrition figures for early phases might be even higher, since an early problem might be first neglected making a failure only at a later phase (Brown & Superti-Furga, 2003).

Reprinted by permission from the Macmillan Publishers Ltd: Nature Chemical Biology, Bunnage, 2011, Copyright, 2011.

disease should be viewed as an emergent property of a pathobiological system, i.e. a property, which cannot be predicted by studying only the parts of the system, but emerges from the complex interrelationships of all system components. Kola and Bell (2011) arrive to the same conclusion urging the reform of the taxonomy of human disease.

These examples illustrate the extent of non-linearity and interdependence of cellular and organismal responses. To understand these observations and outcomes, we need novel approaches.

Over-reliance on inadequate animal or cellular models of disease has been considered to play a major part in the poor levels of Phase II drug candidate survival-rate. We illustrate the limitations and dangers of model-selection by three examples.

- 41% of the proteins expressed in rat lungs were absent from equivalent cultured cells (Lindsay, 2005).
- Animal strains are often in-bred, and are examined in a young age for diseases having an onset in elderly people (Lindsay, 2005).
- In psychological clinical studies 96% of the patients cover 12% of the world population (Henrich et al., 2010). A more equal coverage is also required by the geographic clustering of rare genetic variants affecting drug efficacy (Nelson et al., 2012).

There is a growing recognition that systems-level thinking may help to overcome many of the current troubles of drug development (Brown & Superti-Furga, 2003; Csermely et al., 2005; Lindsay, 2005; Korcsmáros et al., 2007; Henney & Superti-Furga, 2008; Hopkins, 2008; Westerhoff et al., 2008; Bunnage, 2011; Chua & Roth, 2011; Farkas et al., 2011; Penrod et al., 2011; Begley & Ellis, 2012). As a sign of this, leading systems biologists aim to construct a computer replica of the whole human body, called as the 'silicon human' by 2038 (Kolodkin et al., 2012).

In fact, systems-level thinking characterized drug development until the 1970s, when mechanistic drug-targets were unknown. Until the late 1970s even the concept of the receptor was not based on sequence and structural data, but on the chemical similarities of ligands exerting similar pharmacological actions (Brown & Superti-Furga, 2003; Keiser et al., 2010). It was only after the early 1980s, that the focus shifted from physiological observations to the molecular level (Pujol et al., 2010).

The renewal of systems-based thinking in drug discovery was helped by the following three factors. 1.) The development of robust high-throughput platforms to gather large amounts of comparable molecular data. 2.) The assembly and availability of curated databases integrating the knowledge of the field. 3.) The emergence of interdisciplinary research to understand these data (Arrell & Terzic, 2010). Most of the

current largest pharmaceutical firms are products of horizontal mergers between two or more large drug companies which have been taking place since 1989. Though larger companies have the advantage to fund and sustain a broader range of larger research programs, the development of large firms and research enterprises was often considered to decrease flexible responses to novel development opportunities (Austin, 2006; Gros, 2012). An increased efficiency needs coordinated networking of large drug development firms, biotechnological companies and research institutions (Hasan et al., 2012; Heemskerck et al., 2012). Moreover, systems-level thinking needs a new behavior code of sharing data and approaches. This new alliance is characterized by the following behavior.

- In systems-level drug development, quality and not quantity of data is a key issue. A reliable data pipeline must be assembled using appropriate standards and quality control-metrics keeping in mind the needs of systems biology. This is all the more important, since it may also overcome the unreliability problems which surfaced recently, when Amgen tried to reproduce data from 53 published pre-clinical studies of potential anticancer drugs, and it failed in all but 6 cases (11% reproducibility rate), or Bayer Health Care could reproduce only 25% of previously published preclinical studies (Henney & Superti-Furga, 2008; Prinz et al., 2011; Begley & Ellis, 2012; Landis et al., 2012).
- Sharing of systems-level results led to a fast development of predictive toxicology, which is a key step of a more efficient progress (Henney & Superti-Furga, 2008).

Datasets are growing to dimensions, where the three billion nucleotides that comprise the human genome (International Human Genome Sequencing Consortium, 2004; ENCODE Project Consortium, 2012) became millionths of the ~1 petabyte data we had in 2008 (Schadt et al., 2009), which have grown well over 1 exabyte (billion times billion bytes) by 2012. These magnitudes require appropriate computational tools to understand them. Through this review we hope to convince the Reader that network description and analysis offer novel tools, which can help us to understand the complexity of human disease and enable the integration of knowledge towards a more efficient combat strategy for healthier life.

### 1.2. Molecular networks as efficient tools in the description of cellular and organism behavior

Complexity can be described through the rather simple saying that 'in a complex system the whole is more than the sum of its parts: cutting

a horse to two will not result in two small horses' (Kolodkin et al., 2012; San Miguel et al., 2012). Newman (2011) summarized a number of excellent sources to study complexity. A recent summary listed the following hallmarks of complex systems and their behavior: many heterogeneous interacting parts; multiple scales; combinatorial explosion of possible states; complicated transition laws; unexpected or unpredicted emergent properties; sensitivity to initial conditions; path-dependent dynamics; networked hierarchical connectivity; interaction of autonomous agents; self-organization, collective shifts; non-equilibrium dynamics; adaptivity to changing environments; co-evolving subsystems; ill-defined boundaries and multilevel dynamics (San Miguel et al., 2012). Though this list is certainly still incomplete, and not all of its parts are characterizing the complex systems of drug discovery, the list shows the tremendous difficulties we face when trying to understand complex structures and their behavior. The same report (San Miguel et al., 2012) listed the following major challenges of complex system studies:

- data gathering by large-scale experiments, data sharing and data assembly using mutually agreed curation rules, management of huge, distributed, dynamic and heterogeneous databases;
- moving from data to dynamical models going beyond correlations to cause–effect relationships, understanding the relationship between simple and comprehensive models with appropriate choices of variables, ensemble modeling and data assimilation, modeling the 'systems of systems of systems' with many levels between micro and macro; and
- formulating new approaches to prediction, forecasting, and risk, especially in systems that can reflect on and change their behavior in response to predictions and in systems, whose apparently predictable behavior is disrupted by apparently unpredictable rare or extreme events.

Due to the complexity of the cells, organisms and diseases, extreme reductionism often fails in drug design. However, the other extreme, taking into account all possible variables of all possible components, is neither feasible, nor doable. Fortunately we do not have to challenge the impossible when thinking on complexity in drug design for two major reasons. On the one hand, the structure of complex systems is not only complicated, but also modular, and has a number of degenerate segments. This enables us to identify the most important system segments as we will show in Section 2. On the other hand, complex systems often determine a state space, which is also modular, and has a surprisingly low number of major attractors. In fact, this is what makes the discrimination of phenotypes possible at all. In other words: complexity has a side of simplicity. As fortunate 'side-effects' of the attractor-segmented, modular state space, many of the emergent properties of complex systems tolerate a number of errors in the individual data determining them. The above features of drug design-related complex systems make those descriptions successful, which are 'complex' themselves, meaning that they are neither too simplistic, nor go too much into details (Bar-Yam et al., 2009; Csermely, 2009; Huang et al., 2009; Mar & Quackenbush, 2009; Kolodkin et al., 2012). In agreement with these considerations, mathematical systems theory states that "the scale and complexity of the solution should match the scale and complexity of the problem" (Bar-Yam, 2004).

Network-approach is a description, which provides a good compromise between extreme reductionism and the 'knowledge of everything'. We are by far not alone sharing this view. Diseases have been perceived as network perturbations (Huang et al., 2009; Del Sol et al., 2010). In recent years network analysis became an increasingly acclaimed method in drug design (Hopkins, 2008; Ma'ayan, 2008; Pawson & Linding, 2008; Berger & Iyengar, 2009; Schadt et al., 2009; Baggs et al., 2010; Fliri et al., 2010; Lowe et al., 2010; Pujol et al., 2010). In agreement with the expert-opinions, network-applications show a steady increase of drug design-related publications (Fig. 3). We summarize the major network types (detailed in Section 3), network analysis types (detailed in Section 2), drug design areas helped by network studies (detailed in

Section 4) and the four key areas of drug design described in detail as the examples in Section 5 in Fig. 4.

We will detail the definition and types of networks in Section 2.1. The applicability of network analysis in drug design is determined by the following major factors: 1.) proper definition of network nodes, edges and edge weights; 2.) data quality and carefully defined, uniformly applied data inclusion criteria; 3.) data refinement by genetic variability, aging, environmental effects and compounding pathologies such as bacterial or viral infections (Arrell & Terzic, 2010; Kolodkin et al., 2012). However, we will not cover details of data acquisition, since this topic fits better into the broader area of systems biology, which is not the subject of the current review.

Networks are often viewed via their mathematical representations, i.e. graphs. However, this often proves to be an over-simplification in drug design for two major reasons. 1.) Network nodes of cellular systems are not exact 'points', as in graph theory, but macromolecules, having a network structure themselves, as we will show in Section 3.2.2) Network nodes have a lot of attributes in the rich biological context of the cell. 3.) Network dynamics is crucial in order to understand the complexity of diseases and the action of drugs (Pujol et al., 2010). Therefore, it is often useful to include edge directions, signs (activation or inhibition), conditionality (an edge is active only, if one of its nodes has another edge) and a number of dynamically changing quantitative measures in network descriptions. However, it is important to warn here that we should not include too many details in network descriptions, since we may shift our description from optimal towards the 'knowledge of everything'. Including more and more details in network science may lead to the trap of 'over-complication', where the beauty and elegance of the approach are lost. This may lead to the decline of the use of network description and analysis (similarly to the over-use of the explanatory power and decline of chaos theory, fractals, and many other approaches before).

The optimal simplicity of networks is also important, since networks give us a visual image. We summarize a rather long list of network visualization techniques in Table 1 showing the rich variety of approaches to solve this important task. A detailed comparison of some methods was described in several reviews (Suderman & Hallett, 2007; Pavlopoulos et al., 2008; Gehlenborg et al., 2010; Fung et al., 2012). A good visualization method provides a pragmatic trade-off between highlighting the biological concept and comprehensibility. Trying several methods is often advisable, since sampling scale and/or bias may lead to subjective interpretations of the network images obtained.

Correct visualization of networks is not only important for making a pleasing image. The right hemisphere of our brain works with images, and has the unique strength of pattern recognition. This complements the logical thinking of the left hemisphere. Regretfully, our

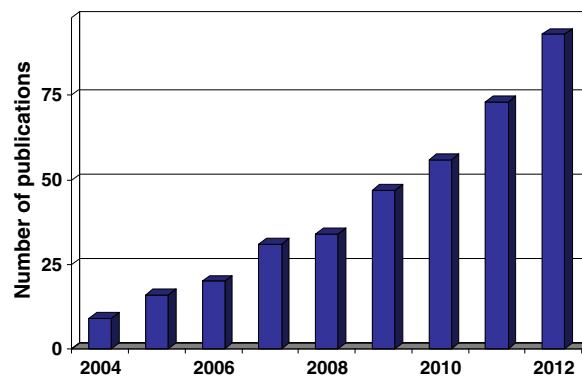


Fig. 3. Network-application in drug-design related publications. Data are from PubMed using the query of "network AND drug" for title and abstract words. The number of publications in 2012 is an extrapolation.

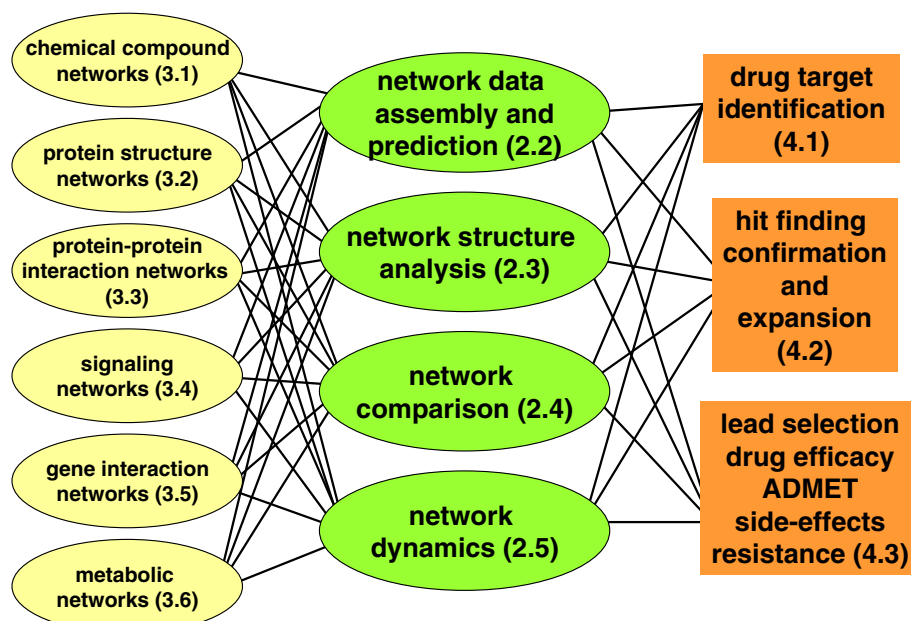


Fig. 4. Uses of network description and analysis in drug design. Numbers in parentheses refer to section numbers of this review.

logical thinking can deal with 5 to 6 independent pieces of information at the same time as an average. However, the complexity of human disease requires an information-handling capacity, which is by magnitudes higher than that of logical thinking. Pattern recognition by the right hemisphere copes with this complexity. This is why we also need to see networks, and may not only measure them. Besides the 'optimal simplicity', visualization is another advantage of networks over data-mining and other very useful, but highly detailed approaches (Csermely, 2009). To illustrate the network description and analysis in drug design, we compare the classic view and the network view of drug action in Fig. 5.

As we have described in the previous paragraphs, network description and analysis offer a wide range of possibilities to understand the complexity of human disease and to develop novel drugs. As an example of the richness of networks, the 'semantic web' covers practically every conceptual entity appearing in the world-wide-web (Chen et al., 2009a). In the current review we cannot cover all. Therefore, with the exception of the network of human diseases described in Section 1.3, we will restrict ourselves to molecular networks ranging from the networks of chemical compounds and of protein structures to the various networks of the macromolecules constituting the cells. We will not cover the following areas, where we list a few reviews and papers of special interest:

- networked particles in drug delivery (Rosen & Elman, 2009; Luppi et al., 2010; Bysell et al., 2011);
- network of plants as resources of herbal remedies and traditional medicines (Saslis-Lagoudakis et al., 2012);
- cytoskeletal networks or membrane organelle networks (Michaelis et al., 2005; Escribá et al., 2008; Gombos et al., 2011);
- inter-neuronal, inter-lymphocyte and other intercellular networks including extracellular matrix, cytokine, endocrine or paracrine networks (Jerne, 1974, 1984; Cohen, 1992; Werner, 2003, 2005; Small, 2007; Jung et al., 2011; Acharyya et al., 2012; Margineanu, 2012);
- the ecological networks of the microorganisms living in human gut, oral cavity, skin, etc. and their interconnected networks with human cells (Ben-Jacob et al., 2012; Clemente et al., 2012; Mueller et al., 2012);

- social networks and their potential effects on spreading of epidemics, as well as disease-related habits such as drug abuse, smoking, over-eating, etc. (Christakis & Fowler, 2011);
- network-related modeling methods, such as: neural network models, differential equation networks, network-related Markov chain methods, Boolean networks, fuzzy logic-based network models, Bayesian networks and network-based data mining models (Huang, 2001; Ideker & Lauffenburger, 2003; Winkler, 2004; Fayos & Fayos, 2007; Fernandez et al., 2011).

At the end of the Introduction we will illustrate network thinking by showing the richness and usefulness of network representations of human diseases.

### 1.3. The networks of human diseases

Several diseases, such as cancer, or complex physiological processes, such as aging, were described as a network phenomenon quite a while ago (Kirkwood & Kowald, 1997; Hornberg et al., 2006; Söti & Csermely, 2007). In this section we will not detail disease-related molecular networks (such as interactomes, or signaling networks changing in disease), since this will be the subject of Section 3. We will describe the large variety of options to build up the networks of human diseases, where diseases are nodes of the network, and will show how network-assembled bio-data can be used to predict novel disease biomarkers including novel disease-related genes.

#### 1.3.1. Network representations of diseases and their therapies

In the network description, sets of interlined data need first to be structured by defining 'nodes'. This might already be rather difficult, as we will show in detail in Section 2.1. However, the definition of edges, i.e. connections between the nodes, may be an especially demanding task. Networks of human diseases provide a very good example, since a large number of data categories are related to the concept of disease enabling the construction of a large variety of networks (Goh et al., 2007; Rzhetsky et al., 2007; Feldman et al., 2008; Spiró et al., 2008; Hidalgo et al., 2009; Barabási et al., 2011; Zhang et al., 2011a; Janjic & Przulj, 2012).

Some of the major disease-related categories are shown in Fig. 6. Human disease can be conceptualized as a phenotype, i.e. an emergent property of the human body as a complex system (Kolodkin et al.,

**Table 1**  
Network visualization resources.

Name	Website	References
Arena3D	<a href="http://arena3d.org">http://arena3d.org</a>	Secrier et al., 2012
ArrayXPath	<a href="http://www.snubi.org/software/ArrayXPath">http://www.snubi.org/software/ArrayXPath</a>	Chung et al., 2005
AVIS	<a href="http://actin.pharm.mssm.edu/AVIS2">http://actin.pharm.mssm.edu/AVIS2</a>	Berger et al., 2007
BioLayout Express 3D	<a href="http://www.biolayout.org">http://www.biolayout.org</a>	Freeman et al., 2007; Theocharidis et al., 2009
BiologicalNetworks	<a href="http://biologicalnetworks.net">http://biologicalnetworks.net</a>	Kozhenkov & Baitaluk, 2012
BioTapestry	<a href="http://www.biotapestry.org">http://www.biotapestry.org</a>	Longabaugh, 2012
BisoGenet	<a href="http://bio.cigb.edu.cu/bisogenet-cytoscape">http://bio.cigb.edu.cu/bisogenet-cytoscape</a>	Martin et al., 2010
CellDesigner	<a href="http://www.celldesigner.org">http://www.celldesigner.org</a>	Kitano et al., 2005
Cell Illustrator	<a href="http://www.cellillustrator.com">http://www.cellillustrator.com</a>	Nagasaki et al., 2011
CFinder	<a href="http://www.cfinder.org">http://www.cfinder.org</a>	Adamcsek et al., 2006
Cytoscape	<a href="http://www.cytoscape.org">http://www.cytoscape.org</a>	Smoot et al., 2011
GenePro	<a href="http://wodaklab.org/geneopro">http://wodaklab.org/geneopro</a>	Vlasblom et al., 2006
GeneWays	<a href="http://anya.igsb.anl.gov/GeneWays/GeneWays.html">http://anya.igsb.anl.gov/GeneWays/GeneWays.html</a>	Rzhetsky et al., 2004
GEOMli	<a href="http://sydney.edu.au/engineering/it/~visual/valacon/geomi/">http://sydney.edu.au/engineering/it/~visual/valacon/geomi/</a>	Ahmed et al., 2006
Gephi	<a href="http://gephi.org">http://gephi.org</a>	Bastian et al., 2009
Graphviz	<a href="http://www.graphviz.org">http://www.graphviz.org</a>	Gansner & North, 2000
Gridlayout	<a href="http://kurata21.bio.kyutech.ac.jp/grid/grid_layout.htm">http://kurata21.bio.kyutech.ac.jp/grid/grid_layout.htm</a>	Li & Kurata, 2005
Guess	<a href="http://graphexploration.cond.org/index.html">http://graphexploration.cond.org/index.html</a>	Adar, 2006
Hive Plots	<a href="http://www.hiveplot.com">http://www.hiveplot.com</a>	Krzywinski et al., 2012
Hybridlayout	<a href="http://www.cadlive.jp/hybridlayout/hybridlayout.html">http://www.cadlive.jp/hybridlayout/hybridlayout.html</a>	Inoue et al., 2012
Hyperdraw	<a href="http://www.bioconductor.org/packages/release/bioc/html/hyperdraw.html">http://www.bioconductor.org/packages/release/bioc/html/hyperdraw.html</a>	Murrell, 2012
IM Browser	<a href="http://proteome.wayne.edu/PIMdb.html">http://proteome.wayne.edu/PIMdb.html</a>	Pacifico et al., 2006
IPath	<a href="http://pathways.embl.de">http://pathways.embl.de</a>	Yamada et al., 2011
JNets	<a href="http://www.manchester.ac.uk/bioinformatics/jnets">http://www.manchester.ac.uk/bioinformatics/jnets</a>	Macpherson et al., 2009
KGML-ED	<a href="http://kgml-ed.ipk-gatersleben.de">http://kgml-ed.ipk-gatersleben.de</a>	Klukas & Schreiber, 2007
LEDA	<a href="http://www.algorithmic-solutions.com/leda/about/index.htm">http://www.algorithmic-solutions.com/leda/about/index.htm</a>	Mehlhorn & Näher, 1999
MAVisto	<a href="http://mavisto.ipk-gatersleben.de">http://mavisto.ipk-gatersleben.de</a>	Schwobbermeyer & Wunschiers, 2012
Medusa	<a href="http://coot.embl.de/medusa">http://coot.embl.de/medusa</a>	Hooper & Bork, 2005
ModuLand	<a href="http://www.linkgroup.hu/modules.php">www.linkgroup.hu/modules.php</a>	Szalay-Bekó et al., 2012
Multilevel Layout	<a href="http://code.google.com/p/multilevellayout">http://code.google.com/p/multilevellayout</a>	Tuikkala et al., 2012
NAVIGaTOR	<a href="http://ophid.utoronto.ca/navigator">http://ophid.utoronto.ca/navigator</a>	Brown et al., 2009
NetMiner	<a href="http://www.netminer.com/index.php">http://www.netminer.com/index.php</a>	
Network Workbench	<a href="http://nwb.cns.iu.edu">http://nwb.cns.iu.edu</a>	NWB Team, 2006
Ondex	<a href="http://www.ondex.org">http://www.ondex.org</a>	Kohler et al., 2006
Osprey	<a href="http://biodata.mshri.on.ca/osprey/servlet/Index">http://biodata.mshri.on.ca/osprey/servlet/Index</a>	Breitkreutz et al., 2003
Pajek	<a href="http://pajek.imfm.si/doku.php">http://pajek.imfm.si/doku.php</a>	Batagelj & Mrvar, 1998
PathDraw	<a href="http://rospath.ewha.ac.kr/toolbox/PathwayViewerFrm.jsp">http://rospath.ewha.ac.kr/toolbox/PathwayViewerFrm.jsp</a>	Paek et al., 2004
Pathway Tools	<a href="http://bioinformatics.ai.sri.com/ptools">http://bioinformatics.ai.sri.com/ptools</a>	Karp et al., 2010
PATIKA	<a href="http://www.patika.org">http://www.patika.org</a>	Dogrusoz et al., 2006
PaVESy	<a href="http://pavesy.mpimp-golm.mpg.de/PaVESy.htm">http://pavesy.mpimp-golm.mpg.de/PaVESy.htm</a>	Ludemann et al., 2004
PhyloGrapher	<a href="http://www.atgc.org/PhyloGrapher">http://www.atgc.org/PhyloGrapher</a>	
PIMWalker	<a href="http://pimr.hybrigenics.com">http://pimr.hybrigenics.com</a>	Meil et al., 2005
PIVOT	<a href="http://acgt.cs.tau.ac.il/pivot">http://acgt.cs.tau.ac.il/pivot</a>	Orlev et al., 2004
PolarMapper	<a href="http://kdbio.inesc-id.pt/software/polarmapper">http://kdbio.inesc-id.pt/software/polarmapper</a>	Goncalves et al., 2009
ProteinNetVis	<a href="http://graphics.cs.brown.edu/research/scviz/proteins/home.htm">http://graphics.cs.brown.edu/research/scviz/proteins/home.htm</a>	Jianu et al., 2010
Proteolens	<a href="http://bio.informatics.iupui.edu/proteolens">http://bio.informatics.iupui.edu/proteolens</a>	Huan et al., 2008
RedeR	<a href="http://bioconductor.org/packages/release/bioc/html/RedeR.html">http://bioconductor.org/packages/release/bioc/html/RedeR.html</a>	Castro et al., 2012
RING	<a href="http://protein.bio.unipd.it/ring">http://protein.bio.unipd.it/ring</a>	Martin et al., 2011
SoNIA	<a href="http://www.stanford.edu/group/sonia">http://www.stanford.edu/group/sonia</a>	Bender-deMoll & McFarland, 2006
Transcriptome-Browser	<a href="http://tagc.univ-mrs.fr/tbrowser">http://tagc.univ-mrs.fr/tbrowser</a>	Lepoivre et al., 2012
UCSF structureViz	<a href="http://www.cgl.ucsf.edu/cytoscape/structureViz">http://www.cgl.ucsf.edu/cytoscape/structureViz</a>	Morris et al., 2007
VANTED	<a href="http://vanted.ipk-gatersleben.de">http://vanted.ipk-gatersleben.de</a>	Rohn et al., 2012
VisANT	<a href="http://visant.bu.edu">http://visant.bu.edu</a>	Hu et al., 2009
VitaPad	<a href="http://sourceforge.net/projects/vitapad">http://sourceforge.net/projects/vitapad</a>	Holford et al., 2005
WebInterViewer	<a href="http://interviewer.inha.ac.kr">http://interviewer.inha.ac.kr</a>	Han et al., 2004b
yFiles	<a href="http://www.yworks.com/en/index.html">http://www.yworks.com/en/index.html</a>	Becker & Rojas, 2001
yWays	<a href="http://www.yworks.com/en/products_yfiles_extensionpackages_ep2.htm">http://www.yworks.com/en/products_yfiles_extensionpackages_ep2.htm</a>	

Summaries of Suderman and Hallett (2007), Pavlopoulos et al. (2008), Gehlenborg et al. (2010) and Fung et al. (2012) compare some of the options above.

2012). Some of the categories, such as symptoms, are related to this phenotype. Many other categories, such as

- disease-related genes (abbreviated as 'disease genes'),
- functions of disease genes (marked as gene ontology);
- the transcriptome (i.e. expression levels of all mRNAs + the cistrome, i.e. DNA-binding transcription factors + the epigenome, i.e. the actual chromatin status of the cell including DNA and histone modifications, as well as their 3D structure)
- the interactome, the signaling network and the metabolome,

are all related to the underlying genotype, i.e. the constituents of the human body related to the etiology of the disease. A third group of categories, such as therapies, drugs and other factors marked as

"environment", represents the effects of the environment (Fig. 6). Connections (uniformly defined, data-encoded relationships) between any two of these categories define a so-called bipartite network, where two different types of nodes are related to each other. Moreover, more than two categories may also form a network, which is called as a multi-partite network (Goh et al., 2007; Yildirim et al., 2007; Nacher & Schwartz, 2008; Spiró et al., 2008; Li et al., 2009a; Bell et al., 2011; Wang et al., 2011a).

We have three options for the visualization of bipartite networks. We will illustrate this in the example of the network of human diseases and human genes shown to be associated with a particular disease in Fig. 7 (Goh et al., 2007). We may include both types of nodes and all their connections to the visual image as shown in the center of

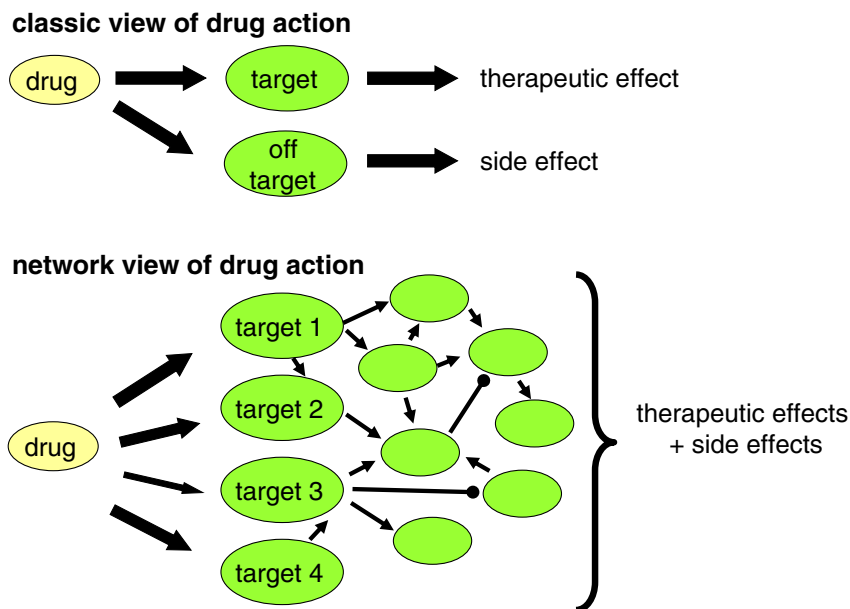


Fig. 5. Classic and network views of drug action. Made after the basic idea of Berger and Iyengar (2009).

Fig. 7. However, the selection of only a single node type results in a simpler network representation, which is easier to understand. We have two projections of the full, bipartite network as shown on the two sides of Fig. 7. In the first type of projection we connect two human diseases, if there is a human gene, which is participating in the etiology of both diseases (left side of Fig. 7). Edge weight may be derived here from the number of genes connecting the two diseases. Alternatively, we may construct a network of human genes, which are connected, if there is at least one human disease, where they both belong (right side of Fig. 7; Goh et al., 2007). Similar

projections can be made with any category-pairs, or multiple category-sets of Fig. 6.

### 1.3.2. The human disease network

The landmark study of Goh et al. (2007) provided the first network map of the genetic relationship of 516 human diseases. This approach used the “shared gene formalism” recognizing that diseases sharing a gene or genes likely have a common genetic basis. Later, this concept was extended with the “shared metabolic pathway formalism” recognizing that enzymatic defects affecting the flux of “reaction A” in a

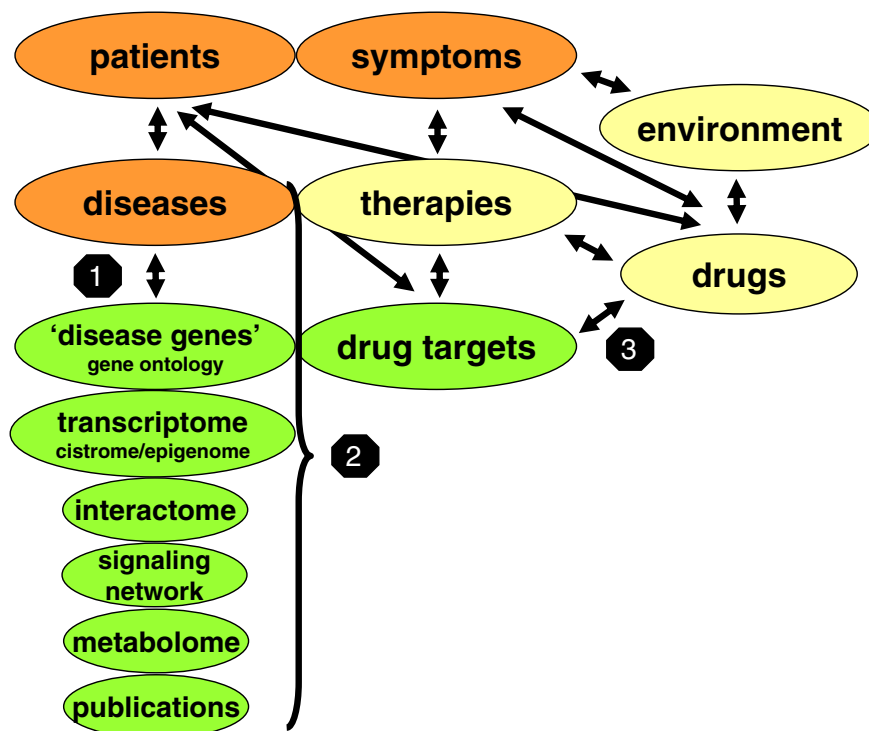
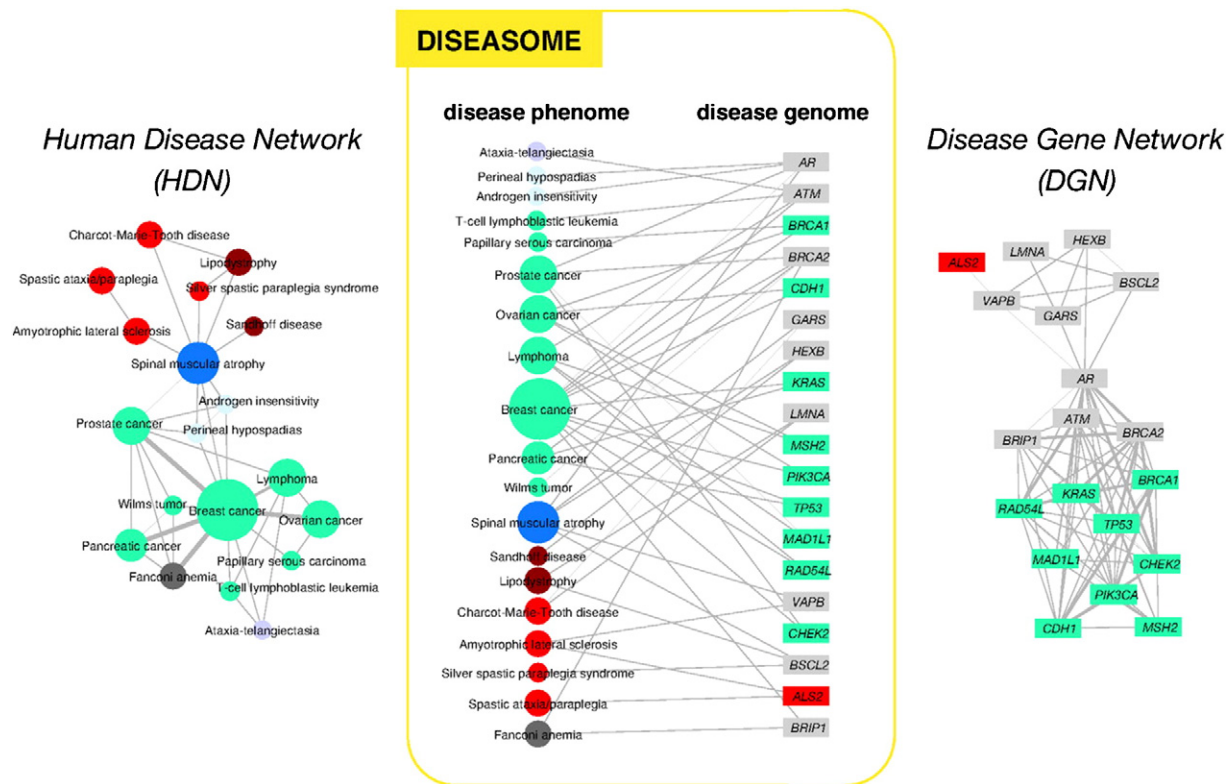


Fig. 6. Options for network representations of disease-related data. The figure summarizes some of the options to assess disease-related data using network description and analysis. Each ellipse represents a type of data. Arrows stand for possible network representations. 1: Human disease networks discussed in this section and in Table 2. 2: Additional network-related data helping the identification of disease-related human genes (acting like possible drug targets) detailed in Table 3. 3: Drug target networks discussed in Section 4.1.3.



**Fig. 7.** Two projections of the human disease network. On the middle of the figure a segment of the bipartite network of human diseases and related human genes is shown. On the projection on the left side two diseases are connected, if they have at least one common gene. On the projection on the right side two genes are connected, if they have at least one common disease.

Reproduced with permission from Goh et al. (2007); Copyright, 2007, National Academy of Sciences, U.S.A.

metabolic pathway will lead to disease-conditions that are known to be associated with the metabolites situated downstream of “reaction A” in the same metabolic pathway. The shared metabolic pathway formalism proved to be better predictor of metabolic diseases than the shared gene formalism. Another approach is based on the “disease comorbidity formalism” connecting diseases, which have a co-occurrence in patients exceeding a predefined threshold. Subsequently, many other studies incorporated a number of other data including gene-expression levels, protein–protein interactions, signaling components, such as microRNAs, tissue-specificity, and a number of environmental effects including drug treatment and other therapies to construct disease similarity networks (Barabási et al., 2011; Goh & Choi, 2012; Janjic & Przulj, 2012). We summarize the disease-network types using two, three or more different datasets in Table 2. We will summarize drug target networks in Section 4.1.3.

Various data-associations listed in Table 2 enrich each other, as it has been shown in the example of the orphan diseases, Tay–Sachs disease and Sandhoff syndrome, which did not share any known disease genes in 2011, but were connected in a literature co-occurrence based network. The connection of the two diseases was in agreement with the shared metabolic pathway of their mutated genes. Zhang et al. (2011a) listed several other examples for such mutual enrichment of various datasets. Comparing Table 2 with Fig. 6 reveals several combinations of data, which have not been used to form human disease networks yet. We expect further advance in this rapidly growing field.

As take home messages from the studies listed in Table 2, we summarize the following observations.

- The intuitive assumption that “hubs (defined here as nodes with many more neighbors than average in the human interactome) play a major role in adult diseases” often fails due to the embryonic lethality of these key genes. In agreement with this, orphan diseases (which are often life-threatening or chronically debilitating, and

affect less than 6.5 patients per 10,000 inhabitants) tend to be hubs, and are often associated with essential genes. Similarly, diseases having somatic mutations, such as cancer, have a central position in the human interactome. Germ-line mutations leading to more common diseases tend to be located in the functional periphery (but not in the utmost periphery) of the human interactome (Goh et al., 2007; Feldman et al., 2008; Barabási et al., 2011; Zhang et al., 2011a).

- Disease-related genes tend to be tissue specific, with the notable exception of most cancer-related genes, which are not overexpressed in the tissues from which the tumors emanate (Goh et al., 2007; Jiang et al., 2008; Lage et al., 2008; Barabási et al., 2011).
- Disease-related genes have a smaller than average clustering coefficient avoiding densely connected local structures (Feldman et al., 2008). Low clustering coefficient was successfully applied as a discriminatory feature in the prediction of disease-related genes (Sharma et al., 2010a).
- Disease-related genes tend to form overlapping disease modules in protein–protein interaction networks showing even a 10-fold increase of physical interactions relative to random expectation (Gandhi et al., 2006; Goh et al., 2007; Oti & Brunner, 2007; Feldman et al., 2008; Jiang et al., 2008; Stegmaier et al., 2010; Bauer-Mehren et al., 2011; Loscalzo & Barabasi, 2011; Xia et al., 2011). Overlaps of disease modules are also characteristic to comorbidity networks (Rzhetsky et al., 2007; Hidalgo et al., 2009).
- Genes bridging disease modules in the human interactome may provide important points of interventions (Nguyen & Jordan, 2010; Nguyen et al., 2011). Genes involved in the aging process often occupy such bridging positions (Wang et al., 2009).
- Diseases that share disease-associated cellular components (genes, proteins, metabolites, microRNAs, etc.) show phenotypic similarity and comorbidity (Lee et al., 2008a; Barabási et al., 2011).
- The above findings are recovered, if we go one level deeper in the network hierarchy than the human interactome, to the level of protein domains and their interactions (Sharma et al., 2010a; Song &



Table 2

Human disease-related networks and network datasets.

Type of related data (types of network nodes) <sup>a</sup>	Name and additional description, website	References <sup>b</sup>
<ul style="list-style-type: none"> <li>• Disease</li> <li>• Disease related genes</li> <li>• Disease</li> <li>• Disease-related genes</li> <li>• Interactome</li> <li>• Publication</li> </ul>	Human disease network (Cytoscape plug-in DisGeNET: <a href="http://ibi.imim.es/DisGeNET/DisGeNETweb.html">http://ibi.imim.es/DisGeNET/DisGeNETweb.html</a> )	Goh et al., 2007; Feldman et al., 2008; Bauer-Mehren et al., 2010; Stegmaier et al., 2010; Zhang et al., 2011a
<ul style="list-style-type: none"> <li>• Disease</li> <li>• Disease-related genes</li> <li>• Interactome</li> <li>• Publication</li> </ul>	Gene-based, interactome-enriched and scientific publication based human disease networks	
<ul style="list-style-type: none"> <li>• Disease</li> <li>• Interactome module</li> <li>• mRNA changes</li> <li>• Disease</li> <li>• mRNA changes at the transcriptome level</li> </ul>	Disease-responsive interactome module-based human disease network (disease correlations based on disease-induced changes in mRNA expression of interactome modules)	Suthram et al., 2010
<ul style="list-style-type: none"> <li>• Disease</li> <li>• mRNA changes at the transcriptome level</li> </ul>	A Bayesian network-based disease-responsive transcriptome analysis to construct a human disease network	Huang et al., 2010a
<ul style="list-style-type: none"> <li>• Drugs</li> <li>• Disease</li> <li>• Disease-related genes</li> <li>• Interactome</li> <li>• Protein/DNA interaction</li> <li>• Tissue</li> <li>• Drug</li> </ul>	iCTNet: a Cytoscape plug-in to construct an integrative network of diseases, associated genes, drugs and tissues ( <a href="http://www.cs.queensu.ca/ictnet">http://www.cs.queensu.ca/ictnet</a> )	Wang et al., 2011b
<ul style="list-style-type: none"> <li>• Disease</li> <li>• Disease-related genes</li> <li>• Interactome</li> </ul>	Biomine: an integrated bio-entity network with more than a million entities and 8 million edges ( <a href="http://biomine.cs.helsinki.fi">http://biomine.cs.helsinki.fi</a> )	Eronen & Toivonen, 2012
<ul style="list-style-type: none"> <li>• Gene Ontology terms</li> <li>• Disease</li> <li>• Expression patterns</li> <li>• MicroRNA targets</li> <li>• Network modules of interactome, transcriptome</li> </ul>	PAGED: an integrated bio-entity network with more than a million entities from 20 organisms ( <a href="http://bio.informatics.iupui.edu/PAGED">http://bio.informatics.iupui.edu/PAGED</a> )	Huang et al., 2012b
<ul style="list-style-type: none"> <li>• Disease</li> <li>• Disease-related genes</li> <li>• Interactome</li> <li>• Protein gene regulation pathways</li> <li>• Gene Ontology terms</li> <li>• Small molecule (drug)</li> <li>• Species</li> </ul>	An integrated bio-entity network	Bell et al., 2011
<ul style="list-style-type: none"> <li>• Disease</li> <li>• Adjacent members of metabolic pathways</li> </ul>	Metabolic pathway-corrected human disease network	Lee et al., 2008a
<ul style="list-style-type: none"> <li>• Disease</li> <li>• MicroRNA</li> </ul>	MicroRNA/disease association-based disease network obtained from publication data	Lu et al., 2008
<ul style="list-style-type: none"> <li>• Patient</li> <li>• Disease</li> <li>• Disease</li> </ul>	Disease comorbidity network	Rzhetsky et al., 2007; Hidalgo et al., 2009
<ul style="list-style-type: none"> <li>• Disease</li> <li>• Environmental factor</li> <li>• Disease-related genes</li> </ul>	Etiome: a database + clustering analysis of environmental + genetic (= etiological) factors of human diseases	Liu et al., 2009

<sup>a</sup> Here we included only those networks and datasets, which contained human diseases. Drug target networks and network datasets will be summarized in Section 4.1.3.

<sup>b</sup> References containing direct network analysis are marked with italic. All other references are referring to datasets, which are potential sources of future network representations.

Lee, 2012). Diseases occurring more frequently are associated with longer proteins (Lopez-Bigas & Ouzounis, 2004; Lopez-Bigas et al., 2005). Disease-associated proteins tend to have ‘younger’ folds, developed later in evolution, and have a smaller ‘family’ of similar folds. These protein folds are less designable (i.e. a smaller number of possible representations by different amino acid sequences) weakening the robustness against mutations, and the fitness of the hosting organism during evolution (Wong & Frishman, 2006).

- Going one level higher in the network hierarchy than the human interactome, to the level of comorbidity networks, patients tend to develop diseases in the vicinity of diseases they already had (Rzhetsky et al., 2007; Hidalgo et al., 2009; Barabási et al., 2011).
- Disease-hubs of comorbidity networks show a higher mortality than less well connected diseases, and are often successors of more peripheral diseases. The progression of diseases is different for patients of different genders and ethnicities (Lee et al., 2008a; Hidalgo et al., 2009; Barabási et al., 2011).

Human disease networks are expected to reveal more on the inter-relationships of diseases using both additional data-associations and

novel network analysis tools, listed in Section 2. These advances will not only enrich our integrated view on human diseases, but will also lead to the following potential uses of human disease networks:

- better classification of diseases (e.g. for putatively useful drugs and therapies) and predictions for understudied or unknown diseases;
- disease diagnosis and identification of disease biomarkers as described in detail in Section 1.3.3;
- identification of drug target candidates (including multi-target drugs, drug repositioning, etc.) as described in detail in Section 4.1;
- help in hit finding and expansion as described in detail in Section 4.2;
- enrich background data for lead optimization (including ADME, side-effects and toxicity, etc.) as described in detail in Section 4.3.

An increasing number of publications describe various molecular networks characterizing the cellular state in a certain type of disease. We have not included their direct description in this section, since here we only review the networks of the diseases as network nodes. In Section 5 we will summarize the drug-design related applications of these molecular networks in case of four disease families: infections, cancer, diabetes and neurodegenerative diseases. In the next section

**Table 3**  
Network-based predictions of disease-related genes as biomarkers.

• Type of prediction methods <sup>a</sup> • Type of data used	Name and additional description, website	References
<ul style="list-style-type: none"> <li>• Similarity-based</li> <li>• Protein structure descriptor-related QSAR</li> <li>• Interaction-based</li> <li>• (Predicted) interactome</li> </ul>	<p>New disease-related proteins are predicted by their structural similarity to known disease-related proteins</p> <p>New disease-related genes are predicted by their interactome neighborhood</p>	<p>Vilar et al., 2009</p> <p>Krauthammer et al., 2004; Chen et al., 2006a; Oti et al., 2006; Xu &amp; Li, 2006</p> <p>Guo et al., 2011</p>
<ul style="list-style-type: none"> <li>• Iterative summary of interactome and disease neighborhood</li> <li>• Disease similarity network, interactome</li> <li>• Semantic similarity score</li> <li>• Semantic similarity networks of diseases and related genes</li> </ul>	<p>Measures the neighborhood association in both the interactome and disease similarity networks and iteratively calculates the similarity of the node to diseases</p> <p>Calculates a semantic similarity score between gene ontology terms as well as human genes associated with them</p>	<p>Jiang et al., 2011</p>
<ul style="list-style-type: none"> <li>• Summarized network neighborhood of several candidate genes</li> <li>• Disease, gene-descriptions, disease related genes, interactome, mRNA co-expressions, pathways</li> <li>• Shortest path length</li> </ul>	<p>Constructs an integrative network and predicts candidate genes by their network closeness to known disease-related genes; Prioritizer: <a href="http://129.125.135.180/prioritizer">http://129.125.135.180/prioritizer</a></p>	<p>Franke et al., 2006</p>
<ul style="list-style-type: none"> <li>• Disease, gene-descriptions, disease related genes, interactome, mRNA co-expressions, pathways</li> <li>• User-defined path distance from known disease-related genes</li> <li>• Up to 10 integrated interactomes</li> </ul>	<p>Uses a maximum expectation gene cover algorithm finding small gene sets to predict associated new disease-related genes</p>	<p>Karni et al., 2009</p>
<ul style="list-style-type: none"> <li>• Interaction-based</li> <li>• Disease-related mutations, domain-domain resolved interactome</li> <li>• Interaction-based</li> <li>• Disease-related mutations, 3D structurally resolved interactome</li> <li>• Clustering</li> <li>• Disease-related genes, interactome</li> <li>• Closeness</li> <li>• Disease-related genes, disease network, interactome</li> </ul>	<p>New disease-related genes are predicted by their interactome closeness to known disease-related proteins; Genes2Networks: <a href="http://actin.pharm.mssm.edu/genes2networks">http://actin.pharm.mssm.edu/genes2networks</a></p> <p>New disease-related genes are predicted by their association to previously known disease-related genes at protein-protein domains affected by the disease-associated mutations of the known disease related gene</p> <p>New disease-related genes are predicted by their association to previously known disease-related genes at 3D modeled protein-protein interfaces affected by the disease-associated mutations of the known disease related gene</p> <p>New disease-related genes are predicted by their common protein-protein interaction network module with previous disease-related genes</p> <p>Closeness of unrelated proteins is calculated in the interactome from protein products of disease-related genes, and compared with phenotype similarity profile: large closeness marks a potential new disease-related gene; CIPHER: <a href="http://rulai.cshl.edu/tools/cipher">http://rulai.cshl.edu/tools/cipher</a></p>	<p>Berger et al., 2007</p> <p>Sharma et al., 2010a; Song &amp; Lee, 2012</p> <p>Wang et al., 2012b</p> <p>Navlakha &amp; Kingsford, 2010;</p> <p>Wu et al., 2008</p>
<ul style="list-style-type: none"> <li>• Random walk</li> <li>• Disease-related genes, disease network, interactome</li> </ul>	<p>Random walks in the interactome are started from protein products of disease-related genes: frequent visits of a previously unrelated protein mark a potential new disease-related gene; Cytoscape plug-in GPEC: <a href="http://sourceforge.net/p/gpec">http://sourceforge.net/p/gpec</a></p>	<p>Kohler et al., 2008; Chen et al., 2009b; Le &amp; Kwon, 2012</p>
<ul style="list-style-type: none"> <li>• Iterative network propagation</li> <li>• Disease-related genes, disease network, interactome</li> </ul>	<p>Iterative steps of information flow from disease-related and between interacting proteins: after convergence a large flow of a previously unrelated protein marks potential new disease-related gene; Cytoscape plug-in PRINCIPLE/PRINCE: <a href="http://www.cs.tau.ac.il/~bnet/software/PrincePlugin">http://www.cs.tau.ac.il/~bnet/software/PrincePlugin</a></p>	<p>Vanunu et al., 2010; Gottlieb et al., 2011b</p>
<ul style="list-style-type: none"> <li>• Random walk with re-starts in both networks</li> <li>• Disease-related genes, disease network, interactome</li> </ul>	<p>Random walk in both the interactome and the disease networks: number of frequent visits marks candidate genes</p>	<p>Li &amp; Patra, 2010</p>
<ul style="list-style-type: none"> <li>• NetworkBlast algorithm to align interactome and disease networks</li> <li>• Disease-related genes, disease network, interactome</li> </ul>	<p>After alignment of the interactome and disease networks finds high scoring subnetworks (bi-modules); candidate genes have the highest scoring bi-modules</p>	<p>Wu et al., 2009a</p>
<ul style="list-style-type: none"> <li>• Information flow with statistical correction</li> <li>• Disease-related genes, interactome</li> <li>• Topological network similarity</li> <li>• Disease-related genes</li> <li>• Neighborhood similarity</li> <li>• Disease-related genes, interactome, expression patterns</li> </ul>	<p>Statistically corrects random walk-based prediction with the degree distribution of the network; DADA: <a href="http://compbio.case.edu/dada">http://compbio.case.edu/dada</a></p> <p>Calculates neighborhood similarity in the interactome and prioritizes candidate genes; VAVIEN: <a href="http://diseasegenes.org">http://diseasegenes.org</a></p>	<p>Erten et al., 2011a</p> <p>Erten et al., 2011b</p>
<ul style="list-style-type: none"> <li>• Calculates expression weighted neighborhood similarity (using Katz centrality or other methods) in the interactome</li> </ul>	<p>Calculates expression weighted neighborhood similarity (using Katz centrality or other methods) in the interactome</p>	<p>Zhao et al., 2011b; Wu et al., 2012</p>
<ul style="list-style-type: none"> <li>• Semantic-based centrality</li> <li>• Disease-related genes, interactome, pathways</li> <li>• Direct neighbor-based Bayesian predictor</li> <li>• Disease-related genes, disease network, interactome, pathways</li> <li>• Genetic linkage analysis of gene network clusters</li> <li>• Disease-related genes, text mining-based associations (binding, phosphorylation, methylation, etc.)</li> </ul>	<p>Calculates data-type weighted centrality in the integrated network and uses it as a rank of candidate genes</p> <p>Constructs candidate protein complexes in a virtual pull-down experiment, and scores candidates by measuring the similarity between the phenotype in the complex and disease phenotype</p> <p>Calculates genetic linkage analysis of connected clusters in a text mining-derived direct interaction network</p>	<p>Gudivada et al., 2008</p> <p>Lage et al., 2007</p> <p>Iossifov et al., 2008</p>
<ul style="list-style-type: none"> <li>• Random forest learning</li> <li>• Disease, disease related genes, disease networks, single-nucleotide polymorphisms (SNPs)</li> <li>• Random walk, iterative network propagation (PRINCE/PRINCIPLE)</li> </ul>	<p>Predict deleterious SNPs and disease genes using the random forest learning method, uses interactomes and deleterious SNPs to predict disease-related genes by random forest learning</p> <p>A Cytoscape plug-in to construct an integrative network of diseases, associated genes, drugs and tissues; ICTNet: <a href="http://www.cs.queensu.ca/ictnet">http://www.cs.queensu.ca/ictnet</a></p>	<p>Care et al., 2009</p> <p>Wang et al., 2011b</p>
<ul style="list-style-type: none"> <li>• Disease, disease related genes, interactome, protein/DNA interaction, tissue, drug</li> <li>• Machine learning</li> <li>• Disease, disease related genes, gene annotations, interactome, expression levels, sequences</li> </ul>	<p>Integrative methods using similarities of neighbors or shortest paths in multiple data sources including interactomes; Endeavour: <a href="http://esat.kuleuven.be/endeavour">http://esat.kuleuven.be/endeavour</a>; Phenopred: <a href="http://www.phenopred.org">http://www.phenopred.org</a></p>	<p>Radivojac et al., 2008; Tranchevent et al., 2008; Linghu et al., 2009; Costa et al., 2010</p>

**Table 3** (continued)

• Type of prediction methods <sup>a</sup> • Type of data used	Name and additional description, website	References
• Rank coherence with target disease and unrelated disease networks • Disease, disease related genes, gene annotations, interactome, expression levels, genome-wide association studies	Calculates rank coherences between the integrated network characteristic to the target disease and unrelated diseases; rcNet: <a href="http://phegenex.cs.umn.edu/Nano">http://phegenex.cs.umn.edu/Nano</a>	Hwang et al., 2011

<sup>a</sup> The table summarizes methods using networks as data representations. We excluded those methods, like neural network or Bayesian network-based methods, which decipher associations between various, not network-assembled data. Several methods are included in the excellent reviews of Wang et al. (2011a) and Doncheva et al. (2012a).

we will illustrate the help of network analysis in the diagnosis and therapy of human diseases by the network-based identification of disease biomarkers.

### 1.3.3. Network-based identification of disease biomarkers

Network-based identification of disease related genes was suggested by relatively early studies (Krauthammer et al., 2004; Chen et al., 2006a; Franke et al., 2006; Gandhi et al., 2006; Oti et al., 2006; Xu & Li, 2006). In the last few years several network-based methods have been developed helping the identification of genes related to a particular disease as reviewed by the excellent summaries of Wang et al. (2011a) and Doncheva et al. (2012a). Table 3 summarizes methods for prediction of disease-related genes using networks as data representations. We excluded those network-related methods, like those neural network-based or Bayesian network-based methods, which decipher associations between various, not network-assembled data. Network prediction methods, which can also be used for prediction of disease-associated genes will be discussed in Section 2.2.2.

Most of the methods listed in Table 3 identify novel disease-related genes as disease biomarkers. Several network-based methods outperform former, sequence-based methods in the identification of novel, disease-related genes. Methods including non-local information of network topology usually perform better than methods based on local network properties. As a general trend the more information the method includes, the better prediction it may achieve. However, with the multiplication of datasets, biases and circularity may also be introduced, which will lead to an overestimation of the performance. Moreover, it is difficult to dissect the performance-contribution of the datasets and the prediction method itself. Additionally, each type of dataset may require a different method for optimal analysis. Therefore, the separate analysis of each data source was suggested with a subsequent combination of the ranking lists using rank aggregation algorithms. This procedure also facilitates backtracking the origin of the most relevant information. Functional GO-term annotations usually bring crucially important information to the analysis. The inclusion of interactome edge-based disease perturbations may improve the performance of these methods even further in the future (Kohler et al., 2008; Navlakha & Kingsford, 2010; Sharma et al., 2010a; Vanunu et al., 2010; Jiang et al., 2011; Wang et al., 2011a; Cho et al., 2012; Doncheva et al., 2012a). Importantly, several of the methods in Table 3 are not only able to diagnose known diseases, but may also identify important features of understudied or unknown diseases (Huang et al., 2010a; Wang et al., 2011a).

'Disease-related gene-hunting' became a very powerful area of medical studies. However, Erler and Linding (2010) warned that network models, and not their individual nodes, should be used as biomarkers, since thresholds and changes of individual nodes (such as the protein phosphorylation at a certain site) may be related to entirely different outcomes in different network contexts of different patients. We will summarize the concepts treating networks (and their segments) as drug targets in Section 4.1.7.

Very similar methods to those listed in Table 3 may be applied to network-based identification of disease-related signaling network, such as phosphorylation or microRNA profiles, or metabolome profiles. As

part of these approaches, metabolic network analysis was applied to identify metabolites, which may serve as biomarkers of a certain disease (Fan et al., 2012). Shlomi et al. (2009) identified 233 metabolites, whose concentration was elevated or reduced as a result of 176 human inborn dysfunctional enzymes affecting of metabolism. Their network-based method can provide a 10-fold increase in biomarker detection performance. Mass spectrometry phosphoproteome analysis combined with signaling networks and bioinformatics sources like NetworKIN and NetPhorest may provide biomarker profiles of several diseases such as cancer or cardiovascular disease (Linding et al., 2007; Yu et al., 2007a; Jin et al., 2008; Miller et al., 2008; Ummanni et al., 2011; Savino et al., 2012).

## 2. An inventory of network analysis tools helping drug design

Even the best network analytical methods will fail, if applied to a network constructed with a crude definition. Therefore, we start this section listing the major points of network definition including network-related questions of data collection, such as sampling, prediction and reverse engineering. The latter two methods are important network-related tools to find novel drug target candidates. We will continue and conclude this section by listing an inventory of the major concepts used in the analysis of network topology, comparison and dynamics evaluating their potential use in drug design. The section will give just the essence of the methods, and will provide the interested Reader a number of original references for further information.

### 2.1. Definition(s) and types of networks

To define a network we have to define its nodes and edges (Barabási & Oltvai, 2004; Boccaletti et al., 2006; Zhu et al., 2007; Csermely, 2009; Lovász, 2012). Network nodes are the entities building up the complex system represented by the network. Nodes are often called as vertices, or network elements. Classical, graph-type network descriptions do not consider the original character of nodes. (A node of such a graph will be "ID-234", which is characterized by its contact structure only.) Thus node definition requires a clear sense of those node properties, which discriminate network nodes from other entities, and make them 'equal'. Recently, node-weights were successfully applied to characterize the node structure of a network in a simple form (Wiedermann et al., 2013). In the case of molecular networks, where nodes are amino acids, proteins or other macromolecules such discrimination is rather easy. However, subtle problems may still remain. For example, should we include extracellular proteins as well? If not, what happens, if an extracellular protein is just about to be secreted? What if it is engulfed by the cell and internalized? Node definition may become especially difficult in the case of complex data structures, like those we mentioned in Section 1.3. Accurate node definitions are time consuming, but lead to benefits at subsequent stages.

Network edges are often called interactions, connections, or links. In the molecular networks discussed in this review edges represent physical or functional interactions of two network nodes (Zlatic et

al., 2009). However, in hypergraph representations meta-edges often connect more than two nodes. Edge definition often inherently contains a threshold determined by the detection limit and by the time-window of the observation. Two nodes may become connected, if the sensitivity and/or duration of detection are increased. A number of recent publications explored the effect of time-window changes on the structure of social networks (Krings et al., 2012; Perra et al., 2012). Several concepts of network dynamics detailed in Section 2.5 are inherently related to time-window of detection. As an example, the distinction of the popular date hubs (Han et al., 2004a), i.e. hubs changing their partners over time, clearly depends on the time-window of observation.

Weights of network edges may give an answer to the “where-to-set-the-detection-threshold” dilemma offering a continuous scale of interactions. Edge weights represent the intensity (strength, probability, affinity) of the interaction. Edges may also be directed, where a sequence of action and/or a difference in node influence are included in the edge definition. Lovász (2012) gives an excellent summary of the basic dilemmas of network definition problems.

However, we have many more options than defining network nodes, edges, weights and directions. Recent network descriptions started to explore the options to include edge reciprocity (Squartini et al., 2012), or to preserve multiple node attributes (Kim & Leskovec, 2011). Moreover, in reality networks are seldom directed in an unequivocal way. (When CEOs and VPs are talking to each other, it is not always the case that CEOs influence VPs, and VPs do not influence CEOs.) However, to date, a continuous scale of edge direction has not been introduced to molecular networks. Edges may also be colored, where different types of interactions are discriminated. A special subset of colored networks is signed networks, where edges are either positive (standing for activation) or negative (representing inhibition). Edges may also be conditional, i.e. being only active, if one of their nodes accommodated another edge previously. There are a number of potential uses of these network representations e.g. in signaling, or in genetic interaction networks.

As a closing remark, the definition of edges often hides one of two fundamentally different concepts. Network connections may either restrict the connected nodes (this is the case, where connections represent physical contacts), or may enrich connected nodes (this is the case, where connections represent channels of transport or information transmission). These constraint-type or transmission-type network properties may appear in the same network, where they may be simplified to activation or inhibition like those in signal transduction networks. Though there were initial explorations of the differences of constraint-type and transmission-type network properties (Guimera et al., 2007a), an extended application of this concept is missing.

## 2.2. Network data, sampling, prediction and reverse engineering

Lovász (2012) gives an excellent summary of the network sampling problem. In most biological systems data coverage has technical limitations, and experimental errors are rather prevalent. As part of these uncertainties and errors, not all of the possible interactions are detected, and a large number of false-positives may also appear (Zhu et al., 2007; De Las Rivas & Fontanillo, 2010; Sardu & Washburn, 2011). However, it is often a question of judgment, whether the investigator believes that only ‘high-fidelity’ interactions are valid, and discards all other data as potential artifacts, or uses the whole spectrum of data considering low-confidence interactions as low affinity and/or low probability interactions (Csermely, 2004, 2009). The highest quality interactions are reliable, but may not be representative of the whole network (Hakes et al., 2008). The unavailability of complete datasets can be circumvented by a number of methods which 1.) help the correct sampling of networks; 2.) enable the prediction of nodes/edges and 3.) infer network structure from the behavior of the complex

system by reverse engineering. We will discuss these methods in this section.

### 2.2.1. Problems of network incompleteness, network sampling

Since complex networks are not homogenous, their segments may display different properties than the whole network (Han et al., 2005; Stumpf et al., 2005; Tanaka et al., 2005; Stumpf & Wiuf, 2010; Annibale & Coolen, 2011; Son et al., 2012). Therefore, the use of a representative sample of the network is a key issue. In the last few years several methods became available to assess whether the available part of an unknown complete network is a representative sample. These methods also allow the extrapolation of the partially available network data to the total dataset (Wiuf et al., 2006; Stumpf et al., 2008). Radicchi et al. (2011) introduced a GloSS filtering technique preserving both the weight distribution and network topology. Recently a comparison of several (re)-sampling methods was given (Mirshahvalad et al., 2013; Wang, 2012). Guimera and Sales-Pardo (2009) provided a method to detect missing interactions (false negatives) and spurious interactions (false positives). Riera-Fernandez et al. (2012) gave numerical quality scores to network edges based on the Markov–Shannon entropy model. However, data purging methods should be applied with caution, since unexpected edges of ‘creative nodes’ may also be identified as ‘spurious’ edges, and may be removed (Csermely, 2008; Lü & Zhou, 2011). Network sampling methods were recently reviewed by Ahmed et al. (2012).

### 2.2.2. Prediction of missing edges and nodes, network predictability

Prediction of missing edges and nodes is not only important to assess network reliability, but can also be used for predictions of e.g. hitherto undetected interactions of disease-related proteins, or extension of drug target networks helping drug design (Spiró et al., 2008). In Section 1.3.3 and Tables 2 and 3 we already listed several methods for the efficient prediction of new edges and nodes from complex human disease-related datasets. Prediction is not only a discovery tool, but it also helps to avoid the unpredictable, which is considered as dangerous. However, as we will see at the end of this section, in complex systems the least predictable constituents are the most exciting ones.

Lü and Zhou (2011) provided an excellent review of edge prediction. Referring to this paper for details here we will summarize only the major points of this field.

- Edges can be predicted by the properties of their nodes, e.g. protein sequences, or domain structures (Smith & Sternberg, 2002; Li & Lai, 2007; Shen et al., 2007; Hue et al., 2010).
- The similarity of the edge neighborhood in the network is widely used in edge prediction. Edge neighborhood may be restricted to the common neighbors of the connected nodes, may include all first neighbors, all first and second neighbors, cliques, the nodes’ network modules, or the whole network. Consequently, similarity indices may be local (like the Adamic–Adar index, common neighbors index, hub promoted index, hub suppressed index, Jaccard index, Leicht–Holme–Newman index, preferential attachment index, resource allocation index, Salton index, or the Sørensen index) mesoscopic (like the local path index or the local random walk index), or global (like the average commute time index, cosine-based index, Katz index, Leicht–Holme–Newman index, matrix forest index, random walk with restart index, or the SimRank index). Edge neighborhood may be compared by using the network degree, preferential attachment methods, fitness values, community structure, network hierarchy, a stochastic bloc model, a probabilistic model, or by using hypergraphs (Albert & Albert, 2004; Liben–Nowell & Kleinberg, 2007; Yan et al., 2007a; Guimera & Sales-Pardo, 2009; Lu et al., 2009; Zhou et al., 2009; Chen et al., 2012a; Eronen & Toivonen, 2012; Hu et al., 2012; Musmeci et al., in press; Yan & Gregory, 2012; Liu et al., 2013). It is important to note that methods may perform differently, if the missing edge is in a dense network core or in a sparsely connected network

periphery (Zhu et al., 2012a). The optimal method also depends on the average length of shortest paths in the network. Edge prediction methods often require a large increase in computational time to achieve a higher accuracy (Lü & Zhou, 2011).

- Edge prediction can be performed by comparing the network to an appropriately selected model network, to a similar real world network, or to an ensemble of networks (Liben-Nowell & Kleinberg, 2007; Clauset et al., 2008; Nepusz et al., 2008; Xu et al., 2011a; Gutfraind et al., 2012).
- Edges can also be predicted by the analysis of sequential snapshots of network topology (also called as network dynamics, or network evolution, see Section 2.5; Hidalgo & Rodriguez-Sickert, 2008; Lü & Zhou, 2011). In network time-series older events might have less influence on the formation of a new edge than newer ones. Additionally, all network evolution models can be used as edge-predictors. However, one has to keep in mind that network evolution models always include guesses about the factors influencing the generation of a novel edge (Lü & Zhou, 2011).

Edge prediction of drug–target networks allows the discovery of new drug target candidates and the repositioning of existing drugs (van Laarhoven et al., 2011). Prediction methods may combine several data-sources, like mRNA expression patterns, genotypic data, DNA–protein and protein–protein interactions (Zhu et al., 2008; Pandey et al., 2010). Dataset combination may help the precision of edge prediction. However, prediction of directed, weighted, signed, or colored edges of these combined datasets is still a largely unsolved task (Lü & Zhou, 2011).

Node prediction is even more difficult, than edge prediction (Getoor & Diehl, 2005; Liben-Nowell & Kleinberg, 2007). Predicted nodes may occupy structural holes, i.e. bridging positions between multiple network modules (Burt, 1995; Csermely, 2008), or may be identified by methods, like chance-discovery. Chance-discovery uses an iterative annealing process, and extends the dense clusters observed at lower annealing ‘temperatures’ (Maeno & Ohsawa, 2008). In fact, the well developed methodology of the identification of disease-related genes that we detailed in Section 1.3 can be regarded as a node prediction problem, and may give exciting clues for node prediction in networks other than those of disease-related data.

The predictability of network edges is not only a function of data coverage and network structure, but also depends on network dynamics. The mistaken identification of unexpected edges as spurious edges (Lü & Zhou, 2011), and the better predictability of edges in dense cores than those in network periphery (Zhu et al., 2012a) are both related to the inherent unpredictability caused by network dynamics. As an example, the edge-structure of date hubs, where hubs change their neighbors (Han et al., 2004a), is certainly less predictable than that of party hubs, i.e. hubs preserving a rather constant neighborhood. Date hubs mostly reside in inter-modular positions (Han et al., 2004a; Komurov & White, 2007; Kovács et al., 2010). Predictability is also related to network rigidity and flexibility (Gáspár & Csermely, 2012): an edge or node in a more flexible network position is less predictable than others situated in a rigid network environment.

Bridging positions are often more flexible and less predictable than intra-modular edges. If a node is connecting multiple, distant modules with approximately the same, low intensity, and continuously changing its position, like the recently described ‘creative nodes’ do (Csermely, 2008), its predictability will be exceptionally low. A shift towards lower predictability (higher network flexibility) is often accompanied by an increased adaptation capability at the system level. Moreover, a complex system lacking flexibility is unable to change, to adapt and to learn (Gyurkó et al., in press). Thus it is not surprising that highly unpredictable, ‘creative’ nodes characterize all complex systems. Importantly, these highly unpredictable nodes help in delaying critical transitions of the systems, i.e. postponing market crash, ecological disaster or death (Csermely, 2008; Scheffer et al., 2009; Farkas et al., 2011; Sornette & Osorio, 2011; Dai et al., 2012). In fact, the most unpredictable nodes are the most exciting

nodes of the system having a hidden influence on the fate of the whole system at critical situations. The prediction of their unpredictable behavior remains a major challenge of network science.

### 2.2.3. Prediction of the whole network, reverse engineering, network-inference

There are situations, when the network is so incomplete that we do not know anything on the network structure. However, we often have a detailed knowledge of the behavior of the complex system encoded by the network. The elucidation of the underlying network from the emergent system behavior is called reverse engineering or network-inference.

In a typical example of reverse engineering we know the genome-wide mRNA expression pattern and its changes after various perturbations (including drug action, malignant transformation, development of other diseases, etc.), but we have no idea of the gene–gene interaction network, which is causing the changes in mRNA expression pattern. As a rough estimate, a network of 10,000 genes can be predicted with reasonable precision using less than a hundred genome-wide mRNA datasets. Network prediction can be greatly helped using previous knowledge, e.g. on the modules of the predicted network. The correct identification of the relatedness of mRNA expression sets (position in time series, tissue-specificity, etc.) may often be a more important determinant of the final precision of network prediction than the precise measurement of the mRNA expression levels. Models of network dynamics, probabilistic graph models and machine learning techniques are often incorporated in reverse engineering methods. Some of these approaches, like Bayesian methods, require a rather intensive computational time. Therefore, computationally less expensive methods such as the copula method, or the simultaneous expression model with Lasso regression were also introduced. The topology of the predicted network often determines the type of the best method. This is one reason, why combination of various methods (or the use of iterative approaches) may outperform individual methodologies (Liang et al., 1998a; Akutsu et al., 1999; Ideker et al., 2000; Kholodenko et al., 2002; Yeung et al., 2002; Segal et al., 2003; Tegnér et al., 2003; Friedman, 2004; Tegnér & Björkegren, 2007; Cosgrove et al., 2008; Kim et al., 2008; Ahmed & Xing, 2009; Stokic et al., 2009; Marbach et al., 2010, 2012; Yip et al., 2010; Pham et al., 2011; Schaffter et al., 2011; Altay, 2012; Crombach et al., 2012; Kotera et al., 2012). Jurman et al. (2012a) designed a network sampling stability-based tool to assess network reconstruction performance.

Reverse engineering techniques were successfully applied to reconstruct drug-affected pathways (Gardner et al., 2003; di Bernardo et al., 2005; Chua & Roth, 2011; Gosline et al., 2012). Besides the identification of gene regulatory networks from the transcriptome, reverse engineering methods may also be used to identify signaling networks from the phosphorome or signaling network (Kholodenko et al., 2002; Sachs et al., 2005; Zamir & Bastiaens, 2008; Eduati et al., 2010; Prill et al., 2011), metabolic networks from the metabolome (Nemenman et al., 2007), or drug action mechanisms and drug target candidates from various datasets (Gardner et al., 2003; di Bernardo et al., 2005; Lehár et al., 2007; Lo et al., 2012; Madhamshettiwar et al., 2012).

Though the number of reverse-engineering methods has been doubled every two years, 1.) the inclusion of non-linear system dynamics, of multiple data sources and of multiple methods; 2.) distinguishing between direct and indirect regulations; 3.) a better discrimination between causal relationships and coincidence; as well as 4.) network prediction in case of multiple regulatory inputs per node remain major challenges of the field (Tegnér & Björkegren, 2007; Marbach et al., 2010).

### 2.3. Key segments of network structure

In this section we will give a brief summary of the major concepts and analytical methods of network structure starting from local

network topology and proceeding towards more and more global network structures. Selection of key network positions as drug target options has a major dilemma. On the one hand, the network position has to be important enough to influence the diseased body; on the other hand, the selected network position must not be so important that its attack would lead to toxicity. The successful solution of this dilemma requires a detailed knowledge on the structure and dynamics of complex networks.

### 2.3.1. Local topology: hubs, motifs and graphlets

A minority of nodes in a large variety of real world networks is a hub, i.e. a node having a much higher number of neighbors than average. Real world networks often have a scale-free degree distribution providing a non-negligible probability for the occurrence of hubs, as it was first generalized to real world networks by the seminal paper of Barabási and Albert (1999). If hubs are selectively attacked, the information transfer deteriorates rapidly in most real world networks. This property made hubs attractive drug targets (Albert et al., 2000). However, some of the hubs are essential proteins, and their attack may result in increased toxicity. This narrowed the use of major hubs as drug targets mostly to antibiotics, to other anti-infectious drugs and to anticancer therapies. In agreement with these, on average, targets of FDA-approved drugs tend to have more connections than peripheral nodes, but fewer connections than hubs (Yildirim et al., 2007). Cancer-related proteins have many more interaction partners than non-cancer proteins making the targeting of cancer-specific hubs a reasonable strategy in anti-cancer therapies (Jonsson & Bates, 2006). Besides the direct count of interactome neighbors algorithms have been developed to identify hubs using Gene Ontology terms (Hsing et al., 2008). Going one level deeper in the network hierarchy, amino acids serving as hubs of protein structure networks play a key role in intra-protein information transmission (Pandini et al., 2012), and may provide excellent target points of drug interactions.

The emerging picture of using hubs as drug targets can be summarized by two opposite effects. On the one hand, hubs are so well connected that their attack may lead to cascading effects compromising the function of a major segment of the network; on the other, nodes with limited number of connections are at the 'ends' of the network, and their modulation may have only limited effects (Penrod et al., 2011). There are several important remarks refining this conclusion.

- Not all hubs are equal. Weighted and directed networks are extremely important in discriminating between hubs. A hub having 20 neighbors connected with an equal edge-weight is different from a hub having the same number of 20 neighbors having a highly uneven edge-structure of a single, dominant edge and 19 low intensity edges. A sink-hub with 20 incoming edges is not at all the same than a source-hub with the same number 20 outgoing edges. Soluble proteins possess more contacts on average than membrane proteins (Yu et al., 2004a) warning that the hub-defining threshold of neighbors cannot be set uniformly.
- Hub-connectors, i.e. edges or nodes connecting major hubs also offer very interesting drug targeting options (Korcsmáros et al., 2007; Farkas et al., 2011).
- Not all peripheral nodes are unimportant. There are peripheral nodes called 'choke points', which uniquely produce or consume an important metabolite. The inhibition of 'choke points' often leads to a lethal effect (Yeh et al., 2004; Singh et al., 2007).
- Importantly, interdependent networks, i.e. at least two interconnected networks, were shown to be much more vulnerable to attacks than single network structures (Buldyrev et al., 2010). We have several interdependent networks in our cells, such as the networks of signaling proteins and transcription factors, or the interactome of membrane proteins and the network of the interacting nuclear, plasma, mitochondrial and endoplasmic reticulum membranes. The excessive

vulnerability of interdependent networks should make us even more cautious in the selection of drug target nodes. The options of edgetic drugs, multi-target drugs and allo-network drugs, we will describe in Section 4.1.6 (Nussinov et al., 2011), may circumvent the worries and problems related to the single and direct targeting of network nodes with drugs.

Network motifs are circuits of 3 to 6 nodes in directed networks that are highly overrepresented as compared to randomized networks (Milo et al., 2002; Kashtan et al., 2004). Graphlets are similar to motifs but are defined as undirected networks (Przulj et al., 2006). Motifs proved to be efficient in predicting protein function, protein–protein interactions and development of drug screening techniques (Bu et al., 2003; Albert & Albert, 2004; Luni et al., 2010; Cloutier & Wang, 2011). Rito et al. (2010) made an extensive search for graphlets in protein–protein interaction networks and concluded that interactomes may be at the threshold of the appearance of larger motifs requiring 4 or 5 nodes. Such a topology would make interactomes both efficient having not too many edges and robust harboring alternative pathways.

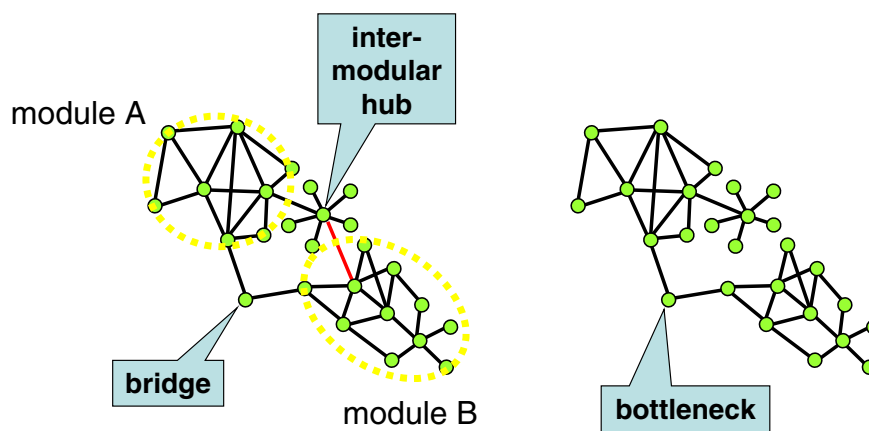
### 2.3.2. Broader network topology: modules, bridges, bottlenecks, hierarchy, core, periphery, choke points

Network modules (or in other words: network communities) are the primary examples of mesoscopic network structures, which are neither local, nor global. Modules represent groups of networking nodes, and are related to the central concept of object grouping and classification. Modules of molecular networks often encode cellular functions. Moreover, the exploration of modular structure was proposed as a key factor to understand the complexity of biological systems. Therefore, module determination gained much attention in recent years. Modules of molecular networks are formed from nodes, which are more densely connected with each other than with their neighborhood (Girvan & Newman, 2002; Fortunato, 2010; Kovács et al., 2010; Koch, 2012; Szalay-Bekó et al., 2012). In Section 1.3 we introduced disease modules, i.e. modules of disease-related genes in protein–protein interaction networks (Goh et al., 2007; Oti & Brunner, 2007; Jiang et al., 2008; Suthram et al., 2010; Bauer-Mehren et al., 2011; Loscalzo & Barabasi, 2011; Nacher & Schwartz, 2012). These node-related properties influence the modular functions, making them attractive network drug-targets (Cho et al., 2012). However, the determination of network modules proved to be a notoriously difficult problem resulting in more than two hundred independent modularization methods (Fortunato, 2010; Kovács et al., 2010).

Modules of molecular networks have an extensive (often called pervasive) overlap, which was recently shown to be denser than the center of the modules in some social networks (Palla et al., 2005; Ahn et al., 2010; Kovács et al., 2010; Yang & Leskovec, 2012). This reflects the economy of our cells using a protein in more than one function. Modules with sparse edge structure also characterize protein–protein interaction networks (Srihari & Leong, 2012). Modules of real world networks were shown to form a 'very small world' having an average distance of 3 from each other (Li & Li, 2013). Inter-modular nodes are attractive drug targets.

Bridges connect two neighboring network modules (Fig. 8). Bridges may also be identified by k-shell analysis (Reppas & Lawyer, 2012). Bridges usually have fewer neighbors than hubs, and are independently regulated from the nodes belonging to both modules, which they connect. This makes them attractive as drug targets, since they may display lower toxicity, while the disruption of information flow between functional network modules could prove to be therapeutically effective (Hwang et al., 2008). Proteins involved in the aging process are often bridges (Wang et al., 2009). Proteins bridging disease modules may provide important points of interventions (Nguyen & Jordan, 2010; Nguyen et al., 2011).

Inter-modular hubs form a special class of inter-modular nodes (Fig. 8). Date hubs, i.e. hubs having only a single or few binding sites



**Fig. 8.** Bridge, inter-modular hub and bottleneck. The network on the left side of the figure has two modules (modules A and B marked by the yellow dotted lines), which are connected by a bridge and by an inter-modular hub. By the removal of the red edge from the network on the left side, the former bridge obtains a unique and monopolistic role connecting modules A and B, and is therefore called as a bottleneck. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and frequently changing their protein partners, were shown to occupy an inter-modular position as opposed to party hubs residing mostly in modular cores (Han et al., 2004a; Kim et al., 2006; Komurov & White, 2007; Kovács et al., 2010). Party hubs tend to have higher affinity binding surfaces than date hubs (Kar et al., 2009). Inter-modular hubs usually have a regulatory role (Fox et al., 2011), and are mutated frequently in cancer (Taylor et al., 2009).

Nodes occupying a unique and monopolistic inter-modular position have been termed 'bottlenecks' (Fig. 8), because almost all information flowing through the network must pass through these nodes. This makes bottlenecks more effective drug targets than bridges (Yu et al., 2007b). In agreement with this concept, hub-bottlenecks were shown to be preferential targets of microRNAs (Wang et al., 2011c) and play an important role in cellular re-programming (Buganim et al., 2012). However, inhibition of bottlenecks often compromises network integrity, restricting their use as drug targets to anti-infectious and (in the case of cancer-specific bottlenecks) anti-cancer therapies (Yu et al., 2007b). In agreement with this proposition, cancer proteins tend to be inter-modular hubs of cancer-specific networks offering an important target option (Jonsson & Bates, 2006).

Nodes connecting more than two modules are in modular overlaps. Overlapping nodes occupy a network position, which can provide more subtle regulation than bridges or bottlenecks. Modular overlaps are primary transmitters of network perturbations, and are key determinants of network cooperation (Farkas et al., 2011). Overlapping nodes play a crucial role in cellular adaptation to stress. In fact, changes in the overlap of network modules were suggested to provide a general mechanism of adaptation of complex systems (Mihalik & Csermely, 2011; Csermely et al., 2012). Modular overlaps (called cross-talks between signaling pathways) are most prevalent in humans, if compared to *Caenorhabditis elegans* or *Drosophila* (Korcsmáros et al., 2010). All these make modular overlaps especially attractive drug targets (Farkas et al., 2011). As we described earlier, 'creative nodes' are in the overlap of multiple modules belonging roughly equally to each module. These nodes play a prominent role in regulating the adaptivity of complex networks, and are lucrative network targets (Csermely, 2008; Farkas et al., 2011).

Despite the important role of hierarchy in network structures (Ravasz et al., 2002; Liu et al., 2012; Mones et al., 2012), the exploration of network hierarchy is largely missing from network pharmacology. Ispolatov and Maslov (2008) published a useful program to remove feedback loops from regulatory or signaling networks, and reveal their remaining hierarchy (<http://www.cmth.bnl.gov/~maslov/programs.htm>). Hartsperger et al. (2010) developed HiNO using an improved, recursive approach to reveal network hierarchy (<http://mips.helmholtz-muenchen.de/hino>). The hierarchical map approach of Rosvall and

Bergstrom (2011) used the shortest multi-level description of a random walk (<http://www.tp.umu.se/~rosvall/code.html>). A special class of hierarchy-representation and visualization uses the hierarchical structure of modules, i.e. the concept that modules can be regarded as meta-nodes and re-modularized, until the whole network coalesces into a single meta-node. Methods like Pyramabs (<http://140.113.166.165/pyramabs.php>; Cheng & Hu, 2010) or the Cytoscape (Smoot et al., 2011) plug-in, ModuLand (<http://linkgroup.hu/modules.php>; Szalay-Bekó et al., 2012) are good examples of this powerful approach.

Network hierarchy has recently been involved as a key factor of network controllability (Liu et al., 2012; Mones et al., 2012), which will be discussed in Section 2.3.4 in detail. However, not all hierarchical networks are 'autocratic', where top nodes have an unparalleled influence. Horizontal contacts of middle-level regulators play a key role in gene regulatory networks. Moreover, such a 'democratic network character' increases markedly in human gene regulation (Bhardwaj et al., 2010).

Similarly, the discrimination between network core and periphery has been published quite a while ago (Guimera & Amaral, 2005) and was extended recently to network modules (Li & Li, 2013), but its applications are largely missing from the field of drug design. As an example of the possible benefits, choke points were identified as those peripheral nodes that either uniquely produce or consume a certain metabolite (including here signal transmitters and membrane lipids too). Efficient inhibition of choke points may cause either a lethal deficiency, or toxic accumulation of the metabolite (Yeh et al., 2004; Singh et al., 2007).

### 2.3.3. Network centrality, network skeleton, rich-club and onion-networks

Network centrality measures span the entire network topology from local to global. Centrality is related to the concept of importance. Central nodes may receive more information, and may have a larger influence on the networking community. Thus it is not surprising that dozens of network centrality measures have been defined. Several centrality measures are local, like the number of neighbors (the network degree), or related to the modular structure, like bridging centrality, community centrality, or subgraph centrality. Centrality measures, like betweenness centrality (the number of shortest paths traversing through the node), random walk related centralities (like the PageRank algorithm of Google), or network salience are based on more global network properties. Recently a number of centrality measures have been defined based on network dynamics (Freeman, 1978; Estrada & Rodríguez-Velázquez, 2005; Estrada, 2006; Hwang et al., 2008; Kovács et al., 2010; Du et al., 2012; Ghosh & Lerman, 2012; Grady et al., 2012; Grassler et al., 2012;

Joseph & Chen, 2012; Mantzaris et al., in press). Global network centrality calculations may be faster, assessing only network segments and using network compression (Sariyüce et al., 2012). Network module-based centralities are related to the determination of bridges and overlaps (Hwang et al., 2008; Kovács et al., 2010), while betweenness centrality is used for the definition of bottlenecks (Yu et al., 2007b). Both are important drug target candidates as we discussed in the previous section. As an additional example, high betweenness centrality hubs were shown to dominate the drug–target network of myocardial infarction (Azuaje et al., 2011).

The network skeleton is an interconnected subnetwork of high centrality nodes. Network skeletons may contain hubs (we call this a ‘rich-club’; Colizza et al., 2006; Fig. 9), may consist of high betweenness centrality nodes (Guimera et al., 2003), or may comprise inter-connected centers of network modules (Kovács et al., 2010; Szalay-Bekő et al., 2012). Network skeletons may be densely interconnected forming an inner core of the network, or may be truly skeleton-like traversing the network like a highway. In both network skeleton representations nodes participating in the network skeleton form the ‘elite’ of the network, like the respective persons in social networks (Avin et al., 2011). Network skeleton nodes are attractive drug target candidates. As an example of this, Milenkovic et al. (2011) defined a dominating set of nodes as a connected network subgraph having all residual nodes as its neighbor. They showed that the dominating set (especially if combined with a network-module type centrality measure called as graphlet degree centrality measuring the summative degree of neighborhoods extending to 4 layers of neighbors) captures disease-related and drug target genes in a statistically significant manner. It will be interesting to see, whether the recently defined intra-modular dominating sets (Li & Li, 2013) also possess similar features. Nicosia et al. (2012) defined a subset of nodes (called controlling sets), which can assign any prescribed set of centrality values to all other nodes by cooperatively tuning the weights of their out-going edges. Nacher and Schwartz (2008) identified a rich-club of drugs serving as a core of the drug–therapy network composed of drugs and established classes of medical therapies.

Network assortativity characterizes the preferential attachment of nodes having similar degrees to each other. Network cores (such as rich-clubs, Fig. 9) may or may not be a part of an assortative network. In a disassortative network low degree, peripheral network nodes are connected to the network core and not to each other. These core–periphery networks have a nested structure (Fig. 9). If peripheral nodes are connected to each other and form consecutive rings around the core, we call the network ‘onion-type’ (Fig. 9). Nested networks were shown to characterize ecosystems and trade networks, while onion-networks are especially resistant against targeted attacks (Saavedra et al., 2011; Schneider et al., 2011; Wu & Holme, 2011). Despite the exciting features of nested and onion networks, these network

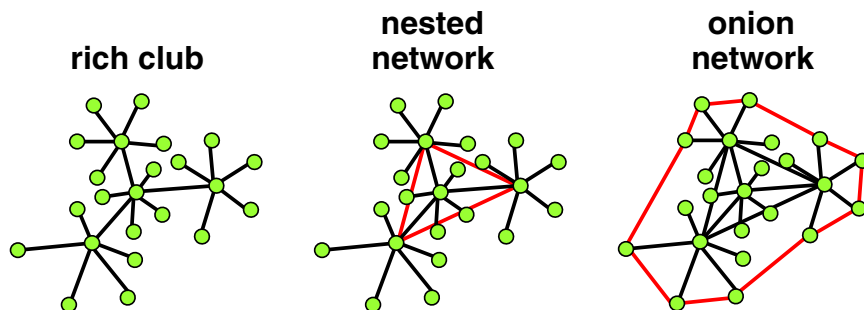
characteristics have not been assessed yet in disease-related, or drug design related-studies.

### 2.3.4. Global network topology: small worlds, network percolation, integrity, reliability, essentiality and controllability

The global topology of most real world networks is characterized by the small world property first generalized in the landmark paper of Watts and Strogatz (1998). Nodes of small worlds are connected well—as it was popularized by the proverbial “six degrees of separation” meaning that members of the social network of Earth can reach each other using 6 consecutive contacts (edges) as an average. In fact, modern web-based social networks, like Facebook, are an even smaller world having an average shortest path of 4.74 edges (Blackstrom et al., 2011).

Percolation is a broader term of global network topology than small worldness, since it refers to the connectedness of network nodes, i.e. the presence of a connected, giant network component. Sequential attacks on network nodes can induce a progressive and dramatic decrease of network percolation. Despite being a sensitive measure, to date the concept of percolation has not been extended to characterize network modules and other non-global structures of molecular networks (Antal et al., 2009). Percolation is related to network integrity and network reliability; that is, related to how much of the network remains connected if a network node or edge fails. In the case of directed networks the connection of sources or sinks can be calculated separately (Gertsbakh & Shprungin, 2010). The network efficiency measure of Latora and Marchiori (2001) is a widely used criterion to judge the integrity of a network. As noted before, intentional attack of hubs can be deleterious to most real world networks (Albert et al., 2000). The effect of a single attack of the largest hub in gene transcription networks can be substituted by a surprisingly low number of partial attacks, which is making the multi-target approaches listed in Section 4.1.5 a viable option from the network point of view (Agoston et al., 2005; Csermely et al., 2005).

In the case of anti-infectious or anti-cancer agents we would like to destroy the network of the parasite or of the malignant cell. In other words we need to predict essential proteins as targets of these therapeutic approaches. This makes network integrity a key measure to judge the efficiency of drug target candidates in these fields. Prediction of essential proteins is also important to forecast the toxicity of other drugs. The number of neighbors in protein–protein interaction networks is an important network measure of essentiality (Jeong et al., 2001). Later more global network measures were also shown to contribute to the prediction of node essentiality (Chin & Samanta, 2003; Estrada, 2006; Yu et al., 2007b; Missiuro et al., 2009; Li et al., 2011a; Song & Singh, 2013). Edge weights and directions may also significantly alter the determination of attack efficiency (Dall’Astra et al., 2006; Yu et al., 2007b). Finally, the constraints of



**Fig. 9.** Rich club, nested network and onion network. The figure illustrates the differences between a network having a rich club (left side), having a highly nested structure (middle) and developing an onion-type topology (right side). Note that the connected hubs of the rich club became even more connected by adding the 3 red edges on the middle panel. Connection of the peripheral nodes by an additional 10 red edges on the right panel turns the nested network to an onion network having a core and an outer layer. Note that the rich club network already has a nested structure, and both the nested network and the onion network have a rich club. Larger onion networks have multiple peripheral layers. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



metabolic networks define different contexts of essentiality exemplified by choke points, i.e. proteins uniquely producing or consuming a certain metabolite (Yeh et al., 2004; Singh et al., 2007). We will describe metabolic network essentiality in Section 3.6.2 in detail.

The most recent aspect of global network topology is similar to essentiality in the sense that it is also related to the influence of nodes on network behavior. However, here node influence is not judged on a 'yes/no scale', i.e. by whether the organism survives the malfunction of the node, but based on the more subtle scale of changing cell behavior. In this way node influence studies are closely related to network dynamics as we will detail in Section 2.5. Network centrality measures, or the dominating set of network nodes we mentioned before, are also related to the influence of selected nodes on others. Recent publications added network controllability, i.e. the ability to shift network behavior from an initial state to a desired state, to the repertoire of network-related measures of node influence. From these initial studies central nodes emerged as key players of network control (Cornelius et al., 2011; Liu et al., 2011, 2012; Banerjee & Roy, 2012; Cowan et al., 2012; Mones et al., 2012; Nepusz & Vicsek, 2012; Wang et al., 2012a; Pósfai et al., 2013). It is important to note that control here is a weak form of control, since we do not want to control how the system reaches the desired state (San Miguel et al., 2012). Despite of the clear applicability of network controllability to drug design (i.e. finding the nodes, which can shift molecular networks of the cell from a malignant state to a healthy state) there were only a few studies testing various aspects of this rich methodology in drug design (Xiong & Choe, 2008; Luni et al., 2010). Development of drug-related applications of network influence and control models is an important task of future studies.

#### 2.4. Network comparison and similarity

As we summarized in Section 2.2, uncovering network similarities is useful to predict nodes and edges. Alignment of networks from various species identifies interologs corresponding to conserved interactions between a pair of proteins having interacting homologs in another organism, or the analogous regulogs in regulatory networks, signalogs in signal transduction networks and phenologs as disease associated-genes. Thus, network comparison may uncover novel protein functions and disease-specific changes. All these greatly help drug design (Yu et al., 2004b; Sharan et al., 2005; Leicht et al., 2006; Sharan & Ideker, 2006; Zhang et al., 2008; McGary et al., 2010; Korcsmáros et al., 2011). However, the great potential to uncover network similarities comes with a price: network comparison is computationally expensive, and remains one of the greatest challenges of the field.

Lovász (2009, 2012) gives an excellent summary of the network similarity problem including a number of network similarity measures such as edit distance (the number of edge changes required to get one network from another), sampling distance (measuring the similarity by an ensemble of random networks), cut distance and similarity distance. A later study also used an interesting combined distance metrics of the edit and spectral distances (Jurman et al., 2012b). Similarity measures based on the comparison of the top-k nodes were recently described (Amin et al., 2012; Lee et al., 2012a). Similarity indices may be local (comparing the closest neighborhood of selected nodes), mesoscopic (which are usually based on local walks), or global (often involving extensive, network-wide walks). Edge neighborhood may be compared by using the modular structure, hypergraphs, network hierarchy, a stochastic bloc model, or a probabilistic model. Comparison may also use an ensemble of random, scale-free or other model networks, and the distribution of the best fitting ensemble. Reviews of Sharan and Ideker (2006), Zhang et al. (2008) and Lü and Zhou (2011) give further details of the methodology used in the comparison of molecular networks.

A specific example of network comparison is the comparison of network descriptions of chemical structures, which we will summarize in

Section 3.1. Table 4 summarizes a few major methods and related web-sites to compare molecular networks. Quite a few methods compare small subnetworks to larger ones. Sometimes the "small subnetwork" is small, containing only 3 to 5 nodes. This reduces the problem of finding a motif in a larger network (also called as network querying). Recent methods 1.) include an expansion process, which explores the network structure beyond the direct neighborhood; 2.) compress the network to meta-nodes, then align this representative network and finally refine the alignment; 3.) use k-hop network coloring to speed up the comparison of the traditional coloring techniques of neighboring nodes, or 4.) extend the comparison using multiple types of networks and functional information (Table 4; Ay et al., 2011; Ay et al., 2012; Berlingerio et al., 2012; Gulsoy et al., 2012). Despite the extensive progress in the field, additional work is needed to develop efficient comparison methods for large molecular networks and multiple network datasets. A widely used area of network comparison is the assessment of two time points, or a time series of a changing network, which will be discussed in the next section.

#### 2.5. Network dynamics

In this section, which concludes the inventory of network analytical concepts and methods, we will summarize the approaches describing network dynamics. First we will list the methods describing the temporal changes of networks, then we describe the usefulness of network perturbation analysis in drug design, and finally we will draw attention to the potential use of spatial games to assess the influence of nodes on network cooperation. Description of network dynamics is a fast developing field of network science holding great promise to renew systems-based thinking in drug design.

##### 2.5.1. Network time series, network evolution

As we mentioned in Section 2.1 summarizing the key points of network definition, the time-window of observation is crucial for the detection of contacts between network nodes. The duration of observation becomes even more important, when describing the temporal changes of networks, which is also often called network evolution. (It is important to note that the concept of network evolution usually has no connection to the Darwinian concept of natural selection.) The order of network edge development has key consequences in directed networks which differ from network topology measures, like shortest path, or small world. As an interesting example of these changes, in the  $A \rightarrow B \rightarrow C$  connection pattern, A cannot influence C, if the  $B \rightarrow C$  contact preceded the  $A \rightarrow B$  contact. Such effects may slow down the propagation of signals by an order of magnitude (Tang et al., 2010; Pfitzner et al., 2012).

The description of temporal changes of network structures is related to the difficult concept and methodology of network comparison and similarity described in the preceding section. Following the early summary of Dorogovtsev and Mendes (2002) on network evolution, Holme and Saramäki (2011) had an excellent review on network time-series re-defining a number of static network parameters, such as connectivity, diameter, centrality, motifs and modules, to accommodate temporal changes. The prediction algorithms described in Section 2.2 can be used to predict edges that may appear at later time points in evolving networks (Lü & Zhou, 2011). Prediction may work backwards, to infer past structures of a current network identifying core-nodes around which the network was organized (Navlakha & Kingsford, 2011). Recently, a method to test the reversibility of changes in network time-series was published (Donges et al., 2012). However, most network time description studies have concentrated on neuronal or social networks offering many, albeit yet untested, possibilities for drug design.

Recently a number of centrality measures were introduced describing central nodes of dynamically changing networks (Joseph & Chen, 2012; Mantzaris et al., in press). The development of network

**Table 4**  
Comparison methods of molecular networks.

Name <sup>a</sup>	Network type(s) <sup>b</sup>	Description and website	References
AlignNemo	Protein–protein interaction networks	Uncovers subnetworks of proteins and uses an expansion process, which gradually explores the network beyond the direct neighborhood. <a href="http://sourceforge.net/p/alignnemo">http://sourceforge.net/p/alignnemo</a>	Ciriello et al., 2012a
Differential dependency network analysis	Transcriptional networks	A set of conditional probabilities is proposed as a local dependency model, and a learning algorithm is developed to show the statistical significance of the local structures. <a href="http://www.cbil.ece.vt.edu/software.htm">http://www.cbil.ece.vt.edu/software.htm</a>	Zhang et al., 2009
Graphcrunch2, C-GRAAL, MI-GRAAL	Multiple networks	Compares networks with random networks. Additionally, clusters nodes based on their topological similarities in the compared networks. <a href="http://bio-nets.doc.ic.ac.uk/graphcrunch2">http://bio-nets.doc.ic.ac.uk/graphcrunch2</a> ; <a href="http://bio-nets.doc.ic.ac.uk/MI-GRAAL">http://bio-nets.doc.ic.ac.uk/MI-GRAAL</a>	Kuchaiev and Przulj, 2011; Kuchaiev et al., 2011; Memisevic & Przulj, 2012
IsoRankN (IsoRank Nibble)	Metabolic networks	Uses spectral clustering on the induced graph of pair-wise alignment scores. <a href="http://isorank.csail.mit.edu">http://isorank.csail.mit.edu</a>	Liao et al., 2009
MetaPathway-Hunter	Metabolic networks	Finds tree-like pathways in metabolic networks. <a href="http://www.cs.technion.ac.il/~olegro/metapathwayhunter">http://www.cs.technion.ac.il/~olegro/metapathwayhunter</a>	Pinter et al., 2005
MNAligner	Molecular networks	Combines molecular and topological similarity using integer quadratic programming, enabling the comparison of weighted and directed networks and finding cycles beyond tree-like structures. <a href="http://doc.aporc.org/wiki/MNAligner">http://doc.aporc.org/wiki/MNAligner</a>	Li et al., 2007
Module Preservation	Measures module preservation in different datasets	Uses several module comparison statistics based on the adjacency matrix, or on the basis of pair-wise correlations between numeric variables. <a href="http://www.genetics.ucla.edu/labs/horvath/CoexpressionNetwork/ModulePreservation">http://www.genetics.ucla.edu/labs/horvath/CoexpressionNetwork/ModulePreservation</a>	Langfelder et al., 2011
NeMo	Gene co-expression networks	Detects frequent co-expression modules among gene co-expression networks across various conditions. <a href="http://zhoulab.usc.edu/NeMo">http://zhoulab.usc.edu/NeMo</a>	Yan et al., 2007b
NetAlign	Protein–protein interaction networks	Aligns conserved network substructures. <a href="http://netalign.ustc.edu.cn/NetAlign">http://netalign.ustc.edu.cn/NetAlign</a>	Liang et al., 2006
NetAligner	Protein–protein interaction networks + pathways	Compares whole interactomes, pathways and protein complexes of 7 organisms. <a href="http://netaligner.irbbarcelona.org">http://netaligner.irbbarcelona.org</a>	Pache et al., 2012
NetMatch	Cytoscape plug-in for molecular networks	Finds subgraphs of the original network connected in the same way as the querying network. Can also handle multiple edges, multiple attributes per node and missing nodes. <a href="http://baderlab.org/Software/NetMatch">http://baderlab.org/Software/NetMatch</a>	Ferro et al., 2007
PathBLAST	Search of smaller linear pathways	Finds smaller linear pathways in protein–protein interaction networks. <a href="http://www.pathblast.org">http://www.pathblast.org</a>	Kelley et al., 2004
PINALOG	Protein–protein interaction network	Combines information from protein sequence, function and network topology. <a href="http://www.sbg.bio.ic.ac.uk/~pinalog">http://www.sbg.bio.ic.ac.uk/~pinalog</a>	Phan & Sternberg, 2012
Rahnuma	metabolic networks	Represents metabolic networks as hypergraphs and computes all possible pathways between two or more metabolites. <a href="http://portal.stats.ox.ac.uk:8080/rahnuma">http://portal.stats.ox.ac.uk:8080/rahnuma</a>	Mithani et al., 2009

<sup>a</sup> The summaries of Sharan and Ideker (2006) and Zhang et al. (2008) describe and compare some of the methods above.

<sup>b</sup> The network type is indicating the primary network, where the method has been tested. However, most methods are applicable to other types of molecular networks.

modules gained considerable attention in network evolution studies, since this representation concentrates on the functionally most relevant changes in the network structure. Network modules may grow, contract, merge, split, be born or die. Some of the modules display a much larger stability than others. The intra-modular nodes of these modules bind to each other with high affinity and to nodes outside the module with low affinity. Interestingly, small modules (of say less than 10 nodes) seem to persist better, if they have a dense contact structure, while larger modules survive better, if they have a dynamic, fluctuating membership (Palla et al., 2007; Fortunato, 2010). Mucha et al. (2010) developed the technique of multislice networks, which monitor the module development of nodes with multiple types of edges. Taylor et al. (2009) showed that altered modularity of hubs had a prognostic value in breast cancer and suggested cancer-specific inter-modular hubs as drug targets in cancer therapies.

Detailed analyses identified change points, i.e. short periods where large changes of modular structure can be observed (Falkowski et al., 2006; Sun et al., 2007; Rosvall & Bergstrom, 2010). The alluvial diagram (applying the visualization technique of Sankey diagrams) introduced by Rosvall and Bergstrom (2010; Fig. 10) illustrates the temporal changes of network modules particularly well. Recently the addition of routes between nodes of the network, called the accessibility graph, was also used successfully to describe network time series (Lentz et al., 2012).

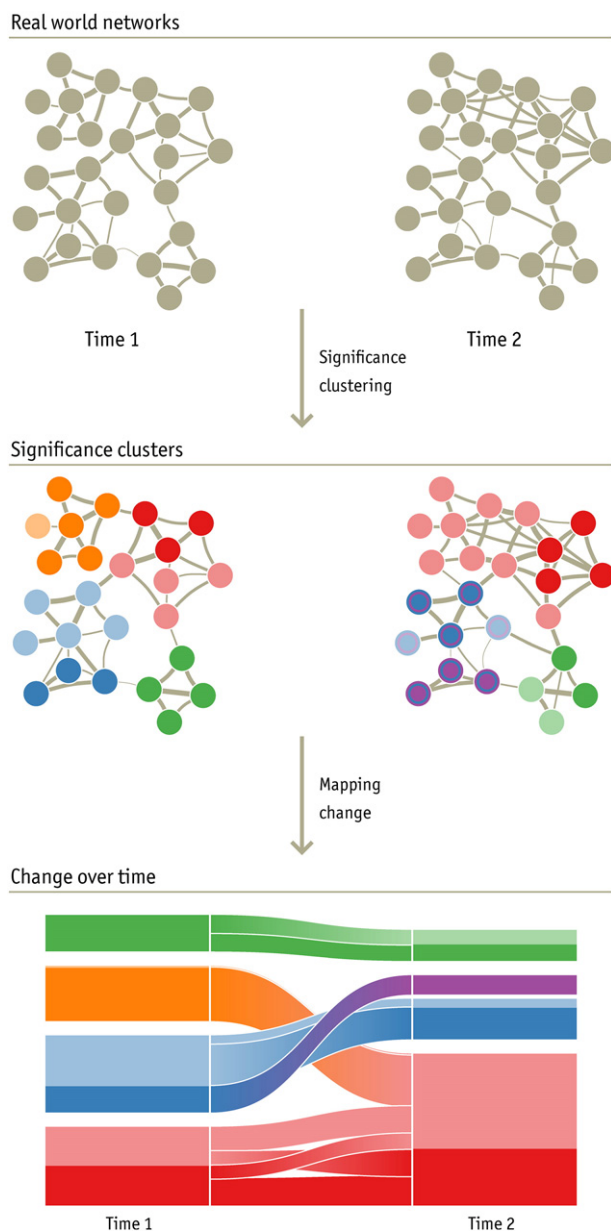
Dramatic changes of network structure called “topological phase transitions” occur when the resources needed to the maintain network contacts diminish, or environmental stress becomes much larger. Networks may develop a hierarchy, a core or a central hub as the relative costs of edge-maintenance increase. Under extreme circumstances, the network may disintegrate to small subgraphs, which corresponds to the death of the complex organism encoded by the formerly

connected network (Derényi et al., 2004; Csermely, 2009; Brede, 2010). Change points and topological phase transitions have not been assessed in disease, or in other therapeutically interesting situations showing an abrupt change, such as apoptosis, and thus provide an exciting field of future drug-related studies.

Going beyond the changes of system structure, network descriptions may also be applied to describe changes of systems-level emergent properties. In these descriptions nodes represent phenotypes of the complex system in the state-space, and edges are the transitions or similarities of these phenotypes. This approach is used in the network representations of energy landscapes (or fitness landscapes) resulting in transition networks, and in the recurrence-based time series analysis resulting in correlation networks, cycle networks, recurrence networks or visibility graphs (Doye, 2002; Rao & Cafilisch, 2004; Donner et al., 2011). Recently a method to compare two visibility graphs, i.e. two network time series was published (Mehraban et al., 2013).

### 2.5.2. Network robustness and perturbations

In the network-related scientific literature perturbations often mean the complete deletion of a network node. However, in drug action the complete inhibition of a molecule is seldom achieved. Therefore, when summarizing network perturbations, we will concentrate on the transient changes of network-encoded complex systems. Transient perturbations play a major role in signaling and in the development of diseases. The action of drugs can be perceived as a network perturbation nudging pathophysiological networks back into their normal state (Gardner et al., 2003; di Bernardo et al., 2005; Ohlson, 2008; Antal et al., 2009; Huang et al., 2009; Lum et al., 2009; Baggs et al., 2010; del Sol et al., 2010; Chua & Roth, 2011). Therefore, studies addressing perturbation dynamics have a key importance in drug design.



**Fig. 10.** Alluvial diagram illustrating the temporal changes of network communities. Each block represents a network module with a height corresponding to the module size. Modules are ordered by size (in case of a hierarchical structure within their super-modules). Darker colors indicate module cores. Modules having a non-significant difference are closer to each other. The height of the changing fields in the middle of the representation corresponds to the number of nodes participating in the change. To reduce the number of crossovers, changes are ordered by the order of connecting modules. To make the visualization more concise transients are passing through the midpoints of the entering and exiting modules and have a slim waist. Note the split of the blue module, and the merge of the orange and red modules. Reproduced with permission from Rosvall and Bergstrom (2010). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Robustness is an intrinsic property of cellular networks that enables them to maintain their functions in spite of various perturbations. Enhanced robustness is a property of only a very small number of all possible network topologies. Cellular networks both in health and in disease belong to this extreme minority. Drug action often fails due to the robustness of disease-affected cells or parasites. In contrast, side-effects often indicate that the drug hits an unexpected point of fragility of the affected networks (Kitano, 2004a, 2004b, 2007; Ciliberti et al., 2007). Robustness analysis was used to reveal primary drug targets

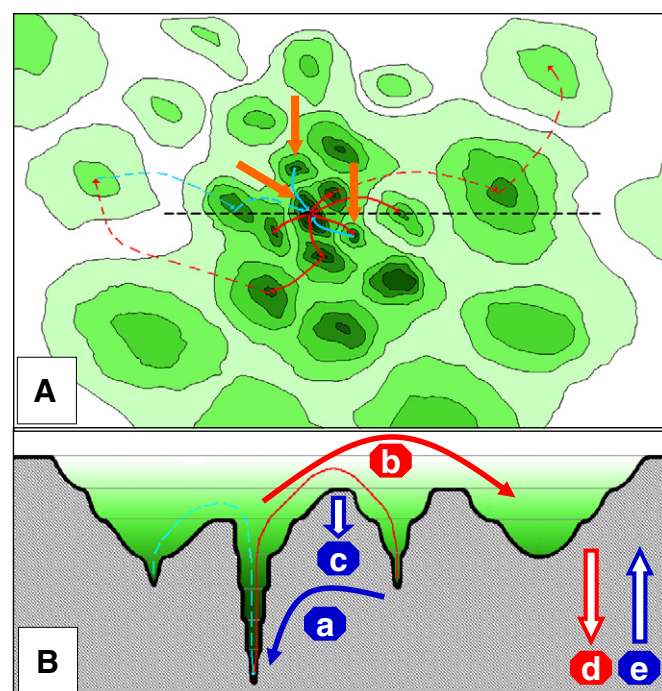
and to characterize drug action (Hallén et al., 2006; Moriya et al., 2006; Luni et al., 2010).

Cellular robustness can be caused by a number of mechanisms.

- Network edges with large weights often form negative or positive feedbacks helping the cell to return to the original state (attractor) or jump to another, respectively.
- Network edges with small weights provide alternative pathways, give flexible inter-modular connections disjoining network modules to block perturbations and buffer the changes by additional, yet unknown mechanisms. These 'weak links' grossly outnumber the 'strong links' participating in feedback mechanisms. Therefore, the two mechanisms have comparable effects at the systems level.
- Finally, robustness of molecular networks also depends on the robustness of their nodes, e.g. the stability of protein structures (Csermely, 2004, 2009; Kitano, 2004a, 2004b, 2007).

We summarize the possible mechanisms through which drugs can overcome cellular robustness in Fig. 11 (letters in the list correspond to symbols of the figure).

- Drugs may activate a regulatory feedback helping disease-affected cells to return to the original equilibrium.
- Drugs may activate a positive feedback and push disease-affected cells to a new state.



**Fig. 11.** Mechanisms of drug action changing cellular robustness. Panel A shows a 2-dimensional contour plot of the stability landscape of healthy and diseased phenotypes. Healthy states are represented by the central and the adjacent two minima marked with the large orange arrows, while all additional local minima are diseased states. Darker green colors refer to states with larger stability. Thin blue and red arrows mark shifts to healthy and diseased states, respectively. Dashed arrows refer to less probable changes. Panel B illustrates mechanisms of drug action on cellular robustness. The valleys and hills are a vertical representation of the stability-landscape shown on Panel A along the horizontal dashed black line. Blue symbols represent drug interactions with disease-prone or disease-affected cells, while red symbols refer to drug effects on cancer cells or parasites. (a) Counteracting regulatory feedback; (b) positive feedback pushing the diseased cell or parasite to another trajectory; (c) a transient decrease of a specific activation energy enabling a shift back to healthy state; (d) 'error-catastrophe': drug action diminishing many activation energies at the same time, causing cellular instability, which leads to cell death; (e) general increase in activation energies leading to the stabilization of healthy cells to prevent their shift to diseased phenotype. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

- c. Drugs may transiently lower a specific activation energy helping disease-affected cells to return to the healthy state.
- d. Drugs may decrease activation energies and thus destabilize malignant or infectious cells causing an ‘error catastrophe’ and activating cell death.
- e. Drugs may increase activation energies and thus stabilize healthy cells preventing their shift to the diseased phenotype (Csermely, 2004, 2009; Kitano, 2004a, 2004b, 2007).

If cellular robustness is conquered, critical transitions, i.e. large unexpected changes, may also occur. Critical transitions are often responsible for unexplained cases of excessive drug side-effects and toxicity. Lack of stabilizing negative feedbacks, excessive positive feedbacks, accumulating cascades may all lead to extreme events characterizing critical transitions (San Miguel et al., 2012).

Recently increasing attention focused on finding, predicting and influencing extreme events, i.e. outliers of common (e.g. scale-free) statistics, also called ‘dragon kings’ (Tajer & Poor, 2012; de S. Cavalcante et al., 2013). The detection of early warning signals of these critical transitions (such as a slower recovery after perturbations, increased self-similarity of the behavior, or increased occurrence of extreme behavior) was shown to characterize different complex systems, such as ecosystems, the market, climate change, or population of yeast cells (Scheffer et al., 2009; Farkas et al., 2011; Sornette & Osorio, 2011; Dai et al., 2012). Boettiger and Hastings (2012) emphasized the importance of using correct statistics in order not to ‘over-examine’ systems which are known to undergo critical transitions. Prediction and control of critical changes (delay/prevention in the case of normal cells and induction/acceleration in the case of malignant or infecting cells) may be an especially important area of future drug-related network studies.

The number of possible regulatory combinations for a given gene increases dramatically with an increase in input-complexity and network size. For example, with 100 genes and 3 inputs per gene there are a million input combinations for each gene in the network resulting in 10,600 different network wiring diagrams (Tegnér & Bjorkegren, 2007). The complexity of precise network perturbation models increases even more with system size. Therefore, it is not surprising that most studies of network dynamics described small networks with at most a few dozens of nodes. As an example of this, the Tide software analyzes the combined effects and optimal positions of drug-like inhibitors or activators using differential equations of reaction pathways up to 8 components (Schulz et al., 2009). Karlebach and Shamir (2010) presented an algorithm determining the smallest perturbations required for manipulating a network of 14 genes. Perturbations of Boolean networks, where nodes may only have an ‘on’ or ‘off’ mode, describe the dynamics of 20 to 50 nodes. These models often incorporate activating, inhibiting, or conditional edges, too (Huang, 2001; Shmulevich et al., 2002; Gong & Zhang, 2007; Abdi et al., 2008; Azuaje et al., 2010; Saadatpour et al., 2011; Wang & Albert, 2011; Garg et al., 2012). To help these studies a versatile, publicly available software library, BooleanNet (<http://booleannet.googlecode.com>) was developed by Albert et al. (2008). PATHLOGIC-S (<http://sourceforge.net/projects/pathlogic/files/PATHLOGIC-S>) offers a scalable Boolean framework for modeling cellular signaling (Fearnley & Nielsen, 2012).

Systems-level molecular networks have a size in the range of thousand to ten-thousand nodes. At this level of system complexity the optimal selection of the perturbation model becomes a key issue. At this system size the highly anisotropic perturbation propagation inside protein structures is usually neglected (we will detail the possibilities to construct atomic resolution interactomes in Section 4.1.6 on allo-network drugs; Nussinov et al., 2011). In current network perturbation models of larger systems delays, differences in individual dissipation patterns, effects of water or molecular crowding are also neglected (Antal et al., 2009).

We summarized an early promising approach of systems-level perturbation studies in Section 2.2.3 on reverse engineering. Here

perturbations were assessed by systems-level mRNA expression profiles and the perturbed network was reconstructed from the output data (Liang et al., 1998a; Akutsu et al., 1999; Ideker et al., 2000; Kholodenko et al., 2002; Yeung et al., 2002; Segal et al., 2003; Tegnér et al., 2003; Friedman, 2004; Tegnér & Bjorkegren, 2007; Ahmed & Xing, 2009; Stokic et al., 2009; Marbach et al., 2010; Yip et al., 2010; Schaffter et al., 2011; Altay, 2012; Crombach et al., 2012; Kotera et al., 2012). Reverse engineering techniques were successfully applied to reconstruct drug-induced system perturbations (Gardner et al., 2003; di Bernardo et al., 2005; Chua & Roth, 2011).

Maslov and Ispolatov (2007) used the mass action law to calculate the effect of a two-fold increase in the expression of single protein on the free concentration of other proteins in the yeast interactome. Despite an exponential decay of changes, there were a few highly selective pathways, where concentration changes propagated to a larger distance (Maslov & Ispolatov, 2007). This and other models of network dynamics have been used in various publicly available algorithms including:

- the system dynamics modeling tool BIOCHAM using Boolean, differential, stochastic models and providing among others bifurcation diagrams (<http://contraintes.inria.fr/biocham>; Calzone et al., 2006);
- the random walk-based ITM-Probe, also available as a Cytoscape plug-in ([http://www.ncbi.nlm.nih.gov/CBBresearch/Yu/mn/itm\\_probe/doc/cytoitmprobe.html](http://www.ncbi.nlm.nih.gov/CBBresearch/Yu/mn/itm_probe/doc/cytoitmprobe.html); Stojmirović & Yu, 2009; Smoot et al., 2011);
- the mass action-based Cytoscape plug-in, PerturbationAnalyzer ([http://chianti.ucsd.edu/cyto\\_web/plugins/displayplugininfo.php?name=PerturbationAnalyzer](http://chianti.ucsd.edu/cyto_web/plugins/displayplugininfo.php?name=PerturbationAnalyzer); Li et al., 2010a; Smoot et al., 2011);
- a user-friendly, Matlab-compatible, versatile network dynamics tool, Turbine supplying a communication vessels propagation model, but handling any user-defined dynamics, and enabling the user to simulate real world networks that include 1 million nodes and 10 million edges per GByte of free system memory, exporting and converting numerical data to a visual image using an inbuilt viewer function ([www.linkgroup.hu/Turbine.php](http://www.linkgroup.hu/Turbine.php); Farkas et al., 2011);
- Conedy, a Python-interfaced C++ program capable to handle various dynamics including differential equations and oscillators (<http://www.conedy.org>; Rothkegel & Lehnertz, 2012).

Studying perturbations of larger networks Adilson Motter and colleagues developed an exciting model of compensatory perturbations showing that surprisingly, a debilitating effect can often be compensated by another inhibitory effect in a complex, cellular system (Motter et al., 2008; Motter, 2010; Cornelius et al., 2011). Perturbation dynamics of signaling networks was extensively analyzed including close to 10 thousand phosphorylation events in an experimental study of yeast cells (Bodenmiller et al., 2010). As we described in Section 2.2.3 on reverse engineering, perturbation studies are often used to reconstruct networks. As examples of this, the signaling network of T lymphocytes was reconstructed using single cell perturbations (Sachs et al., 2005), and the perturbations of 21 drug pairs were predicted from the reconstituted network of phospho-proteins and cell cycle markers of a human breast cancer cell line (Nelander et al., 2008). As another example, a perturbation amplitude scoring method was developed to test the biological impact of drug treatments, and was assessed using the transcriptome of colon cancer cells treated with the CDK cell cycle inhibitor, R547 (Martin et al., 2012).

Despite their complexity and robustness, cellular networks have their ‘Achilles-heel’. Hitting it, a perturbation may cause dramatic changes in cell behavior. Stem cell reprogramming is a well-studied example of these network-reconfigurations (Huang et al., 2012a), where special bottleneck proteins may play a pivotal role (Buganim et al., 2012). As another example of ‘streamlined’ cellular responses, effects of multiple drug-combinations on protein levels can be quite

accurately described by the linear superposition of drug-pair effects (Geva-Zatorsky et al., 2010).

Recent perturbation studies identified key nodes governing network dynamics. Central nodes, such as hubs, or inter-modular overlaps and bridges were shown to serve as highly efficient mediators of perturbations (Cornelius et al., 2011; Farkas et al., 2011). Network oscillations can be governed by a few central nodes forming a small network skeleton (Liao et al., 2011). Targets of viral proteins were shown to be major perturbators of human networks (de Chasseay et al., 2008; Navratil et al., 2011). Perturbation mediators are often at cross-roads of cellular pathways. These key nodes bind multiple partners at shared binding sites. These shared binding sites can be identified as hot spot residues in protein structures (Ozbabacan et al., 2010). The fast-developing field of viral marketing identified influential spreaders of information at network cores and at other central network positions (Kitsak et al., 2010; Valente, 2012). Spreader proteins may be excellent targets of anti-infectious or anti-cancer therapies. Conversely, drugs against other diseases need to avoid these central proteins affecting a number of cellular functions. The identification of influential spreaders may provide important analogies of future drug target studies.

### 2.5.3. Network cooperation, spatial games

Spatial games, i.e. social dilemma games (such as the well known Prisoners' Dilemma, hawk-dove or ultimatum games) played between neighboring network nodes, provide a useful model of cooperation (Nowak, 2006). In a recent review Foster (2011) described the 'sociobiology of molecular systems' and provided convincing evidence how molecular networks determine social cooperation. We argue that cooperation of proteins and other macromolecules may offer an important description of cellular complexity. This view is based on the delicate dynamics of protein-protein interactions, which proceed via mutual selection of the binding-compatible conformations of the two protein partners. As the two proteins approach each other, they signal their status to the other via the hydrogen-bonded network of water molecules. Binding is achieved by a complex set of consecutive conformational adjustments. These concerted, conditional steps were called a 'protein dance', and can be perceived as rounds of a repeated game (Kovács et al., 2005; Csermely et al., 2010).

The stepwise encounter of protein molecules can be modeled as a series of rounds in common social dilemma games. In hawk-dove games the more rigid binding partner (corresponding to the drug) can be modeled as a hawk, while the more flexible binding partner (corresponding to the drug target) will be the dove. The hawk/dove encounter corresponds to an induced-fit-like scenario, where the conformational change of the dove is much larger than that of the hawk. The game is won by drug (hawk), since its enthalpy gain is not accompanied by an entropy cost. On the contrary, the flexible drug target loses several degrees of freedom during binding (Kovács et al., 2005; Chettaoui et al., 2007; Schuster et al., 2008; Antal et al., 2009; Csermely et al., 2010). In agreement with our previous description (Csermely et al., 2010) we note that induced-fit is conceived here as an extremity of an extended conformational selection model, where one of the partners (in this case the drug) is much more rigid than the other.

If we model drug binding with the ultimatum game, the drug and its target want to share the free energy decrease as a common resource. The drug proposes how to divide the sum between the two partners, and the target can either accept or reject this proposal, i.e. bind the drug or not (Kovács et al., 2005; Chettaoui et al., 2007; Schuster et al., 2008; Antal et al., 2009; Csermely et al., 2010).

Extending the above drug-binding scenario to the network level of the whole cell spatial game models are not only important to provide an estimate of systems-level cooperation, but are able to predict, which protein can most efficiently destroy the existing cooperation of the cell. This is a very helpful model of drug action in anti-infectious or anti-cancer therapies. Game models also identify those proteins, which are the most efficient to maintain cellular cooperation. This

provides a useful model of drug efficiency in maintaining normal functions of diseased cells. Recently a versatile program, called NetworGame ([www.linkgroup.hu/NetworGame.php](http://www.linkgroup.hu/NetworGame.php)) was made publicly available for simulating spatial games using any user-defined molecular networks and identifying the most influential nodes to establish, maintain or break cellular cooperation. Nodes having an exceptional influence in these cellular games may be promising targets of future drug development efforts (Farkas et al., 2011).

### 2.6. Limitations of network-related description and analysis methods

After completing a large inventory of network-related description and analysis methods here we list some of the major limitations of this approach.

- The first and foremost limitation of all network-related methods is data quality. Network description is a tool, which depends on the accuracy and coverage of input data. The entire dataset must have the same, well defined quality control. Currently we often lack these high-quality datasets especially on system dynamics (Henney & Superti-Furga, 2008; Prinz et al., 2011; Begley & Ellis, 2012; Landis et al., 2012).
- The definition of network nodes and edges often restricts network descriptions to well-defined connections. Hypergraph descriptions, which may overcome this problem are not wide-spread, and well documented yet. Despite the flexibility offered by hypergraphs, network descriptions often reduce the dimensions of the available information. When this is performed without a deep knowledge of the problem and the system, it may result in significant information loss. We described these difficulties of network construction in Sections 1.2 and 2.1.
- Visualization of networks improved over the years (see Table 1), but there is still room for development of 3D, large-capacity, zoom-in-type network visualization tools.
- Since the information spread is not only local in the cell, network analysis concentrating only local topological signatures, such as hubs or motifs may not use the full potency of network description. Definition of network modules or central nodes resulted in a large number of methods, but has reached a consensus neither in the applicability nor in overcoming the limitations of these methods. Additional studies on the identification of influential network nodes and edges are needed.
- Network-based prediction methods also need considerable improvement.
- Comparison of dynamically changing networks needs additional methods. Generally, the dynamics of complex systems is not adequately described by networks yet.

Applications of network description and analysis to molecular problems started less than 15 years ago. Therefore, several limitations arise from the current state-of-the-art of network related studies. Others may indeed pose limits for the use of network analysis. We need a number of well-based comparative studies in the future to understand the areas, where the use of network description and analysis gives the most efficient help in drug design. The following sections of our review try to help this clarification process.

### 3. The use of molecular networks in drug design

In this section we will describe molecular networks starting from networks of chemical substances, followed by protein structure networks (i.e. networks of amino acids forming 3D protein structures), protein-protein interaction networks, signaling networks, genetic interaction and chromatin networks (i.e. networks of chromatin segments forming the 3D structure of chromatin). We will conclude the section with the description of metabolic networks, i.e. networks of metabolites connected by enzyme reactions. The section will not

give a detailed description of all studies on these networks, but will concentrate only on the most important aspects related to drug development.

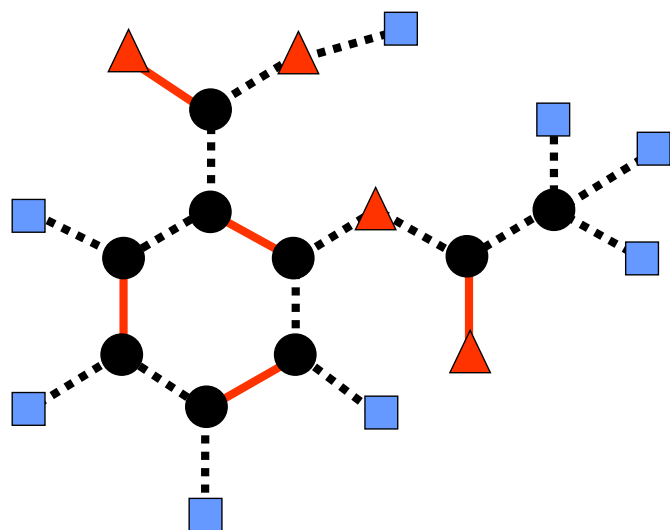
Nodes of the networks above are connected either physically or conceptually. Chemical compound networks are often constructed by connecting two chemical compounds, if there is a chemical reaction to transform one of them to the other. This logic is very similar to that used in the construction of metabolic networks. In another form of chemical compound networks two drugs are considered similar, if they have a common binding protein. This is actually the inverse of drug target networks (where two drug targets are connected, if the same drug binds to them). Substrates and products also have a common binding protein, the enzyme, serving as the edge of metabolic networks. However, drug-related studies on metabolic networks often incorporate knowledge of protein–protein interaction and signaling networks. Therefore, we will summarize metabolic networks separately, at the end of the current section. Similarly, drug target networks often use the rich conceptual context of the drug development process. Therefore, we will re-assess the major features of drug target networks in Section 4.1.3. However, due to the unavoidable overlaps we encourage the Reader to compare the sections on chemical compound, metabolic and drug target networks.

### 3.1. Chemical compound networks

In this section we will summarize all networks which are related to chemical compounds: structural networks, reaction networks, and the large variety of chemical similarity networks. Of all these networks, especially the latter, chemical similarity networks can be used very well in lead optimization and selection of drug candidates. There is a very large variability in the names of these networks in the literature. Therefore, we selected the most discriminative name as the titles of sub-sections, and refer to some other network denotations in the text.

#### 3.1.1. Chemical structure networks

The structure of chemical compounds can be perceived as a network, where labeled (colored) nodes are the atoms constructing the molecule, and labeled (colored) edges are the covalent bonds binding the atoms together (Fig. 12). Chemical structure networks (also called



**Fig. 12.** Example of a chemical structure network. From the network point of view chemical structures are networks with differently labeled (colored) nodes representing different kinds of atoms and differently labeled (colored) edges related to different types of bonds. The chemical structure network representation of aspirin is shown. Black circles, red triangles and blue rectangles represent carbon, oxygen and hydrogen atoms, respectively. Dotted black edges stand for single bonds, while red solid edges represent double bonds. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

chemical graphs) may use multiple edges representing multiple bonds. The core electron structure of the various atoms is often represented as a complete graph. Descriptors of this network structure, such as discrete invariants representing the chemical structure, connectivity indices, topological charge indices, electro-topological indices, shape indices and others are useful for quantitative structure/property and structure/activity (QSPR and QSAR) models (Bonchev & Buck, 2007; García-Domenech et al., 2008; Gonzalez-Diaz et al., 2010a). Molgen (<http://molgen.de>; Baricic & Mackov, 1995) and Modeslab (<http://modeslab.com>; Estrada & Uriarte, 2001) are widely used programs to draw and analyze chemical structure networks. SIMCOMP (<http://www.genome.jp/tools/simcomp>) and SUBCOMP (<http://www.genome.jp/tools/subcomp>) compare chemical structure networks and show the position of results in molecular pathways (Hattori et al., 2010).

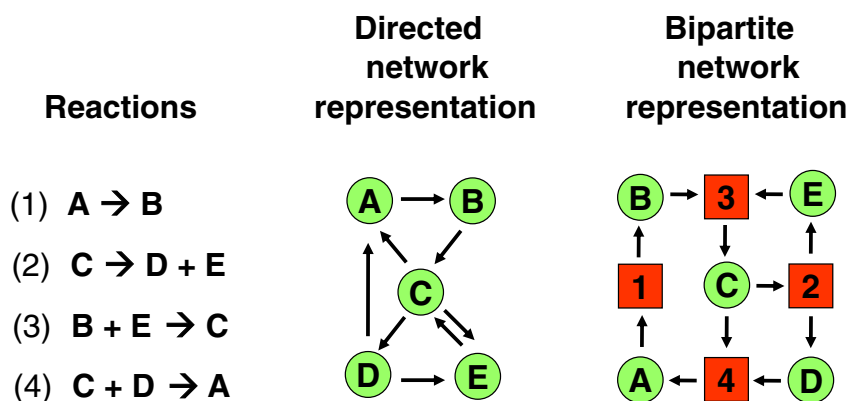
#### 3.1.2. Chemical reaction networks

The mind-boggling set of 1060 chemical compounds that can be created by chemical reactions, defines the so-called chemical space (Kirkpatrick & Ellis, 2004). The size of drug-like chemical space is estimated to be larger than a million compounds (Drew et al., 2012). The increasing costs of experiments and the need for compounds with specific properties increased the efforts to apply new tools for chemical space discovery (Lipinski & Hopkins, 2004). Chemical reactions make the chemical space continuous. Therefore, their network representation serves as a promising tool. Nodes of chemical reaction networks are the chemical compounds and their edges are the reactions transforming them to one another (Christiansen, 1953; Temkin & Bonchev, 1992).

The chemical reaction network, comprising the whole synthetic knowledge of organic chemistry containing 7 million compounds in 2012, was first assembled by Fialkowski et al. (2005). Chemical reaction networks may only contain the participating compounds, or may be bipartite networks, where besides the participating compounds a different type of nodes represents the reactions (Fig. 13). The chemical reaction network is a small-world containing hubs, i.e. compounds, which can be formed and transformed to and from many other compounds. The chemical reaction network contains hubs. Importantly, hub compounds have a lower market price than chemicals involved in a low number of reactions. Moreover, hub molecules are more likely to be prepared via new methodologies, and may also be involved in the synthesis of many new compounds (Grzybowski et al., 2009). The chemical reaction network can be separated into a core, containing over 70% of the top 200 industrial chemicals, and a periphery, which has a tree-like structure, and can be easily synthesized from the core (Bishop et al., 2006). Chemical reaction networks offer help in the design of ‘one-pot’ reactions without the need for isolation, purification and characterization of intermediate structures, and without the production of much chemical waste. Gothard et al. (2012) used 8 filters of 86,000 chemical criteria to identify more than 1 million ‘one-pot’ reaction series. The number of possible synthetic pathways can be astronomical having 1019 routes of just 5 synthetic steps. Network analysis of Kowalik et al. (2012) identified optimal synthetic pathways of single and multiple-target syntheses using a simulated annealing-based network optimization. These optimizations help in the synthesis of drug candidate variants for lead selection.

#### 3.1.3. Similarity networks of chemical compounds: QSAR, chemoinformatics, chemical genomics

Molecular similarity can be viewed as the distance between molecules in a continuous high-dimensional space of numerical descriptors (Johnson & Maggiora, 1990; Bender & Glen, 2004; Eckert & Bajorath, 2007). This high dimensional similarity space is called the chemistry space, which constitutes an important part of chemoinformatics (Faulon & Bender, 2010; Krein & Sukumar, 2011; Varnek & Baskin, 2011). Nodes in similarity networks are most often chemical compounds,



**Fig. 13.** Directed and bipartite network representations of chemical reaction networks. On the middle panel the 4 reactions of the left side of the figure are represented as a directed network of participating compounds. On the bipartite network of the right panel green circles represent chemicals, while red rectangles stand for reactions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.) Figures were adapted from Fialkowski et al. (2005) and from Grzybowski et al. (2009).

but may also be molecular fragments, or molecular scaffolds (Hu & Bajorath, 2010; Hu & Bajorath, 2011). Edge definition is a difficult task in similarity networks. Un-weighted networks can be constructed using a pre-determined similarity threshold, while the extent of similarity may also be used as edge-weight. From the large number of numerical descriptions of similarity listed in Table 5, we will first consider those networks, which are based on simple chemical similarity of the compounds

involved using e.g. the Tanimoto-coefficient for the definition of edges (Rogers & Tanimoto, 1960; Tanaka et al., 2009; Bickerton et al., 2012). We will call these networks chemical similarity networks.

Chemical similarity networks are also small-worlds possessing hubs with a modular structure. Similarity hubs may be used as priority starting points in fragment-based drug design. If hubs become non-hits, many fragment-combinations can be excluded as candidates, under the

**Table 5**  
Chemical compound similarity networks.

Basis of chemical compound similarity	References
<i>Chemical compound similarity networks</i>	
Chemical similarity based on e.g. the Tanimoto-coefficient	Tanaka et al., 2009; Bickerton et al., 2012
QSAR-related similarity networks (a freely available program to mine structure–activity and structure–selectivity relationship information in compound datasets, SARANEA: <a href="http://www.limes.uni-bonn.de/forschung/abteilungen/Bajorath/labwebsite/downloads/saranea/view">http://www.limes.uni-bonn.de/forschung/abteilungen/Bajorath/labwebsite/downloads/saranea/view</a> )	Estrada et al., 2006; Gonzalez-Diaz & Prado-Prado, 2008; Hert et al., 2008; Prado-Prado et al., 2008; Wawer et al., 2008; Bajorath et al., 2009; Garcia et al., 2009; Prado-Prado et al., 2009; Gonzalez-Diaz et al., 2010a; Lounkine et al., 2010; Peltason et al., 2010; Prado-Prado et al., 2010; Wawer et al., 2010; Iyer et al., 2011a; Iyer et al., 2011b; Iyer et al., 2011c; Krein & Sukumar, 2011; Wawer & Bajorath, 2011a; Wawer & Bajorath, 2011b
BioAssay network: bioassay data of chemical compounds from PubChem	Zhang et al., 2011b
Similarity of protein binding sites	Paolini et al., 2006; Keiser et al., 2007; Hert et al., 2008; Park & Kim, 2008; Adams et al., 2009; Keiser et al., 2009; Hu et al., 2011
Network of drug–receptor pairs with multitarget QSAR	Vina et al., 2009
CARLSBAD: a Cytoscape plug-in for connecting common chemical patterns to biological targets via small molecules <a href="http://carlsbad.health.unm.edu">http://carlsbad.health.unm.edu</a>	
Drug–target network combined with the chemical structure network of the drug and the protein structure network of its target giving quality-scores of drug–target networks	Riera-Fernandez et al., 2012
Similarity of mRNA expression profiles extended with disease mRNA expression profiles: Connectivity Map <a href="http://www.broadinstitute.org/cmap">http://www.broadinstitute.org/cmap</a>	Lamb et al., 2006; Iorio et al., 2009; Huang et al., 2010a
Side-effect similarity of drugs	Campillos et al., 2008
Protein–protein interaction network topology of the target neighborhood (a database of more than 700,000 chemicals, 30,000 proteins and their over 2 million interactions integrated to a human protein–protein interaction network having over 400,000 interactions, ChemProt: <a href="http://www.cbs.dtu.dk/services/ChemProt">http://www.cbs.dtu.dk/services/ChemProt</a> )	Hansen et al., 2009; Li et al., 2009a; Taboreau et al., 2011; Edberg et al., 2012
<i>Integrated bio-entity relationship datasets and networks</i>	
• Structural similarity, QSAR, gene–disease interactions, biological processes, drug absorption, distribution, metabolism and excretion (ADME) data and toxicity mechanisms	Brennan et al., 2009
• Integrated semantic network of chemogenomic repositories, Chem2Bio2RDF <a href="http://cheminfv.informatics.indiana.edu:8080">http://cheminfv.informatics.indiana.edu:8080</a>	Chen et al., 2010a
• Drug therapeutic and chemical similarity with protein–protein interaction network data: drugCIPHER	Zhao & Li, 2010
• Protein–protein interactions, protein/gene regulations, protein–small molecule interactions, protein–Gene Ontology relationships, protein–pathway relationships and pathway–disease relationships: bio-entity network (IBN)	Bell et al., 2011
• Phenotype/single-nucleotide polymorphism (SNP) associations, protein–protein interactions, disease–tissue, tissue–gene and drug–gene relationships: integrated Complex Traits Networks, iCTNet Cytoscape plug-in, <a href="http://flux.cs.queensu.ca/ictnet">http://flux.cs.queensu.ca/ictnet</a>	Wang et al., 2011b
• Protein–protein interactions, protein–small molecule interactions, associations of interactions with pathways, species, diseases and Gene Ontology terms with the user-selected integration of manually curated and/or automatically extracted data: integrated molecular interaction database, IMID, <a href="http://integrativebiology.org">http://integrativebiology.org</a>	Balaji et al., 2012

assumption that molecules similar to non-hits are also non-hits. This strategy was shown to explore the chemistry space in much less trials than random selection or the selection of cluster centers (Tanaka et al., 2009). Well connected fragments can also be used in library design and in fragment-based database searches. The top 10% of most frequently occurring molecular segments accounts for the majority of overall fragment occurrences, thus, storing a relatively small number of fragments can cover a large portion of the searching space (Benz et al., 2008). Chemical similarity networks were shown to be a very useful description of the diversity and drug-likeness of bioactive compounds against various drug targets (Bickerton et al., 2012).

Molecular similarity is particularly important in medicinal chemistry. This is due to the 'similar property principle' which states that similar molecules have similar biological activity (Johnson & Maggiora, 1990). This principle also serves as a basis of most quantitative structure–activity relationship (QSAR) modeling methods (note that we will use the term, QSAR to describe structure activity relationships in general). However, the relationship between chemical similarity and biological activity is not always straightforward (Martin et al., 2002), which necessitates the use of sophisticated approaches in drug design, such as the multi-component similarity networks listed in Table 5.

In QSAR-related similarity networks (also called network-like similarity graphs) nodes are often color-coded according to their biological action potency value ( $\text{pIC}_{50}$  or  $\text{pK}_i$ ), and scaled in size based on their contribution to the QSAR landscape features such as 'activity cliffs' or smooth regions. Near activity cliffs, small changes in molecular structure induce large changes in biological activity, while in smooth regions of the QSAR landscape changes in chemical structure only result in small or gradual changes in activity. QSAR-related similarity networks contain more information than chemical similarity networks. In contrast, chemical similarity networks were found to be topologically robust to the methods for representing and comparing chemical information. The choice of molecular representation (molecular descriptors) may change the interpretation of QSAR landscapes, where the appropriate selection of similarity (distance) cut-offs was proven to be crucial. If the cut-off value was too low, there were many isolated nodes; if the cut-off was too high, QSAR-related similarity networks became overcrowded and less useful for predictions. QSAR-related similarity networks are small worlds and contain hubs. Subsets of compounds related by different local QSARs are often organized in small communities (also called as clusters). High centrality nodes form 'chemical bridges' between various compound communities providing important QSAR information. These nodes can be used for 'hopping' between sub-networks having different chemical characteristics. Searching for nodes with high centrality and a closer look at their properties may contribute to the discovery of new drug candidates and uncover new directions for mechanistic-, scaffold- or target-hopping approaches (Gonzalez-Diaz & Prado-Prado, 2008; Hert et al., 2008; Prado-Prado et al., 2008, 2009, 2010; Wawer et al., 2008, 2010; Bajorath et al., 2009; Gonzalez-Diaz et al., 2010a; Peltason et al., 2010; Iyer et al., 2011a, 2011b, 2011c; Krein & Sukumar, 2011; Wawer & Bajorath, 2011a, 2011b). SARANEA (<http://www.limes.uni-bonn.de/forschung/abteilungen/Bajorath/labwebsite/downloads/saranea/view>) is a freely available program to mine structure–activity and structure–selectivity relationship information in compound datasets (Lounkine et al., 2010). Recently, group-based QSAR models were introduced, providing promising approaches in multi-target drug design (Ajmani & Kulkarni, 2012). Methods for the systematic comparison of molecular descriptors, such as that introduced by Bender et al. (2009), are very useful to guide future work—including network-related applications.

Dehmer et al. (2009) showed the usefulness of network complexity analysis in the determination of topological descriptor uniqueness. We demonstrate the usefulness of QSAR-related similarity network descriptors on chirality, since the different enantiomers of drug

candidates can exhibit large differences in activity. Using complex networks, Garcia et al. (2009) investigated the drug–drug similarity relationship of more than 1600 experimentally unexplored, chiral 3-hydroxy-3-methyl-glutaryl coenzyme A inhibitor derivatives with a potential to lower serum cholesterol preventing cardiovascular disease. Inclusion of chirality in network description may guide synthesis efforts towards new chiral derivatives of potentially high activity. QSAR-related similarity networks including chiral information of G protein-coupled receptor ligands identified that opposing chiralities induced alterations in molecular mechanism (Iyer et al., 2011b).

Another important application of QSAR-related similarity networks is the molecular fragment network of human serum albumin binding defined by Estrada et al. (2006). The identification of polar 'emphatic' fragments anchoring chemicals to serum albumin and hydrophobic fragments determining albumin binding was an important step in network-related prediction of bioavailability.

Interestingly, a similar growth mechanism was found in the evolution of chemical reaction networks (Fialkowski et al., 2005; Grzybowski et al., 2009) and QSAR-related similarity networks (Iyer et al., 2011a). Growth was predominantly observed around a few hubs that emerged early in the growth process, and did not reach whole segments of the network until a very late phase of development. Analyzing evolving datasets can be very important to identify over-sampled regions containing redundant compound structure information, or yet unexplored regions in the chemical reaction network or QSAR-related similarity network.

The 'similar property principle' stating that similar molecules have similar biological activity (Johnson & Maggiora, 1990) can be reversed, and used for the construction of similarity networks, which means that compounds having a similar biological action are similar. Compounds or compound scaffolds can be connected using the similarity of their protein binding sites, as well as of the protein domains or the entire proteins harboring these sites. The emerging network defined the 'pharmacological space'. Hub ligands of this network were bridges between different ligand clusters. The network representation proved to be useful for identifying drug chemotypes, and for the probabilistic modeling of yet undiscovered biological effects of chemical compounds (Paolini et al., 2006; Keiser et al., 2007, 2009; Yildirim et al., 2007; Hert et al., 2008; Park & Kim, 2008; Yamanishi et al., 2008, 2011; Adams et al., 2009; Bleakley & Yamanishi, 2009; Hu et al., 2011; Tabei et al., 2012). Using the above datasets He et al. (2010) encoded chemical compounds with functional groups and proteins with biological features of 4 major drug target classes, and worked out a prediction of drug–target interactions using the maximum relevance minimum redundancy method. Riera-Fernandez et al. (2012) gave quality-scores of drug–target network edges using the combined information of the chemical structure network of the drug and the protein structure network of its target.

An important approach to compare the similarity of chemical compounds is to construct the network of drug–therapy interactions, where drugs are connected, if they are used in the same therapy class of the five hierarchical Anatomical Therapeutic Chemical (ATC) classification levels. Average paths in this drug–therapy network are shorter than 3 steps. Distant therapies are separated by a surprisingly low number of chemical compounds. Inter-modular, bridging and otherwise central drugs in the drug–therapy network may have more indications than currently known, thus drug–therapy network data may be useful for drug-repositioning (Nacher & Schwartz, 2008). Text mining may be an important method to enrich drug–therapy networks in the future (Ruan et al., 2004).

mRNA expression patterns were the first system-wide descriptors of drug effects enabling target clustering, target identification, and prediction of the mechanism of action of new compounds (Marton et al., 1998; Hughes et al., 2000; Lamb et al., 2006; Iorio et al., 2009; Chua & Roth, 2011). Huang et al. (2010a) connected mRNA expression profiles with a disease diagnosis database. Using a Bayesian



learning algorithm they could query drug-treatment related mRNA expression profile and decipher drug similarity not only to each other, but also to specific disease and disease classes.

As we will discuss in detail in Sections 4.1.5 and 4.3.5, drugs seldom have a single effect. Based on this, the chemical similarity of drugs may be derived from their side-effects, describing a broader repertoire of drug action than the effect related to the original target. Campillos et al. (2008) connected drugs sharing a certain degree of side-effect similarity. This network uncovered shared targets of unrelated drugs and forms an important network method for drug repositioning.

Going one level further in systems-level abstraction, similarity of compounds can be measured by comparing the topological similarity of their target neighborhoods in protein–protein interaction networks (Hansen et al., 2009; Edberg et al., 2012). Li et al. (2009a) concluded from the investigation of an Alzheimer's Disease-related dataset, that the combination of curated drug–target databases and literature mining data outperformed both datasets when used alone. Systems-level inquiries are helped by ChemProt (<http://www.cbs.dtu.dk/services/ChemProt>), a database of more than 700,000 chemicals, 30,000 proteins and their over 2 million interactions integrated to a human protein–protein interaction network having over 400,000 interactions (Taboureau et al., 2011).

Baggs et al. (2010) encouraged the inclusion of network readouts (like transcriptome, proteome, phosphoproteome, metabolome and epigenetic system-wide datasets) in QSAR methods leading to QNSAR (quantitative network structure–activity relationships). In agreement with this suggestion, in recent years an increasing number of complex databases were published, where network reconstitution was used to predict biologically meaningful clusters of datasets, novel drug-candidate molecules, new drug applications, unexpected drug–drug interactions, drug side-effects and toxicity. We list these datasets in Table 5. As noted by Vina et al. (2009), increased reliance on indirect data similarities may compromise accuracy, but may also enable the exploration of those segments of the data association landscape, where no direct alignments were available. The aggregative assessment of multiple (and system-wide) datasets helps to pick up those similarities, which are the most relevant despite the many uncertainties of the individual data or their associations.

Utilizing the rich repertoire of the assessment of network topology and dynamics, listed in Section 2, will be helpful for predicting future directions in compound optimization, or redirecting research efforts to unexplored or more fruitful regions of chemical space. Moreover, detailed analysis of complex similarity networks is useful for predicting new targets of existing drugs, i.e. multi-target drug identification and drug repositioning. Finally, assessment of similarity networks can be used as an efficient predictor of drug specificity, efficacy, ADME, resistance, side-effects, drug–drug interactions and toxicity.

### 3.2. Protein structure networks

Proteins are the major targets of drug action, and therefore the description of their structure and dynamics has a crucial importance in the determination of drug binding sites, as well as in prediction of drug effects at the sub-molecular level. In this section we will show how protein structure networks help the characterization of disease-related proteins, the understanding of drug action mechanisms and drug targeting.

#### 3.2.1. Definition and key residues of protein structure networks

In most protein structure network representations (also called amino acid networks, residue interaction networks, or protein meta-structures) nodes are the amino acid side chains. Though occasionally protein structure network nodes are defined as the atoms of the protein, the side-chain representation is justified by the concerted movement of side-chain atoms. Edges of protein structure networks are defined using

the physical distance between amino acid side-chains. Distances are usually measured between C $\alpha$  or C $\beta$  atoms, but in some representations the centers of mass of the side chains are calculated, and distances are measured between them. Edges of unweighted protein structure networks connect amino acids having a distance below a cut-off distance, which is usually between 4 and 8.5 Å (Artymiuk et al., 1990; Kannan & Vishveshwara, 1999; Greene & Higman, 2003; Bagler & Sinha, 2005; Böde et al., 2007; Krishnan et al., 2008; Vishveshwara et al., 2009; Doncheva et al., 2011, 2012b; Csermely et al., 2012; Di Paola et al., in press). A detailed study compared the effect of various C $\alpha$ –C $\alpha$  contact assessments, such as the atom distance criteria, the isotropic sphere chain and the anisotropic ellipsoid side-chain models, as well as of the selection of various cut-off distances. The study showed that the atom distance criteria model was the most accurate description having a moderate computational cost. The best amino acid pair specific cut-off distances varied between 3.9 and 6.5 Å (Sun & He, 2011). In protein structure networks with weighted edges, edge weight is usually inversely proportional to the distance between the two amino acid side-chains (Artymiuk et al., 1990; Kannan & Vishveshwara, 1999; Greene & Higman, 2003; Bagler & Sinha, 2005; Böde et al., 2007; Krishnan et al., 2008; Vishveshwara et al., 2009; Doncheva et al., 2011, 2012b; Csermely et al., 2012; Di Paola et al., in press).

Web-servers have been established to convert Protein Data Bank 3D protein structure files into protein structure networks, and to provide their network analysis. The RING server (<http://protein.bio.unipd.it/ring>) gives a set of physico-chemically validated amino acid contacts (Martin et al., 2011), and imports it to the widely used Cytoscape platform (Smoot et al., 2011) enabling their network analysis using the tool-inventory described in Section 2. Recently a specific, Cytoscape-linked (Smoot et al., 2011) tool-kit for protein structure network assessment, RINalyzer (<http://www.rinalyzer.de>) was published. The program is complemented by a protein structure determination module, called RINerator (<http://rinalyzer.de/rindata.php>), which determines the protein structure networks and stores pre-determined protein structure networks of Protein Data Bank 3D protein structure files. The RINalyzer program was also linked to the NetworkAnalyzer software (<http://med.bioinf.mpi-inf.mpg.de/netanalyzer>; Assenov et al., 2008) allowing the comparison of protein structure networks and the extension of their analysis to protein–protein interaction networks (Doncheva et al., 2011, 2012b).

Protein structure networks are “small worlds”. This is very important for the fast transmission of drug-induced conformational changes, since in the small-world of protein structure networks all amino acids can communicate with each other by taking only a few steps. Path-length analysis of individual amino acid side-chains was shown to be effective in predicting whether the protein, or its segment, is disordered. In protein structure networks we may find considerably less large hubs than in other networks. However, the existing smaller hubs still play an important role in protein structures, since these ‘micro-hubs’ were shown to increase the thermodynamic stability of proteins (Kannan & Vishveshwara, 1999; Greene & Higman, 2003; Atilgan et al., 2004; Bagler & Sinha, 2005; Brinda & Vishveshwara, 2005; Del Sol et al., 2006, 2007; Alves & Martinez, 2007; Krishnan et al., 2008; Konrat, 2009; Morita & Takano, 2009; Estrada, 2010; Csermely et al., 2012). Protein structure networks possess a rich club structure with the exception of membrane proteins, where hubs form disconnected, multiple clusters (Pabuwal & Li, 2009).

Protein structure networks have modules, which often encode protein domains (Xu et al., 2000; Guo et al., 2003; Delvenne et al., 2010; Delmotte et al., 2011; Szalay-Bekő et al., 2012). High-centrality segments of protein structure networks (i.e. hubs, or nodes with high closeness or betweenness centralities) having a low clustering coefficient participate in hem-binding (Liu & Hu, 2011). High-centrality, inter-modular bridges play a key role in the transmission of allosteric changes as we will describe in the next section.

Evolutionary conservation patterns of amino acids in related protein structures identified protein sectors (Halabi et al., 2009; McLaughlin et al., 2012). A similar concept has been published by Jeon et al. (2011), who determined that co-evolving amino acid pairs are often clustered in flexible protein regions. Protein sectors are sparse networks of amino acids spanning a large segment of the protein. Protein sectors are collective systems operating rather independently from each other. Segments of protein sectors are correlated with protein movements related to enzyme catalysis, and sector-connected surface sites are often places of allosteric regulation (Reynolds et al., 2011).

### 3.2.2. Key network residues determining protein dynamics

Understanding protein dynamics is a key step in the prediction of drug-induced changes and the identification of novel types of drug binding sites. Several questions related to protein dynamics, such as the mechanism of allosteric changes gained much attention in the last century (Fischer, 1894; Koshland, 1958; Straub & Szabolcsi, 1964; Závodszy et al., 1966; Tsai et al., 1999; Goodey & Benkovic, 2008; Csermely et al., 2010; Szilágyi et al., in press), but have not been completely elucidated yet.

Our current understanding indicates that an allosteric change extends across a spectrum: at one end there is a switch-type conformational change, where signaling appears focused on a small number of amino acids (Fig. 14A). At the other end, allosteric signaling may involve a large number of amino acids. In both cases allosteric signals propagate through multiple trajectories, with different distributions and weights. Pathways often converge at inter-domain boundaries (Fig. 14B). While protein segments involved in switch-type allosteric changes may be more rigid, protein segments harboring multiple trajectories may be more flexible. Convergence points in these latter proteins may mark the more flexible inter-domain regions. Switch-

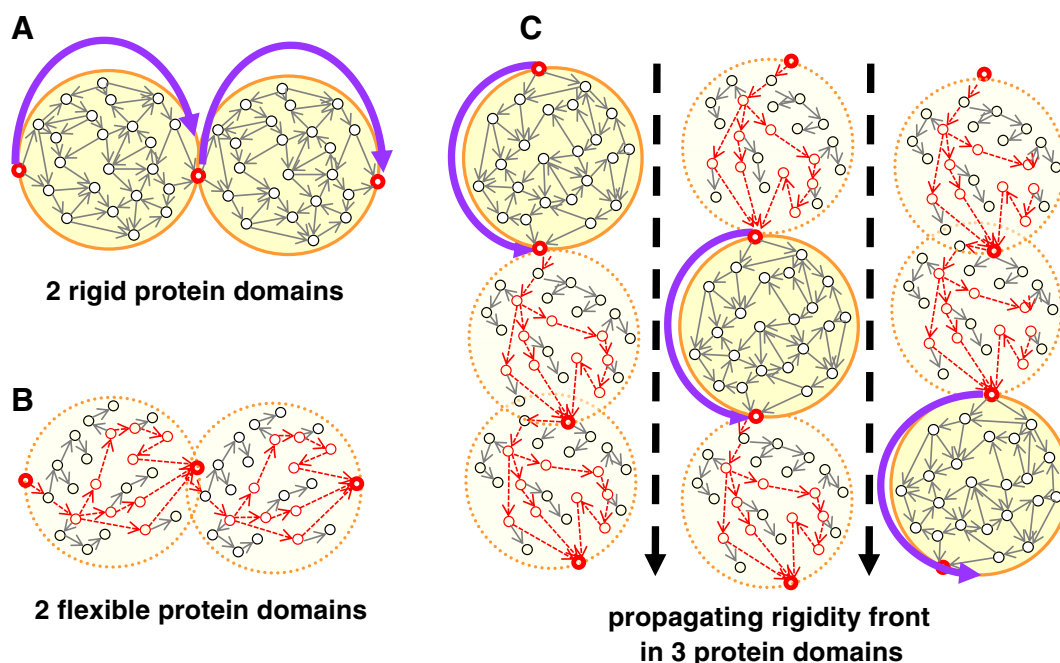
type mechanism is typical of multidomain proteins, and is often expressed in all-or-none observable consequences.

Disordered protein regions serve the need for large flexibility, and therefore are often used, especially in the human proteome (Csermely et al., 2012; Tompa, 2012). Protein structure network description may be a useful method to describe the complexity of protein structures, which are neither very rigid, nor very flexible. However, protein structure networks may not adequately describe the dynamics of very rigid and extremely flexible (e.g. disordered) protein structures (Szilágyi et al., in press).

Hinges connecting relatively rigid protein segments often play a decisive role in switch-type changes. Hinges may be co-localized with independent dynamic segments, which are situated in the stiffest parts of the protein, and harbor spatially localized vibrations of non-linear origin, like those of discrete breathers. Independent dynamic segments exchange their energy largely via a direct energy transfer, which is in agreement with a switch-type behavior (Daily et al., 2008; Piazza & Sanejouand, 2008, 2009; Csermely et al., 2010, 2012).

In contrast, in allosteric systems, where signaling involves a large number of amino acids, signals propagate using multiple trajectories (Fig. 14B). These multiple trajectories often converge at inter-modular residues of protein structure networks (Pan et al., 2000; Chennubhotla & Bahar, 2006; Ghosh & Vishveshwara, 2007, 2008; Tang et al., 2007; Daily et al., 2008; Sethi et al., 2009; Tehver et al., 2009; Vishveshwara et al., 2009; Csermely et al., 2012; Gasper et al., 2012).

A given protein may have a spectrum of the above mechanisms for the propagation of conformational changes. In agreement with this behavior, discrete breathers were shown to be present at the interface between monoatomic and diatomic granular chain models (Hoogeboom et al., 2010). If certain protein segments become more rigid, the mechanism may shift towards the first, switch-type signal transduction mechanism. This can be conceptualized as the propagation



**Fig. 14.** Saltatoric signal transduction along a propagating rigidity-front: a possible mechanism of allosteric action in protein structures. Panel A shows two rigid modules of protein structure networks (corresponding to protein segments or domains). Such modules have little overlap, behave like billiard balls, and transmit signals 'instantaneously' (illustrated with the violet arrows). Panel B shows two flexible modules. These modules have a larger overlap, and transmit signals via a slower mechanism using multiple trajectories, which converge at key, bridging amino acids situated in modular boundaries. Panel C combines rigid and flexible modules in a hypothetical model of rigidity front propagation of the allosteric conformational change. In the 3 snapshots of this illustration of protein dynamics (organized from left to right) the 3 protein segments become rigid from top to bottom. Consecutive 'rigidization' of protein segments both induces similar changes in the neighboring segment, and accelerates the propagation of the allosteric change within the rigid segment. Rigidity front propagation may use sequential energy transfers (illustrated by the violet arrows), and may increase the speed of the allosteric change approaching that of an instantaneous process (Piazza & Sanejouand, 2009; Csermely et al., 2010, 2012). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

of a 'rigidity-front', which we recently proposed as a mechanism of allosteric signaling (Csermely et al., 2012). Panel C of Fig. 14 shows an illustrative mechanism of rigidity front propagation. Consecutive 'rigidization' of protein segments both induces similar changes in the neighboring segment, and accelerates the propagation of the allosteric change within the rigid segment. Rigidity front propagation may use sequential energy transfers (illustrated by the violet arrows; Piazza & Sanejouand, 2009; Csermely et al., 2010), and thus may increase the speed of the allosteric change (Csermely et al., 2012). The rigidity front propagation model combines elements of 'rigidity propagation' (Jacobs et al., 2001, 2003; Rader & Brown, 2011) with the 'frustration front' concept of Zhuravlev and Papoian (2010), with the dynamic pre-stress of proteins as large as a few 100 pN (Edwards et al., 2012), and is in agreement with the recent proposal of Dixit and Verkhivker (2012) suggesting an interactive network of minimally frustrated (rigid) anchor sites (hot spots) and locally frustrated (flexible) proximal recognition sites to play a key role in allosteric signaling. We will describe the use of allosteric signal propagation mechanisms to design allo-network drugs (Nussinov et al., 2011) in Section 4.1.6.

Protein structure networks may be efficiently used to identify key amino acids involved in intra-protein signal transmission. In these studies topological network analysis was often combined with the assessment of evolutionary conservation, elastic network models and/or normal mode analysis. Inter-modular nodes, hinges, loops and hubs were particularly important in information transmission (Chennubhotla & Bahar, 2006, 2007; Zheng et al., 2007; Chennubhotla et al., 2008; Tehver et al., 2009; Liu & Bahar, 2010; Liu et al., 2010a; Park & Kim, 2011; Su et al., 2011; Dixit & Verkhivker, 2012; Ma et al., 2012a; Pandini et al., 2012). The examination of a hierarchical representation of protein structure networks showed a key function of top level, 'superhubs' in allosteric signaling (Ma et al., 2012a). xPyder provides an interface between the widely employed molecular graphics system, PyMOL and the analysis of dynamical cross-correlation matrices (<http://linux.btbs.unimib.it/xpyder>; Pasi et al., 2012).

The incorporation of novel network centrality measures (described in Section 2.3.3) and network dynamics (described in Section 2.5) will enrich our knowledge of the mechanism of conformational changes (including allostery).

### 3.2.3. Disease-associated nodes of protein structure networks

Proteins related to more frequently occurring diseases tend to be longer than average (Lopez-Bigas et al., 2005). Disease-related proteins have a smaller 'designability'; that is, their folds can be built up from fewer variants than the average. In other words: disease-related proteins have a constrained structure, which may explain, why they have debilitating mutations (Wong & Frishman, 2006). Disease-associated mutations (single-nucleotide polymorphisms) often occur at sites having a high local or global centrality in the protein structure network, and are enriched by 3-fold at the interaction interfaces of proteins associated with the disorder (Akula et al., 2011; Li et al., 2011b; Wang et al., 2012b). Recently, a machine learning method has been developed to predict the disease-association of a single-nucleotide polymorphism using the network neighborhood of the mutation site (Li et al., 2011b).

### 3.2.4. Prediction of hot spots and drug binding sites using protein structure networks

Key functional residues are very useful in the identification of drug binding sites as we will discuss in Section 4.2. Usual drug binding sites are cavity-type, and overlap with substrate or allosteric ligand binding. High centrality residues of protein structure networks were shown to participate in ligand binding (Liu & Hu, 2011). Protein structure network position-based scores improved the rigid-body docking algorithm of pyDock (Pons et al., 2011). Protein structure comparison

can also be used for the identification of chemical scaffolds of potential drug candidates (Konrat, 2009).

Binding sites of edgetic drugs modifying protein–protein interactions (see Section 4.1.2) are large and flat, and have been considered as non-druggable for a long time. Hot spots are those residues of these alternative drug binding sites, which provide a key contribution (>2 kcal/mol) to the decrease in binding free energy. Hot spots tend to cluster in tightly packed, relatively rigid hydrophobic regions of the protein–protein interface also called hot regions. Hot spots and hot regions are very helpful; they aid drug design, since 1.) they constitute small focal points of drug binding, which can be predicted within the large and flat binding-interface; 2.) these focal points are relatively rigid, helping rigid docking and molecular dynamics simulations (Clackson & Wells, 1995; Bogan & Thorn, 1998; Keskin et al., 2005, 2007; Ozbabacan et al., 2010). Hot spots can be predicted as central nodes of protein structure networks (del Sol & O'Meara, 2005; Liu & Hu, 2011; Grosdidier & Fernández-Recio, 2012).

The elastic network model-guided molecular dynamics simulation of Isin et al. (2012) showed that different ligands of the  $\beta_2$ -adrenergic receptor prefer different predicted conformers of the receptor. This model predicted a novel allosteric binding site for larger drugs, such as salmeterol. Since different conformations participate in different metabolic and signaling pathways, such conformational modeling will be a powerful tool in the determination of novel drug binding sites and in the analysis of refined drug action mechanisms.

Despite these advances, the use of the predictive power of protein structure networks is surprisingly low in the determination of drug binding sites. We believe that the arsenal of network analytical and network dynamics methods we listed in Sections 2.3 and 2.5 and their application to protein structure networks will provide a much greater help in the identification of drug binding sites in the future.

## 3.3. Protein–protein interaction networks (network proteomics)

Protein–protein interaction networks are one of the most promising network types to predict drug action or identify new drug target candidates. In this section we will summarize the major properties of protein–protein interaction networks and will assess their use in the characterization and prediction of disease-related proteins and drug targets.

### 3.3.1. Definition and general properties of protein–protein interaction networks

Protein–protein interaction networks (PPI-networks) are often called interactomes—especially if they cover genome-wide data. Nodes of protein–protein interaction networks are proteins, and network edges are their direct, physical interactions. Protein–protein interaction networks are probability-type networks; that is, the edge weights reflect the probability of the actual interaction. Interactome edge weights are often calculated as confidence scores. Interaction probability includes protein abundance, interaction affinity, and also co-expression levels, co-localization in subcellular compartments, etc. (De Las Rivas & Fontanillo, 2010; Jessulat et al., 2011; Sardiú & Washburn, 2011; Seebacher & Gavin, 2011). Table 6 summarizes a number of major protein–protein interaction datasets concentrating on publicly available, human interactome data. There are several types of protein–protein interaction networks, which we list below.

- Even though protein–protein interaction datasets usually cover multiple species, the derived networks, that is, the interactomes are usually species-specific. Interactome subnetworks may be restricted to cell type, to cellular sub-compartment, or to certain temporal segments of cellular life, such as a part of the cell cycle, cell differentiation, malignant transformation, etc. These specializations may be direct, where the interactions of proteins are experimentally measured in the given species, cell type, cellular compartment, or

**Table 6**  
Protein–protein interaction network resources.

Name	Content	Website	References
3did	Domain–domain interaction with 3D data	<a href="http://3did.irbbarcelona.org">http://3did.irbbarcelona.org</a>	Stein et al., 2011
APID	Interactome exploration	<a href="http://bioinfow.dep.usal.es/apid/index.htm">http://bioinfow.dep.usal.es/apid/index.htm</a>	Prieto & De Las Rivas, 2006
atBioNet	Integration of 7 interactomes, protein complex identification	<a href="http://www.fda.gov/ScienceResearch/BioinformaticsTools/ucm285284.htm">http://www.fda.gov/ScienceResearch/BioinformaticsTools/ucm285284.htm</a>	Ding et al., 2012
BioGRID	Integrated protein–protein interaction data	<a href="http://thebiogrid.org">http://thebiogrid.org</a>	Stark et al., 2011
BioProfiling	Inference of network data from expression patterns	<a href="http://www.bioprofiling.de">http://www.bioprofiling.de</a>	Antonov et al., 2009
DIP	Experimental protein–protein interaction data	<a href="http://dip.doe-mbi.ucla.edu">http://dip.doe-mbi.ucla.edu</a>	Salwinski et al., 2004
DomainGraph	Cytoscape plug-in for domain–domain interaction analysis	<a href="http://domaingraph.bioinf.mpi-inf.mpg.de">http://domaingraph.bioinf.mpi-inf.mpg.de</a>	Emig et al., 2008
DOMINE	Domain–domain interaction data	<a href="http://domine.utdallas.edu">http://domine.utdallas.edu</a>	Yellaboina et al., 2011
Estrella	Detection of mutually exclusive protein–protein interactions	<a href="http://bl210.caspar.it/ESTRELLA/help.php">http://bl210.caspar.it/ESTRELLA/help.php</a>	Sanchez Claros & Tramontano, 2012
HAPPI	Human protein–protein interaction data	<a href="http://discern.uits.iu.edu:8340/HAPPI">http://discern.uits.iu.edu:8340/HAPPI</a>	Chen et al., 2009c
HINT	High quality human protein–protein interaction data	<a href="http://hint.yulab.org">http://hint.yulab.org</a>	Das & Yu, 2012
HPRD	Human protein–protein interaction data	<a href="http://www.hprd.org">http://www.hprd.org</a>	Goel et al., 2012
Hubba	Identification of hubs (potentially essential proteins)	<a href="http://hub.iis.sinica.edu.tw/Hubba">http://hub.iis.sinica.edu.tw/Hubba</a>	Lin et al., 2008
IntAct	Curated protein–protein interaction data	<a href="http://www.ebi.ac.uk/intact/main.xhtml">http://www.ebi.ac.uk/intact/main.xhtml</a>	Kerrien et al., 2012
IntNetDB	Human protein–protein interaction data	<a href="http://hanlab.genetics.ac.cn/sys">http://hanlab.genetics.ac.cn/sys</a>	Xia et al., 2006
IRView	Protein interacting regions	<a href="http://ir.hgc.jp">http://ir.hgc.jp</a>	Fujimori et al., 2012
MiMI	Protein interaction information	<a href="http://mimi.ncbi.org">http://mimi.ncbi.org</a>	Gao et al., 2009
MINT	Protein–protein interactions in refereed journals	<a href="http://mint.bio.uniroma2.it/mint">http://mint.bio.uniroma2.it/mint</a>	Licata et al., 2012
NeAT (Network Analysis Tools)	Interactome analysis	<a href="http://rsat.ulb.ac.be/neat/">http://rsat.ulb.ac.be/neat/</a>	Brohee et al., 2008
NetAligner	Interactome comparison	<a href="http://netaligner.irbbarcelona.org">http://netaligner.irbbarcelona.org</a>	Pache et al., 2012
PanGIA	A Cytoscape plug-in for integration of physical and genetic interactions into hierarchical module maps	<a href="http://prosecco.ucsd.edu/PanGIA">http://prosecco.ucsd.edu/PanGIA</a>	Srivastava et al., 2011
Pathwaylinker	Combines protein–protein interaction and signaling data	<a href="http://PathwayLinker.org">http://PathwayLinker.org</a>	Farkas et al., 2012
PINA	Interactome analysis	<a href="http://cgb.garvan.unsw.edu.au/pina">http://cgb.garvan.unsw.edu.au/pina</a>	Wu et al., 2009b; Cowley et al., 2012
PIPs	Human protein–protein interaction prediction	<a href="http://www.compbio.dundee.ac.uk/www-pips">http://www.compbio.dundee.ac.uk/www-pips</a>	McDowall et al., 2009
PPIsearch	Search of homologous protein–protein interactions across many species	<a href="http://gemdock.life.nctu.edu.tw/ppisearch">http://gemdock.life.nctu.edu.tw/ppisearch</a>	Chen et al., 2009d
STRING	Integrated protein–protein interaction data	<a href="http://string.embl.de">http://string.embl.de</a>	Szklarczyk et al., 2011
UniHI	Human protein–protein interaction and drug target data	<a href="http://www.unihi.org">http://www.unihi.org</a>	Chaurasia & Futschik, 2012

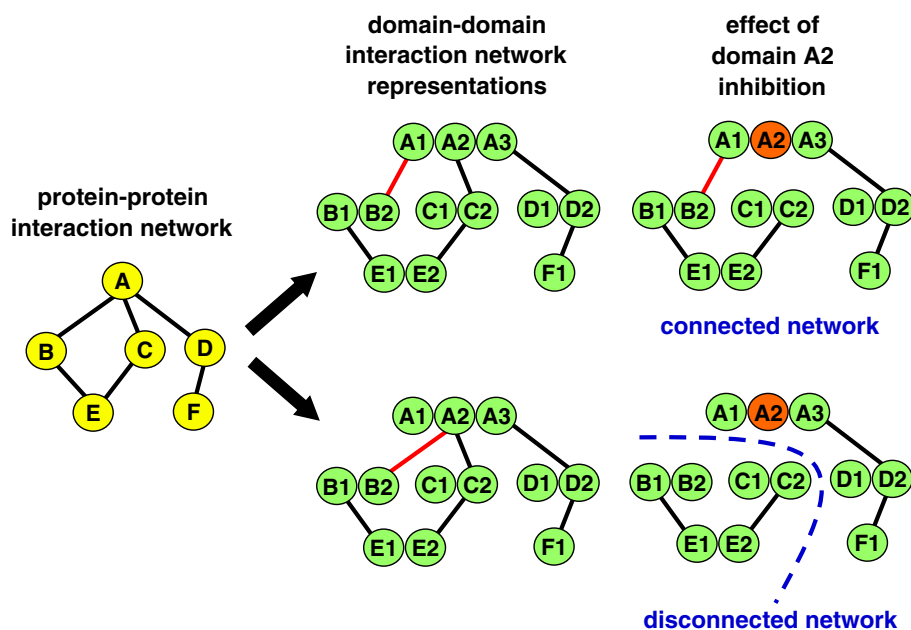
The table is focused on recently available public databases or web-servers applicable to human protein–protein interaction data and/or to drug design. Network visualization tools were listed in Table 1. A collection of protein–protein interaction network analysis web-tools can be found in recent reviews (Ma'ayan, 2008; Moschopoulos et al., 2011; Ma & Gao, 2012; Sanz-Pamplona et al., 2012). The Reader may find a more extensive list of web-sites in recent collections (<http://ppi.fli-leibniz.de>; Seebacher & Gavin, 2011).

condition. In many cases the specializations are indirect, where the presence of the actual proteins and/or the intensity of the protein–protein interactions are estimated from mRNA expression levels. Disease-specific or drug treatment-related interactomes hold promise for future drug development efforts (De Las Rivas & Fontanillo, 2010; Jessulat et al., 2011; Sardiú & Washburn, 2011; Seebacher & Gavin, 2011).

- Protein–protein interaction networks may be refined to networks of interacting protein domains, called domain networks or DDI-networks. Domain networks can be a better representation of drug action, deciphering domain-specific inhibition or activation (Fig. 15). Current lists of possible domain–domain interactions predict millions of novel, potential protein–protein interactions. However, not all domain–domain interactions may occur in the cellular context due to hindrances, binding competition or subcellular localization. Domain–domain interactions and their networks were used both to score protein–protein interactions (bottom–up approach) and to predict domain composition and interactions from interactome data (top–down approach) (Deng et al., 2002; Ng et al., 2003; Moon et al., 2005; Santonico et al., 2005; Emig et al., 2008; Prieto & De Las Rivas, 2010; Stein et al., 2011; Yellaboina et al., 2011).
- Atomic resolution interactomes expand protein–protein interaction networks with the protein structure networks of each interacting node aiming to construct the 3D structure of the whole interactome, and discriminating between parallel and sequential interactions (Kim et al., 2006; Prieto & De Las Rivas, 2010; Bhardwaj et al., 2011; Clarke et al., 2012; Pache & Aloy, 2012; Sanchez Claros & Tramontano, 2012). It is important to note that atomic level interactomes will never reach the real 3D complexity of the cell, since protein–protein interactions are probabilistic, reflecting an average of the possible interactions.

Protein–protein interaction data can be obtained using various high-throughput methods (such as protein fragment complementation assays, or affinity purification combined with mass spectrometry), text mining or prediction techniques. For details on the increasing number of methodologies the Reader is referred to recent reviews on the subject (Chautard et al., 2009; De Las Rivas & Fontanillo, 2010, 2012; Jessulat et al., 2011; Sardiú & Washburn, 2011; Seebacher & Gavin, 2011; Gonzalez & Kahn, 2012). Prediction methods were also summarized in Sections 1.3.3 and 2.2.2, as well as in Tables 2 and 3. Data-quality is a major problem of interactomes. Sampling bias, missing interactions and false positives are all important factors influencing the robustness of interactome results. Evolutionary conservation rates of interactions are often low (Lewis et al., 2012). High-quality data are more reliable, but are not necessarily representative of whole interactomes. Some of these problems may be circumvented by using confidence scores calculated by various methods, such as the summative, network topology-based or Bayesian network-based models. Since different methods have different biases, composite scores taking multiple data-types into account perform better (Hakes et al., 2008; Sanchez Claros & Tramontano, 2012). The size of the human interactome has been estimated to have 650,000 interactions (Stumpf et al., 2008). Though a recent report (Havugimana et al., 2012) added 14 thousand high-confidence interactions to the growing list of human interactome edges, currently we are still far from deciphering the full complexity of this richness.

Table 6 lists a number of web-resources used for interactome analysis. A collection of protein–protein interaction network analysis web-tools can be found in recent reviews (Ma'ayan, 2008; Moschopoulos et al., 2011; Ma & Gao, 2012; Sanz-Pamplona et al., 2012). Protein–protein interaction networks are small worlds, have hubs and a well developed, hierarchical modular structure. These interactomes do not



**Fig. 15.** The effect of more detailed representation of protein–protein interaction networks in representation of drug mechanism action. The left side of the figure shows a hypothetical protein–protein interaction network (yellow nodes). The middle panels show two representations of the very same network as a domain–domain interaction network (green nodes). Note that on the middle top panel the edge marked with red connects domains A1 and B2 while on the middle bottom panel the same edge connects domains A2 and B2. Note that these two representations cannot be discriminated at the protein–protein interaction level (shown on the left side marked with green nodes). If domain A2 (highlighted with red) is inhibited by a drug (and there is limited domain–domain interaction in protein A), this single edge-change leaves the sub-interactome in the right top panel intact. On the contrary, in the right bottom panel, the inhibition of domain A2 leads to the dissociation of the sub-network. An atomic level resolution of the interactome can discriminate even more subtle changes as we will discuss in Section 4.1.6 on allo-network drugs (Nussinov et al., 2011). The figure is re-drawn from Figure 2 of Santonico et al. (2005) with permission. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

possess such an extensive rich-club as the social elite, i.e. hubs do not form dense clusters with each other (Maslov & Sneppen, 2002; Colizza et al., 2006; De Las Rivas & Fontanillo, 2010; Sardiú & Washburn, 2011). Soluble proteins tend to possess more connections than membrane proteins (Yu et al., 2004a). Steric hindrances severely limit the maximum number of simultaneous interactions. Tsai et al. (2009) warned that large interactome hubs may often be a result of aggregated data not taking into account protein conformations, posttranslational modifications, isoforms, expression differences and localizations. Another possibility to increase binding partners is sequential binding, which results in the formation of date hubs (as opposed to party hubs binding their partners simultaneously). Date hubs are often singlish-interface hubs as opposed to party-hubs, which are multi-interface hubs. Multi-interface hubs display a greater degree of conformational change than singlish-interface hubs (Han et al., 2004a; Kim et al., 2006; Bhardwaj et al., 2011). Interestingly, natural product drugs were shown to target proteins having a higher number of neighbors than targets of synthetic drugs (Dancik et al., 2010).

Interactome modules overlap with each other, since most proteins are members of multiple protein complexes. Modules often correspond to major cellular functions (Palla et al., 2005). Refined modularization methods define modular cores containing only a few proteins, which occupy a central position of the interactome module. Major function of core proteins often reflects a consensus function of the whole module (Kovács et al., 2010; Szalay-Bekó et al., 2012; Srihari & Leong, 2013). Importantly, protein complexes may also form sparsely interconnected network modules, which often escape traditional detection methods (Srihari & Leong, 2012). Date hubs occupy inter-modular positions as opposed to party-hubs, which are in modular centers. Multi-component hubs (which, similarly to date-hubs, bridge multiple dense local network components) were enriched in regulatory proteins. Bridges and other inter-modular nodes play a key role in drug action (Han et al., 2004a; Komurov & White, 2007; Kovács et al., 2010; Fox et al., 2011; Szalay-Bekó et al., 2012).

As we discussed in Section 2.3.4, interactome hubs were shown to be an important predictor of essentiality (Jeong et al., 2001). Hub Objects Analyzer (Hubba) is a web-based service for exploring potentially essential nodes of interactomes assessing the maximum neighborhood component (Lin et al., 2008). Single-component hubs (i.e. hubs in the middle of a stable network neighborhood) were shown to be more essential than multi-component hubs, i.e. hubs connecting multiple dense network regions (Fox et al., 2011). Essential proteins associate with each other more closely than the average, and tend to be more promiscuous in their function. Many of these essential genes are house-keeping genes with high and less fluctuating expression levels (Jeong et al., 2001; Yu et al., 2004c). Later more global network measures, such as bottlenecks or more globally central proteins were also shown to contribute to the determination of essential nodes (Chin & Samanta, 2003; Estrada, 2006; Yu et al., 2007b; Missiuro et al., 2009; Li et al., 2011a).

The recent work of Hamp and Rost (2012) uncovered that variability of protein–protein interactions is much more frequent than previously thought. Besides single-nucleotide polymorphisms, alternative splicing, addition of N- or C-terminal tags, partial proteolysis and other post-translational modifications (such as phosphorylation), changes in protein expression patterns may dramatically re-configure protein complexes. Dynamic changes of protein–protein interactions, such as co-expression based clustering are key determinants of the disease state as we discuss in the next section. Importantly, interactome analysis has not been adequately extended to the assessment of interactome dynamics, and currently the application of the tools listed in Section 2.5 is largely missing. The human proteome is enriched in disordered proteins causing dynamically fluctuating, ‘fuzzy’ interaction patterns (Tompa, 2012). As an initial example of these studies the yeast interactome was shown to develop more condensed and more separated modules after heat shock and other types of stresses than under optimal growth conditions. Importantly, yeast cells preserved a few inter-modular bridges during stress and

developed novel, stress-specific bridges containing key proteins in cell survival (Mihalik & Csermely, 2011).

### 3.3.2. Protein–protein interaction networks and disease

Most human diseases are oligogenic or polygenic affecting a whole set of proteins and their interactions. In the last decade several genome-wide datasets became available to characterize disease-related patho-mechanisms. mRNA expression patterns, genome-wide association studies (GWAS) of disease-associated single-nucleotide polymorphisms (SNPs) and disease-related changes in posttranslational modifications (such as the phospho-proteome) are just three of the most widely used datasets, which may also include system-wide changes of subcellular localization. All this information can be incorporated in protein–protein interaction networks as changes in edge configuration and weights (Zanzoni et al., 2009; Coulombe, 2011).

As we described in Section 1.3, disease-associated proteins do not generally act as interactome hubs, with the important exception of somatic mutations, such as those occurring in cancer, where disease-associated multi-interface hubs form an inter-connected rich club (Jonsson & Bates, 2006; Goh et al., 2007; Feldman et al., 2008; Kar et al., 2009; Barabási et al., 2011; Zhang et al., 2011a). Disease-related proteins have a smaller clustering coefficient than average, which was used for the prediction of novel disease-related genes (Feldman et al., 2008; Sharma et al., 2010a).

Disease-related proteins form overlapping disease modules. Suthram et al. (2010) identified 59 core modules out of the 4620 modules of the human interactome, which were affected by mRNA changes in more than half of the 54 diseases examined. These core modules were often targeted by drugs, and drugs affecting the core modules were more often multi-target drugs than those acting on ‘peripheral’ modules, which changed their mRNA levels only in a few specific diseases. Bridges and additional types of overlaps between disease-related interactome modules may provide important points of interventions (Nguyen & Jordan, 2010; Nguyen et al., 2011).

### 3.3.3. The use of protein–protein interaction networks in drug design

Uncovering the estimated ~650,000 interactions of the human interactome (Stumpf et al., 2008) is an ongoing, key step in network-related drug design efforts (see Rual et al., 2005; Stelzl et al., 2005; Chautard et al., 2009; Burkard et al., 2011; De Las Rivas & Prieto, 2012; Havugimana et al., 2012 and databases of Table 6). Databases like ChemProt: <http://www.cbs.dtu.dk/services/ChemProt> including 700,000 chemicals and 2 millions of interactions of their target proteins in various species (Taboureau et al., 2011) provide a great help in this process. However, interactome complexity goes much beyond the inventory of contacts and binding partners, and includes expression level-induced, posttranslational modification-induced (such as phosphorylation-dependent), cellular environment-induced (such as calcium-dependent) and protein domain-dependent variations (Santonico et al., 2005). We illustrate the latter in Fig. 15.

Drug targets have a generally larger number of neighbors than average. In agreement with assumptions related to disease-associated protein contacts described in the previous section, the larger number of neighbors comes mostly from the contribution of middle-degree nodes; but not hubs. Drug targets in cancer are exceptions having a more defined hub-structure. Drug target proteins have a lower clustering coefficient than other proteins. Drug targets often occupy a central position in the human interactome bridging two or more modules. Nodes having an intermediate number of neighbors have an extensive contact structure. Targeting these non-hub nodes (with the exception of infectious diseases and cancer) is crucial to avoid unwanted side-effects. As opposed to targets of withdrawn drugs having a too large network influence, drug target proteins perturb the interactome in a controlled manner (Hase et al., 2009; Zhu et al., 2009; Bultinck et al., 2012; Yu & Huang, 2012).

Properties of the interactome topology were used to predict and score novel drug target candidates using mainly machine learning techniques (Zhu et al., 2009; Zhang & Huan, 2010; Yu & Huang, 2012). Network neighborhood similarity to the drug targets test-set proved to be a good predictor of additional targets (Zhang & Huan, 2010). This feature may actually show the limits of machine learning-based approaches: since current drug targets are often similar to each other (Cokol et al., 2005; Yildirim et al., 2007; Iyer et al., 2011a), machine learning techniques may not be useful to extend the current drug target inventory to surprisingly novel hits.

Modulation of specific protein–protein interactions provides a much higher specificity to restore disease pathology to the normal state than targeting a whole protein. We will describe methods for the design of such ‘edgetic drugs’ in Section 4.1.2. Conceptually, it is much easier to develop inhibitors of protein–protein interactions than agents for increasing binding affinity or stability. The latter option together with the inclusion of interactome dynamics using the tools listed in Section 2.5 is a very promising future trend of the field. As a recent advance to explore drug-induced interactome dynamics, Schlecht et al. (2012) investigated the changes in the yeast interactome upon the addition of 80 diverse small molecules. Their method could identify novel protein–protein contacts specifically disrupted by the addition of drugs such as the immunosuppressant, FK506.

### 3.4. Signaling, microRNA and transcriptional networks

“Representations of signaling networks in many textbooks and even in some of the most recent review articles have in their simplicity a striking similarity to children’s drawings” (Lewitzky et al., 2012; Fig. 16). The complex representation of a signaling network is constructed by upstream and downstream subnetworks. The upstream subnetwork contains the intertwined network of signaling pathways, while the downstream, regulatory part contains DNA transcription factor binding sites and microRNAs (Fig. 17). As we will show in the following subsections, both subnetworks are highly structured, are linked to each other, and are very important in drug discovery. The systems-level exploration and understanding of signaling networks significantly facilitate drug target identification,

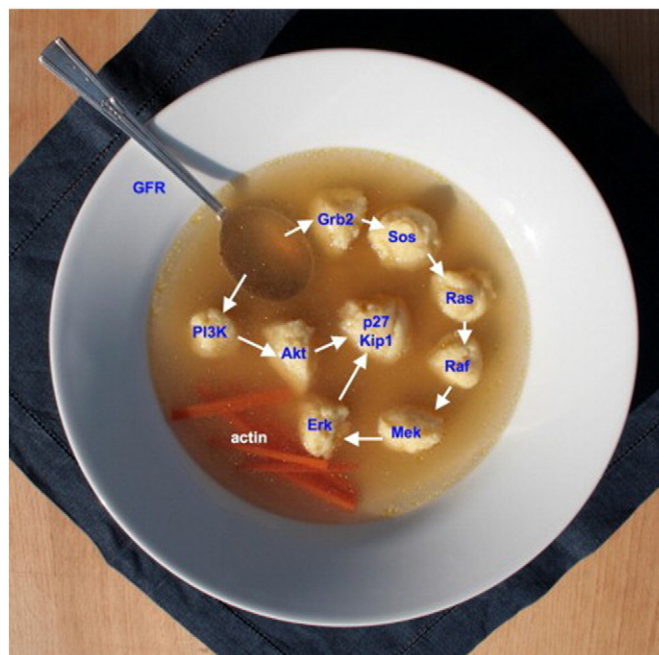
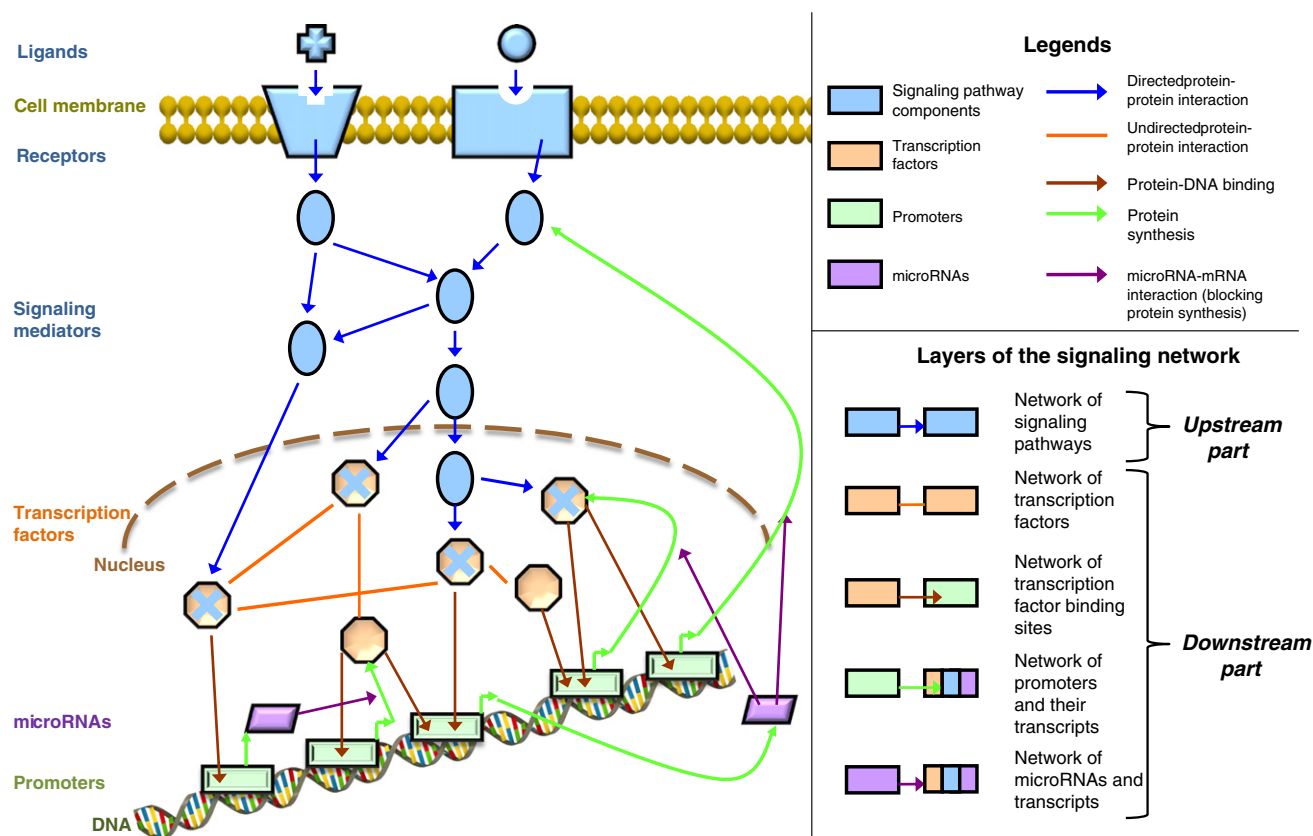


Fig. 16. The dumpling soup representation of growth factor initiated signaling. Reproduced with permission from Lewitzky et al. (2012).



**Fig. 17.** Structure of the signaling network. The figure illustrates the major components of the signaling network including the upstream part of the signaling pathways and their cross-talks and the downstream part of gene regulation network. The gene regulation network contains the subnetworks of transcription factors, their DNA-binding sites and regulating microRNAs. Directed protein–protein interactions may encode enzyme reactions, such as phosphorylation events, while undirected protein–protein interactions participate (among others) in formation of scaffold and adaptor complexes.

target selection in pathological networks and the avoidance of unwanted side-effects. At the end of the section we will also point out important network features that make signaling-related drug discovery a challenging task.

### 3.4.1. Organization and analysis signaling networks

Signaling pathways, the functional building blocks of intracellular signaling, transmit extracellular information from ligands through receptors and mediators to transcription factors, which induce specific gene expression changes. Signaling pathways constitute the upstream part of signaling networks (Fig. 17). Over the past decade, it has been realized that signaling pathways are highly structured, and are rich in cross-talks, where cross-talk was defined as a directed physical interaction between pathways (Papin et al., 2005; Fraser & Germain, 2009). As the number of input signals (ligands/receptors) and output components (transcription factors) is limited, cross-talks between pathways can create novel input/output combinations contributing to the functional diversity and plasticity of the signaling network (Kitano, 2004a). However, cross-talks have to be precisely regulated to maintain output specificity (meaning that inputs preferentially activate their own output) and input fidelity (meaning that outputs preferentially respond to their own input). Regulation of cross-talks to prevent ‘leaking’ or ‘spillover’ can be achieved using different mediation mechanisms, such as scaffolding proteins, cross-pathway inhibitions, kinetic insulation, and the spatial and temporal expression patterns of proteins (Freeman, 2000; Bhattacharyya et al., 2006; Kholodenko, 2006; Behar et al., 2007; Haney et al., 2010; Lewitzky et al., 2012).

The regulatory subnetwork (gene regulatory network) constitutes the downstream part of the signaling pathway-network (Lin et al., 2012). The gene regulatory network can be separated into the

transcriptional and the post-transcriptional levels. At the transcriptional level, transcription factors bind specific regions of DNA sequences (called transcription factor binding sites, or response elements) regulating their mRNA expression. Horizontal contacts of middle-level regulators play a key role in gene regulatory networks, especially in human cells. The human transcription factor regulatory network has a basic architecture, which is independent from the cell type and is complemented by cell type specific segments (Bhardwaj et al., 2010; Gerstein et al., 2012; Neph et al., 2012).

MicroRNAs (miRNAs or miRs) are key players of gene regulatory networks, and regulate gene expression by binding to complementary sequences (i.e. microRNA binding-sites) on target mRNAs. MicroRNA binding may suspend or permanently repress the translation of given transcripts (Doench & Sharp, 2004; Guo et al., 2010). In the last decade, it became evident that nearly all human genes can be controlled by at least one microRNA (Lewis et al., 2003), and that mutations in microRNA coding genes often have pathological consequences (Calin & Croce, 2006). Interactome hubs, bottleneck proteins and downstream signaling components, such as transcription factors are regulated by more microRNAs than other nodes (Cui et al., 2006; Liang & Li, 2007; Hsu et al., 2008).

Besides biochemical and molecular biological approaches, reverse engineering of genome-wide transcriptional changes proved to be very efficient for determining signaling networks as we detailed in Section 2.2.3. Signaling networks are small-worlds and possess signaling hubs. Networks (partially due to their pathway structures) have modules, and cross-talking proteins may often be considered as bridges between these modules. In the last decade several resources have been developed to provide signaling pathways, transcription factor and transcription factor binding site information, as well as microRNA networks. We summarize some of these signaling network resources in Table 7. A

**Table 7**

Signaling network resources.

Name	Content	Website	References
IPAVS	Signaling pathway resources	<a href="http://ipavs.cidms.org">http://ipavs.cidms.org</a>	Sreenivasiah et al., 2012
Reactome		<a href="http://reactome.org">http://reactome.org</a>	Croft et al., 2011
NCI Nature–Pathway Interaction Database		<a href="http://pid.nci.nih.gov">http://pid.nci.nih.gov</a>	Schaefer et al., 2009
NetPath	Transcription factor–transcription factor binding information	<a href="http://netpath.org">http://netpath.org</a>	Kandasamy et al., 2010
JASPAR		<a href="http://jaspar.genereg.net">http://jaspar.genereg.net</a>	Portales-Casamar et al., 2010
HTRIdb		<a href="http://www.lbbc.ibb.unesp.br/htri">http://www.lbbc.ibb.unesp.br/htri</a>	Bovolenta et al., 2012
MPromDB		<a href="http://mpromdb.wistar.upenn.edu">http://mpromdb.wistar.upenn.edu</a>	Gupta et al., 2011
PAZAR		<a href="http://pazar.info">http://pazar.info</a>	Portales-Casamar et al., 2007
OregAnno		<a href="http://oreganno.org">http://oreganno.org</a>	Griffith et al., 2008
Expander	Transcription factor and microRNA target prediction from gene expression data	<a href="http://acgt.cs.tau.ac.il/expander">http://acgt.cs.tau.ac.il/expander</a>	Ulitsky et al., 2010
TarBase	mRNA–microRNA target information	<a href="http://diana.cslab.ece.ntua.gr/tarbase">http://diana.cslab.ece.ntua.gr/tarbase</a>	Vergoulis et al., 2012
TargetScan		<a href="http://www.targetscan.org">http://www.targetscan.org</a>	Lewis et al., 2005
PicTar	Integrated resource of microRNA target information	<a href="http://pictar.mdc-berlin.de">http://pictar.mdc-berlin.de</a>	Krek et al., 2005
miRecords		<a href="http://mirecords.biolead.org">http://mirecords.biolead.org</a>	Xiao et al., 2009
miRGen		<a href="http://diana.cslab.ece.ntua.gr/mirgen">http://diana.cslab.ece.ntua.gr/mirgen</a>	Alexiou et al., 2010
TransMir		<a href="http://202.38.126.151/hmdd/mirna/TF">http://202.38.126.151/hmdd/mirna/TF</a>	Wang et al., 2010a
PutMir		<a href="http://www.isical.ac.in/~bioinfo_miu/TF-miRNA.php">http://www.isical.ac.in/~bioinfo_miu/TF-miRNA.php</a>	Bandyopadhyay & Bhattacharyya, 2010
MiRandola		<a href="http://atlas.dmi.unict.it/mirandola">http://atlas.dmi.unict.it/mirandola</a>	Russo et al., 2012
IntegromeDB	Integrated signaling network resources	<a href="http://integromedb.org">http://integromedb.org</a>	Baitaluk et al., 2012
SignalLink 2.0		<a href="http://signalink.org">http://signalink.org</a>	Korcsmáros et al., 2010; Fazekas et al., 2013
TranscriptomeBrowser 3.0		<a href="http://tagc.univ-mrs.fr/tbrowser">http://tagc.univ-mrs.fr/tbrowser</a>	Lepoivre et al., 2012

list of several other pathway databases can be found at PathGuide (<http://pathguide.org>; Bader et al., 2006). A compendium of human transcription factors have been collected and analyzed by Vaquerizas et al. (2009). Experimentally validated microRNA–mRNA interactions are available from TarBase (Vergoulis et al., 2012), while predicted interactions can be accessed at TargetScan and PicTar (Krek et al., 2005; Lewis et al., 2005). To examine the signaling network in a unified fashion, recently a few integrated resources, like IntegromeDB, TranscriptomeBrowser 3.0 and SignalLink 2.0, have been developed allowing the examination of all layers from signaling pathways to microRNAs through transcription factors (Korcsmáros et al., 2010; Baitaluk et al., 2012; Lepoivre et al., 2012; Fazekas et al., 2013).

There was considerable progress in defining algorithms to identify the downstream components of a signaling network affected by the inhibition of a specific protein or protein set. Such methods identify targets, which inhibit certain outputs of the signaling network, while leaving others intact, redirecting the signal flow in the network (Dasika et al., 2006; Ruths et al., 2006; Pawson & Linding, 2008). This recuperates the output specificity and input fidelity in a drug–target context, where output specificity corresponds to the minimization of side-effects, while input fidelity represents drug efficiency at the signaling network level.

The dynamics of signaling networks is regulated by changes in the abundance of their components, by complex formation, by macromolecular crowding, and subcellular localization (Lewitzky et al., 2012). The assessment of signaling network kinetics is helped by perturbation analysis, differential equation models, constrained fuzzy logic models and Boolean methods. In the latter, the activity of signaling components is represented by 0:1 states connected by directed and conditional edges as we summarized in Section 2.5.2 on network perturbations (Kauffman et al., 2003; Shmulevich & Kauffman, 2004; Berg et al., 2005; Antal et al., 2009; Farkas et al., 2011). There are several excellent methods for the analysis of Boolean networks.

- BooleanNet (<http://booleannet.googlecode.com>) is a versatile, publicly available software library to describe signaling network dynamics using the Boolean description (Albert et al., 2008).
- PATHLOGIC-S (<http://sourceforge.net/projects/pathlogic/files/PATHLOGIC-S>) offers a scalable Boolean framework for modeling cellular signaling (Fearnley & Nielsen, 2012).
- PathwayOracle (<http://old-bioinfo.cs.rice.edu/pathwayoracle>) is a fast simulation program of large signaling networks taking into

account their topology (Ruths et al., 2008a, 2008b).

- Changes in memory effects (i.e. specific decay times of gene products) greatly affected Boolean network behavior (Graudenzi et al., 2011a, 2011b).
- Boolean models can be trained by high-throughput data such as by phosphoproteomics (Videla et al., 2012).
- Recently the CellNOpt algorithm (<http://www.ebi.ac.uk/saezrodriguez/cno>) was introduced, which uses either Boolean logic or constrained fuzzy logic to generate and analyze cell-specific signaling networks (Terfve et al., 2012; Morris et al., 2013).
- The method of Chen et al. (2011) is able to identify sub-pathways and principal components of sub-pathways affected by a drug or disease.

The elucidation of signaling network dynamics can be greatly helped by quantitative phosphoproteomics (White, 2008). Reaction network analysis pointed out receptor-related events and not kinase-related events as rate-limiting factors in IL-12 signaling (Klinke & Finley, 2012). In an interesting study on signaling dynamics, Cheong et al. (2011) assessed the amount of information transduced by the TNF-related signaling network in the presence of cellular noise. They found that signaling bottlenecks may have a crucial influence on signaling capacity in the presence of noise. Negative feedback can reduce noise, increasing signal capacity in the short term (30 min after stimulation). However, negative feedback behaves as a double-edged sword, and also reduces the dynamic range of the signal input reducing the capacity in the longer run (4 h after stimulation). Perturbation modeling suggested that cancer-related proteins may have a larger than usual signaling capacity (Serra-Musach et al., 2012). Signal transmission capacity analysis has many unsolved questions. Currently, we do not understand the decision making limits of the vast majority of signaling systems. Information-loss, and information-integration all affect the information handling capacity, which is presumed to be minimally sufficient (Brennan et al., 2012). Future, network-related analysis of signaling dynamics faces the important task of finding an optimal ratio of large-scale signaling network topology and refined kinetic details to find answers to these questions.

#### 3.4.2. Drug targets in signaling networks

Understanding the structure and dynamics of signaling networks is used more and more often in drug discovery (Pawson & Linding,



2008). Drugs having a similar pharmacological profile reach similarly discrete positions in signaling networks (Fliri et al., 2009). In signaling networks of healthy cells a distinctive role was suggested for proteins in the junctions of signaling pathways. These proteins were termed 'critical nodes' by Taniguchi et al. (2006) as exemplified by the PI3 kinase, AKT and IRS isoforms in insulin signaling. Proteins forming a bridge between signaling modules (e.g., SHC, SRC and JAK2) have a track record as targets of drug action (Hwang et al., 2008; Gardino & Yaffe, 2011).

Studies on pathologically altered signaling networks can uncover possible drug targets, whose malfunction is involved in the etiology of the disease. For example, driver mutations of tumorigenesis affect a limited number of central pathways (Tomlinson et al., 1996; Ali & Sjoblom, 2009). Targeting of these specific pathways may prevent tumor growth. However, the development of aggressive tumor cells causes a systems-level change of the signaling network causing the appearance of angiogenic and metastatic capabilities, as well as the deregulation of cellular metabolism and avoidance of the immune system (Hanahan & Weinberg, 2000, 2011; Hornberg et al., 2006; Papatsoris et al., 2007). Changes of cross-talking (i.e. multi-pathway) proteins are key steps in disease-induced rewiring of the signaling network, e.g. transforming a 'death' signal into a 'survival' signal (Hanahan & Weinberg, 2000; Hornberg et al., 2006; Kim et al., 2007; Torkamani & Schork, 2009; Mimeault & Batra, 2010; Farkas et al., 2011). Multi-pathway proteins show a significant change in their expression level in hepatocellular carcinoma (Korcsmáros et al., 2010).

We have more than 500 types of post-translational modifications offering a rich repertoire for signaling (Lewitzky et al., 2012). Traditionally, among these, protein kinases are the most targeted proteins of the cellular signaling network (Pawson & Linding, 2008). However, the similarity of ATP-binding pockets poses a significant challenge in kinase targeting. Kinase domains and their target motifs (i.e., specific amino acid sequences in the substrate proteins) can be accessed in resources such as Phosphosite (Hornbeck et al., 2012), NetworKIN and NetPhorest (Linding et al., 2008; Miller et al., 2008). Kinase regulatory domain associations, kinase-associated scaffold proteins and multi-site docking proteins often direct subcellular localizations, and play a key role determining signaling kinetics and substrate specificity (Remenyi et al., 2006; Pálffy et al., 2012). Scaffolding proteins are flexible; and they, and multi-site docking proteins may increase network flexibility (by allowing integrative cross-talks; Lewitzky et al., 2012). However, our systems-level knowledge on these undirected protein-protein interactions is rather limited. Disruption of kinase-centered sub-interactomes and/or remodeling of kinase-centered protein complexes are focus areas of drug design (Brehme et al., 2009; Bandyopadhyay et al., 2010; Li et al., 2012a).

Protein phosphatases play a dominant role in determining the spatio-temporal behavior of protein phosphorylation systems (Herzog et al., 2012; Sacco et al., 2012; Nguyen et al., 2013). Despite their promising effect, only a few protein tyrosine phosphatases are currently used as therapeutic targets (Alonso et al., 2004). The development of phosphatase-related drugs is more complicated than that of kinase targeting drugs, since i.) the high-level of homology between phosphatase domains limits the development of selective compounds; ii.) contrary to kinases, phosphatase substrate specificity is achieved through docking of the phosphatase complex at a site distant from the dephosphorylated amino acid (Roy & Cyert, 2009; Shi, 2009); (iii) the targeted sequences are highly charged, and many of the interacting compounds are not hydrophobic enough to cross the membrane (Barr, 2010). Despite these difficulties, phosphatase-targeting holds great promise in signaling-related drug design.

In the last decade, microRNAs have been recognized as highly promising intervention points of the signaling network. Though the use of antisense nucleotides comes with great challenges in pharmacological availability, microRNA targeting affects mRNA clusters having a rather specific effect at the transcriptome level (Gambari et al.,

2011). Down- or up-regulation of microRNAs is implicated in more than 270 diseases according to the Human MicroRNA Disease Database (<http://202.38.126.151/hmdd/mirna/md>; Lu et al., 2008) including cardiovascular, neurodegenerative diseases, viral infections and various types of cancer (McDermott et al., 2011). Identification of new microRNA targets may be helped by the microRNA clusters associated with the same disease (Lu et al., 2008), or with expression modules (Bonnet et al., 2010).

MicroRNA targeting is typically a systems-level endeavor. Most microRNAs have a number of targets, and are in an intensive cross-talk with transcription factors (Lin et al., 2012) forming a highly cross-reacting, and cross-regulated network. The microRNA network has hierarchical layers, and hundreds of 'target hubs', each potentially subject to massive regulation by dozens of microRNAs (Shalgi et al., 2007). Cancer cells, as opposed to normal cells, have disjoint microRNA networks, where major hubs of normal cells are down-regulated, and cancer-specific novel microRNA hubs emerge (Volinia et al., 2010). MicroRNA-regulated drug targets were shown to preferentially interact with each-other, and tend to form hub-bottlenecks of the human interactome (Wang et al., 2011c). Unwanted, network-level side-effects of microRNA targeting may be predicted using databases, such as SIDER (<http://sideeffects.embl.de>; Kuhn et al., 2010), web-services, such as PathwayLinker (<http://PathwayLinker.org>; Farkas et al., 2012) or the other integrated resources listed in Table 7. Miravirsin, a locked nucleic acid-modified antisense oligonucleotide, targets the liver-expressed microRNA-122 and is in phase-II clinical trial for treatment of hepatitis C virus infection (Lindow & Kauppinen, 2012).

#### 3.4.3. Challenges of signaling network targeting

Signal transduction is highly context-specific: it depends on the gene expression patterns, mRNA stability, protein synthesis, and degradation conditions. Certain signaling modules (such as regulation of apoptosis) seem to share evolutionary traits with others, while other signaling proteins developed more independently. These variabilities necessitate a precise knowledge of the actual status of the signaling network in the disease condition and in the affected patient population (Cui et al., 2009; Davis et al., 2012; Hamp & Rost, 2012; Kirouac et al., 2012).

As we have shown in the preceding section, cross-talking (i.e., multi-pathway, bridge) proteins are in a critical position of the signaling network providing a very efficient set of potential drug targets (Korcsmáros et al., 2007, 2010; Hwang et al., 2008; Kumar et al., 2008; Spiró et al., 2008). However, numerous drug developmental failures were caused by undiscovered or underestimated cross-talk effects (Rajasethupathy et al., 2005; Jia et al., 2009). Cross-talking proteins may have opposite roles in healthy and diseased (such as in malignant) cells. Moreover, targeting of cross-talking proteins may significantly affect the systems-level stability (robustness) of healthy or diseased cells (Kitano, 2004b, 2007). Targeting proteins in negative feedback loops may suppress the inhibitory effect of the feedback loop, and thereby activate the targeted pathway (Sergina et al., 2007). Feedback loops are not always direct, and can exist at multiple levels of a pathway. In conclusion, targeting multi-pathway and feedback loop proteins requires a particularly detailed knowledge of signaling network responses (Berger & Iyengar, 2009; Barabási et al., 2011).

Systems-level properties are also needed to assess the development of drug resistance and drug toxicity.

- Development of drug resistance is often a result of a systems-level response of signaling networks involving mutation of key signaling proteins (such as multi-pathway proteins), or of the activation of alternative pathways due to system-robustness (Kitano, 2004a; Logue & Morrison, 2012). As a specific form of drug resistance, many anticancer drugs induce stress response/survival pathways directly, or indirectly, by producing a stressful environment (Tomida

& Tsuruo, 1999; Chen et al., 2006b; Tiligada, 2006). Thus, systems-level approaches that combine anti-tumor drugs and stress response targeting may increase therapeutic efficiency (Rocha et al., 2011; Tentner et al., 2012).

- Hepatotoxicity is a major cause of drug development failures in the pre-clinical, clinical and post-approval stages (Kaplowitz, 2001). Hepatic cytotoxicity responses are regulated by a multi-pathway signaling network balance of intertwined pro-survival (AKT) and pro-death (MAPK) pathways. Importantly, therapeutic modulation of cross-talks between these pathways as well as specific pathway inhibitors could antagonize drug-induced hepatotoxicity (Cosgrove et al., 2010).

We will address network-based assessment of drug toxicity and drug resistance in Sections 4.3.3 and 4.3.6 in more detail.

### 3.5. Genetic interaction and chromatin networks

In this section we will describe the drug-related aspects of genetic interaction networks. Genetic interaction networks are related to gene regulatory networks. However, here gene–gene interactions are often indirect. Chromatin networks encode 3D interactions between distant DNA-segments of the chromatin structure, and may be regarded as a specific representation of genetic interaction networks. While genetic interaction networks already helped drug design, chromatin interaction networks are recent developments holding a great promise for future studies.

#### 3.5.1. Definition and structure of genetic interaction networks

The most stringent (and traditional) description of a genetic interaction comes from comparing the phenotypes of the individual single mutants with the phenotype of the double mutant. We can distinguish between negative and positive genetic interactions: if the fitness of the double mutants is worse than the additive effect of the two single mutants, then the genes have a negative interaction. Conversely, if the fitness of the double mutants is better than expected, the two genes interact positively. A severe type of negative interactions is synthetic lethality, when the two single mutants are viable, but their double mutant becomes lethal. Genes of negative (i.e. aggravating) interactions may operate in parallel processes, while those of positive (i.e., alleviating or epistatic) interactions may function in the same process (Guarente, 1993; Hartman et al., 2001; Dixon et al., 2009). The complexity of the genetic interaction network is illustrated well by compensatory perturbations, where a debilitating effect can be compensated by another inhibitory effect (Motter, 2010; Cornelius et al., 2011).

Most comprehensive genome-wide studies were performed in inbred model systems, such as yeast and worm, as well as in isogenic populations of cultured cells derived from fruit flies and mammals. It is plausible that many genetic interactions identified in these unicellular organisms can be relevant for all other eukaryotes (Tong et al., 2004; Dixon et al., 2008, 2009; Roguev et al., 2008; Costanzo et al., 2010). However, comparison between orthologous genes of yeast and worm found less than 5% of synthetic lethal genetic interactions to be conserved (Byrne et al., 2007). Furthermore, most of the human disease genes are metazoan-specific. Despite the widespread specificity, there are some genetic interactions (such as those of DNA repair enzymes, which are commonly mutated in cancer), which are conserved from yeast to humans (McManus et al., 2009).

Besides mutational studies, system-wide assessments of output signals, such as transcriptomes, allowed the phenotype analysis of thousands of perturbations inferring a genetic interaction network. We reviewed the reverse engineering methods allowing network inference in Section 2.2.3. It appears that no single inference method performs optimally across all datasets. Therefore, the integration of predictions from multiple inference methods seems to give the best

results (Marbach et al., 2012). The genetic interaction network obtained by direct mutational studies, or by reverse engineering methods exhibited dense local neighborhoods, while highly correlated profiles delineated specific pathways defining gene function, and were used for pre-clinical drug prioritization (Xiong et al., 2010). Recently, an algorithm called HotNet was introduced to identify genetic interaction network clusters (<http://compbio.cs.brown.edu/software.html>; Vandin et al., 2012).

Mapping of genetic interactions to protein–protein interaction or to signaling networks may uncover the underlying mechanisms. Consequently, we can define ‘between-pathway’, ‘within-pathway’ and ‘indirect’ types of genetic interactions. With this approach, ~40% of the yeast synthetic lethal genetic interactions were mapped to physical pathway models identifying 360 between-pathway and 91 within-pathway models (Kelley & Ideker, 2005). Synthetic lethal gene pairs were found mostly close to each other (often within the same modules), while rescuing genes were often in alternative pathways and/or modules (Hintze & Adami, 2008). Combinations of genetic interactions with mRNA expression patterns, interactome data, gene–drug interactions, or with chemical compound similarity measures (for details see the compendium of Table 5) offered a great help in the identification of drug targets and drug-affected genes (Parsons et al., 2006; Hansen et al., 2009; Gosline et al., 2012).

Measurement of time-series of genome-wide mRNA expression patterns after drug treatment of 95 genotyped yeast strains led to the identification of novel genetic interaction network relationships including novel feedback loops and transcription factor binding sites (Yeung et al., 2011). BioLayout Express gives an integrative network visualization and analysis of gene expression data (<http://www.biolayout.org>; Freeman et al., 2007; Theocharidis et al., 2009). SteinerNet provides integrated transcriptional, proteomic and interactome data to assess regulatory networks (<http://fraenkel.mit.edu/steinernet/>; Tuncbag et al., 2012). Genetic interaction networks may also be defined in a more general manner, where any types of interactions, such as correlated expression levels, interacting protein products, or co-participation in a disease etiology or drug action, may form an edge between two genes serving as nodes of the network (Schadt et al., 2009). Genome-wide association studies (GWAS) identified single-nucleotide polymorphism (SNP) derived gene–gene association networks revealing novel between-pathway models (Cowper-Salari et al., 2011; Fang et al., 2011; Hu et al., 2011; Li et al., 2012b).

#### 3.5.2. Chromatin networks and network epigenomics

An underlying molecular mechanism establishing genetic interaction networks is the network of long-range interactions of the 3D chromatin structure. Recent methodologies based on proximity ligation with next generation sequencing (abbreviated as Hi-C or ChIA-PET) enabled the construction of a functionally associated, long-range contact network of the human chromatin structure. This chromatin network contains functional modules, and has a rich club of hub–hub interactions (Fullwood et al., 2009; Lieberman-Aiden et al., 2009; Dixon et al., 2012; Li et al., 2012c; Sandhu et al., 2012).

The chromatin network determines cancer-associated chromosomal alterations (Fudenberg et al., 2011). Moreover, the chromatin network configuration was shown to be grossly altered by the overexpression of ERG, an oncogenic transcription factor activated primarily in prostate cancers (Rickman et al., 2012). The structure of the chromatin network is largely determined by inheritable epigenetic factors, such as histone posttranslational modifications, DNA silencers and nascent RNA scaffolds (Schreiber & Bernstein, 2002; Moazed, 2011; Pujadas & Feinberg, 2012). Chromatin networks are an exciting and fast-developing area of network-studies, which will provide promising tools to predict drug–drug interactions, drug side-effects and system-wide effects of anti-cancer and other drugs inducing chromatin reprogramming.

### 3.5.3. Genetic interaction networks as models for drug discovery

The genetic interaction network of yeast can be used as a system for rational ranking of potential new antifungal targets; it may also shed light on human drug mechanisms of action, since several human drugs specifically inhibit the orthologous proteins in yeast (Hartwell et al., 1997; Cardenas et al., 1999; Hughes, 2002). The identification of 16 genes, whose inactivation suppressed the defects in the retinoblastoma tumor suppressor pathway in another widely used model system, *C. elegans*, could point out potential targets for pharmaceutical intervention or prevention of human retinoblastoma-linked tumors (Lee et al., 2008b). Extending this methodology, McGary et al. (2010) defined orthologous phenotypes, or 'phenologs', which can be regarded as evolutionarily conserved outputs that arise from the disruption of a set of genes. The phenolog approach identified non-obvious equivalences between mutant phenotypes in different species, establishing a yeast model for angiogenesis defects, a worm model for breast cancer, mouse models of autism, and a plant model for the neural crest defects associated with the Waardenburg syndrome (McGary et al., 2010).

Many pharmacologically interesting genes, such as nuclear hormone receptors and GPCRs occur in large families containing paralogs, i.e. duplicated homologous genes. Though model organisms can significantly help us to understand how human genes interact with each other, it is important to keep in mind that paralogs often do not have the same function. This problem can be circumvented by targeting paralog-sets, which makes the identification of paralogs and their functions a key point in multi-target drug design (Searls, 2003).

Wang et al. (2012c) gave an interesting example for the use of genetic interaction networks in the assessment of the effects of drug combinations. They showed that drug combinations have significantly shorter effect radius than random combinations. Drug combinations against diseases affecting the cardiovascular and nervous systems have a more concentrated effect radius than immuno-modulatory or anti-cancer agents.

### 3.6. Metabolic networks

In this section we will describe metabolic networks, i.e. networks of major metabolites connected by enzyme reactions, which transform them to each other. Metabolic networks are the biochemically constrained subsets of the chemical reaction networks we summarized in Section 3.1.2. After the description of the structure and properties of metabolic networks we will summarize their use in drug targeting with special reference to the identification of essential reactions as potential drug targets in infectious diseases and in cancer.

#### 3.6.1. Definition and structure of metabolic networks

In a metabolic network, each node represents a metabolite. Two nodes are connected, if there is a biochemical reaction that can transform one into the other. Edges of metabolic networks represent both reactions and the enzymes that catalyze them. (We note that metabolic networks may also have another projection, where nodes are the enzymes and edges are the metabolites connecting them, but this projection is seldom used, since it is less relevant to biological processes.) Metabolic processes may be represented as hypergraphs, where edges connect multiple nodes. An edge may correspond to multiple reactions both in the forward direction and in the opposite direction. Moreover, some reactions occur spontaneously, and therefore have no associated enzymes. Most metabolites are fairly general, but the biochemical reaction structure connecting them is often rather special to the given organism (Guimera et al., 2007b; Ma & Goryanin, 2008; Chavali et al., 2012).

Reconstruction of metabolic networks became a highly integrative process, which applies genome sequences, enzyme databases, and specifies the network using transcriptome and proteome data (Kell, 2006; Ma & Goryanin, 2008). In the last decade several metabolic

networks, such as those of *Escherichia coli*, yeast and humans have been assembled. Moreover, recently bacterium-, strain-, tissue- and disease-specific metabolic networks were reconstituted. However, it should be kept in mind that metabolic network data are often still incomplete, and often reflect optimal growth conditions (Edwards & Palsson, 2000a; Forster et al., 2003; Duarte et al., 2007; Ma et al., 2007; Shlomi et al., 2008, 2009; Folger et al., 2011; Holme, 2011; Chavali et al., 2012; Szalay-Bekő et al., 2012).

Metabolic networks have a small-world character, possess hubs, and display a hierarchical bow-tie structure similar to other directed networks, such as the world-wide-web. Metabolic networks have a hierarchical modular structure (Jeong et al., 2001; Wagner & Fell, 2001; Ravasz et al., 2002; Ma & Zeng, 2003; Ma et al., 2004; Guimera & Amaral, 2005; Zhao et al., 2006). Correlated reaction sets (Co-sets) are representations of metabolic network modules encoding reaction-groups with linked fluxes. Hard-coupled reaction sets (HCR-sets) are those subgroups of Co-sets, where consumption/production rates of participating metabolites are 1:1. Since all reactions of a HCR-set changes, if any of its reactions is targeted, HCR-sets help in prioritizing potential drug target lists (Papin et al., 2004; Jamshidi & Palsson, 2007; Xi et al., 2011). Metabolic networks have a core and a periphery (Almaas et al., 2004, 2005; Guimera & Amaral, 2005; Guimera et al., 2007b). Core and periphery may also be discriminated in the non-topological sense that genes and gene pairs of the 'core' are essential under many environmental conditions, while those of the 'periphery' are needed under some environmental conditions (Papp et al., 2004; Pál et al., 2006; Harrison et al., 2007).

Metabolic control analysis (MCA) is good for smaller networks where kinetic parameters are known, while flux balance analysis (FBA), flux-variability analysis (FVA) and elementary flux mode analysis are very useful methods to characterize systems-level metabolic responses (Fell, 1998; Cascante et al., 2002; Klamt & Gilles, 2004; Chavali et al., 2012). Resendis-Antonio (2009) integrated high throughput metabolome data describing transient perturbations in a red blood cell metabolic network model. This approach may be applicable for the modeling and metabolome-wide understanding of drug-induced metabolic changes (Fan et al., 2012). In Table 8 we list resources to define and analyze metabolic networks.

#### 3.6.2. Essential enzymes of metabolic networks as drug targets in infectious diseases and in cancer

Metabolic networks help in the identification essential proteins. This requires a systems-level approach, since an essential metabolite might be produced by several pathways (Palumbo et al., 2007). As an example of metabolic robustness, the early work of Edwards and Palsson (2000b) showed that the flux of even the tricarboxylic acid cycle can be reduced to 19% of its optimal value without significantly influencing the growth of *E. coli*. When designing a drug against a metabolic network of an infectious organism or against cancer cells, many parameters should be kept in mind. We list a few of them here.

- Network topology analysis is not enough to predict essential enzymes, since it does not indicate whether the topologically important enzymes are active under specific conditions. Moreover, metabolic fluxes are determined by gene expression levels (representative to the affected tissue, or cell status) and by signaling- or interaction-related activation/inhibition of pathway enzymes. Genes that are not identified correctly as essential genes are usually connected to fewer reactions and to less over-coupled metabolites, and/or their associated reactions are not carrying flux in the given condition (Becker & Palsson, 2008; Chavali et al., 2012; Kim et al., 2012).
- Infectious organisms often take advantage of the metabolism of the host requiring the analysis of integrated parasite–host metabolic networks (Fatumo et al., 2011).
- Essentiality is not a yes/no variable: essentiality of a given reaction

**Table 8**  
Metabolic network resources.

Name	Content	Website	References
KEGG	Metabolic pathway resource	<a href="http://kegg.jp">http://kegg.jp</a>	Kanehisa et al., 2012
MetaCyc		<a href="http://metacyc.org">http://metacyc.org</a>	Caspi et al., 2012
HumanCyc		<a href="http://humancyc.org">http://humancyc.org</a>	Romero et al., 2005
SMPDB	Small Molecule (e.g., drug) Pathway Database	<a href="http://smpdb.ca">http://smpdb.ca</a>	Frolkis et al., 2010
HMDB	Human Metabolome Database	<a href="http://hmdb.ca">http://hmdb.ca</a>	Wishart et al., 2009
BRENDA	Comprehensive enzyme data resource	<a href="http://brenda-enzymes.info">http://brenda-enzymes.info</a>	Scheer et al., 2011
YeastNet	Yeast metabolic network	<a href="http://comp-sys-bio.org/yeastnet">http://comp-sys-bio.org/yeastnet</a>	Herrgard et al., 2008
iMAT another metabolic network construction and analysis tools	Several metabolic network construction and analysis tools	<a href="http://www.cs.technion.ac.il/~tomersh/methods.html">http://www.cs.technion.ac.il/~tomersh/methods.html</a>	Shlomi et al., 2008; Zur et al., 2010
ModelSEED and its Cytoscape plug-in, CytoSEED	Genome level metabolic network reconstruction and analysis	<a href="https://github.com/ModelSEED">https://github.com/ModelSEED</a> , <a href="http://sourceforge.net/projects/cytoseed">http://sourceforge.net/projects/cytoseed</a>	Henry et al., 2010; DeJongh et al., 2012
Markov Chain Monte Carlo modeling	Bayesian inference method to uncover perturbation sites in metabolic pathways	<a href="ftp://anonymous@dbkweb.mib.man.ac.uk/pub/Bioinformatics_BJ.zip">ftp://anonymous@dbkweb.mib.man.ac.uk/pub/Bioinformatics_BJ.zip</a>	Jayawardhana et al., 2008
PyNetMet	Python library of tools for the analysis of metabolic models and networks	<a href="http://pypi.python.org/pypi/PyNetMet">http://pypi.python.org/pypi/PyNetMet</a>	Gamermann et al., 2012

depends on the environment of the infectious organism, or cancer cells. Therefore, metabolic network-based drug-design should incorporate environment interactions and stressor effects (Guimera et al., 2007b; Jamshidi & Palsson, 2007; Ma & Goryanin, 2008; Kim et al., 2012).

- A promising current trend, metabolic interactions of bacterial communities, such as the gut microbiome, are also important factors to consider (Chavali et al., 2012; Kim et al., 2012).
- Drug targets against infectious organisms or against cancer should be specific for the target itself or for its drug binding site, or for its network-related consequences of targeting (Guimera et al., 2007b; Ma & Goryanin, 2008; Chavali et al., 2012). This highlights the importance of comparing metabolic network pairs.
- Finally, network-analysis offers a great help to predict side-effects (Guimera et al., 2007b). We will detail network-methods of side-effect prediction in Section 4.3.5.

Enzymes catalyzing a single chemical reaction on one particular substrate are frequently essential (Nam et al., 2012). Through the analysis of metabolic network structure, choke points were identified as reactions that either uniquely produce or consume a certain metabolite. Efficient inhibition of choke points may cause either a lethal deficiency, or toxic accumulation of metabolites in infectious organisms (Yeh et al., 2004; Singh et al., 2007). Later choke point analysis was combined with load point analysis (identification of nodes with a high ratio of  $k$ -shortest paths to the number of nearest neighbor edges providing many alternative metabolic pathways) and with comparison of the metabolic networks of pathogenic and related non-pathogenic strains. Such methods can test multiple knock-outs on a high throughput manner predicting effective drug combinations (Fatumo et al., 2009, 2011; Perumal et al., 2009).

Guimera et al. (2007b) developed a network modularity-based method for target selection in metabolic networks. They systematically analyzed the effect of removing edges from the metabolic networks of *E. coli* and *Helicobacter pylori* quantifying the effect by the difference in growth rate. In both bacteria, essential reactions (edges) mostly involved satellite connector metabolites that participate in a small number of biochemical reactions, and serve as bridges between several different modules (Guimera et al., 2007b).

Essential and non-essential genes propagate their deletion effects via distinct routes. Flux selectivity of a deletion of a metabolic reaction was used to design the appropriate type and concentration of the inhibitor (Gerber et al., 2008). Recently, several iterative methods have been constructed, sequentially identifying a set of enzymes whose inhibition can produce the expected inhibition of targets with reduced side-effects in human and *E. coli* metabolic networks (Lemke et al., 2004; Sridhar

et al., 2007, 2008; Song et al., 2009). Ma et al. (2012b) assembled 'damage lists' of reactions affected by deleting other reactions using flux-balance analysis. They showed that the knockout of an essential gene mainly affects other essential genes, whereas the knockout of a non-essential gene only interrupts other non-essential genes. Genes sharing the same 'damage list' tend to have the same level of essentiality.

A subset of genes and gene pairs may be essential under various environmental conditions, while most genes are essential only under a certain environmental condition. In yeast environmental condition-specificity accounts for 37–68% of dispensable genes, while compensation by duplication and network-flux reorganization is responsible for 15–28 and 4–17% of yeast dispensable genes, respectively (Papp et al., 2004; Blank et al., 2005). Almaas et al. (2005) suggested the use of metabolic network cores to identify drug targets. Barve et al. (2012) identified a set of 124 superessential reactions required in all metabolic networks under all conditions. They also assigned a superessentiality index for thousands of reactions. Superessentiality of the 37 reactions catalyzed by enzymes having a very low homology to human genes (Becker et al., 2006; Aditya Barve & Andreas Wagner, personal communication) can provide substantial help in drug target selection, since the index is not highly sensitive to the chemical environment of the pathogen.

An interesting approach to narrow metabolic networks to essential components is to identify essential metabolites. As examples of this process, in two pathogenic organisms a total of 221 or 765 metabolites were narrowed to 9 or 5 essential metabolites, respectively, after the removal of the currency metabolites, i.e. those present in the human metabolic network and those participating in reactions catalyzed by enzymes having human homologs. Enzymes that catalyze reactions involved in the production or consumption of these essential metabolites may be considered as drug targets. Moreover, structural analogs of essential metabolites may be considered as drug candidates for experimental evaluation (Kim et al., 2010, 2011a, 2011b).

Using a comparison of metabolic networks, Shen et al. (2010) provided a blueprint of strain-specific drug selection combining metabolic network analysis with atomistic level modeling. They deduced common antibiotics against *E. coli* and *Staphylococcus aureus*, and ranked more than a million small molecules identifying potential antimicrobial scaffolds against the identified target enzymes.

The analysis of disease-specific metabolic networks is a key step to find 'differentially essential' genes (Murabito et al., 2011). Analysis of cancer-specific human metabolic networks led Folger et al. (2011) to predict 52 cytostatic drug targets, of which 40% were targeted by known anticancer drugs, and the rest were new target-candidates. Their method also predicted combinations of synthetic lethal drug targets and potentially selective treatments for specific cancers. We will describe network-related anti-infection and anti-cancer strategies in more detail in Sections 5.1 and 5.2.

### 3.6.3. Metabolic network targets in human diseases

Many human diseases cause a metabolic deficiency rather than overproduction making the recovery of a specific metabolic reaction a widely used drug-development strategy (Ma & Goryanin, 2008). Systems-level assessment may lead to the development of successful combined-therapies, such as the combination of Niacin, an inhibitor of cholesterol transportation, with Lovastatin, an inhibitor of the cholesterol synthesis pathway to reduce blood cholesterol level (Gupta & Ito, 2002). As a very interesting approach, a flux-balance analysis model was developed to predict compensatory deletions (also called as synthetically viable gene pairs, or synthetic rescues), where a debilitating effect can be compensated by another inhibitory effect (Motter et al., 2008; Motter, 2010; Cornelius et al., 2011). Since inhibition is often a pharmacologically more feasible intervention than activation, this approach opens novel possibilities for drug design to restore disease-induced malfunctions. The approach of Jamshidi and Palsson (2008) to describe temporal changes of metabolic networks is an example of what seems to be a very promising future direction to study the process of disease progression, and to design disease-stage specific drug treatment protocols.

## 4. Areas of drug design: an assessment of network-related added-value

In this section we will highlight the added-value of network related methods in major steps of the drug design process. Fig. 18 illustrates various stages of drug development starting with target identification, followed by hit finding, lead selection and optimization including various methods of chemoinformatics, drug efficiency optimization, ADMET (drug absorption, distribution, metabolism, excretion and toxicity) studies, as well as optimization of drug–drug interactions, side-effects and resistance. Table 9 summarizes a few major data-sources and web-services, which can be used efficiently in network-related drug design studies.

### 4.1. Drug target prioritization, identification and validation

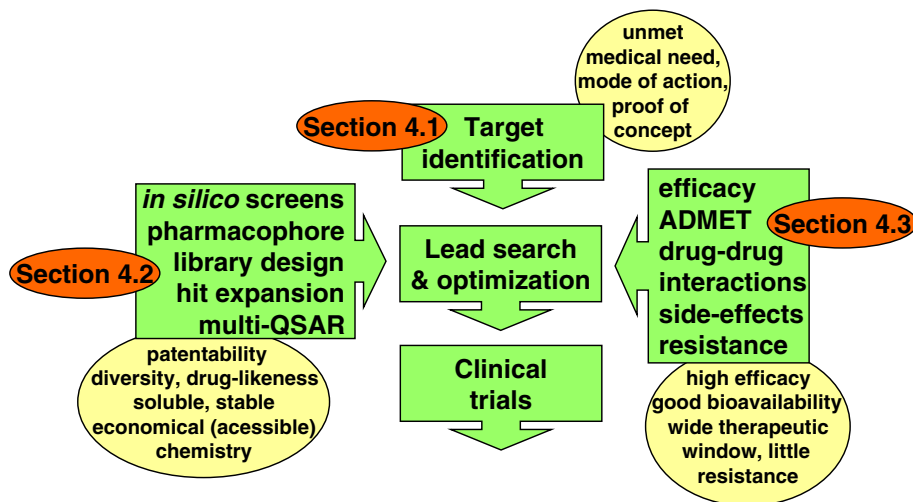
Network-based drug target prioritization and identification are essentially a top–down approach, where system-wide effects of putative targets are modeled to help in the identification of novel network drug targets. These network drug targets are non-obvious from a traditional magic-bullet type analysis aiming to find the single most important

cause of a given disease. Network node-based drug target prediction may highlight non-obvious hits, and edge-targeting may make these hits even more specific. Drug target networks allow us to see the system-wide target landscape and, combined with other network methods, help drug repositioning. Multi-target drug design needs the integration of drug effects at the system level. The new concept of allo-network drugs may identify non-obvious drug targets, which specifically influence the major targets causing fewer side-effects than direct targeting. Finally, treating the whole cellular network (or its segment) as a drug target, gives a conceptual synthesis of network description and analysis in drug design.

#### 4.1.1. Two strategies of network-based drug targeting: the central hit and the network influence strategies

Here we propose that our current knowledge discriminates two network-based drug identification strategies. We name the first strategy the central hit strategy. This strategy is useful to find drug target candidates in anti-infectious and in anti-cancer therapies. The second strategy is named network influence strategy. This strategy uses systems-level knowledge to find drug target candidates in therapies of polygenic, complex diseases (Fig. 19). In the central hit strategy our aim is to damage the network integrity of the infectious agent or of the malignant cell in a selective manner. For this, detailed knowledge of the structural differences of host/parasite or healthy/malignant networks can help. In the network influence strategy we would like to shift back the malfunctioning network to its normal state. For this, an understanding of network dynamics both in healthy and diseased states is required. Knowledge of the existing drug targets of the particular disease also helps.

System destruction of the central hit strategy finds hubs and central nodes of various networks (the latter are called load-points in metabolic networks), and uses the methods listed in Section 3.6.2 to find essential enzymes of metabolic networks (Jeong et al., 2001; Chin & Samanta, 2003; Agoston et al., 2005; Estrada, 2006; Guimera et al., 2007b; Yu et al., 2007b; Fatumo et al., 2009, 2011; Missiuro et al., 2009; Perumal et al., 2009; Li et al., 2011a). In addition, choke points of metabolic networks, i.e. proteins uniquely producing or consuming a certain metabolite are also excellent targets in anti-infectious therapies (Yeh et al., 2004; Singh et al., 2007). In the case of directed, hierarchical networks nodes at the top of the hierarchy should be attacked by the central hit strategy. Liu et al. (2012) suggested the random upstream attack to find high position nodes in hierarchical networks. Recent work on connections of essential



**Fig. 18.** The drug development process. Green boxes illustrate the major stages of the drug development process starting with target identification, followed by hit finding, hit confirmation and hit expansion leading to lead selection/optimization and concluded by clinical trials. Lead search and lead optimization are helped by various methods of chemoinformatics (left side), drug efficiency optimization, ADMET (drug absorption, distribution, metabolism, excretion and toxicity) studies, as well as optimization of drug–drug interactions, side-effects and resistance (right side). Yellow ellipses summarize a few major optimization criteria, while orange ellipses refer to the subsections of Section 4 discussing the given drug development stage. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

**Table 9**  
Drug-design related resources.

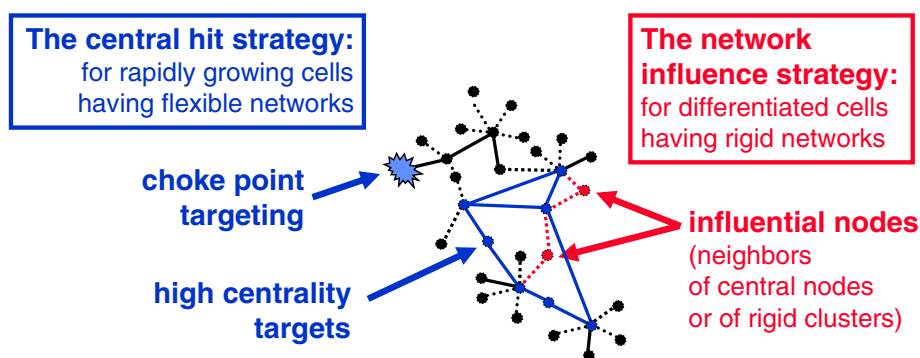
Name	Content	Website	References
<i>Section 4.1. Drug target prioritization, identification and validation</i>			
Pubchem	Repository of small molecule biological activities	<a href="http://pubchem.ncbi.nlm.nih.gov">http://pubchem.ncbi.nlm.nih.gov</a>	Wang et al., 2010b
chEMBLdb	Chemical properties and biological activities of drug-like molecules	<a href="https://www.ebi.ac.uk/chembl/db">https://www.ebi.ac.uk/chembl/db</a>	Gaulton et al., 2012
DailyMed	Drug package insert texts	<a href="http://dailymed.nlm.nih.gov">http://dailymed.nlm.nih.gov</a>	de Leon, 2011
DrugBank	Integrated drug and drug target information resource	<a href="http://drugbank.ca">http://drugbank.ca</a>	Knox et al., 2011
PharmGKB		<a href="http://pharmgkb.org">http://pharmgkb.org</a>	Thorn et al., 2010
Therapeutic Target Database		<a href="http://bid.nus.edu.sg/group/cjttd/TTD_HOME.asp">http://bid.nus.edu.sg/group/cjttd/TTD_HOME.asp</a>	Zhu et al., 2012b
MATADOR		<a href="http://matador.embl.de">http://matador.embl.de</a>	Günther et al., 2008
Supertarget		<a href="http://insilico.charite.de/supertarget">http://insilico.charite.de/supertarget</a>	Hecker et al., 2012
KEGG DRUG		<a href="http://genome.jp/kegg/drug">http://genome.jp/kegg/drug</a>	Kanehisa et al., 2012
TDR	Drug targets of neglected tropical diseases	<a href="http://tdrtargets.org">http://tdrtargets.org</a>	Agüero et al., 2008
PDTD–Potential Drug Target Database	Information on drug targets	<a href="http://dddc.ac.cn/pdtd">http://dddc.ac.cn/pdtd</a>	Gao et al., 2008
DTome	Drug–target network construction tool	<a href="http://bioinfo.mc.vanderbilt.edu/DTome">http://bioinfo.mc.vanderbilt.edu/DTome</a>	Sun et al., 2012
My-DTome	Myocardial infarction-related drug target interactome	<a href="http://my-dtome.lu">http://my-dtome.lu</a>	Azuaje et al., 2011
PROMISCUOUS	Interactome-based database for drug-repurposing	<a href="http://bioinformatics.charite.de/promiscuous">http://bioinformatics.charite.de/promiscuous</a>	von Eichborn et al., 2011
MANTRA	mRNA expression profile-based server for drug-repurposing	<a href="http://mantra.tigem.it">http://mantra.tigem.it</a>	Iorio et al., 2010
CDA	Combinatorial drug assembler and drug repositioner (mRNA expression profiles, signaling networks)	<a href="http://cda.i-pharm.org">http://cda.i-pharm.org</a>	Lee et al., 2012c
<i>Section 4.2.1. Hit finding for ligand binding sites</i>			
STITCH 3	Integrated network resource of chemical–protein interactions	<a href="http://stitch.embl.de">http://stitch.embl.de</a>	Kuhn et al., 2012
CARLSBAD	Cytoscape plug-in connecting common chemical patterns to biological targets via small molecules	<a href="http://carlsbad.health.unm.edu">http://carlsbad.health.unm.edu</a>	
BindingDB	Binding affinity data for almost a million protein–ligand pairs	<a href="http://bindingdb.org">http://bindingdb.org</a>	Liu et al., 2007a
BioDrugScreen	Structural protein ligand interactome + scoring system	<a href="http://biodrugscreen.org">http://biodrugscreen.org</a>	Li et al., 2010c
CREDO	a protein–ligand interaction database including a wide range of structural information	<a href="http://www-cryst.bioc.cam.ac.uk/databases/credo">http://www-cryst.bioc.cam.ac.uk/databases/credo</a>	Schreyer & Blundell, 2009
<i>Section 4.2.2. Hit finding for protein–protein interaction hot spots</i>			
TIMBAL	A curated database of ligands inhibiting protein–protein interactions	<a href="http://www-cryst.bioc.cam.ac.uk/databases/timbal">http://www-cryst.bioc.cam.ac.uk/databases/timbal</a>	Higueruelo et al., 2009
Dr. PIAS–Druggable Protein–protein Interaction Assessment System	Machine learning-based web-server to judge, if a protein–protein interaction is druggable	<a href="http://drpias.net">http://drpias.net</a>	Sugaya & Furuya, 2011; Sugaya et al., 2012
<i>Section 4.3. Drug efficiency, ADMET, drug–drug interactions, side-effects and resistance</i>			
Supertarget	Drug metabolism information	<a href="http://insilico.charite.de/supertarget">http://insilico.charite.de/supertarget</a>	Hecker et al., 2012
KEGG DRUG		<a href="http://genome.jp/kegg/drug">http://genome.jp/kegg/drug</a>	Kanehisa et al., 2012
ACToR	Integrated toxicity resource	<a href="http://actor.epa.gov">http://actor.epa.gov</a>	Judson et al., 2012
DITOP	Drug-induced toxicity related protein database	<a href="http://bioinf.xmu.edu.cn:8080/databases/DITOP/index.html">http://bioinf.xmu.edu.cn:8080/databases/DITOP/index.html</a>	Zhang et al., 2007
DCDB	Drug combination database	<a href="http://www.cls.zju.edu.cn/dcdb">http://www.cls.zju.edu.cn/dcdb</a>	Liu et al., 2010b
DTome	Adverse drug–drug interactions	<a href="http://bioinfo.mc.vanderbilt.edu/DTome">http://bioinfo.mc.vanderbilt.edu/DTome</a>	Sun et al., 2012
KEGG DRUG		<a href="http://genome.jp/kegg/drug">http://genome.jp/kegg/drug</a>	Kanehisa et al., 2012
SIDER	Drug side-effect resource	<a href="http://sideeffects.embl.de">http://sideeffects.embl.de</a>	Kuhn et al., 2010
DRAR–CPI	Drug-binding structural similarity based server for adverse drug reaction and drug repositioning	<a href="http://cpi.bio-x.cn/drar">http://cpi.bio-x.cn/drar</a>	Luo et al., 2011
DvD	An R/Cytoscape plug-in assessing system-wide gene expression data to predict drug side effects and drug repositioning	<a href="http://www.ebi.ac.uk/saezrodriguez/DVD">http://www.ebi.ac.uk/saezrodriguez/DVD</a>	Pacini et al., 2013
NPC	Approved and experimental drugs useful for drug-repositioning	<a href="http://tripod.nih.gov/npc">http://tripod.nih.gov/npc</a>	Huang et al., 2011b
SePreSA	Binding pocket polymorphism-based serious adverse drug reaction predictor	<a href="http://sepresa.bio-x.cn">http://sepresa.bio-x.cn</a>	Yang et al., 2009a; Yang et al., 2009b
<i>Section 4.3. Drug efficiency, ADMET, drug–drug interactions, side-effects and resistance (continuation)</i>			
SADR–Gengle	PubMed record text mining-based data on 6 serious adverse drug reactions	<a href="http://gengle.bio-x.cn/SADR">http://gengle.bio-x.cn/SADR</a>	Yang et al., 2009c

reactions and on supressential reactions (where the latter are needed in all organisms) suggests that essential reactions form a core of metabolic networks (Barve et al., 2012; Ma et al., 2012b). Cytostatic drug targets have also been identified through analysis of cancer-specific human metabolic networks (Folger et al., 2011). Recent anti-cancer strategies mostly use the cancer-specific targeting of signaling networks as we will describe in detail in Section 5.2.

At the protein structure level the central hit strategy may mostly target active sites. Targeting allosteric regulatory sites may be shared by both strategies. These, cavity-like binding sites are easier to target than the flat, ‘hot-spot’-type protein–protein interfaces mostly involved in the network influence strategy (Keskin et al., 2007;

Ozbabacan et al., 2010). We will discuss the network-based identification of ligand binding sites in Sections 4.2.1 and 4.2.2.

Network-based methods of the network influence strategy are much less developed than those of the central hit strategy. Using the network influence strategy we need to conquer system robustness to push the cell back from the attractor of the diseased state to that of the healthy state, which is a difficult task—as we summarized in Section 2.5.2 on network dynamics. Nodes with intermediate connection numbers located in vulnerable points of disease-related networks (such as in inter-modular, bridging positions) driving disease-specific network traffic are preferred targets of the network influence strategy (Kitano, 2004a, 2004b, 2007; Ciliberti et al., 2007;



**Fig. 19.** Illustrative figure on the two major strategies to find network nodes as drug targets. The central hit strategy (represented by dark blue symbols) is useful to find drug targets against infectious agents or in anti-cancer therapies. These cells are presumed to have flexible networks. The central hit strategy targets central nodes (often forming a core of the network) or 'choke points', which are peripheral nodes uniquely producing or consuming a cellular metabolite. The network influence strategy (represented by red symbols) is needed to use the systems-level knowledge to find the targets in therapies of polygenic, complex diseases. The differentiated cells of these diseases are presumed to have rigid networks. Targeting central nodes here may cause an 'over-excitation' of the system leading to side-effects and toxicity. Thus the network influence strategy targets nodes, which are neither hubs nor otherwise central nodes themselves, but occupying strategically important disease-specific network positions able to influence central nodes. (In a typical case the network influence strategy targets are neighbors of central nodes exerting an indirect influence on the central nodes often representing the 'real targets'.) Solid lines represent network edges with high weight, while dashed lines represent network edges with low weight. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Antal et al., 2009; Hase et al., 2009; Zanzoni et al., 2009; Fliri et al., 2010; Cornelius et al., 2011; Farkas et al., 2011; Yu & Huang, 2012). In signaling networks preferred nodes of the network influence strategy inhibit certain outputs of the signaling network, while leaving others intact redirecting the signal flow in the network (Dasika et al., 2006; Ruths et al., 2006; Pawson & Linding, 2008). The network influence strategy often targets network segments, e.g. disease-modules (Cho et al., 2012), as we will discuss in Section 4.1.7. Edgetic drugs (Section 4.1.2.), multi-target drugs (Section 4.1.5.) and allo-network drugs (Section 4.1.6.) are all very promising forms of the network influence strategy.

It is important to note that the central hit strategy is applied against rapidly growing cells, such as those of infectious agents or cancer cells. Networks of these cells are in an 'exploratory phase', have larger entropy (West et al., 2012), and are presumably much more flexible, than the networks of differentiated cells. Efficient targeting of flexible systems having a high plasticity needs the targeting of their central nodes/edges. This is the major mode through which the central hit strategy operates.

In contrast, the network influence strategy is applied to differentiated cells. Networks of these cells are in an 'optimized phase', where the optimum was re-set for a diseased cell to a different attractor of its state-space than that of healthy cells. Networks of differentiated cells are presumably much more rigid than the networks of rapidly growing, undifferentiated or dedifferentiated cells. This is true all the more, since diseases are typically recognized by their late-appearing manifestations (Loscalzo and Barabasi, 2011), where the corresponding network already reached the disease-specific, rigid state. Targeting the central nodes/edges of systems having a low plasticity may easily 'over-saturate' the system leading to a change, which becomes too large to be selective, and causes side-effects and toxicity. Therefore, the network influence strategy often needs an indirect approach, where e.g. neighbors of the real target are targeted (allo-network drugs), or multiple targets are targeted 'mildly' (multi-target drugs), and their indirect and/or superposing effects lead to the reconfiguration of diseased network state back to normal.

In general, the strategy-pair we described here for drug action appears valid, and is a key for efficient modification of molecular structures, as well as neuronal and social systems at their different states besides the cellular networks discussed here. Thus, the best targeting strategy depends on the extent of network rigidity. Flexible, plastic systems are 'under-defined', and dissipate the perturbations well. These plastic systems may generally require a well-defined attack in

the form of targeting of their central nodes/edges. Note that central nodes form the most rigid segments of these flexible, plastic networks, which transmit the attack with the best efficiency. On the other hand, rigid systems are 'well-defined', and transmit (but not dissipate) the perturbations well. Thus optimal modification of rigid systems may be achieved by an indirect, 'under-defined' attack of the neighbors of their central nodes or rigid clusters. Note that nodes connecting rigid clusters (or being neighbors of central nodes) constitute the most flexible segments of these rigid networks, which dissipate the perturbation best, thus their attack does not over-excite the network. Moreover, rigid systems often contain multiple rigid clusters, and are often multi-modular (Mihalik & Csermely, 2011; Gáspár & Csermely, 2012). A single central hit may not optimally target a multi-modular, multi-centered system, which makes the network influence strategy even more efficient.

Network effects of existing drugs (e.g. in the form of drug target networks detailed in Section 4.1.3) may help in finding disease-specific network control-points. Drugs showing an mRNA expression profile that is strongly anticorrelated with a disease expression profile might actually reverse some of the disease effects and can be used for drug-repurposing. On the contrary, positively correlated profiles may reveal side effects (Dudley et al., 2011; Sirota et al., 2011; Iskar et al., 2012; Pacini et al., 2013). Reverse-engineering methods finding the underlying network structure from complex dynamic system output data (such as genome-wide mRNA expression patterns, signaling network or metabolome, see Section 2.2.3), as well as discriminating the primary targets from secondarily affected network nodes help in identifying control nodes directing network dynamics (Gardner et al., 2003; di Bernardo et al., 2005; Hallén et al., 2006; Lamb et al., 2006; Xing & Gardner, 2006; Lehár et al., 2007; Madhamshettiwar et al., 2012). The identification of disease-specific control-points of network dynamics will be an exciting task in the near future.

Disease-specificity may well be hierarchical. Suthram et al. (2010) identified 59 modules out of the 4620 modules of the human interactome, which are dysregulated in at least half of the 54 diseases tested, and were enriched in known drug targets. Influence-cores of the interactome, signaling, metabolic and other networks may be involved in the regulation of many more diseases than the connection-core (e.g. hub containing rich club) or periphery of these networks.

Potential methods to find influential nodes redirecting perturbations, affecting cellular cooperation or asserting network control have been described in Sections 2.3.4, 2.5.2 and 2.5.3 (Xiong & Choe, 2008; Antal et al., 2009; Kitsak et al., 2010; Luni et al., 2010; Farkas

et al., 2011; Liu et al., 2011, 2012; Banerjee & Roy, 2012; Cowan et al., 2012; Mones et al., 2012; Nepusz & Vicsek, 2012; Valente, 2012; Wang et al., 2012a). Influential nodes may have a hidden influence, like those highly unpredictable, ‘creative’ nodes, which may delay critical transitions of diseased cells (see Sections 2.2.2 and 2.5.2 for more details; Csermely, 2008; Scheffer et al., 2009; Farkas et al., 2011; Sornette & Osorio, 2011; Dai et al., 2012). Finding sets of influence-core nodes with fewer side-effects, or periphery nodes specifically influencing an influence-core or connection-core node, will be the subject of Sections 4.1.5 and 4.1.6 on multi-target drugs and allo-network drugs (Nussinov et al., 2011), respectively.

#### 4.1.2. Edgetic drugs: edges as targets

Perturbations of selected network edges give a grossly different result than the partial inhibition (or deletion) of the whole node. Development of drugs targeting network edges (recently called: edgetic drugs) has a number of advantages (Arkin & Wells, 2004; Keskin et al., 2007; Sugaya et al., 2007; Dreze et al., 2009; Zhong et al., 2009; Schlecht et al., 2012; Wang et al., 2012b).

- Many disease-associated proteins, e.g. p53, were considered non-tractable for small-molecule therapeutics, since they do not have an enzyme activity. In these cases edgetic drugs may offer a solution.
- Edgetic drugs are advantageous, since targeting network edges, i.e. protein–protein interaction, signaling or other molecular networks, is more specific than node targeting. This becomes particularly useful, when a protein simultaneously participates in two complexes having different functions, where only one of these functions is disease-related, like in the case of the mammalian target of rapamycin, mTOR (Huang et al., 2004; Agoston et al., 2005; Ruffner et al., 2007; Zhong et al., 2009; Wang et al., 2012b).
- Due to its larger selectivity, edge targeting may provide an efficient solution in targeting networks of multigenic diseases described as the network influence strategy in the preceding section. Edge targeting may also be used in the central hit strategy (targeting whole network-encoded systems) in the case of cancer, where selectivity may be more limited than in targeting of infectious agents. Importantly, the selectivity of edgetic drugs is not unlimited: hitting frequent interface motifs in a network may be as destructive as eliminating hubs. However, “interface-attack” may affect functional changes better than the attack of single proteins (Engin et al., 2012).

Edgetic drug development has inherent challenges. Interacting surfaces lack small, natural ligands, which may offer a starting point for drug design. Moreover, protein–protein binding sites involve large, flat surfaces, which are difficult to target. However, these flat surfaces often contain hot spots, which cluster to hot regions corresponding to a smaller set of key residues, which may be efficiently targeted by a drug of around 500 Da (Keskin et al., 2007; Wells & McClendon, 2007; Ozbabacan et al., 2010). We showed the usefulness of protein structure networks in finding hot spots in Section 3.2.4, and will summarize the possibilities to define edgetic drug binding sites in Section 4.2.2.

In one of the few systematic studies on edgetic drugs, Schlecht et al. (2012) constructed an assay to identify changes in the yeast interactome in response to 80 diverse small molecules, including the immunosuppressant FK506, which specifically inhibited the interaction between aspartate kinase and the peptidyl-prolyl-cis–trans isomerase, Fpr1. Sugaya et al. (2007) provided an *in silico* screening method to identify human protein–protein interaction targets. Edgetic perturbation of a *C. elegans* Bcl-2 ortholog, CED-9, resulted in the identification of a new potential functional link between apoptosis and centrosomes (Dreze et al., 2009). The TIMBAL database is a hand curated assembly of small molecules inhibiting protein–protein interactions (<http://www-cryst.bioc.cam.ac.uk/databases/timbal>; Higuero et al., 2009).

The Dr. PIAS server offers a machine learning-based assessment if a protein–protein interaction is druggable (<http://drpias.net>; Sugaya & Furuya, 2011; Sugaya et al., 2012).

Current development of edgetic drugs is mostly concentrated on protein–protein interaction networks. (We note here that most metabolic network-related drugs are by definition ‘edgetic drugs’, since in these networks target-enzymes constitute the edges between metabolites.) Signaling networks and gene interaction networks (including chromatin interaction networks) are promising fields of edgetic drug development. Scaffolding proteins and signaling mediators are particularly attractive targets of edgetic drug design efforts (Klussmann & Scott, 2008). In conclusion of this section, we list a few other future aspects of edgetic drug design.

- To date, the preferential topology of edge-targets in the human interactome has not been systematically addressed. Thus, currently we do not know, if indeed such a preference exists. Similarly, little attention has been paid to systematic studies of edge-weights, i.e. binding affinity-related drug target preference. Low-affinity binding is easier to disrupt, but interventions may not be that efficient. Disruption of high-affinity interactions may be more challenging (Keskin et al., 2007).
- Those interactions of intrinsically disordered proteins, which couple binding with folding, display a large decrease in conformational entropy, which provides a high specificity and low affinity. This pair of features is useful for regulation of protein–protein interactions and signaling, and this mechanism is widely used in human cells. Coupled binding and folding interactions often involve localized, small hydrophobic interaction surfaces, which provide a feasible targeting option in edgetic drug design (Cheng et al., 2006).
- Both low-probability interactions and interactions of intrinsically disordered proteins involve transient binding complexes. Modulation of these transient edges by ‘interfacial inhibition’ (Pommier & Cherfilis, 2005; Keskin et al., 2007) may be an option in future edgetic drug design.
- Edgetic drugs are usually inhibiting interactions (Gordo & Giralt, 2009). Stabilization of specific interactions is an area of great promise in drug design as we will discuss in Section 4.1.6 on allo-network drugs (Nussinov et al., 2011). Since the changed cellular environment in diseases often induces protein unfolding, general stabilizers of protein–protein interactions in normal cells, such as chemical chaperones, or chaperone inducers and co-inducers (Vigh et al., 1997; Söti et al., 2005; Papp & Csermely, 2006; Crul et al., 2013) offer an exciting therapeutic area of network-wide restoration of protein–protein interactions.

#### 4.1.3. Drug target networks

A broader representation of drug target networks are the protein-binding site similarity networks, where network edges between two proteins are defined by not only common, FDA-approved drugs, but also by a wide variety of common natural ligands and chemical compounds, as well as by binding site structural similarity measures. We list a few approaches to construct such protein binding site similarity networks below.

- Protein binding site networks can be constructed by large-scale experimental studies. One of these systematic studies examined naturally binding hydrophobic molecule profiles of kinases and proteins of the ergosterol biosynthesis in yeast using mass spectrometry. Hydrophobic molecules, such as ergosterol turned out to be potential regulators of many unrelated proteins, such as protein kinases (Li et al., 2010b).
- Protein binding site similarity networks may be constructed using a simplified representation of binding sites as geometric patterns, or numerical fingerprints. Here similarities are ranked by similarity scores based on the number of aligned features (Kellenberger et al., 2008).



- Pocket frameworks encoding binding pocket similarities were also used to create protein binding site similarity networks (Weisel et al., 2010). Pocket frameworks are reduced, graph-based representations of pocket geometries generated by the software PocketGraph using a growing neural gas approach. Another pocket comparison method, SMAP-WS combines a pocket finding shape descriptor with the profile-alignment algorithm, SOIPPA (Ren et al., 2010).
- Enzyme substrate and ligand binding sites have been compared using cavity alignment. Clustering of cavity space resembles most the structure of chemical ligand space and less that of sequence and fold spaces. Unexpected links of consensus cavities between remote targets indicated possible cross-reactivity of ligands, suggested putative side-effects and offered possibilities for drug repositioning (Zhang & Grigorov, 2006; Liu et al., 2008a; Weskamp et al., 2009).
- Andersson et al. (2010) proposed a method avoiding geometric alignment of binding pockets and using structural and physicochemical descriptors to compare cavities. This approach is similar to QSAR models of comparison detailed in Section 3.1.3.

Identifying clusters of proteins with similar binding sites may help drug repositioning, and could be a starting point for designing multi-target drugs as we will describe in the following two sections. Binding site similarities help in finding appropriate chemical molecules for new drug target candidates as described in Section 4.2. However, designing drugs for a group of targets with similar binding sites is challenging due to low specificity as exemplified by the drug design efforts against the ATP binding sites of protein kinases. Construction and analysis of protein binding site similarity networks in these cases can be helpful to identify proteins, whose active sites are different enough to be targeted selectively. Using 491 human protein kinase sequences, Huang et al. (2010b) constructed similarity networks of kinase ATP binding sites. The recent tyrosine kinase target, EphB4 belonged to a small, separated cluster of the similarity network supporting the experimental results of selective EhpB4 inhibition.

Signaling components, particularly membrane receptors and transcription factors form a major segment of drug target networks. Drug target networks are bipartite networks having drugs and their targets as nodes, and drug–target interactions as edges. These networks can be projected as drug similarity networks (where two drugs are connected, if they share a target). We summarized these projections as similarity networks in Section 3.1.3. In the other projection of drug–target networks, nodes are the drug targets, which are connected, if they both bind the same drug (Keiser et al., 2007, 2009; Ma'ayan et al., 2007; Yildirim et al., 2007; Hert et al., 2008; Yamanishi et al., 2008; Bleakley & Yamanishi, 2009; van Laarhoven et al., 2011). We describe the drug development applications of this projection in the remaining part of this section.

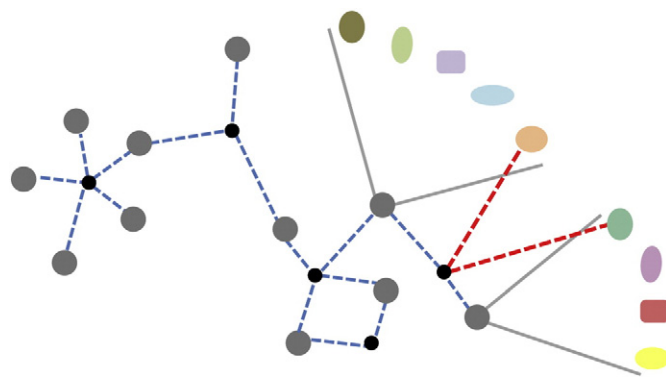
Drug target networks are particularly useful for comparisons of drug target proteins, since such a network comparison can be more informative pharmacologically than comparing protein sequences or protein structures. Drug target networks are modular: many drug targets are clustered by ligand similarity even though the targets themselves have minimal sequence similarity. This is a major reason, why drug target networks were successfully used to predict and experimentally verify novel drug actions (Keiser et al., 2007, 2009; Ma'ayan et al., 2007; Yildirim et al., 2007; Hert et al., 2008; Yamanishi et al., 2008, 2010; Bleakley & Yamanishi, 2009; van Laarhoven et al., 2011; Nacher & Schwartz, 2012; Mei et al., 2013).

Chen et al. (2012a) merged protein–protein similarity, drug similarity and drug–target networks and applied random walk-based prediction on this meta-network to predict drug–target interactions. Riera-Fernandez et al. (2012) developed a Markov–Shannon entropy-based numerical quality score to measure connectivity quality of drug–target networks extended by both the chemical structure networks of the drugs and the protein structure networks of their targets. As we will detail in Section 4.1.6 on allo-network drugs

(Nussinov et al., 2011), the integration of protein structure networks and protein–protein interaction networks may significantly enhance the success-rate of drug target network-based predictions of novel drug target candidates. Importantly, many drugs do not target the actual disease-associated proteins but proteins in their network-neighborhood (Yildirim et al., 2007; Keiser et al., 2009). Drugs having a target less than 3 or more than 4 steps from a disease-associated protein in human signaling networks have significantly more side-effects, and fail more often (Wang et al., 2012c). This substantiates the importance of the targeting of ‘silent’, ‘by-stander’ proteins further, which may influence the disease-associated targets in a selective manner (Section 4.1.6; Nussinov et al., 2011).

We listed a number of drug target databases and resources useful for the construction of drug–target networks in Table 9 at the beginning of Section 4. However, a refined representation of a drug target network should also include protein conformations (Fig. 20). Drugs may favor or disfavor certain protein conformations, and therefore this information is important for a more detailed understanding of drug action (Isin et al., 2012).

Indirect drug target networks may also be constructed using available data on human diseases, patients, their symptoms, therapies, or the systems-level effects of drug-induced perturbations (see Fig. 6 in Section 1.3.1; Spiró et al., 2008). Recently, several approaches extended drug/target datasets. Vina et al. (2009) assessed drug/target interaction pairs in a multi-target QSAR analysis enriching the dataset with chemical descriptors of targets and affinity scores of drug–target interactions. Wang et al. (2011b) assembled the Cytoscape (Smoot et al., 2011) plug-in of the integrated Complex Traits Networks (iCTNet, <http://flux.cs.queensu.ca/ictnet>) including phenotype/single-nucleotide polymorphism (SNP) associations, protein–protein interactions, disease–tissue, tissue–gene and drug–gene relationships. Balaji et al. (2012) compiled the integrated molecular interaction database (IMID, <http://integrativebiology.org>) containing protein–protein interactions, protein–small molecule interactions, associations of interactions with pathways, species, diseases and Gene Ontology terms with user-selected integration of manually curated and/or automatically extracted data. These and other complex approaches to drug target networks (Yamanishi et al., 2010, 2011; Savino et al., 2012; Tabei et al., 2012; Takarabe et al., 2012) will lead to the development of prediction techniques of novel drug targets, and improve drug efficiency, as well as ADMET, drug–drug interaction, side-effect and resistance profiles.



**Fig. 20.** A refined representation of a drug target network includes protein conformations. In current drug target network representations drug targets (gray circles) are interpreted as single entities connected through drugs (black circles). In these representations protein conformations preferred or dispreferred by a certain drug are ignored. For a more complete understanding of the interactions of drugs to their targets a target should be represented by its different functionally relevant conformations (differently colored shapes within gray line enclosed areas). Drug targets that are represented by single structures are connected to drugs by blue dashed lines. The target conformations that preferentially bind a drug are connected by red dashed lines. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.) Reproduced by permission from Isin et al. (2012).

#### 4.1.4. Network-based drug repositioning

Drug repositioning (or drug repurposing) aims to find a new therapeutic modality for an existing drug, and thus provides a cost-efficient way to enrich the number of available drugs for a certain therapeutic purpose. Drug repurposing uses a compound having a well-established safety and bioavailability profile, and a proven formulation and manufacturing process, as well as a well-characterized pharmacology. Most drug repositioning efforts use large screens of existing drugs against a multitude of novel targets (Chong & Sullivan, 2007). The pharmacological network approach asks, given a pattern of chemistry in the ligands, which targets a particular drug may bind (Kolb et al., 2009)? Here we list network-based methods mobilizing and efficiently using systems-level knowledge for rational drug repositioning.

- Analysis of common segments of protein–protein interaction and signaling networks affected by different drugs or participating in different diseases may reveal unexpected cross-reactions suggesting novel options for drug repurposing (Bromberg et al., 2008; Kotelnikova et al., 2010; Ye et al., 2012). As an example of these efforts, PROMISCUOUS (<http://bioinformatics.charite.de/promiscuous>) offers a web-tool for protein–protein interaction network-based drug-repositioning (von Eichborn et al., 2011).
- As an extension of the above approach, analysis of the complex drug similarity networks, by modularization, edge-prediction or by machine learning methods, described in Section 3.1.3 (see Table 5 there), may show unexpected links between remote drug targets indicating possible cross-reactivity of existing drugs with novel targets (Zhang & Grigorov, 2006; Liu et al., 2008a; Weskamp et al., 2009; Zhao & Li, 2010; Gottlieb et al., 2011a; Chen et al., 2012a; Cheng et al., 2012a, 2012b; Lee et al., 2012b). Network-based comparison of drug-induced changes in gene expression profiles (combined with disease-induced gene expression changes, disease–drug associations, interactomes, or signaling networks), was used to suggest unexpected, novel uses of existing drugs (Hu & Agarwal, 2009; Iorio et al., 2010, MANTRA server, <http://mantra.tigem.it>; Kotelnikova et al., 2010; Suthram et al., 2010; Gottlieb et al., 2011a; Luo et al., 2011, DRAR-CPI server, <http://cpi.bio-x.cn/drar>; Jin et al., 2012; Lee et al., 2012c, CDA server: <http://cda.i-pharm.org>; Pacini et al., 2013, DvD server as a Cytoscape plug-in: <http://www.ebi.ac.uk/saezrodriguez/DvD>).
- Genome-wide association studies (GWAS) may also be used to construct drug-related networks helping drug repositioning including in a personalized manner (Zanzoni et al., 2009; Coulombe, 2011; Cowper-Salari et al., 2011; Fang et al., 2011; Hu et al., 2011; Li et al., 2012b; Sanseau et al., 2012). Important future applications may use the comparison of phosphoproteome and metabolome data to reveal further drug repositioning option, including personalized drug application protocols.
- Drug target networks (including drug-binding site similarity networks and drug–target–disease networks) summarized in the preceding section help in drug repositioning. Modularization or edge prediction of these networks may reveal novel applications of existing drugs (Keiser et al., 2007, 2009; Ma'ayan et al., 2007; Yildirim et al., 2007; Hert et al., 2008; Yamanishi et al., 2008; Bleakley & Yamanishi, 2009; Kinnings et al., 2010; Mathur & Dinakarandian, 2011; van Laarhoven et al., 2011; Daminelli et al., 2012; Nacher & Schwartz, 2012).
- Central drugs of drug–therapy networks, where two drugs are connected, if they share a therapeutic application (Nacher & Schwartz, 2008), such as inter-modular drugs connecting two otherwise distant therapies, may reveal novel drug indications. Drug–disease networks have also been constructed and used for this purpose (Yildirim et al., 2007; Qu et al., 2009). Moreover, disease–disease networks (Goh et al., 2007; Rzhetsky et al., 2007; Feldman et al., 2008; Spiró et al., 2008; Hidalgo et al., 2009; Barabási et al., 2011; Zhang et al., 2011a) and the other disease and drug-related

network representations we listed in Section 1.3.1 (see Fig. 6 there; Spiró et al., 2008) may also be used for drug repositioning. Edge prediction methods (detailed in Section 2.2.2) and network-based machine learning methods may also be applied to these networks to uncover novel drug–therapy associations.

- Tightly interacting modules of drug–drug interaction networks (Yeh et al., 2006; Lehár et al., 2007) may also reveal unexpected, novel therapeutic applications.
- Side-effects of drugs, summarized in Section 4.3.5, may often reveal novel therapeutic areas. Shortest path, random walk and modularity analysis of side-effect similarity networks offers a number of novel options for network-based drug repositioning (Campillos et al., 2008; Yamanishi et al., 2010; Oprea et al., 2011; Takarabe et al., 2012).

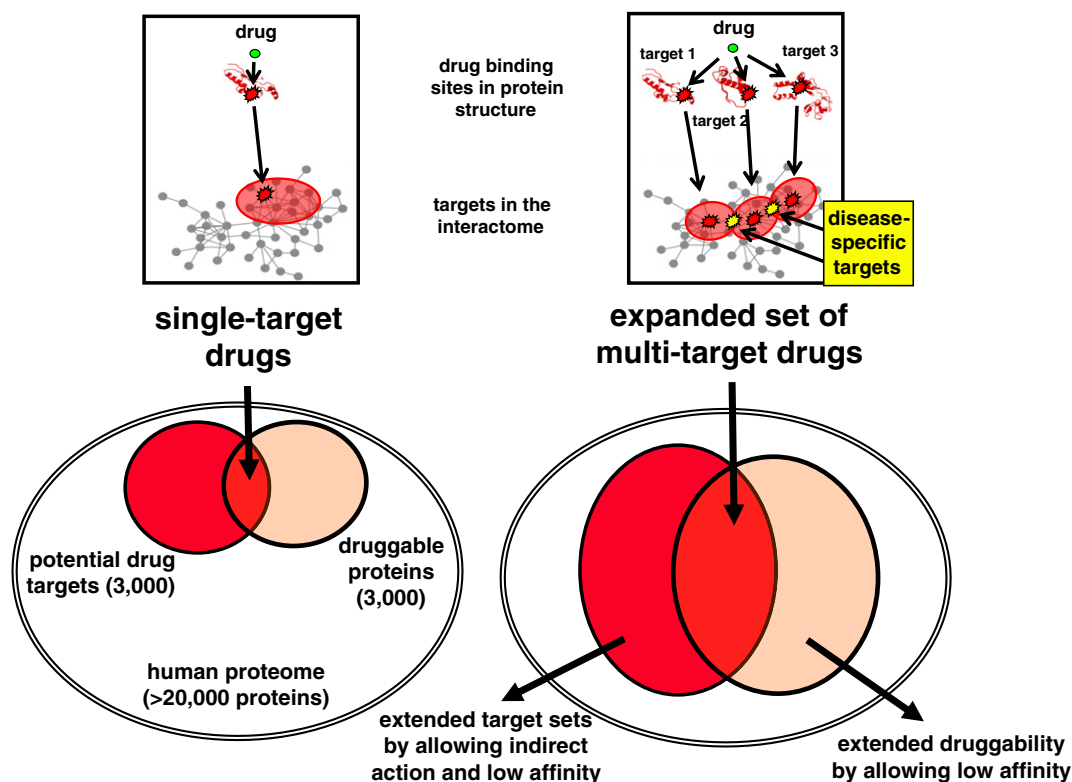
Network-related datasets and methods to reveal drug–drug interactions (Section 4.3.4), or drug side-effects (Section 4.3.5) may all give clues for drug re-positioning. Drug repositioning also has challenges, such as validation of the drug candidate from incomplete and outdated data, and the development of novel types of clinical trials (Mei et al., 2012). However, most network-based methods helping drug repositioning may also be used to predict multi-target drugs, an area we will summarize in the next section.

#### 4.1.5. Network polypharmacology: multi-target drugs

Robustness of molecular networks may often counteract drug action on single targets thus preventing major changes in systems-level outputs despite the dramatic changes in the target itself (see Section 2.5.2; Kitano, 2004a; Kitano, 2004b; Papp et al., 2004; Pál et al., 2006; Kitano, 2007; Tun et al., 2011). Moreover, most cellular proteins belong to multiple network modules in the human interactome, signaling or metabolic networks (Palla et al., 2005; Kovács et al., 2010; Wang et al., 2012d). As a consequence, efficient targeting of a single protein may influence many cellular functions at the same time. In contrast, efficient restoration of a particular cellular function to that of the healthy state (or efficient cell damage in anti-cancer strategies) can often be accomplished only by a simultaneous attack on many proteins, wherein the targeting efficiency on each protein may only be partial. These target sets preferentially contain proteins with intermediate number of neighbors having an intermediate level of influence of their own (Hase et al., 2009; Wang et al., 2012d).

The above systems-level considerations explain the success of polypharmacology, also called as modular pharmacology, i.e. the development and use of multi-target drugs (Fig. 21; Ginsburg, 1999; Csermely et al., 2005; Mencher & Wang, 2005; Millan, 2006; Hopkins, 2008; Wang et al., 2012d). The goal of polypharmacology is “to identify a compound with a desired biological profile across multiple targets whose combined modulation will perturb a disease state” (Hopkins, 2008). Multiple targeting is a well-established strategy. Snake or spider venoms, plant defense strategies are all using multi-component systems. Traditional medicaments and remedies often contain multi-component extracts of natural products. Combinatorial therapies are used with great success to treat many types of diseases, including AIDS, atherosclerosis, cancer and depression (Borisy et al., 2003; Keith et al., 2005; Dancey & Chen, 2006; Millan, 2006; Yeh et al., 2006; Lehár et al., 2007). Importantly, more than 20% of the approved drugs are multi-target drugs (Ma'ayan et al., 2007; Yildirim et al., 2007; Nacher & Schwartz, 2008). Many drugs interact with more than one transporter, which increases the complexity of polypharmacology (Kell et al., 2013). Moreover, multi-target drugs have an increasing market-value (Lu et al., 2012). Multi-target drugs possess a number of beneficial network-related properties, which we list below.

- Multi-target drugs can be designed to act on a selected set of primary targets influencing a set of key, therapeutically relevant secondary targets.



**Fig. 21.** Multi-target drugs are target multipliers. The top left panel and the red circle of the bottom left part of the figure show the targets of single-target drugs situated in pharmacologically interesting pathways and the hits of chemical proteomics, which represent those proteins, which can interact with druggable molecules. (The numbers are only approximate, and in case of the human proteome contain only the non-redundant proteins.) The overlap between the two sets constitutes the 'sweet spot' of drug discovery (Brown & Superti-Furga, 2003). On the right side of the figure the expansion of the 'sweet spot' is shown by multi-target drugs. The top left part illustrates the action of multi-target drugs. Yellow asterisks highlight the indirect targets, where the changes initiated by the multiple primary targets are superposed. It is a significant advantage, if these targets are disease-specific. On the bottom left part the indirect targets of multi-target action and the allowed low affinity binding of multi-target drugs both expand the number of pharmacologically relevant targets, while low-affinity binding enlarges the number of druggable proteins. The overlap of the two groups (the 'sweet spot') displays a dramatic increase. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

- Multiple targeting may need a compromise in binding affinity. However, even low-affinity binding multi-target drugs are efficient: in our earlier study a 50% efficient, partial, but multiple attack on a few sites of *E. coli* or yeast genetic regulatory networks caused more damage than the complete inhibition of a single node (Agoston et al., 2005; Csermely et al., 2005).
- Via the above, 'indirect' targeting, and via their low affinity binding, multi-target drugs may avoid the presently common dual-trap of drug-resistance and toxicity (Lipton, 2004; Csermely et al., 2005; Lehár et al., 2007; Zimmermann et al., 2007; Ohlson, 2008; Savino et al., 2012).
- Due to their low affinity binding, multi-target drugs may often stabilize diseased cells, which sometimes may be at least as beneficial as their primary therapeutic effect (Csermely et al., 2005; Korcsmáros et al., 2007; Csermely, 2009; Farkas et al., 2011).

In summary, multi-target drugs offer a magnification of the 'sweet spot' of drug discovery, where the 'sweet spot' represents those few hundred proteins, which are both parts of pharmacologically important pathways, and are druggable (Brown & Superti-Furga, 2003). The resulting beneficial effects have two reasons. First, both indirect and partial targeting by multi-target drugs expand the number of possible targets. Second, low affinity binding eases druggability constraints, and allows the targeting of partially hydrophilic binding sites by orally-deliverable, hydrophobic molecules. These two effects cause a remarkable increase of the drug targets situated in the overlap region of the potential target and druggable pools. Thus, multi-target drugs are, in fact, target multipliers (Fig. 21; Keith & Zimmermann, 2004; Csermely et al., 2005; Korcsmáros et al., 2007).

We list a number of network-related methods below to find target-sets of multi-target drugs by systems-level, rational multi-target design.

- Network efficiency (Latora & Marchiori, 2001), or critical node detection (Boginski & Commander, 2009) may serve as a starting measure to judge network integrity after multi-target action (Agoston et al., 2005; Csermely et al., 2005; Li et al., 2011c). Pathway analysis of molecular networks gives a more complex picture, and may reveal multiple intervention points affecting pathway-encoded functions, utilizing pathway cross-talks, or switching off compensatory circuits of network robustness. Network methods allow the identification of target sets, which disconnect signaling ligands from their downstream effectors with the simultaneous preservation of desired pathways (Dasika et al., 2006; Ruths et al., 2006; Lehár et al., 2007; Jia et al., 2009; Hormozdiari et al., 2010; Kotelnikova et al., 2010; Pujol et al., 2010). Deconvolution of network dynamics showing interrelated dynamics modules, such as those of elementary signaling modes (Wang & Albert, 2011), is a promising approach for future multi-drug design efforts.
- Experimental testing of drug combinations may uncover unexpected effects in drug–drug interactions, which may be used for selection of multi-target sets (Borisy et al., 2003; Keith et al., 2005; Dancy & Chen, 2006; Yeh et al., 2006; Lehár et al., 2007; Jia et al., 2009; Liu et al., 2010b). Combination therapies may also be designed using network methods, such as the minimal hitting set method (Vazquez, 2009), or a complex method taking into account adjacent network position and action-similarity (Li et al., 2011d). Recently, several iterative algorithms were developed to find optimal target combinations restricting the search to a few combinations out

of the potential search space of several millions to billions of combinations (Calzolari et al., 2008; Wong et al., 2008; Small et al., 2011; Yoon, 2011; Zhao et al., 2011a). Pritchard et al. (2013) demonstrated the existence of simple and predictable combination mechanisms using RNA interference signatures. Network-based search algorithms may improve this search efficiency even further in the future. Drug combinations against diseases affecting the cardiovascular and nervous systems have a more concentrated effect radius in the human genetic interaction network than that of immuno-modulatory or anti-cancer agents (Wang et al., 2012e). Network methods were applied to predict and avoid unwanted drug–drug interaction effects and the emergence of multi-drug resistance as we will describe in Sections 4.3.4 and 4.3.6, respectively.

- Side-effect networks connect drugs by the similarity of their side-effects. Shortest path and random walk analysis, as well as the identification of tight clusters, bridges and bottlenecks of these networks (Campillos et al., 2008; Yamanishi et al., 2010; Oprea et al., 2011; Takarabe et al., 2012) combined with the selective optimization of side activities (Wermuth, 2006) may be used to design multi-target drugs.
- The combined similarity networks of chemical molecules including drug targets, various molecular networks (such as interactomes or signaling networks), system-wide biological data (such as mRNA expression patterns) and medical knowledge (such as disease characterization) listed in Tables 5 and 9 (Lamb et al., 2006; Paolini et al., 2006; Brennan et al., 2009; Hansen et al., 2009; Iorio et al., 2009; Li et al., 2009a; Huang et al., 2010a; Zhao & Li, 2010; Azuaje et al., 2011; Bell et al., 2011; Taboureau et al., 2011; Wang et al., 2011b; Balaji et al., 2012; Edberg et al., 2012) may all be used for multi-target drug design using modularization method-, similarity score-, network inference-, Bayesian network- or machine learning-based clustering (Hopkins et al., 2006; Hopkins, 2008; Chen et al., 2009e; Xiong et al., 2010; Yang et al., 2010; Hu et al., 2011; Takigawa et al., 2011; Yabuuchi et al., 2011; Cheng et al., 2012a, 2012b; Lee et al., 2012c; Nacher & Schwartz, 2012; Yu et al., 2012). We note that some of the above methods describe and interpret multi-target action and thus need further development for multi-target prediction.
- Multiple perturbations of interactomes, signaling networks or metabolic networks may uncover alternative target sets causing a similar systems-level perturbation to that of the original target set. Differential analysis of networks in healthy and diseased states may enable an even more efficient prediction (Antal et al., 2009; Farkas et al., 2011). Such perturbation studies were successfully applied to smaller, well-defined networks before using differential equation sets and disease-state specific Monte Carlo simulated annealing (Yang et al., 2008). Assessment of network oscillations may reveal central node sets governing the dynamic behavior (Liao et al., 2011).
- Recent advances in establishing the controllability conditions of large networks and in defining complex network hierarchy measures (Cornelius et al., 2011; Liu et al., 2011, 2012; Banerjee & Roy, 2012; Cowan et al., 2012; Mones et al., 2012; Nepusz & Vicsek, 2012; Wang et al., 2012a; Yazicioglu et al., 2012) may uncover multiple target sets, as shown by the assessment of the controllability of smaller networks (Luni et al., 2010). Controlling sets, which can assign any prescribed set of centrality values to all other nodes by cooperatively tuning the weights of their out-going edges (Nicosia et al., 2012) may also be promising in the identification of multi-target sets.
- Appropriate reduction of the definition of dominant node sets, i.e. sets of nodes reaching all other nodes of the network, may also be used to determine target sets of multi-target drugs (Milenkovic et al., 2011). Minimal dominant node set determination was recently shown to be equal to finding minimal transversal sets of hypergraphs (i.e. a hitting set of a hypergraph, which has a nonempty intersection with each edge; Kanté et al., 2011), which extends this technique to the powerful hypergraph description, where an edge may connect

any groups of nodes and not only two nodes. Definition and determination of appropriately limited dominant edge-sets (Milenkovic et al., 2011) constitute a powerful approach for multi-target identification.

- Analysis of transport between multiple sources and sinks in directed networks (Morris & Barthelemy, 2012), such as in signaling networks or in metabolic networks may reveal preferred source sets (encoding target sets of multi-target drugs) preferentially affecting pre-defined sink sets (encoding the desired effects). Throughflow centrality has been recently defined as an important measure of such network configurations (Borrett, 2012). Methods to find conceptually similar seed sets of information spread in social networks (Shakarian & Paulo, 2012) may also be applied to find multi-target drug sets.
- Recently highly powerful methods were published to design and optimize multitarget ligands for polypharmacology profiles (Ajmani & Kulkarni, 2012; Besnard et al., 2012). Besnard et al. (2012) tested 800 automatically designed multi-target ligands of G-protein coupled receptors, and found that 75% of them had a correctly predicted polypharmacology profile. This area will remain an exciting priority of multi-target drug design efforts.

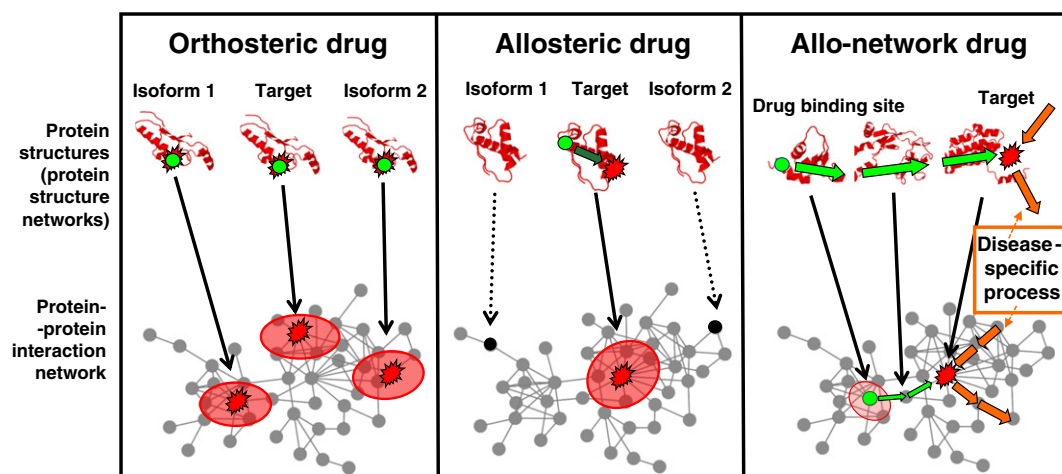
Some of the above methodologies (such as those based on chemical similarity networks) result in target sets, where lead design is a more feasible process. Target sets, which are highly relevant at the systems-level, but have diverse binding site structures may require the identification of a set of indirect targets selectively influencing the desired target set, but posing a more feasible lead development task. We will describe the network-based identification of such indirect targets in the next section describing allo-network drugs (Nussinov et al., 2011). We note that almost all methods finding target sets of multi-target drugs can be used for drug repositioning summarized in the preceding section. Moreover, all these methods are related to the *in silico* prediction of drug–drug interactions (detailed in Section 4.3.4) and side-effects (summarized in Section 4.3.5).

#### 4.1.6. Allo-network drugs: a novel concept of drug action

Allosteric drugs (binding to allosteric effector sites; Fig. 22) are considered to be better than orthosteric drugs (binding to active centers; Fig. 22) due to 4 reasons. 1.) The larger variability of allosteric binding sites than that of active centers causes less allosteric drug-induced side-effects than that of orthosteric drugs. 2.) Allosteric drugs allow the modulation of therapeutic effects in a tunable fashion. 3.) In most cases the effect of allosteric drugs requires the presence of endogenous ligand making allosteric action efficient exactly at the time when the cell needs it. 4.) Allosteric drugs are non-competitive with the endogenous ligand. Therefore, their dosage can be low (DeDecker, 2000; Rees et al., 2002; Goodey & Benkovic, 2008; Lee & Craik, 2009; Nussinov et al., 2011; Nussinov & Tsai, 2012).

We summarized our current knowledge on allosteric action (Fischer, 1894; Koshland, 1958; Straub & Szabolcsi, 1964; Závodszky et al., 1966; Tsai et al., 1999; Jacobs et al., 2003; Goodey & Benkovic, 2008; Csermely et al., 2010; Zhuravlev & Papoian, 2010; Rader & Brown, 2011; Dixit & Verkhivker, 2012; Szilágyi et al., *in press*) from the point of view of protein interaction networks in Section 3.2.2. In that section we described the rigidity front propagation model as a possible molecular mechanism of the propagation of allosteric changes (Fig. 14; Csermely et al., 2012).

The concept of allosteric drugs can be broadened to allo-network drugs, whose effects can propagate across several proteins via specific, inter-protein allosteric pathways of amino acids activating or inhibiting the final target (Fig. 22; Nussinov et al., 2011). Earlier data already pointed to an allo-network type drug action. Inter-protein propagation of allosteric effects (Bray & Duke, 2004; Fliri et al., 2010) and its possible use in drug design (Schadt et al., 2009) were mentioned sporadically in the literature. Moreover, drug–target network studies revealed that in more than half of the established



**Fig. 22.** Comparison of orthosteric, allosteric and allo-network drugs. Top parts of the three panels illustrate the protein structures of the primary drug targets showing the drug binding site as a green circle. Bottom parts of the panels illustrate the position of the primary targets in the human interactome. Red ellipses illustrate the ‘action radius’, i.e. the network perturbation induced by the primary targets. In the top part of the middle panel the allosteric drug binds to an allosteric site and affects the pharmacologically active site of the target protein (marked by a red asterisk) via the intra-protein allosteric signal propagation shown by the dark green arrow. In the top part of the right panel the signal propagation (illustrated by the light green arrows) extends beyond the original drug binding protein, and via specific interactions affects two neighboring proteins in the interactome. The pharmacologically active site is also marked by a red asterisk here. Orange arrows illustrate an intracellular pathway of propagating conformational changes, which is disease-specific in case of successful allo-network drugs. Allo-network drugs allow indirect and specific targeting of key proteins by a primary attack on a ‘silent’ protein, which is not involved in major cellular pathways. Targeting ‘silent’, ‘by-stander’ proteins, which specifically influence pharmacological targets, not only expands the current list of drug targets, but also causes much less side-effects and toxicity. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Adapted with permission from Nussinov et al. (2011).

922 drug–disease pairs drugs do not target the actual disease-associated proteins, but bind to their 3rd or 4th neighbors. However, the distance between drug targets and disease-associated proteins was regarded as a sign of palliative drug action (Yildirim et al., 2007; Barabási et al., 2011), and the expansion of the concept of allosteric drug action to the interactome level has been formulated only recently (Nussinov et al., 2011). Interestingly, targeting neighbors was found to be more influential on the behavior of social networks than direct targeting (Bond et al., 2012).

Allo-network drug action propagates from the original binding site to the interactome neighborhood in an anisotropic manner, where propagation efficiency is highly directed and specific. Binding sites of promising allo-network drug targets are not parts of ‘high-intensity’ intracellular pathways, but are connected to them. These intracellular pathways are disease-specific in the case of promising allo-network drugs (Fig. 22). Thus allo-network drugs can achieve specific, limited changes at the systems level with fewer side-effects and lower toxicity than conventional drugs. Allosteric effects can be considered at two levels: 1.) small-scale events restricted to the neighbors or interactome module of the originally affected protein; 2.) propagation via large cellular assemblies over large distances (i.e. hundreds or even thousands of Angstroms; Nussinov et al., 2011; Szilágyi et al., in press). Drugs with targets less than 3 steps (or more than 4 steps) from a disease-associated protein were shown to have significantly more side-effects, and failed more often (Wang et al., 2012c); however, rational drug design in recent years proceeded in the opposite direction, identifying drug targets closer to disease-associated proteins than earlier (Yildirim et al., 2007). The above data argue that reversing this trend may be more productive. Allo-network drugs point exactly to this direction.

Databases of allosteric binding sites (Huang et al., 2011a; <http://mdl.shsmu.edu.cn/ASD>) help the identification possible sites of allo-network drug action. However, allo-network drugs may also bind to sites, which are not used by natural ligands. For the identification of allo-network drug targets and their binding sites, first the interactome has to be extended to atomic level (amino acid level) resolution. For this, docking of 3D protein structures and the consequent

connection of their protein structure networks are needed. Thus allo-network drug targeting requires the integration of our knowledge on protein structures, molecular networks, and their dynamics focusing particularly on disease-induced changes. We conclude this section by listing a few possible methods to define allo-network drug target sites.

- A general strategy for the identification of allosteric sites may involve finding large correlated motions between binding sites. This can reveal which residue–residue correlated motions change upon ligand binding, and thus can suggest new allosteric sites (Liu & Nussinov, 2008) even in integrated networks of protein mega-complexes.
- Reverse engineering methods (Tegnér & Bjorkegren, 2007) allow us to discriminate between ‘high-intensity’ and ‘low-intensity’ communication pathways both in molecular and atomic level networks, and thus may provide a larger safety margin for allo-network drugs.
- As we summarized in Section 3.2.2, network-based analysis of perturbation propagation is a fruitful method to identify intra-protein allosteric pathways (Pan et al., 2000; Chennubhotla & Bahar, 2006; Ghosh & Vishveshwara, 2007, 2008; Tang et al., 2007; Daily et al., 2008; Goodey & Benkovic, 2008; Sethi et al., 2009; Tehver et al., 2009; Vishveshwara et al., 2009; Park & Kim, 2011; Csermely et al., 2012; Ma et al., 2012a). A successful candidate for the inter-protein allosteric pathways involved in allo-network drug action disturbs network perturbations specific to a disease state of the cell at a site distant from the original drug-binding site. Perturbation analysis (see Section 2.5.2; Antal et al., 2009; Farkas et al., 2011) applied to atomic level resolution of the interactome in combination with disease specific protein expression patterns may help the identification of such allo-network drug targets.
- Central residues play a key role in the transmission of allosteric changes (Section 3.2.2; Chennubhotla & Bahar, 2006; Chennubhotla & Bahar, 2007; Zheng et al., 2007; Chennubhotla et al., 2008; Tehver et al., 2009; Liu & Bahar, 2010; Liu et al., 2010a; Su et al., 2011; Park & Kim, 2011; Dixit & Verkhivker, 2012; Ma et al., 2012a;

- Pandini et al., 2012). We may use a number of centrality measures (Kovács et al., 2010), including perturbation-based or game-theoretical assumptions (see Sections 2.5.2 and 2.5.3; Farkas et al., 2011), to find the level of importance of proteins and pathways in interactomes, in signaling networks and important amino acids in their extensions to atomic level resolution (Szalay-Bekő et al., 2012).
- At both the molecular network level and its extension to atomic level resolution we may subtract network hierarchy (Ispolatov & Maslov, 2008; Jothi et al., 2009; Cheng & Hu, 2010; Hartsperger et al., 2010; Rosvall & Bergstrom, 2011; Liu et al., 2012; Mones et al., 2012; Szalay-Bekő et al., 2012) to assess the importance of various nodes (proteins and/or amino acids), or we may find nodes or edges controlling the network by the application of recently published methods (Cornelius et al., 2011; Liu et al., 2011, 2012; Banerjee & Roy, 2012; Cowan et al., 2012; Mones et al., 2012; Nepusz & Vicsek, 2012; Wang et al., 2012a; Pósfai et al., 2013).
  - Combination of evolutionary conservation data proved to be an efficient predictor of intra-protein signaling pathways (Tang et al., 2007; Halabi et al., 2009; Joseph et al., 2010; Jeon et al., 2011; Reynolds et al., 2011). Similar approaches may be extended to protein neighborhoods helping to find starting sites for allo-network drug action.
  - Disease-associated single-nucleotide polymorphisms (SNPs; Li et al., 2011b) and/or mutations (Wang et al., 2012b) may be part of the propagation pathways of allosteric effects. In-frame mutations are enriched in interaction interfaces (Wang et al., 2012b), and provide a potential dataset to assess the existence of allo-network drug binding sites.

Targeting disease-induced dynamical changes in molecular networks may also be focused on transient interactions specific to disease. Thus allo-network drugs might also provide a novel solution to uncompetitive, 'interfacial' drug action (Pommier & Cherfil, 2005; Keskin et al., 2007). When available, current drugs aim to directly inhibit protein–protein interactions (Gordo & Giral, 2009). We note that the methods above are suitable to find allo-network drugs, which stabilize/restore/activate a protein, its function or one (or more) of its interactions. The methods we listed here are suitable for finding both primary targets of allo-network drugs in molecular networks and allo-network drug binding sites in the amino acid networks of involved proteins. We will describe additional network-related methods to find binding sites of allo-network drugs proper in Section 4.2.

#### 4.1.7. Networks as drug targets

The last two sections on multi-target drugs and allo-network drugs already demonstrated the utility of network-based thinking in the determination of drug-targets. In this closing section on drug target identification we summarize the ideas considering key segments of networks as drug targets.

Considering molecular networks as targets has gained increasing support in recent papers on systems-level drug design (Brehme et al., 2009; Schadt et al., 2009; Zanzoni et al., 2009; Baggs et al., 2010; Pujol et al., 2010; Cho et al., 2012; Erler & Linding, 2012). As we defined in the starting section on drug target identification, from the network point of view it is important to discriminate between two strategies: 1.) the central hit strategy aiming to destroy the network of infectious agents or cancer cells and 2.) the network influence strategy using the systems-level knowledge to find drug target candidates in therapies of polygenic, complex diseases (see Fig. 19 and Section 4.1.1 for further details). Here we list a few major characteristics of both strategies.

Optimal network targeting of the central hit strategy:

- finds hubs and other central nodes or edges of molecular networks or identifies choke points of metabolic networks, i.e. proteins uniquely producing or consuming a certain metabolite;
- finds unique targets of infectious agent- or cancer-specific networks.

Optimal network targeting of the network influence strategy:

- shifts disease-specific changes of cellular functions back to their normal range (Kitano, 2007);
- applies precise targeting of selected network pathways, protein complexes, network segments, nodes or edges avoiding highly influential nodes and edges of molecular networks in healthy cells but converging drug effects at specific pathway sites of diseased cells;
- uses multiple or indirect targeting;
- takes into consideration tissue specificity.

Optimal network targeting of both the central hit and network influence strategies:

- incorporates patient- and disease stage-specific data (such as single-nucleotide polymorphisms, metabolome, phosphoproteome or gut microbiome data) ADMET-related data, side-effect- and drug resistance-related data as detailed in the next section.

We believe that the arsenal of network (re)construction and network analysis methods we listed in this review may offer help and promise for the prediction of novel, systems-level drug targeting possessing the characteristics detailed above.

## 4.2. Hit finding, expansion and ranking

Following target selection discussed in the preceding section, here we will discuss the added-value of network-related methods in the search, confirmation and expansion of hit molecules. Several steps in this process, such as pharmacophore identification, network-based QSAR models, building of a hit-centered chemical library, hit expansion, as well as other network-related methods of chemoinformatics and chemical genomics, have already been discussed in Section 3.1.3. Therefore, the Reader is asked to compare Section 3.1.3 with the current chapter. Here we will first summarize the help provided by network description and analysis in the determination of ligand binding sites, as applicable to network nodes as drug targets. We will continue with network methods to find hot spots, which reside in protein interfaces, and are targets of edgetic drugs. We will conclude the section by a summary of network-related approaches in hit expansion and ranking.

### 4.2.1. In silico hit finding for ligand binding sites of network nodes

Node targeting aims to find a selective, drug-like (low molecular weight, possibly hydrophobic) molecule that binds with high affinity to the target (Lipinski et al., 2001). There are two main network-based approaches for the identification of ligand binding sites. A 'bottom-up approach' uses protein structure networks (see Section 3.2 in detail), while a 'top-down approach' reconstructs binding site features from binding site similarity networks (Section 4.1.3).

For in silico hit prediction, a logical first step is to find pockets (cavities, clefts) on the protein surface. Medium-sized proteins have 10 to 20 cavities. Ligands often bind to the largest surface cavities of this ensemble (Laskowski et al., 1996; Liang et al., 1998b; Nayal & Honig, 2006). Using a protein structural approach, Coleman and Sharp (2010) identified a hierarchical tree of protein pockets using the travel depth algorithm that computes the physical distance a solvent molecule would have to travel from a given protein surface point to the convex hull of the surface. Using the similarity network approach, pocket similarity networks have been constructed, and their small-world character, hubs and hierarchical modules were identified. Pocket groups were found to reflect functional separation (Liu et al., 2008a, 2008b), and may be used for hit identification. However, shape information alone is insufficient to discriminate between diverse binding sites, unless combined with chemical descriptors (<http://proline.physics.iisc.ernet.in/pocketmatch>; Yeturu & Chandra, 2008; <http://proline.physics.iisc.ernet.in/pocketalign>; Yeturu & Chandra, 2011). CAVER (<http://caver.cz>; Chovancova et al., 2012)

uses molecular dynamics simulation to predict intra-protein transport pathways.

Protein structure networks (Section 3.1) were relatively seldom used so far to predict ligand binding sites. However, high-centrality segments of protein structure networks were shown to participate in ligand binding (Liu & Hu, 2011). Evolutionary conservation patterns of amino acids in related protein structures identified protein sectors related to catalytic and allosteric ligand binding sites (Halabi et al., 2009; Jeon et al., 2011; Reynolds et al., 2011). Protein structure networks were extended, incorporating ligand atoms, participating ions and water molecules and chemical properties aiming to find network motifs representing a favorable set of protein–ligand interactions used for as a scoring function (Xie & Hwang, 2010; Kuhn et al., 2011). Protein structure network comparison was demonstrated to be useful for the identification of chemical scaffolds of potential drug candidates (Konrat, 2009).

Similarity clusters or network prediction methods of binding site similarity networks (also called as pocket similarity networks, or cavity alignment networks; Zhang & Grigorov, 2006; Kellenberger et al., 2008; Liu et al., 2008a; Park & Kim, 2008; Andersson et al., 2010; Weskamp et al., 2009; Xie et al., 2009a; BioDrugScreen, <http://biodrugscreen.org>; Li et al., 2010c; Reisen et al., 2010; Ren et al., 2010; Weisel et al., 2010) can be used to predict binding site topology of yet unknown proteins. The complex drug target network, PDTD (<http://dddc.ac.cn/pdtd>) incorporating 3D active site structures and the web-server TarFishDock enables simultaneous target and target-site prediction of new chemical entities (Gao et al., 2008). The versatile protein–ligand interaction database, CREDO (<http://www-cryst.bioc.cam.ac.uk/databases/credo>; Schreyer & Blundell, 2009) and the extensive protein–ligand databases, STITCH (<http://stitch.embl.de>; Kuhn et al., 2012) and BindingDB (<http://bindingdb.org>; Liu et al., 2007a) offer an important help to search for potential targets and identify their binding sites.

#### 4.2.2. *In silico* hit finding for edgetic drugs: hot spots

Edgetic drugs (Section 4.1.2) modify protein–protein interactions. Protein–protein interaction binding sites were considered for a long time as “non-druggable”, since they are large and flat. However, Clackson and Wells (1995) discovered hot spots of binding surfaces, which are residues providing a contribution to the decrease in binding free energy of larger than 2 kcal/mol. Bogan and Thorn (1998) proposed that hot spots are surrounded by hydrophobic regions excluding water from the hot spot residues. Hot spots are often populated by aromatic residues, and tend to cluster in hot regions, which are tightly packed, relatively rigid hydrophobic regions of the protein–protein interface. Hot spots and hot regions help in finding hits, since 1.) they constitute small focal points of drug binding, which can be predicted within the large and flat binding-interface; 2.) these focal points are relatively rigid, and help in docking. An inhibitor needs to cover 70 to 90 atoms at the protein–protein interaction site, which corresponds to the ‘Lipinski-conform’ (Lipinski et al., 2001) ~500 Da molecular weight. Several small molecules were found, which are able to compete with the natural binding partner very efficiently (Keskin et al., 2005, 2007; Wells & McClendon, 2007; Ozbabacan et al., 2010). Druggable hot regions have a concave topology combined with a pattern of hydrophobic and polar residues (Kozakov et al., 2011).

Hot spots can be predicted as central nodes of protein structure networks (del Sol & O’Meara, 2005; Liu & Hu, 2011; Grosdidier & Fernández-Recio, 2012). In agreement with this, disease-associated mutations (single-nucleotide polymorphisms) are enriched by 3-fold at the interaction interfaces of proteins associated with the disorder, and often occur at central nodes of the protein structure network (Akula et al., 2011; Li et al., 2011b; Wang et al., 2012b). Using this knowledge, the pyDock protein–protein interaction docking algorithm was improved by protein structure network-based scores (Pons et al., 2011). Intra-protein energy fluctuation pathways were proposed to help in the prediction of hot spot localization (Erman,

2011). Recently the use of associative-memory, water-mediated coarse-grained protein folding model, AWSEM was also demonstrated to predict protein binding surfaces well (Zheng et al., 2012a).

Hit identification of edgetic drugs is helped by the TIMBAL database containing ligands inhibiting protein–protein interactions (<http://www-cryst.bioc.cam.ac.uk/databases/timbal>; Higuero et al., 2009). The machine learning-based technique of the Dr. PIAS server assesses if a protein–protein interaction is druggable (<http://drpias.net>; Sugaya & Furuya, 2011; Sugaya et al., 2012). Despite the considerable progress of this field in the last decade, we are still at the beginning of using network-related knowledge to identify edgetic drug binding sites. Network-related methods for hot spot and hot region identification are also promising, if applied to aptamers, peptidomimetics or proteomimetics.

#### 4.2.3. Network methods helping hit expansion and ranking

An important step of hit confirmation is the check of the chemical amenability of the hit, i.e. the feasibility up-scaling costs of its synthesis. Core and hub positions or other types of centrality of the hits in the chemical reaction network (Section 3.1.2; Fialkowski et al., 2005; Bishop et al., 2006; Grzybowski et al., 2009) are all predictors of good chemical tractability. Moreover, a simulated annealing-based network optimization uncovered optimal synthetic pathways of selected hits (Kowalik et al., 2012). In the case of multiple hits, hit clustering can be performed by modularization of their chemical similarity networks described in Section 3.1.3. Hubs and clusters of hit-fragments in chemical similarity networks may be used for hit-specific expansion of existing compound libraries (Benz et al., 2008; Tanaka et al., 2009). QSAR-related similarity networks and the other complex similarity networks we listed in Table 5 help the lead development and selection efforts we will detail in the next section.

Hit cluster should usually conform to the Lipinsky-rules of drug-like molecules (Lipinski et al., 2001) restricting the hit-range to small and hydrophobic molecules with a certain hydrogen-bond pattern. Leeson and Springthorpe (2007) warned that systematic deviations from these rules may have a dangerous impact on drug design, increasing late-attributions due to side-effects and/or toxicity. However, natural compounds also contain a set of ‘anti-Lipinsky’ molecules, which form a separate island in the chemical descriptor space having a higher molecular weight and a larger number of rotatable bonds (Ganesan, 2008). The network-related methods predicting the efficiency, ADME, toxicity, interactions, side-effect and resistance occurrence detailed in the next section may help in decreasing the risk of non-conform hit and lead molecules, and highlight issues of drug safety in an early phase of drug development.

#### 4.3. Lead selection and optimization: drug efficacy, drug absorption, distribution, metabolism, excretion, toxicity, drug interactions, side-effects and resistance

Following hit selection and expansion discussed in the preceding section network-related methods may also help the lead selection process. Various aspects of lead selection such as drug toxicity, side-effects and drug–drug interactions are tightly interrelated. The incorporation of personalized data, such as genome-wide association studies/single-nucleotide polymorphisms (GWAS/SNPs), signaling network or metabolome data into the complex network structures which help lead selection may not only predict well the pharmacogenomic properties of the lead, but also help patient profiling in clinical trials, as well as therapeutic guideline determination of the marketed product.

##### 4.3.1. Networks and drug efficacy, personalized medicine

Drug efficacy is the theoretical efficiency of drug action not taking into account the effects in practice, such as patient compliance.

Efficacy is a highly personalized efficiency measure of drug action, which heavily depends on multiple factors including the genetic background (e.g. single-nucleotide polymorphisms and other genetic variants assessed in genome-wide association studies), network robustness and the ADME properties (see next section; Kitano, 2007; Barabási et al., 2011; Yang et al., 2012). Single-nucleotide polymorphisms (SNPs) may alter the interaction properties of at least 20% of the nodes in the human interactome (Davis et al., 2012), and were recently shown to be a reason for the unexpectedly high variability of protein–protein interactions (Hamp & Rost, 2012). A number of studies assessed the effects of SNPs on changing the underlying properties of interactomes and gene–gene association networks (Akula et al., 2011; Cowper-Salari et al., 2011; Fang et al., 2011; Hu et al., 2011; Li et al., 2011b, 2012b; Wang et al., 2012b), which may change drug efficacy both directly or indirectly. The integrated Complex Traits Networks (iCTNet, <http://flux.cs.queensu.ca/ictnet>), including phenotype/single-nucleotide polymorphism (SNP) associations, protein–protein interactions, disease–tissue, tissue–gene and drug–gene relationships, is a rich dataset helping drug efficacy assessments (Wang et al., 2011b).

Incorporation of omics-type data into complex, drug action-related networks will allow the construction of personalized efficacy profiles. Integration of pharmacogenomics, signaling network or metabolome data may improve clinical trial design. However, network-related methodologies for complex drug efficacy profiling have not been developed yet. Similarly, analysis of the semantic networks of medical records by text mining and by network analysis techniques is a future tool to improve the assessment of drug efficiency measures, extending the efficacy with patient compliance and other effects occurring in medical practice (Chen et al., 2009a). Network-related models may help in the development of optimal drug dosage and frequency schedules. As an example of this, the study of Li et al. (2011e) uncovered a ‘sweet spot’ of drug efficacy dose and schedule regions by the extension of their model to the genetic regulatory network environment of the drug target. Drug dose and schedule considerations are already parts of the ADME characterization, which we will detail in the next section.

#### 4.3.2. Networks and absorption, distribution, metabolism and excretion, toxicity: drug absorption, distribution, metabolism and excretion

The integration of early ADME (absorption, distribution, metabolism, excretion) profiling to lead selection is an important element of successful drug design. Prediction of ADME properties using structural networks of lead candidates (Kier & Hall, 2005), molecular fragment networks predicting human albumin binding (Estrada et al., 2006), chemical similarity networks (Brennan et al., 2009), as well as drug–tissue networks (Gonzalez-Diaz et al., 2010b), isotope-labeled metabolomes and drug metabolism networks (Martinez-Romero et al., 2010; Fan et al., 2012), non-linear diffusion models of drug partitioning in lipid network structures, such as the stratum corneum (Schumm et al., 2010), multiple binding to transporters (Kell et al., 2013) and complex networks of major cellular mechanisms participating in ADME determination (Ekins et al., 2006), were all important advances which can help in incorporating better ADME complexity into the lead selection process. Despite these methods, there is room to improve ADME prediction and assessment by network techniques. ADME studies are often combined with toxicity assessments (ADMET), which we will detail in the next section. Toxicity is related to side-effects discussed in Section 4.3.5. Drug combinations may have an especially complex ADME profile due drug–drug interaction effects, which will be described in Section 4.3.4.

#### 4.3.3. Networks and drug toxicity

Toxicity plays a different role in drug targets identified using the central hit strategy and the network influence strategy of Section 4.1.1. In the central hit strategy our aim is to kill the cells of

the infectious agent or cancer. Therefore, toxicity is a must here—but it has to be selective to the targeted cells. In the network influence strategy targeting other diseases, toxicity becomes generally avoidable. Toxicity is often a network property depending on the extent of network perturbation and robustness (Kitano, 2004a, 2004b, 2007; Apic et al., 2005; Geenen et al., 2012). Network hubs and the essential proteins described in Section 2.3.4 are less frequently targeted by drugs—with the exception of anti-infective and anti-cancer agents (Jonsson & Bates, 2006; Yildirim et al., 2007). In contrast, those inter-modular bridges, which modulate specific information flows, are preferred drug targets (Hwang et al., 2008). Node centrality in drug-regulated networks correlates with drug toxicity (Kotlyar et al., 2012). All these findings give further support to the utility of network-based toxicity assessments.

Hepatotoxicity is a major reason of drug attritions (Kaplowitz, 2001). The number of network studies addressing this important issue is increasing, and includes cytokine signaling networks related to idiosyncratic drug hepatotoxicity (Cosgrove et al., 2010) and gene–gene interaction networks based on transcriptional profiling (Hayes et al., 2005; Kiyosawa et al., 2010). Importantly, toxicity-related networks should be understood as signed networks containing both toxicity promoting effects and detoxifying effects, such as the glutathione network in liver (Geenen et al., 2012), or hepatic pro-survival (AKT) and pro-death (MAPK) pathways, where specific pathway inhibitors may antagonize drug-induced hepatotoxicity (Cosgrove et al., 2010).

Network-based *in silico* prediction of human toxicity aims to bridge the gap between animal toxicity studies and clinical trials. Toxicity assessment applications of chemical similarity networks (Section 3.1.3; Kier & Hall, 2005; Brennan et al., 2009), as well as the use of association networks between chemicals and toxicity-related proteins or processes (DITOP, <http://bioinf.xmu.edu.cn:8080/databases/DITOP/index.html>; Zhang et al., 2007; Audouze et al., 2010; Iskar et al., 2012) open a number of additional possibilities for network-predictions of human toxicity in the future.

#### 4.3.4. Networks and drug–drug interactions

Drug–drug interactions may often cause highly unexpected effects. As we already described in Section 4.1.5 on network polypharmacology and multi-target drug design, most of the unexpected drug–drug interactions are not due to direct competition for the same binding site, but are caused by the complex interaction structure of molecular networks. Experimental testing of drug–drug interactions may be used to help infer the underlying molecular network structure, and drug–drug interaction networks (Borisy et al., 2003; Yeh et al., 2006; Lehár et al., 2007; Jia et al., 2009) may be used to predict additional drug–drug interactions using network modularization methods.

A drug–drug interaction network was assembled using drug package insert texts. This network was extended by potential mechanisms, such as drug targets or enzymes involved in drug metabolism, and was included in the KEGG DRUG database (<http://genome.jp/kegg/drug>; Takarabe et al., 2008; Takarabe et al., 2011; Kanehisa et al., 2012). Recently text mining rules to refine literature-derived drug–drug interaction networks were proposed (Kolchinsky et al., 2013). Drug–drug interaction networks may be perceived as signed networks containing synergistic or antagonistic interactions (Yeh et al., 2006; Jia et al., 2009), and have hubs, i.e. drugs which are involved in most of the observed interactions (Hu & Hayton, 2011). Many of the drug-related databases listed in Table 9 may help to uncover adverse drug–drug interactions. Besides the KEGG DRUG database mentioned above the DTome (<http://bioinfo.mc.vanderbilt.edu/DTome>; Sun et al., 2012) database also explicitly contains adverse drug interactions. Complex chemical similarity networks and drug–target networks, discussed in Sections 3.1.3 and 4.1.3, respectively, were also used for the prediction of unexpected drug–drug interactions (Zhao & Li, 2010; Yu et al., 2012).



Drugs may affect each other's ADME properties by simple competition, or by more refined network-effects (Jia et al., 2009; Kell et al., 2013), such as the positive synergism of amoxicillin and clavulanate, where clavulanate is an inhibitor of the enzyme responsible for amoxicillin destruction (Matsuura et al., 1980). Drug–herb interactions are important aspects of drug–drug interaction analysis particularly in China, where traditional Chinese medicine is often combined with Western medicine. Here semantic networks and other combined networks of drug and herb effects and targets may offer help in prediction of drug safety (Chen et al., 2009a; Zheng et al., 2012b). Despite the wide variety of approaches listed, network techniques provide many more possibilities in the prediction of drug–drug interaction effects. In practice, all methods listed in Section 4.1.5 on multi-target drugs, such as perturbation, network influence and source/sink analyses, as well as the drug side-effect networks described in the next section may be used for the prediction of drug–drug interactions.

#### 4.3.5. Network pharmacovigilance: prediction of drug side-effects

Discovering unexpected side-effects by experimental methods alone, is a daunting task requiring the screen of a large number of potential off-targets. However, side-effects may have their origin in both single and multi-target. Both are systems-level responses, which allow the prediction of drug off-targets by computational methods (Berger & Iyengar, 2011; Zhao & Iyengar, 2012). In this section we introduce several network-related methods of side-effect identification.

Side-effects may come from the involvement of a single drug target in multiple cellular functions or may involve multiple drug targets. In a study on protein–protein interaction networks two third of the side-effect similarities were related to shared targets, while 5.8% of the side-effect similarities were due to drugs targeting proteins close in the human interactome (Brouwers et al., 2011). This result may reflect both the concentration of side-effects in direct drug targets and the efficiency of those allo-network drugs (Section 4.1.6; Nussinov et al., 2011), whose direct target is not the primary binding site, but a neighboring protein in the interactome.

The previous sections uncovered many network-related strategies to avoid side-effects at the level of target selection. We will summarize only a few major considerations here.

- Avoidance of targeting hubs and high centrality nodes of interactomes, signaling networks and metabolomes is a general network strategy of side-effect reduction, especially when using the network influence strategy of Section 4.1.1 against polygenic diseases such as diabetes. Disease specific, limited network perturbation is a key systems-level requirement to avoid drug adverse effects (Guimera et al., 2007b; Hase et al., 2009; Zhu et al., 2009; Yu & Huang, 2012). Network algorithms focusing the downstream components of node-targeting to a certain network segment are important methods to reduce potential side-effects at the level of target identification (Dasika et al., 2006; Ruths et al., 2006; Pawson & Linding, 2008).
- Iterative methods sequentially identified sets of metabolic network edges corresponding to enzymes, whose inhibition can produce the expected inhibition of targets with reduced side-effects in humans and in *E. coli* (Lemke et al., 2004; Sridhar et al., 2007, 2008; Song et al., 2009).
- Unexpected edges between remote targets in ligand binding site similarity networks (also called as pocket similarity networks, or cavity alignment networks) suggest potential side-effects (Zhang & Grigorov, 2006; Liu et al., 2008a; Weskamp et al., 2009).
- Edgetic drugs (Section 4.1.2) are usually more specific and may have generally less side-effects than node-targeting drugs. However, common protein–protein interaction interface motifs are important indicators of potential side-effects of edgetic drugs (Engin et al., 2012).
- Future analysis may uncover nodes and edges having a major influence on the occurrence of the disease-specific critical network-

transitions mentioned in Section 2.5.2. These influential nodes will most probably represent the 'Achilles-heel' of network in the disease state, and their targeting will induce fewer side-effects than the average.

Side-effect prediction is tightly related to drug–target prediction (Section 4.1) involving the comparison of novel target(s) with those of existing drugs. The selective optimization of side-effects (Wermuth, 2006) is a known lead development technique. Consequently both drug–target interaction networks (Section 4.1.3; Xie et al., 2009b; Yang et al., 2010; Azuaje et al., 2011; Xie & Bourne, 2011; Yang et al., 2011; Takarabe et al., 2012; Yu et al., 2012) and drug–disease networks (Hu & Agarwal, 2009) may be used for the prediction of side-effects. Analysis of drug–disease networks may be extended using pathway analysis (Ye et al., 2012). Complex chemical similarity networks (Section 3.1.3) also use a combination of network-related data including e.g. interactomes for the prediction of off-target effects (Hase et al., 2009; Yamanishi et al., 2010; Zhao & Li, 2010). The web-servers SePreSA (<http://SePreSA.Bio-X.cn>; Yang et al., 2009a) and DRAR-CPI (<http://cpi.bio-x.cn/drar>; Luo et al., 2011) were constructed to show possible adverse drug reactions based on drug–target interactions. Practically all methods listed in Section 4.1.5 on multi-target drugs may be used to predict side-effects. As an example, the Monte Carlo simulated annealing network perturbation method of Yang et al. (2008) correctly predicted the well-known side-effects of non-steroidal anti-inflammatory drugs and the cardiovascular side-effects of the recalled drug, Vioxx. Moreover, side-effect determination may be extended to any complex similarity networks we listed in Table 5 (such as that containing disease-specific genome-wide gene expression data; Huang et al., 2010a; the Cytoscape plug-in DvD program, <http://www.ebi.ac.uk/saezrodriguez/DVD>; Pacini et al., 2013) and to those future network representations, which will include signaling network or metabolome data. These datasets may be used to construct personalized or patient cohort-specific side-effect profiles enabling a better focusing of therapeutic indications and contraindications.

In recent years, many types of side-effect networks, drug target/adverse drug reaction networks or drug target/adverse target networks were constructed.

- Campillos et al. (2008) and later Yamanishi et al. (2010) and Takarabe et al. (2012) combined structural similarity and side-effect similarity to construct a side-effect similarity network of drugs, and used this network to identify novel drug targets for drug repositioning (Section 4.1.4).
- Correlation analysis of drug protein-binding profiles and side-effect profiles revealed the enrichment of drug targets participating in the same biological pathways (Mizutani et al., 2012).
- Text mining of drug package insert text was used for the construction of side-effect networks showing a gross similarity of preclinical and clinical compound profiles (Fliri et al., 2005; Oprea et al., 2011). Text mining of scientific papers may result in an extended drug–target network revealing potential side-effects (Garten et al., 2010).
- A drug–target/adverse drug reaction network was constructed from chemical similarity-based prediction of off-targets and related side-effects of 656 drugs (Lounkine et al., 2012).
- A network of 162 drugs causing at least one serious adverse drug reaction and their 845 targets showed similar target profiles for similar serious adverse drug reactions. The MHC I (Cw\*4) protein was identified and confirmed as the possible target of the sulfonamide-induced toxic epidermal necrolysis adverse effect (Yang et al., 2009b).
- Yang et al. (2009c) used the CitationRank network centrality algorithm and a dataset of gene/serious adverse drug reaction associations (collected by text mining from PubMed records) to identify the association strength of genes with 6 major serious adverse drug reactions (<http://gengle.bio-x.cn/SADR>).

Side-effect similarity networks were used for efficient refinement of primary side-effect identification based on similarities in drug structures (Atias & Sharan, 2011). Network prediction methods detailed in Section 2.2.2 and network modularization methods may help to decipher novel side-effects from side-effect networks in the future.

The side-effect database, SIDER (<http://sideeffects.embl.de>; Kuhn et al., 2010) considerably enhanced side-effect network studies. The SIDER-derived side-effect network was extended by biological processes related to Gene Ontology terms and text mining of PubMed data (Lee et al., 2011). Combination of SIDER data with those on disease-associated genes showed that drugs having a target less than 3 or more than 4 steps away from a disease-associated protein in human signaling networks had significantly more side-effects, and failed more often (Wang et al., 2012c).

Sources of unexpected side-effects can sometimes be focused on a certain tissue or cellular process. Analysis of tissue-specific network dynamics, such as that of the kidney metabolic network revealing hypertensive side-effects (Chang et al., 2010), might be a promising method to predict tissue-specific side-effects. Csoka and Szyf (2009) raised the possibility of epigenetic side-effects, where a drug modifies the chromatin structure, and thus indirectly influences a number of other genes. Similarly, microRNA related side-effects may be developed by the interaction of a drug with the complex microRNA signaling network (Section 3.4), where a change in the transcription of a microRNA may influence a set of rather unrelated proteins and related functions.

#### 4.3.6. Resistance and persistence

In recent years antibiotic resistance became a major threat of human health (Bush et al., 2011). Antibiotic persistence is a form of antibiotic resistance, which is related to a dormant, drug-insensitive subpopulation of bacteria (Rotem et al., 2010). Resistance development against chemotherapeutic agents is a key challenge of anti-cancer therapies (Section 3.4.3; Kitano, 2004a; Logue & Morrison, 2012). Resistance development is involved in the application of the central hit strategy defined in Section 4.1.1 aiming to destroy pathogen- or cancer-related networks.

Ligands may be optimized against resistance by targeting conserved amino acids and main-chain atoms with strong interactions instead of weaker interactions pointing towards mutable residues (Hopkins et al., 2006). Tuske et al. (2004) defined the substrate-envelope for HIV reverse transcriptase as the space occupied by various conformations of naturally occurring ligands and their targets. Lamivudine and zidovudine induced resistance by protruding beyond this substrate-envelope, while tenofovir, which did not have handles projecting beyond the substrate envelope was more resistant against resistance development (Tuske et al., 2004). Protein structure network studies may help in designing more resistant-prone lead molecules.

Development of drug resistance is often a phenomenon involving network-robustness, when the affected cell activates alternative or counter-acting pathways to minimize the consequences of drug action (Kitano, 2007). Oberhardt et al. (2010) offer a comprehensive analysis of metabolic network adaptation of *Pseudomonas aeruginosa* to a host organism during a 44-months period. Co-targeting of an additional crucial point of drug-affected network pathways is an efficient tool to fight against resistance. Drug combinations and multi-target drugs develop less resistance (Zimmermann et al., 2007; Pujol et al., 2010; Rosado et al., 2011; Savino et al., 2012). Analysis of pathogen interactomes involving random walks or known drug resistance-related proteins plus gene expression changes revealed pathways often involved in resistance development helping co-target determination (Raman & Chandra, 2008; Chen et al., 2012b). Resistance-related proteins defined a subset of pathogen interactome, called resistome. Drug-induced gene expression changes and betweenness centralities of their interactions were used as weights of

resistome edges. Resistome hubs may serve as important co-targets (Padiadpu et al., 2010). Differential assessment of molecular networks of normal and resistant pathogens allows more efficient drug resistant target and/or co-target identification (Kim et al., 2010). As we described in Section 3.4.3, the combination of anti-tumor drugs and stress response targeting increases therapeutic efficiency (Rocha et al., 2011; Tentner et al., 2012). The Hebbian learning rule, i.e. the property of neuronal networks to increase edge-weights along frequently used pathways (Hebb, 1949) may be extended to molecular networks, and studied as a possible source of systems-level resistance development.

Importantly, the most efficient synergistic drug combinations typically preferred in clinical settings may develop a faster resistance, which might argue for other, e.g. antagonistic drug combinations (Chait et al., 2007; Hegreness et al., 2008). Synthetic rescues (when the inhibition of a target compensates for the inhibition of another; Section 3.6.3) are good candidates for anti-resistant antagonistic co-target action (Motter, 2010). Network simulation of resistance transmission in bacterial populations also underlined the need for potent antimicrobials and high-enough doses to kill the susceptible population segment as soon as possible (Gehring et al., 2010). Network-related methods to fight drug-resistance may help both anti-infective and anti-cancer strategies as we will describe in the next two sections.

## 5. Four examples of network description and analysis in drug design

In this section we will illustrate the usefulness of network-related methods in drug design with examples of four major threats of human health: infectious diseases, cancer, diabetes (extended to metabolic diseases) and neurodegenerative diseases. The first two disease groups (infectious diseases and cancer) are examples of the central hit strategy aiming to destroy the network of infectious agents or cancer cells (see Section 4.1.1 for a definition of the two strategies). The last two therapeutic areas (diabetes and neurodegenerative diseases) are examples of the network influence strategy aiming to re-wire molecular networks of diseased cells to restore normal function (see Section 4.1.7 for a summary of these two strategies). The sections on the use of network science to combat these diseases will not give a comprehensive summary, but will only highlight a few key solutions and their results.

### 5.1. Anti-viral drugs, antibiotics, fungicides and antihelminthics

Drugs against infectious agents are central hit strategy drugs in the classification that we introduced in Section 4.1.1. Efficient central hit strategy drugs interfere with viral replication or kill the infectious cells with high efficiency (instead of temporal growth inhibition, which may induce resistance), and avoid any toxic effects in humans. We list a number of selected illustrative examples of network approaches for drug target identification against pathogenic agents in Table 10.

Integrated host-pathogen networks proved to be very efficient in complex targeting strategies in the case of viral/host interactomes (Uetz et al., 2006; Calderwood et al., 2007; de Chassey et al., 2008; Chautard et al., 2009; Navratil et al., 2009, 2011; Brown et al., 2011; Prussia et al., 2011; Xu et al., 2011b; Lai et al., 2012; Schleker et al., 2012; Simonis et al., 2012). Complex databases of viral/host interactions were also assembled (Zhang et al., 2005). Anti-viral target proteins often emerge as bridges between host/pathogen and human network modules, as well as hubs or otherwise central proteins of the virus-targeted human interactome (Uetz et al., 2006; Calderwood et al., 2007; de Chassey et al., 2008; Navratil et al., 2011; Lai et al., 2012). Targets of viral proteins were shown to be major perturbators of human networks (de Chassey et al., 2008; Navratil et al., 2011). A machine learning technique with a learning set including viral/host interactome-derived topological and functional

**Table 10**

Illustrative examples of the use of network methods in anti-viral drugs, antibiotics, fungicides and antihelminthics.

Drug design area	Network method	References
<i>Protein–protein interaction networks</i>		
Drug target identification	Identification of (pathogen-specific) hubs as potentially essential proteins ( <a href="http://hub.iis.sinica.edu.tw/Hubba">http://hub.iis.sinica.edu.tw/Hubba</a> ) Identification of clique-forming, high-centrality, or otherwise topologically essential proteins	Lin et al., 2008; Kushwaha & Shakya, 2009 Real et al., 2004; Estrada, 2006; Florez et al., 2010; Milenkovic et al., 2011; Raman et al., 2012; Zhang et al., 2012
<i>Metabolic networks</i>		
Drug target identification	Comparative load point (high-centrality) and choke point (unique reaction) analysis of pathogenic and non-pathogenic bacteria (with the identification of conserved critical amino acids forming similar cavities: UniDrugTarget server, <a href="http://117.211.115.67/UDT/main.html">http://117.211.115.67/UDT/main.html</a> ) Selection of essential metabolites Selection of super-essential reactions	Perumal et al., 2009; Chanumolu et al., 2012 Kim et al., 2011 Barve et al., 2012 Shen et al., 2010
Drug target identification, drug repositioning	Strain-specific anti-infective therapies by comparative metabolic network analysis	
<i>Drug–target, drug–drug similarity and complex dataset networks</i>		
Target identification, drug repositioning	Drug target network of <i>Mycobacterium tuberculosis</i>	Kinnings et al., 2010
Prediction of drug activity against different pathogens	Multi-tasking QSAR drug–drug similarity network analysis	Gonzalez-Diaz & Prado-Prado, 2008; Prado-Prado et al., 2008; Prado-Prado et al., 2009; Prado-Prado et al., 2010; Prado-Prado et al., 2011
Drug target identification	Interactome, signaling network and gene regulation network of <i>Mycobacterium tuberculosis</i>	Vashisht et al., 2012

information identified several formerly validated viral targets of the influenza A virus, and predicted novel drug target candidates (Lai et al., 2012). The combination of viral/host interactome data with siRNA, transcriptome, microRNA, toxicity and other data may significantly extend the prediction efficiency of antiviral targets (Brown et al., 2011).

Analysis of integrated bacterial/fungal/parasite and human metabolic networks also became a widely used tool to predict potential drug target efficiency (Bordbar et al., 2010; Huthmacher et al., 2010; Fatumo et al., 2011; Riera-Fernandez et al., 2012). Chavali et al. (2012) and Kim et al. (2012) offered comprehensive collections of datasets and analyses of antimicrobial drug target identification using metabolic networks. Combinations of the metabolic network and the interactome of *Mycobacterium tuberculosis* were used to identify the most influential network target singletons, pairs, triplets and quadruplets (Raman et al., 2008, 2009; Kushwaha & Shakya, 2010). Multiple targets are useful to prevent the development of resistance (see preceding section; Raman & Chandra, 2008; Chen et al., 2012b). However, recent studies showed that synergistic drug combinations, which are preferred in clinical settings due to their high efficiency, may develop a faster resistance. Therefore, antagonistic drug combinations should also be tried (Yeh et al., 2006; Chait et al., 2007; Hegreiness et al., 2008).

Complex chemical similarity networks including chemical-genetic interactions (i.e. hypersensitivity data of mutant strains for chemical compounds; for additional examples see Table 5) help in the identification of drug targets in anti-infective therapies (Parsons et al., 2006; Hansen et al., 2009). The (random) upstream attack strategy proposed by Liu et al. (2012) may uncover even more influential targets than currently known in directed networks such as metabolic or signaling networks. Mészáros et al. (2011) warned that the selective targeting of bacterial proteins may involve complex domain architecture. Complex similarity networks (Vilar et al., 2009) may allow patient- and disease stage-specific target search in the anti-cancer therapies detailed in the next section. As another approach linking the two central hit strategy-type drug design areas, anti-infective and anti-cancer therapies, the assessment of interactome and transcriptome perturbations by DNA tumor virus proteins highlighted the Notch- and apoptosis-related pathways that also go awry in cancer (Rozenblatt-Rosen et al., 2012).

## 5.2. Anti-cancer drugs

Cancer is a systems-level disease (Hornberg et al., 2006), where the rapidly proliferating system is characterized by an increase of network entropy (West et al., 2012), i.e. an increase of network flexibility and plasticity. This increase in network flexibility may characterize the initial stages of tumor development (such as adenomas) better than later stages of malignant transformation (such as carcinomas; Dezső Módos, Tamás Korcsmáros and Péter Csermely, in preparation). Key aims of anti-cancer pharmacology include the identification of targets and the efficient combination of drugs to overcome the robustness of cancer-specific cellular networks with the least toxicity and resistance development possible (Kitano, 2004a, 2004b, 2007; Werner, 2011; Cheng et al., 2012c; Rivera et al., 2012). Similarly to the anti-pathogenic drugs described in the preceding section, anti-cancer drugs mostly belong to central hit strategy drugs (see Section 4.1.1). However, central targets of cancer cells are more often central components of healthy cells, than those of infectious agents discussed in the preceding section. Therefore, as we will discuss in Section 5.2.5, a few anti-cancer drugs (presumably those, which target the more rigid networks of advanced cancer types) start to resemble the network influence strategy-type drugs.

In the following sections we will show the help of interactomes, metabolic and signaling networks to find cancer-specific drug targets and drug combinations. To illustrate the special importance of network-level thinking in anti-cancer drug design, we start the section with the description of autophagy, which is a very promising area to develop novel anti-cancer drugs—but only if treated in a systems-level context using network description and analysis.

### 5.2.1. Autophagy and cancer—an example for the need of systems-level view

Autophagy (cellular self-degradation) has a highly ambiguous role in cancer. On the one hand, autophagy has tumor suppressing functions a.) by limiting chromosomal instability; b.) by restricting oxidative stress, which is also an oncogenic stimulus; and c.) by promoting oncogene-induced senescence. On the other hand, autophagy is used by tumor cells to escape hypoxic, metabolic, detachment-induced and therapeutic stresses as well as to develop metastasis and dormant tumor cells (Apel et al., 2009; Morselli et al., 2009; White & DiPaola,

2009; Kenific et al., 2010; Chen & Klionsky, 2011). Thus autophagy should be modulated in a cell-specific manner. In cancer cells over-activation of autophagy can induce cell death, while autophagy inhibitors sensitize cancer cells to chemotherapy. In normal cells, autophagy stimulators may be useful for cancer prevention by enhancing damage mitigation and senescence, while autophagy inhibitors can induce tumorigenesis (White & DiPaola, 2009; Ravikumar et al., 2010; Chen & Karantza, 2011). Network analysis of the regulation of autophagy may point out such context-specific intervention points. The recent work of Serra-Musach et al. (2012) showed that in contrast to most cancer-related proteins, proteins involved in autophagy are more 'failure-prone', i.e. can be saturated by perturbations faster. This gives an additional rationale to employ autophagy-related proteins as drug targets corresponding to the 'Achilles-heel' of cancer cells. Network approaches described in all the following sections may be promising for the identification of autophagy-related drug target candidates.

### 5.2.2. Protein–protein interaction network targets of anti-cancer drugs

Cancer-specificity in the anti-cancer drug targets is a primary requirement to avoid toxicity. Target-specificity may increase by selecting cancer-related mutation events or proteins having altered gene expression. In addition, all these data can be combined at the network-level (Pawson & Linding, 2008).

Large-scale sequencing identified thousands of genetic changes in tumors, which were collected in databases, such as COSMIC (<http://www.sanger.ac.uk/genetics/CGP/cosmic>; Forbes et al., 2011) or the Network of Cancer Genes (<http://bio.ieu.eu/ncg/index.html>; D'Antonio et al., 2012). Section 1.3.3 and Tables 2 and 3 listed a number of network-related methods to identify cancer-associated proteins. From the large number of tumor-associated genes, only a few play a key role in tumor pathogenesis (called driver mutations). Driver mutations can be characterized by their pathway association. In many tumors p53, Ras and PI3K are the major signaling pathways containing driver mutations (Li et al., 2009b; Pe'er & Hacoheh, 2011). Genes with co-occurring mutations in the COSMIC database prefer direct signaling interactions. Genes having a less coherent neighborhood in the network of co-occurring mutations tend to have a higher mutation frequency (Cui, 2010). Driver mutations are in cancer-modules and are neighbors of signature genes, whose expression can be used as a prognostic marker of metastasis and survival in breast tumors (Li et al., 2010d). Recently pathway and network reconstitution methods were suggested using patient survival-related mutation data (Vandin et al., 2012).

In human interactomes, proteins with cancer-specific mutations are hubs. They form a rich-club, acting as bridges between modules of different functions, and behave as bottlenecks providing exclusive connections between network segments or are otherwise central nodes (Jonsson & Bates, 2006; Chuang et al., 2007; Sun & Zhao, 2010; Rosado et al., 2011; Xia et al., 2011). Preferential connectedness of cancer-related proteins may contribute to their increased robustness to transmit a large volume of perturbations without being damaged (Serra-Musach et al., 2012). In agreement with the above observations, targets of anticancer drugs have a significantly larger number of neighbors than targets of drugs against other diseases (Hase et al., 2009). Inter-modular interactome hubs were found to associate with oncogenesis better than intra-modular hubs (Taylor et al., 2009). Integration of the interactome, protein domain composition, evolutionary conservation and gene ontology data in a machine learning technique predicted target genes, whose knockdown greatly reduced colon cancer cell viability (Li et al., 2009b). The interactomes of cancer associated cells, such as cancer associated fibroblasts may also highlight important prognostic markers of the disease (Bozóky et al., in press).

Differential gene expression analysis became one of the key approaches to identify genes important in diagnosis and prediction of cancer progression. The Oncomine resource includes more than

18,000 gene expression profiles (<http://oncomine.org>; Rhodes et al., 2007a). Oncomine data were extended by drug treatment signatures and target/reference gene sets providing a network of Molecular Concepts Map (<http://private.molecularconcepts.org>; Rhodes et al., 2007b). Differentially expressed proteins in human cancers were cataloged in the dbDEPC database (<http://lifecenter.sgst.cn/dbdepc/index.do>; He et al., 2012). Gene expression profiles may be used for reverse engineering of cancer specific regulatory networks (Basso et al., 2005; Ergun et al., 2007). Gene expression subnetworks showed increased similarity with the progression of chronic lymphocytic leukemia, suggesting that degenerate pathways converge into common pathways that are associated with disease progression (Chuang et al., 2012).

Interactome nodes may be marked according to their up- or down-regulation in cancer, and may identify clusters of proteins involved in cancer progression, such as in metastasis-formation (Rhodes & Chinnaiyan, 2005; Jonsson et al., 2006; Hernandez et al., 2007). Network analysis measures (e.g. degree, betweenness centrality, shortest path) of integrated interactome and expression data ranked cancer related proteins for target prediction, and showed their central network position (Wachi et al., 2005; Platzer et al., 2007; Chu & Chen, 2008; Mani et al., 2008).

However, altered expression of mRNAs is generally not enough to predict target efficiency (Yeh et al., 2012). mRNAs are often regulated by microRNAs, thus the inclusion of microRNA pattern analysis improves prediction as we will show in the next section. Moreover, the analysis of proteomic changes is also necessary in most cases (Gulmann et al., 2006; Pawson & Linding, 2008). Changes in protein levels may act synergistically (Maslov & Ispolatov, 2007). Starting from this idea, random walk-based interactome analysis identified sub-networks, which were around 'seeds' changing their protein levels in colorectal cancer, and screened these subnetworks using the level of the synergistic dysregulation of the associated mRNAs in colorectal cancer (Nibbe et al., 2010). Inclusion of additional data in the interactome and gene expression datasets, such as protein domain interactions, gene ontology annotations, cancer-related mutations, or cancer prognosis information refined predictions further (Franke et al., 2006; Pujana et al., 2007; Chang et al., 2009; Lee et al., 2009; Wu et al., 2010; Xiong et al., 2010; Yeh et al., 2012).

### 5.2.3. Metabolic network targets of anti-cancer drugs

The metabolism of cancer cells is adapted to meet their proliferative needs in predominantly anaerobic conditions (Warburg, 1956). Network modeling uses a number of cancer-specific pathways of energy metabolism (Resendis-Antonio et al., 2010; Vazquez et al., 2011; Khazaei et al., 2012; Kung et al., 2012). Metabolic networks of several cancer-types such as that of colorectal cancer were constructed recently (Martinez-Romero et al., 2010). Li et al. (2010e) used the k-nearest neighbor model to predict the metabolic reactions of the NCI-60 set (a set of 60 human tumor cell lines derived from various tissues of origin) influenced by approved anti-cancer drugs, and extended their method to suggest possible enzyme targets for anti-cancer drugs. Through the analysis of cancer-specific human metabolic networks Folger et al. (2011) predicted 52 cytostatic drug targets, of which 40% were targeted by known anti-cancer drugs, and the rest were new target-candidates. However, it should be kept in mind that key enzymes of cancer-specific metabolism, such as the PKM2 isoenzyme of pyruvate kinase playing a predominant role in the Warburg-effect (Warburg, 1956; Steták et al., 2007; Christofk et al., 2008; Kung et al., 2012), were also shown to play a direct role in cancer-specific signaling (Gao et al., 2012). Both metabolic and signaling networks are directed networks, where a hierarchy can be established, and where targeting of upstream, more influential nodes may be a fruitful strategy (Liu et al., 2012). We will review the use of signaling networks in anti-cancer therapies in the next section.

#### 5.2.4. Signaling network targets of anti-cancer drugs

Signaling-related anti-cancer therapies increasingly outnumber metabolism-related chemotherapy options. From the network point of view this trend is due to the more developed signaling in humans than in pathogens, and to the increased selectivity of signaling interactions as compared to metabolism-related targeting.

Mass spectrometry can be effectively used for the analysis of post-translational modifications during the progression of cancer. Post-translational modifications, e.g. phosphorylation may change due to changes in the cellular environment and regulation under physiological conditions, but also due to a mutation at the phosphorylation site, or at a protein binding interface regulating kinase or phosphatase activity (Pawson & Linding, 2008). The bioinformatics resources NetworkKIN (<http://networkkin.info>; Linding et al., 2007) and NetPhorest (<http://netphorest.info>; Miller et al., 2008) can help in the analysis cancer-related signaling changes.

Rewiring of cancer-related changes of signaling networks is a primary aim in signal transduction-related anti-cancer therapies (Papatsoris et al., 2007). Cancer-specific changes in gene expression may activate or inactivate noncanonical edges in signal transduction networks (Klinke, 2010). Higher complexity of cancer-specific signaling network was shown to correlate with shorter survival (Breitkreutz et al., 2012). Proteins with cancer-related mutations are often hubs of human signaling network and are enriched in positive signaling regulatory loops (Awan et al., 2007; Cui et al., 2007; Cloutier & Wang, 2011; Li et al., 2012a). Alteration in cross-talking, multi-pathway, inter-modular proteins of signaling networks was proposed to be a key process in tumorigenesis (Hornberg et al., 2006; Taylor et al., 2009; Korcsmáros et al., 2010; Rivera et al., 2012).

The mammalian target of rapamycin (mTOR) is an important example of multi-pathway effects. mTOR has a key role in cell growth and regulation of cellular metabolism. In most tumors, mTOR is mutated, causing a hyper-active phenotype (Zoncu et al., 2011). Though mTOR activity was expected to be a promising therapeutic target, drugs showed poor results in clinical trials. mTOR could not meet node-targeting expectations because of its multi-pathway position, participating in at least two major signaling complexes, mTORC1 and mTORC2 (Huang et al., 2004; Caron et al., 2010; Catania et al., 2011; Pe'er & Hacohen, 2011; Fingar & Inoki, 2012).

Edgetic drugs specifically targeting mTOR interactions may selectively influence cancer-specific mTOR functions (Section 4.1.2; Ruffner et al., 2007). Another example of edgetic anti-cancer therapy options is that of nutlins, which block the interaction between p53 and its negative modulator MDM2 activating the tumor suppressor effect of p53 (Vassilev et al., 2004). Cancer-related proteins have smaller, more planar, more charged and less hydrophobic binding interfaces than other proteins, which may indicate low affinity and high specificity of cancer-related interactions (Kar et al., 2009). These structural features make lead compound development of cancer-related edgetic drugs a challenging task.

MicroRNAs are increasingly recognized as highly promising, non-protein intervention points of the signaling network (see also in Section 3.4). Loss- or gain-of-function mutations of microRNAs have been identified in nearly all solid and hematologic types of cancer (Calin & Croce, 2006; Spizzo et al., 2009). In addition, microRNAs were recently found as a form of intercellular communication (Chen et al., 2012c). Thus, alteration of microRNA content may have an effect on the microenvironment of tumor cells. Drug-induced changes in the expression of specific microRNAs can induce drug sensitivity leading to an increased inhibition of cell growth, of invasion and of metastasis formation (Sarkar et al., 2010). However, microRNAs have a dual role in cancer, acting both as oncogenes targeting mRNAs coding tumor-suppression proteins, or tumor suppressors targeting mRNAs coding oncoproteins (Iguchi et al., 2010; Gambari et al., 2011). This suggests the use of systems-level, network approaches to select microRNA targets.

Combination of cancer-specific mRNA and microRNA expression data may be used to infer cancer-specific regulatory networks (Bonnet et al., 2010). MicroRNAs involved in prostate cancer progression preferentially target interactome hubs (Budd et al., 2012). MicroRNA networks obtained from 3312 neoplastic and 1107 nonmalignant human samples showed the dysregulation of hub microRNAs. Cancer-specific microRNA networks had more disjointed subnetworks than those of normal tissues (Volinia et al., 2010). The fast growing complexity of signaling networks still awaits a comprehensive treatment in anticancer therapies.

#### 5.2.5. Influential nodes and edges in network dynamics as promising drug targets

As we mentioned in the introduction of Section 5, the central hit strategy of anti-cancer drug design would often hit a protein, which is also central in the networks of healthy cells. Therefore, here the more indirect targeting of the network influence strategy may also prove useful. Many targets of anti-cancer therapies are not directly cancer-related (Hornberg et al., 2006; Cheng et al., 2012c). Context can influence network behavior in at least four different ways: a.) the genetic background (e.g., single-nucleotide polymorphisms and other mutations); b.) gene expression changes (caused by e.g. transcription factor, epigenetic or microRNA changes); c.) neighboring cells; and finally d.) exogenous signals (e.g. nutrients or drugs) all providing increment to the patient-specific, context-dependent responses to anti-cancer therapy (Klinke, 2010; Sharma et al., 2010b; Pe'er & Hacohen, 2011). Differential gene expression and phosphorylation studies were already shown to be useful to distinguish among different stages of cancer development in the preceding sections. The next challenging step is to examine the cancer-induced dynamic changes on a network-level.

The examination of differential networks of cancer stages, or networks of drug treated and un-treated cells, is one of the first steps in possible solutions (Ideker & Krogan, 2012). Network level integration of cancer-related changes (such as mutations, gene expression changes, post-translational modifications, etc.) may capture key differences in network wiring (Pe'er & Hacohen, 2011).

Network dynamics may be assessed by the dissipation of perturbations, which can be used for the prioritization of drug target candidates. The early work of White and Mikulecky (1981) used a small network to assess the dynamics of methotrexate action. Sites et al. (2007) studied changes of Ras signaling in cancer using a differential equation model applied to a limited signaling network-set. They concluded that a hypothetical drug preferably binding to GDP-Ras would only induce a cancer-specific decrease in Ras signaling. Shiraishi et al. (2010) identified 6585 pairs of bistable toggle switch motifs in regulatory networks forming a network of 442 proteins. Among the 24 conditions examined, mRNA expression level changes reversed the ON/OFF status of a significantly high number of bistable toggle switches in various types of cancer, such as in lung cancer or in hepatocellular carcinoma. Serra-Musach et al. (2012) found that cancer-related proteins have an increased robustness to transmit a large volume of perturbations without being damaged. Extensions of such investigations to network-wide perturbations (modulated by neighboring cells and exogenous signals) will be an important research area for finding influential nodes/edges serving as drug target candidates.

#### 5.2.6. Drug combinations against cancer

As we have shown in the preceding sections cancer is a systems-level disease, where magic-bullet type drugs may fail. Partially redundant signaling pathways are hallmarks of cancer robustness. Thus an inhibitor of a particular hallmark may not be enough to block the related function. Moreover, when inhibitors of a specific cancer hallmark are used separately, they may even strengthen another hallmark, like certain types of angiogenesis inhibitors increased the rate of metastasis. In most failures of anti-cancer therapies, unwanted off-target effects and undiscovered feedbacks

prevented the desired pharmacological goal. Combination therapies and multi-target drugs may both overcome system robustness and provide less side-effects (see Section 4.1.5; Gupta et al., 2007; Berger & Iyengar, 2009; Wilson et al., 2009; Azmi et al., 2010; Glaser, 2010; Hanahan & Weinberg, 2011).

Cancer-specific subsets of the human interactome can provide a guide for the development of multi-target therapies. Mutually exclusive gene alterations which share the same biological process may define cancer type-related interactome modules (Ciriello et al., 2012b). Other types of cancer-related network modules were identified as sub-interactomes, as in colorectal cancer. These were centered on proteins, which markedly change their levels, and showed a synergistic dysregulation at their mRNA levels (Nibbe et al., 2010). Simultaneous targeting of these modules may be an efficient therapeutic strategy.

Multiple-targets can be identified using cancer-specific metabolic network models. Combinations of synthetic lethal drug targets were predicted in cancer-specific metabolic networks (Moreno-Sanchez et al., 2010; Folger et al., 2011). Ensemble modeling, which exploited a perturbation of known targets in a subset of 58 central metabolic reactions, was used to predict target sets of key enzymes of central energy metabolism (Khazaei et al., 2012).

Potential drug target sets were identified by an algorithm, which calculates the downstream components of a prostate cancer-specific signaling network affected by the inhibition of the target set (Dasika et al., 2006). In the particular example of EGF receptor inhibition, subsequent applications of drug combinations were shown to have a dramatically improved effect. This unmasked an apoptotic pathway, and via complex signaling network effects dramatically sensitized breast cancer cells to subsequent DNA-damage (Lee et al., 2012d). These findings substantiate Kitano's earlier emphasis on the importance of cancer chronotherapy (Kitano, 2004a, 2004b, 2007).

Tumors contain a highly heterogeneous cell population. Drug combinations may act via an intracellular network of a single cell; but also via inhibiting subsets of the heterogeneous population of malignant cells. Cell populations and their drug responses can be perceived as a bipartite graph. Applying minimal hitting set analysis allowed the search for effective drug combinations at the inter-cellular network level (Vazquez, 2009).

As we showed in this section, analysis of network topology and, especially, network dynamics can predict novel anti-cancer drug targets. Incorporation of personalized data, such as mutations, singalome or metabolome profiles to the molecular networks listed in this section may enhance patient- and disease stage-specific drug targeting in anti-cancer therapies.

### 5.3. Diabetes (metabolic syndrome including obesity, atherosclerosis and cardiovascular disease)

Diabetes is the first of our two examples showing the applications of the network influence strategy defined in Section 4.1.1, where therapeutic interventions need to push the cell back from the attractor of the diseased state to that of the healthy state. Diabetes is a multigenic disease tightly related to central obesity, atherosclerosis and cardiovascular disease, a connection also revealed by network representations (Ghazalpour et al., 2004; Lusic & Weiss, 2010; Stegmaier et al., 2010). Here we summarize network-related methods to predict novel drug target candidates in diabetes and related metabolic diseases.

Type 2 diabetes is the most common form of diabetes that is characterized by insulin resistance and relative insulin deficiency. T2D-db is a database of molecular factors involved in type 2 diabetes (<http://t2ddb.ibab.ac.in>; Agrawal et al., 2008) providing useful information for the construction of various diabetes-related networks. Combination of interactome and diabetes-related gene expression data identified the possible molecular basis of several endothelial, cardiovascular and kidney-related complications of diabetes, and revealed novel links between diabetes, obesity, oxidative stress and inflammatory

abnormalities (Sengupta et al., 2009; Mori et al., 2010). Similar studies suggested a network of protein–protein interactions bridging insulin signaling and the peroxisome proliferator-activated receptor-(PPAR)-related nuclear hormone receptor family (Liu et al., 2007b). Refinement of interactome data containing domain–domain interactions combined with the earlier observation that disease-related genes have a smaller than average clustering coefficient (Feldman et al., 2008) led to the prediction of type 2 diabetes-related genes (Sharma et al., 2010a). Inter-modular interactome nodes between type 2 diabetes-, obesity- and heart disease-related proteins may play a key role in the dysregulation of these complex syndromes (Nguyen & Jordan, 2010; Nguyen et al., 2011).

Type 1 diabetes is primarily related to the dysregulation of insulin secretion of pancreatic  $\beta$ -cells, where  $\beta$ -cell dedifferentiation was recently shown to play an important role (Talchai et al., 2012). The  $\beta$ -cell endoplasmic reticulum stress signaling network is an important regulator of this process (Fonseca et al., 2007; Mandl et al., 2009). Integration of interactome and genetic interaction data revealed novel protein network modules and candidate genes for type 1 diabetes (Bergholdt et al., 2007).

Reconstruction of changes of the human metabolic network of skeletal muscle in type 2 diabetes enabled the identification of potential new metabolic biomarkers. Analysis of gene promoters of proteins associated with the biomarker metabolites led to the construction of a diabetes-related transcription factor regulatory network (Zelezniak et al., 2010). Recently an integrated, manually curated and validated metabolic network of human adipocytes, hepatocytes and myocytes was assembled. Several metabolic states, such as the alanine-cycle, the Cori-cycle and an absorptive state, as well as their changes between obese and diabetic obese individuals were characterized (Bordbar et al., 2011). Such studies will highlight key enzymes of metabolic network, where a drug-induced activity and/or regulation change may significantly contribute to the rewiring of the metabolic network to its normal state.

Insulin signaling is in the center of the etiology of metabolic diseases. Several studies highlighted diabetes-responsible segments of the human signaling network enriching and re-focusing the traditionally known insulin signaling pathway. The mammalian target of rapamycin (mTOR) protein is one of the focal points of the insulin signaling network. From the two mTOR-related signaling complexes mentioned in the preceding section, Complex 1 (mTORC1) is a key player in nutrient-related signaling involving the hypothalamus, peripheral organs, adipose tissue differentiation and  $\beta$ -cell dependent insulin secretion (Catania et al., 2011; Fingar & Inoki, 2012). An siRNA knockout screen of 300 genes involved in the lipolysis of 3T3-L1 adipocytes led to the identification of a core, insulin resistance-related sub-network of the insulin signaling pathway highlighting a number of novel genes related to insulin-resistance, such as the sphingosine-1-phosphate receptor-2 (Tu et al., 2009). Reconstruction of the subnetwork of human interactome related to insulin signaling and the determination of its hubs and bottleneck proteins (Durmus Tekir et al., 2010) is an ongoing work, which will uncover many important novel targets of therapeutic interventions in the future. As an additional extension of insulin signaling, recent studies started to uncover the changes and most influential members of the microRNA regulatory network in diabetes (Huang et al., 2010c; Zampetaki et al., 2010). Phosphoproteome-studies help to extend the insulin signaling network further, and to uncover its time-dependent changes (Schmelzle et al., 2006).

Tissue-specific gene expression data identified metabolic disease-specific regulatory network modules, and revealed the involvement of both macrophages and the inflammatome in the pathogenesis of metabolic diseases (Schadt et al., 2009; Lusic & Weiss, 2010; Wang et al., 2012f). These studies show the inter-pathway and inter-organ complexity reached in the network understanding of metabolic disease. In Section 4.1.7 we summarized the network influence strategy; that is, to rewire the cellular networks from their diseased state to healthy state

as a tool to help in successful drug design. This includes avoiding network segments which are essential in healthy cells, and focusing on targeting pathway sites specific to diseased cells, and the use of multiple or indirect targeting. For this, metabolic disease network studies need to apply network dynamics methods such as we listed in Section 2.5. Systematic, network-based identification of edgetic, multi-target and allo-network drugs (see Section 4.1) could also be beneficial. Refined network methods should also incorporate patient- and disease stage-specific data. These are intimately related to the network consequences of aging, which will be described in the next section.

#### 5.4. Promotion of healthy aging and neurodegenerative diseases

Aging is one of the most complex processes of living organisms. Aging was described as a network phenomenon (Kirkwood & Kowald, 1997; Söti & Csermely, 2007; Simkó et al., 2009; Chautard et al., 2010). In the first half of this section we will summarize the few initial network studies on age-related multifactorial changes. Besides cancer and the metabolic syndrome described in the preceding sections, neurodegeneration is one of the major aging-associated diseases. In the concluding part of the section we will describe network-related studies on the prediction of potential drug targets to prevent and slow down various forms of neurodegeneration, such as Alzheimer's, Parkinson's, Huntington's and prion-related diseases.

##### 5.4.1. Aging as a network process

Aging organisms show similar early warning signals of critical phase transitions (i.e.: slower recovery from perturbations, increased self-similarity of behavior and increased variance of fluctuation-patterns) as described for a wide variety of complex systems (Section 2.5.2; Scheffer et al., 2009; Sornette & Osorio, 2011; Dai et al., 2012). Aging can be perceived as an early warning signal of a critical phase transition, where the phase transition itself is death (Farkas et al., 2011). However, this sobering message also has a positive implication: phase transitions of complex systems can be slowed down, postponed, or prevented by nodes having an independent and unpredictable behavior (Csermely, 2008). The identification of these nodes may lead to the discovery of novel molecular agents promoting healthy aging.

The complexity of the aging process is illustrated well by the duality of possible aging-related trends in network changes. Aging-related disorganization causes an increase of non-specific edges, and an aged organism has fewer resources predicting the loss of network edges during aging. Thus, small-worldness may often be lost during the aging process, and the hub-structure may get reorganized. Aging networks are likely to become more rigid, and may have less overlapping modules (Söti & Csermely, 2007; Csermely, 2009; Kiss et al., 2009; Simkó et al., 2009; Gáspár & Csermely, 2012). The longest documented lifespan is currently 122 years achieved by a French woman (Allard et al., 1998). It is currently an open question, whether lifespan has any upper limits. It will be interesting if future aging-related studies of network topology and behavior will predict any upper limit of human lifespan.

Aging-associated genes form an almost fully connected sub-interactome (also called longevity networks; Budovsky et al., 2007), and occupy both hub (Promislow, 2004; Ferrarini et al., 2005; Budovsky et al., 2007; Bell et al., 2009) and inter-modular positions (Xue et al., 2007). Aging-associated genes are concentrated in 4 modules of the yeast interactome (Barea & Bonatto, 2009). Similarly, age-related gene expression changes preferentially affect only a few modules of the human brain and *Drosophila* interactomes (Xue et al., 2007). The sub-interactome of aging-genes can be extended by their neighbors and the related network edges. The extended network provides a target-set to identify novel aging-related genes (Bell et al., 2009). The sub-interactomes of aging-associated genes and major age-related disease genes highly overlap with each other.

Aging-genes bridge other genes related to various diseases (Wang et al., 2009; Wolfson et al., 2009).

Longevity networks are enriched by key signaling proteins (Reja et al., 2009; Simkó et al., 2009; Wolfson et al., 2009; Borklu Yucel & Ulgen, 2011). The complexity of age-related processes is exemplified well by the extensive cross-talks of age-related signaling pathways (de Magalhaes et al., 2012). As an example, the growth hormone-related pathways, the oxidative stress-induced pathway and the dietary restriction pathway all affect the FOXO (Daf-16) transcription factor (Greer & Brunet, 2008). The yeast gene regulatory network was reconstituted by reverse engineering methods using age-associated transcriptional changes. The regulatory network revealed novel aging-associated regulatory components (Lorenz et al., 2009). MicroRNAs play an important role in aging-related signaling events (Chen et al., 2010b). Network analysis will help the identification of critical nodes of age-related signaling. These nodes may serve as potential targets of drugs promoting healthy aging.

During the aging process, the nuclear pore complexes become more permeable (D'Angelo et al., 2009). It is likely that age-induced increase of permeability is a general phenomenon involving other cellular compartments (Simkó et al., 2009) and increasing in the number of non-specific edges of the inter-organelle network.

Though drug development efforts are rapidly increasing in the field, currently there are only a few drugs which directly target the aging process (Simkó et al., 2009; de Magalhaes et al., 2012). To date, it is an open question, if central hit strategy-type or network influence strategy-type drug targeting (aiming to target key network nodes, or aiming to influence aging-related changes, respectively) will be the most efficient route for finding appropriate drugs for the promotion of healthy aging. Most probably the network influence strategy will be the 'winner', and the anti-aging drugs of the future will be multi-target drugs, providing an indirect influence on key processes of aging networks. For this additional studies on aging-related dynamics of molecular networks are needed.

##### 5.4.2. Network strategies against neurodegenerative diseases

As Lipton (2004) remarked, according to some predictions by 2050 the entire economy of the industrialized world could be consumed by the costs of caring for the sick and elderly. Neurodegenerative diseases, such as Alzheimer's, Parkinson's, Huntington's and prion diseases constitute one of the major aging-related disease-class besides cancer and metabolic diseases. Although several symptomatic drugs are available, a disease-modifying agent is still elusive making novel approaches especially valuable (Dunkel et al., 2012; Mei et al., 2012; Funke et al., in press).

We listed the major network-related methods to uncover novel neurodegenerative disease-associated genes, potential drug targets, or for drug repositioning in Table 11. Two major network methodologies emerge, which are widely used in connection with neurodegenerative diseases. One of them constructs, or extends disease-related protein-protein interaction networks and predicts novel disease-associated proteins. This appears a straightforward technique, since a neurodegenerative disease causes a major reconfiguration of cellular protein complexes. The other major method uses network analysis of differentially expressed genes in disease-affected patients or model organisms. This method identifies novel regulatory and signaling components involved in disease progression. Both methods may identify disease-affected pathways, which may be used to construct "heat-maps" identifying novel drug target candidates (Dunkel et al., 2012; Mei et al., 2012).

When summarizing neurodegenerative disease-related network efforts, it was surprising that, besides a few initial attempts in Alzheimer's disease, how little attention was devoted to chemical similarity networks, metabolic networks, signaling networks and drug-target networks in this field. Dysregulated, over-acting signaling pathways have a major contribution to all neurodegenerative diseases, and their network analysis would deserve more attention. A

**Table 11**  
Illustrative examples of network strategies against neurodegenerative diseases.

Type of network	Drug design benefit	References
<i>Alzheimer's disease</i>		
Protein–protein interaction network (extended with drug interactions)	Prediction of novel disease-related genes and novel disease-associated drugs from existing ones by interactome proximity	Krauthammer et al., 2004; Li et al., 2009a; Yang & Jiang, 2010; Hallock & Thomas, 2012; Raj et al., 2012
Differentially co-expressed gene networks of normal and Alzheimer's disease affected patients	Identification of co-expressed gene modules and disease-related transcription factors	Ray et al., 2008; Satoh et al., 2009; Liang et al., 2012
Network of differentially expressed MicroRNAs of Alzheimer's affected patients	Prediction of novel disease-associated signaling pathways and regulators	Satoh, 2012
Drug binding site similarity networks	Prediction of novel drug targets	Yang et al., 2010
Drug target networks of anti-Alzheimer's herbal medicines	Prediction of novel drug targets	Sun et al., in press
<i>Parkinson's disease</i>		
Differentially co-expressed gene networks of normal and Parkinson's disease affected patients	Identification of central disease-associated genes	Moran & Graeber, 2008
<i>Poly-glutamine (polyQ) expansion diseases (Huntington's disease, ataxias)</i>		
Protein–protein interaction network of poly-glutamine proteins (and their known interactome neighbors)	Identification of novel modifiers of disease progression	Goehler et al., 2004; Lim et al., 2006; Kaltenbach et al., 2007; Kahle et al., 2011
<i>Prion disease</i>		
Differentially co-expressed gene networks of normal and prion disease affected mice	Identification of disease-associated pathways and modules	Hwang et al., 2009; Kim et al., 2011b

good anti-neurodegenerative drug is typically a network influence strategy-type drug reconfiguring the distorted pathways of disease-associated networks (Lipton, 2004; Dunkel et al., 2012). Learning more on changes in network dynamics during neurodegenerative disease progression would be a major advance of drug design efforts in this crucially important field.

## 6. Conclusions and perspectives

The value of every drug design technology must be assessed by asking: “How much does the new technology help to solve one of the two central problems: the identification and validation of a disease-specific target or the identification of a molecule that can modify this target in a way that makes therapeutic sense?” (Brown & Superti-Furga, 2003; Drews, 2003). Network description and analysis may offer novel leads in both questions. In this concluding section we will highlight the promises and perspectives of network-aided drug development.

### 6.1. Promises and optimization of network-aided drug development

One of the major promises of network description and analysis is their help to overcome the “one-effect/one-cause/one-target” magic bullet-type drug development paradigm (Ehrlich, 1908). Magic bullets do work—sometimes. When designing “central hit strategy-type” drugs (Sections 4.1.1 and 4.1.7), which target key nodes of the network to eliminate pathogens or malignant cells, eradicating single hit may be beneficial. However, pathogen resistance or unexpected toxicity of anti-cancer drugs (and resistance against them) may dog the outcome. In the development of “network influence strategy-type drugs”, where an efficient reconfiguration of rigid networks needs to be achieved, network dynamics has to be reset from its disease-affected state back to normal (Sections 4.1.1 and 4.1.7). Under these circumstances, the traditional approach of rational drug discovery selecting a single and central target often fails. The paucity of disease-modifying anti-neurodegenerative drugs described in the preceding section is an example for the need for novel approaches in network influence strategy-type drug design.

James Black described well a wide-spread behavior saying that ‘the most fruitful basis for the discovery of a new drug is to start with an old drug’ (Chong & Sullivan, 2007). We started our review

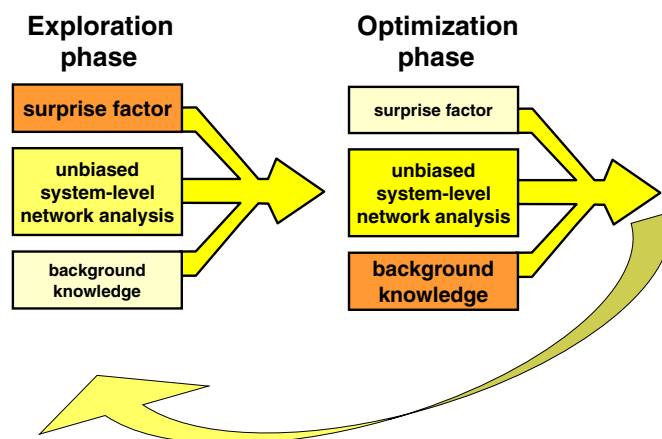
with the statement that ‘business as usual’ is no longer an option in drug industry (Begley & Ellis, 2012). Currently, there is a broad consensus that this state, where the vast majority of new drugs are related to existing ones, needs to be changed (Section 1.1; Cokol et al., 2005; Yildirim et al., 2007; Iyer et al., 2011a).

Nonetheless, the question is how to find ‘surprisingly novel drugs’. The failure of some efforts using the reductionist approach of rational drug design shifted the thinking to the other extreme that ‘we need unbiased research methods to cover complexity’. Indeed, unbiased methods (including network analysis or machine learning) may successfully predict novel drug targets. However, clearly, artificial intelligence may miss ‘true’ surprises (Section 2.2.2). This leads to our first major conclusion, which we summarized in Fig. 23: network description and analysis should be combined with human creativity and background knowledge.

Network analysis helps in comprehending the vast amounts of systems-level data, which accumulated over the last decade. However, network analysis alone is clearly insufficient and has to be complemented with the intuitions coming from background knowledge (Valente, 2010). In this process creativity (marked as the ‘surprise factor’ in Fig. 23), which strives for novelty, and identifies it in networks as the ‘prediction of the unpredictable’ (Section 2.2.2) cannot be overlooked. Combined, the current boom in network dynamics-related methods can help in discovering the key actors in the cellular community, which are the hidden masterminds of cellular changes in health and in disease (Fig. 23).

Our second major conclusion is that a protocol of network-aided drug development would be aided by alternating exploration and optimization phases of drug design (Fig. 23). In the exploration phase background knowledge may be temporarily suppressed. In contrast, in the optimization phase we need to suppress the playfulness and ambiguity tolerance of the exploration phase, and rank the options by rigorous application of background knowledge including all the well-orchestrated rules of the drug development process (Csermely, in press; Gyurkó et al., in press). Importantly, the sequence of exploration and optimization phases may be applied repeatedly, providing a more detailed ‘zoom-in’ of the optimal (drug) target than a single round of exploration/optimization (Fig. 23). The utility of repeated exploration/optimization rounds was shown in a number of examples ranging from thermal cycles of simulated annealing optimization





**Fig. 23.** Optimized protocol of network-aided drug development. The figure illustrates the two major phases of discovery having three segments marked as boxes on the left side of the triple arrows. The “surprise factor” box denotes originality (as the highest level of human creativity), a strong drive to discover the unexpected, including playfulness and ambiguity tolerance. The “unbiased systems-level network analysis” box marks the network methods described in this review. The “background knowledge” box includes all our contextual, background knowledge on diseases, drugs and their actions, as well as the validation procedures guiding our judgment on the quality of the drug discovery process. In the exploration phase the surprise factor is dominant. At this phase background knowledge may be temporarily suppressed. On the contrary, at the optimization phase we need to suppress the surprise factor, and rank our previous options by the rigorous application of our background knowledge. The arrow at the bottom of the figure marks the heretofore not rigorously applied method that the sequence of exploration and optimization phases may be applied repeatedly, which gives a much more precise ‘zoom-in’ to the optimal (drug) target than a single round of exploration/optimization.

(Möbius et al., 1997) to the human learning process (Bassett et al., 2011).

Drug design-related networks became increasingly complex during the past decade. Albert Einstein's saying that “the supreme goal of all theory is to make the irreducible basic elements as simple and as few as possible without having to surrender the adequate representation of a single datum of experience” (Einstein, 1934) (also called ‘Einstein's razor’, extending the Occam's razor theorem advocating only the simplest solution) encourages finding the optimal network representation, which is simple enough, but not too simple. Finding the optimal complexity of network representation in the drug discovery process is an important task. Eventually, the recurring application of exploration and optimization phases shown in Fig. 23 suggests that network data coverage should be extended in consecutive phases separated by recurrent network simplifications based on background knowledge.

Thomas Singer wrote a few years ago “Extrapolation of preclinical data into clinical reality is a translational science and remains an ultimate challenge in drug development.” (Singer, 2007). Addressing this challenge our third and last major conclusion stresses the importance of network prediction of these human data, which are not available experimentally. This includes overcoming three major hurdles (Fig. 2; Brown & Superti-Furga, 2003; Austin, 2006; Bunnage, 2011; Ledford, 2012): 1.) insufficient drug efficacy; 2.) unexpected major adverse effects; 3.) unexpected forms of human toxicity. Network analysis may help ameliorate the efficacy by taking into account patient-, disease stage-, age-specificities (Section 4.3.1); it may help obtain a better prediction of side-effects (Section 4.3.5) and predictive human toxicology (Section 4.3.3; Henney & Superti-Furga, 2008).

Network science is a novel area of biology; and this is particularly the case with respect to drug design. We often lack rigorous comparisons of existing methods, which could have allowed a more critical approach to some of them. It is an ongoing effort of the current years to develop benchmarks, gold-standards and rigorous assessment tools in network science.

## 6.2. Systems-level hallmarks of drug quality and trends of network-aided drug development helping to achieve them

In this closing section we identify the systems-level hallmarks of drug quality, and list the major trends of network-aided drug development helping to achieve them. From the network point of view

we propose two strategies in finding drug targets: 1.) the central hit strategy aiming to destroy the network of infectious agents or cancer cells and 2.) the network influence strategy aiming to shift the network dynamics of polygenic, complex diseases back to normal (Fig. 19; Sections 4.1.1 and 4.1.7). Both strategies converge to the same level of network complexity in hit finding, hit expansion, lead selection and optimization phases. We note that these two strategies appear as general strategies to design the most efficient attack of flexible, plastic systems (using the central hit strategy) or rigid systems (using the network influence strategy) valid from molecular structures, through molecular and cellular networks to social and engineered networks or ecosystems.

Table 12 lists the systems-level hallmarks of drug target identification and validation, hit finding and development, as well as lead selection and optimization. We believe that the systematic application of these systems-level hallmarks will not only help the identification of novel drug targets, but will also streamline the drug design process to be more selective, less attrition-prone and more profitable.

Table 12 also includes the most important network-related drug design trends helping the accomplishment of various systems-level hallmarks. We highlight the development of edgetic drugs (Section 4.1.2), multi-target drugs (Section 4.1.5) and allo-network drugs (Section 4.1.6) among the richness of network strategies to find novel drug targets. We believe that there are a large number of unexplored drug targets, which are the hidden masterminds of cellular regulation. Analysis of network dynamics can help to find them. Incorporation of disease-stage, age-, gender- and human population-specific genetic, metabolome, phosphoproteome and gut microbiome data; the development of human ADME and toxicity network models; and the use of side-effect networks to judge drug safety, may greatly increase the efficiency of the drug development process.

Network-related methods—if applied systematically (and carefully)—will uncover a number of novel drug targets, and will increase the efficiency of the drug development process. Analysis of the structure and dynamics of molecular networks, extended by the network dynamics of constituting proteins and in particular their binding sites, provides a novel paradigm of drug discovery.

## Conflict of interest statement

The authors declare that there are no conflicts of interest.

**Table 12**  
Systems-level hallmarks of drug quality and trends of network-related drug design helping to achieve them.

Systems-level hallmark of drug quality	Network-related drug design trend
<p>Drug target identification</p> <ul style="list-style-type: none"> <li>Using the central hit strategy: drug hits central (or otherwise essential) network nodes, whose efficient inhibition selectively destroys infectious agent or cancer cell</li> <li>Using the network influence strategy: drug hits disease-specific network segments (nodes, edges or their sets), whose manipulation shifts disease-affected functions back to normal</li> </ul> <p>Drug target validation</p> <ul style="list-style-type: none"> <li>Network dynamics-based, disease-specific early and robust human biomarkers are used for drug target validation, drug added-value assessment over current standard care, and translation for later monitoring in clinical trials</li> </ul>	<ul style="list-style-type: none"> <li>Disease stage-related interactome, signaling network, metabolic network data (including protein abundance, human/comparative genetic data and microRNA profiles, optionally combined with protein, RNA and chromatin structure information, as well as with subcellular localization)</li> <li>Network comparison and reverse engineering</li> <li>Disease-specific models of network dynamics (including deconvolution, perturbation, hierarchy, source/sink/seeder analysis and network influence)</li> <li>Drug target, patient and therapy-related networks helping multi-target design and drug repositioning</li> <li>Use of weighted, directed, signed, colored and conditional edges or hypergraph structures</li> <li>Network prediction methods (sensitized for finding the unexpected)</li> <li>In the central hit strategy: network centrality measures; host/parasite, host/cancer network combinations at the local ecosystem level</li> <li>In the network influence strategy: network controllability, influence, dynamic network centrality; compensatory deletions; edgetic, multi-target and allo-network drugs; chrono-therapies (temporal shifts in administration of drug combinations) besides application of the trends listed above</li> <li>Complex chemical similarity (QSAR) networks including chemical similarity networks, multi-QSAR networks, pocket similarity networks, and chemical descriptors of ligand binding sites, integrated bio-entity networks</li> <li>Chemical reaction networks</li> <li>Network analysis of protein structures and correlated segments of protein dynamics</li> <li>Analysis of the substrate envelope to avoid drug resistance development</li> <li>Hot spot, and hot region identification</li> </ul>
<p>Hit finding and development</p> <ul style="list-style-type: none"> <li>Hit finding and ranking are helped by network chemoinformatics</li> <li>Hit expansion and library design are helped by chemical reaction networks</li> </ul>	<p>Besides application of the trends listed above</p> <ul style="list-style-type: none"> <li>Complex chemical similarity (QSAR) networks including chemical similarity networks, multi-QSAR networks, pocket similarity networks, and chemical descriptors of ligand binding sites, integrated bio-entity networks</li> <li>Chemical reaction networks</li> <li>Network analysis of protein structures and correlated segments of protein dynamics</li> <li>Analysis of the substrate envelope to avoid drug resistance development</li> <li>Hot spot, and hot region identification</li> </ul>
<p>Lead selection and optimization</p> <ul style="list-style-type: none"> <li>Optimization of drug efficacy, selection of robust efficacy end-points and patient populations are guided by network pharmacogenomics, as well as by disease-stage, age-, gender- and population-specific metabolome, phosphoproteome and gut microbiome data</li> <li>ADME and toxicity data are 'humanized', side-effect, drug–drug interaction and drug resistance evaluation are helped, as well as indications and contraindications are defined by extensive network data</li> </ul>	<ul style="list-style-type: none"> <li>Network extension by disease-stage, age-, gender- and human population-specific genetic, metabolome, phosphoproteome and gut microbiome data</li> <li>Analysis of semantic networks from medical records</li> <li>Human ADME and toxicity network models</li> <li>Network methods of multi-target drug design to uncover adverse drug–drug interactions</li> <li>Assessment of side-effect networks</li> <li>Antagonistic drug combinations to avoid drug resistance development</li> </ul>

## Acknowledgments

Authors thank Aditya Barve and Andreas Wagner (University of Zürich, Switzerland) for sharing the human homology of enzymes encoding superessential metabolic reactions, Haiyuan Yu, Xiujuan Wang (Department of Biological Statistics and Computational Biology, Weill Institute for Cell and Molecular Biology, Cornell University, Ithaca NY, USA) and Balázs Papp (Szeged Biological Centre, Hungarian Academy of Sciences, Szeged, Hungary) for the critical reading of Sections 1.3.3 and 3.6, respectively. Authors thank Zoltán P. Spiró (École Polytechnique Federale de Lausanne, Switzerland) for help in drawing Fig. 11, and the anonymous referee, members of the LINK-Group ([www.linkgroup.hu](http://www.linkgroup.hu)), as well as more than twenty additional colleagues for reading the original version of the paper and for valuable suggestions. Work in the authors' laboratory was supported by research grants from the Hungarian National Science Foundation (OTKA K83314), by the EU (TÁMOP-4.2.2/B-10/1-2010-0013), by a Bolyai Fellowship of the Hungarian Academy of Sciences (TK) and by a residence at the Rockefeller Foundation Bellagio Center (PC). This project has been funded, in part, with federal funds from the NCI, NIH, under contract HHSN261200800001E. This research was supported, in part, by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

## References

Abdi, A., Tahoori, M. B., & Emamian, E. S. (2008). Fault diagnosis engineering of digital circuits can identify vulnerable molecules in complex cellular pathways. *Sci Signal* 1, ra10.

- Acharyya, S., Oskarsson, T., Vanharanta, S., Malladi, S., Kim, J., Morris, P. G., et al. (2012). A CXCL1 paracrine network links cancer chemoresistance and metastasis. *Cell* 150, 165–178.
- Adamcsek, B., Palla, G., Farkas, I. J., Derenyi, I., & Vicsek, T. (2006). CFinder: locating cliques and overlapping modules in biological networks. *Bioinformatics* 22, 1021–1023.
- Adams, J. C., Keiser, M. J., Basuino, L., Chambers, H. F., Lee, D. S., Wiest, O. G., et al. (2009). A mapping of drug space from the viewpoint of small molecule metabolism. *PLoS Comput Biol* 5, e1000474.
- Adar, E. (2006). GUESS: A language and interface for graph exploration. In R. Grinter, T. Rodden, P. Aoki, E. Cutrell, R. Jeffries, & G. Olson (Eds.), *CHI '06 Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 791–800). New York: Association for Computing Machinery.
- Agoston, V., Csermely, P., & Pongor, S. (2005). Multiple, weak hits confuse complex systems: a transcriptional regulatory network as an example. *Phys Rev E* 71, 051909.
- Agrawal, S., Dimitrova, N., Nathan, P., Udayakumar, K., Lakshmi, S. S., Sriram, S., et al. (2008). T2D-Db: an integrated platform to study the molecular basis of type 2 diabetes. *BMC Genomics* 9, 320.
- Agüero, F., Al-Lazikani, B., Aslett, M., Berriman, M., Buckner, F. S., Campbell, R. K., et al. (2008). Genomic-scale prioritization of drug targets: the TDR Targets database. *Nature Rev Drug Discov* 7, 900–907.
- Ahmed, A., & Xing, E. P. (2009). Recovering time-varying networks of dependencies in social and biological studies. *Proc Natl Acad Sci USA* 106, 11878–11883.
- Ahmed, A., Dwyer, T., Forster, M., Fu, X., Ho, J., Hong, S. -H., et al. (2006). GEOMI: geometry for maximum insight. *Lect Notes Comput Sci* 3843, 468–479.
- Ahmed, N. K., Neville, J., & Kompella, R. (2012). Network sampling: from static to streaming graphs. <http://arxiv.org/abs/1211.3412>
- Ahn, Y. -Y., Bagrow, J. P., & Lehmann, S. (2010). Link communities reveal multi-scale complexity in networks. *Nature* 466, 761–764.
- Ajmani, S., & Kulkarni, S. A. (2012). Application of QSAR for scaffold hopping and lead optimization in multitarget inhibitors. *Mol Inf* 31, 473–490.
- Akula, N., Baranova, A., Seto, D., Solka, J., Nalls, M. A., Singleton, A., et al. (2011). A network-based approach to prioritize results from genome-wide association studies. *PLoS One* 6, e24220.
- Akutsu, T., Miyano, S., & Kuhara, S. (1999). Identification of genetic networks from a small number of gene expression patterns under the Boolean network model. *Pac Symp Biocomput* 1999, 17–28.
- Albert, I., & Albert, R. (2004). Conserved network motifs allow protein–protein interaction prediction. *Bioinformatics* 20, 3346–3352.
- Albert, R., Jeong, H., & Barabasi, A. L. (2000). Error and attack tolerance of complex networks. *Nature* 406, 378–382.
- Albert, I., Thakar, J., Li, S., Zhang, R., & Albert, R. (2008). Boolean network simulations for life scientists. *Source Code Biol Med* 3, 16.

- Alexiou, P., Vergoulis, T., Gleditsch, M., Prekas, G., Dalamagas, T., Megraw, M., et al. (2010). miRGen 2.0: a database of microRNA genomic information and regulation. *Nucleic Acids Res* 38, D137–D141.
- Ali, M. A., & Sjoblom, T. (2009). Molecular pathways in tumor progression: from discovery to functional understanding. *Mol Biosyst* 5, 902–908.
- Allard, M., Lebre, V., Robine, J.-M., & Calment, J. (1998). *Jeanne Calment: From Van Gogh's Time to Ours: 122 Extraordinary Years*. New York: W.H. Freeman.
- Almaas, E., Kovacs, B., Vicsek, T., Oltvai, Z. N., & Barabasi, A. L. (2004). Global organization of metabolic fluxes in the bacterium *Escherichia coli*. *Nature* 427, 839–843.
- Almaas, E., Oltvai, Z. N., & Barabasi, A. L. (2005). The activity reaction core and plasticity of metabolic networks. *PLoS Comput Biol* 1, e68.
- Alonso, A., Sasin, J., Bottini, N., Friedberg, I., Osterman, A., Godzik, A., et al. (2004). Protein tyrosine phosphatases in the human genome. *Cell* 117, 699–711.
- Altay, G. (2012). Empirically determining the sample size for large-scale gene network inference algorithms. *IET Syst Biol* 6, 35–43.
- Alves, N. A., & Martinez, A. S. (2007). Inferring topological features of proteins from amino acid residue networks. *Phys A* 375, 336–344.
- Amin, M. S., Finley, R. L., Jr., & Jamil, H. M. (2012). Top-k similar graph matching using TraM in biological networks. *IEEE/ACM Trans Comput Biol Bioinform* 9, 1790–1804.
- Andersson, C. D., Chen, B. Y., & Linusson, A. (2010). Mapping of ligand-binding activities in proteins. *Proteins* 78, 1408–1422.
- Annibale, A., & Coolen, A. C. C. (2011). What you see is not what you get: how sampling affects macroscopic features of biological networks. *Interface Focus* 1, 836–856.
- Antal, M. A., Bode, C., & Csermely, P. (2009). Perturbation waves in proteins and protein networks: applications of percolation and game theories in signaling and drug design. *Curr Prot Prot Sci* 10, 161–172.
- Antonov, A. V., Dietmann, S., Rodchenkov, I., & Mewes, H. W. (2009). PPI spider: a tool for the interpretation of proteomics data in the context of protein–protein interaction networks. *Proteomics* 9, 2740–2749.
- Apel, A., Zentgraf, H., Buchler, M. W., & Herr, I. (2009). Autophagy—a double-edged sword in oncology. *Int J Cancer* 125, 991–995.
- Apic, G., Ignjatovic, T., Boyer, S., & Russell, R. B. (2005). Illuminating drug discovery with biological pathways. *FEBS Lett* 579, 1872–1877.
- Arkin, M. R., & Wells, J. A. (2004). Small-molecule inhibitors of protein–protein interactions: progressing towards the dream. *Nat Rev Drug Discov* 3, 301–317.
- Arrell, D. K., & Terzic, A. (2010). Network systems biology for drug discovery. *Clin Pharmacol Ther* 88, 120–125.
- Artymiuk, P. J., Rice, D. W., Mitchell, E. M., & Willett, P. (1990). Structural resemblance between the families of bacterial signal-transduction proteins and of G proteins revealed by graph theoretical techniques. *Protein Eng Des Sel* 4, 39–43.
- Assenov, Y., Ramirez, F., Schelhorn, S. E., Lengauer, T., & Albrecht, M. (2008). Computing topological parameters of biological networks. *Bioinformatics* 24, 282–284.
- Atias, N., & Sharan, R. (2011). An algorithmic framework for predicting side effects of drugs. *J Comput Biol* 18, 207–218.
- Atilgan, A. R., Akan, P., & Baysal, C. (2004). Small-world communication of residues and significance for protein dynamics. *Biophys J* 86, 85–91.
- Audouze, K., Juncker, A. S., Roque, F. J., Krysiak-Baltyn, K., Weinhold, N., Taboureau, O., et al. (2010). Deciphering diseases and biological targets for environmental chemicals using toxicogenomics networks. *PLoS Comput Biol* 6, e1000788.
- Austin, O. H. (2006). Research and development in pharmaceutical industry. *A Congressional Budget Office Study*. <http://www.cbo.gov/publication/18176>
- Avin, C., Lotker, Z., Pignolet, Y.-A., & Turkel, I. (2011). From Caesar to Twitter: an axiomatic approach to elites of social networks. <http://arxiv.org/abs/1111.3374>
- Awan, A., Bari, H., Yan, F., Moksong, S., Yang, S., Chowdhury, S., et al. (2007). Regulatory network motifs and hotspots of cancer genes in a mammalian cellular signalling network. *IET Syst Biol* 1, 292–297.
- Ay, F., Kellis, M., & Kahveci, T. (2011). SubMAP: aligning metabolic pathways with subnetwork mappings. *J Comput Biol* 18, 219–235.
- Ay, F., Dang, M., & Kahveci, T. (2012). Metabolic network alignment in large scale by network compression. *BMC Bioinforma* 13, S2.
- Azmi, A. S., Wang, Z., Philip, P. A., Mohammad, R. M., & Sarkar, F. H. (2010). Proof of concept: network and systems biology approaches aid in the discovery of potent anticancer drug combinations. *Mol Cancer Ther* 9, 3137–3144.
- Azuaje, F., Devaux, Y., & Wagner, D. R. (2010). Identification of potential targets in biological signalling systems through network perturbation analysis. *Biosystems* 100, 55–64.
- Azuaje, F. J., Zhang, L., Devaux, Y., & Wagner, D. R. (2011). Drug–target network in myocardial infarction reveals multiple side effects of unrelated drugs. *Sci Rep* 1, 52.
- Bader, G. D., Cary, M. P., & Sander, C. (2006). Pathguide: a pathway resource list. *Nucleic Acids Res* 34, D504–D506.
- Baggs, J. E., Hughes, M. E., & Hogenesch, J. B. (2010). The network as the target. *Wiley Interdiscip Rev Syst Biol Med* 2, 127–133.
- Bagler, G., & Sinha, S. (2005). Network properties of protein structures. *Phys A* 346, 27–33.
- Baitaluk, M., Kozhenkov, S., Dubinina, Y., & Ponomarenko, J. (2012). IntegromeDB: an integrated system and biological search engine. *BMC Genomics* 13, 35.
- Bajorath, J., Peltason, L., Wawer, M., Guha, R., Lajiness, M. S., & Van Drie, J. H. (2009). Navigating structure–activity landscapes. *Drug Discov Today* 14, 698–705.
- Balaji, S., McClendon, C., Chowdhary, R., Liu, J. S., & Zhang, J. (2012). IMID: integrated molecular interaction database. *Bioinformatics* 28, 747–749.
- Bandyopadhyay, S., & Bhattacharyya, M. (2010). PuTmiR: a database for extracting neighboring transcription factors of human microRNAs. *BMC Bioinforma* 11, 190.
- Bandyopadhyay, S., Chiang, C. Y., Srivastava, J., Gersten, M., White, S., Bell, R., et al. (2010). A human MAP kinase interactome. *Nat Methods* 7, 801–805.
- Banerjee, S. J., & Roy, S. (2012). Key to network controllability. <http://arxiv.org/abs/1209.3737>
- Barabási, A. L., & Albert, R. (1999). Emergence of scaling in random networks. *Science* 286, 509–512.
- Barabási, A. L., & Oltvai, Z. N. (2004). Network biology: understanding the cell's functional organization. *Nat Rev Genet* 5, 101–113.
- Barabási, A. L., Gulbahce, N., & Loscalzo, J. (2011). Network medicine: a network-based approach to human disease. *Nat Rev Genet* 12, 56–68.
- Barea, F., & Bonatto, D. (2009). Aging defined by a chronologic-replicative protein network in *Saccharomyces cerevisiae*: an interactome analysis. *Mech Ageing Dev* 130, 444–460.
- Baricic, P., & Mackov, M. (1995). MOLGEN: personal computer-based modeling system. *J Mol Graph* 13(184–189), 198.
- Barr, A. J. (2010). Protein tyrosine phosphatases as drug targets: strategies and challenges of inhibitor development. *Futur Med Chem* 2, 1563–1576.
- Barve, A., Rodrigues, J. F., & Wagner, A. (2012). Superessential reactions in metabolic networks. *Proc Natl Acad Sci USA* 109, E1121–E1130.
- Bar-Yam, Y. (2004). *Making Things Work: Solving Complex Problems in a Complex World*. Cambridge: New England Complex Systems Institute Knowledge Press.
- Bar-Yam, Y., Harmon, D., & de Bivort, B. (2009). Systems biology. Attractors and democratic dynamics. *Science* 323, 1016–1017.
- Bassett, D. S., Wymbs, N. F., Porter, M. A., Mucha, P. J., Carlson, J. M., & Grafton, S. T. (2011). Dynamic reconfiguration of human brain networks during learning. *Proc Natl Acad Sci USA* 108, 7641–7646.
- Basso, K., Margolin, A. A., Stolovitzky, G., Klein, U., Dalla-Favera, R., & Califano, A. (2005). Reverse engineering of regulatory networks in human B cells. *Nat Genet* 37, 382–390.
- Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: an open source software for exploring and manipulating networks. *International AAAI Conference on Weblogs and Social Media*. <http://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154/>
- Batagelj, V., & Mrvar, A. (1998). Pajek—program for large network analysis. *Connections* 21, 47–57.
- Bauer-Mehren, A., Rautschka, M., Sanz, F., & Furlong, L. I. (2010). DisGeNET: a Cytoscape plugin to visualize, integrate, search and analyze gene–disease networks. *Bioinformatics* 26, 2924–2926.
- Bauer-Mehren, A., Bundschuh, M., Rautschka, M., Mayer, M. A., Sanz, F., & Furlong, L. I. (2011). Gene–disease network analysis reveals functional modules in Mendelian, complex and environmental diseases. *PLoS One* 6, e20284.
- Becker, S. A., & Palsson, B. O. (2008). Three factors underlying incorrect *in silico* predictions of essential metabolic genes. *BMC Syst Biol* 2, 14.
- Becker, M. Y., & Rojas, I. (2001). A graph layout algorithm for drawing metabolic pathways. *Bioinformatics* 17, 461–467.
- Becker, D., Selbach, M., Rollenhagen, C., Ballmaier, M., Meyer, T. F., Mann, M., et al. (2006). Robust *Salmonella* metabolism limits possibilities for new antimicrobials. *Nature* 440, 303–307.
- Begley, C. G., & Ellis, L. M. (2012). Drug development: raise standards for preclinical cancer research. *Nature* 483, 531–533.
- Behar, M., Dohlman, H. G., & Elston, T. C. (2007). Kinetic insulation as an effective mechanism for achieving pathway specificity in intracellular signaling networks. *Proc Natl Acad Sci USA* 104, 16146–16151.
- Bell, R., Hubbard, A., Chettier, R., Chen, D., Miller, J. P., Kapahi, P., et al. (2009). A human protein interaction network shows conservation of aging processes between human and invertebrate species. *PLoS Genet* 5, e1000414.
- Bell, L., Chowdhary, R., Liu, J. S., Niu, X., & Zhang, J. (2011). Integrated bio-entity network: a system for biological knowledge discovery. *PLoS One* 6, e21474.
- Ben-Jacob, E., Coffey, D. S., & Levine, H. (2012). Bacterial survival strategies suggest rethinking cancer cooperativity. *Trends Microbiol* 20, 403–410.
- Bender, A., & Glen, R. C. (2004). Molecular similarity: a key technique in molecular informatics. *Org Biomol Chem* 2, 3204–3218.
- Bender, A., Jenkins, J. L., Scheiber, J., Sukuru, S. C., Glick, M., & Davies, J. W. (2009). How similar are similarity searching methods? A principal component analysis of molecular descriptor space. *J Chem Inf Model* 49, 108–119.
- Bender-deMoll, S., & McFarland, D. A. (2006). The art and science of dynamic network visualization. *J Social Struct* 7, 2.
- Benz, R. W., Swamidass, S. J., & Baldi, P. (2008). Discovery of power-laws in chemical space. *J Chem Inf Model* 48, 1138–1151.
- Berg, E. L., Hytopoulos, E., Plavec, I., & Kunkel, E. J. (2005). Approaches to the analysis of cell signaling networks and their application in drug discovery. *Curr Opin Drug Discov Devel* 8, 107–114.
- Berger, S. I., & Iyengar, R. (2009). Network analyses in systems pharmacology. *Bioinformatics* 25, 2466–2472.
- Berger, S. I., & Iyengar, R. (2011). Role of systems pharmacology in understanding drug adverse events. *Wiley Interdiscip Rev Syst Biol Med* 3, 129–135.
- Berger, S. I., Posner, J. M., & Ma'ayan, A. (2007). Genes2Networks: connecting lists of gene symbols using mammalian protein interaction databases. *BMC Bioinforma* 8, 372.
- Bergholdt, R., Stirling, Z. M., Lage, K., Karlberg, E. O., Olason, P. I., Aalund, M., et al. (2007). Integrative analysis for finding genes and networks involved in diabetes and other complex diseases. *Genome Biol* 8, R253.
- Berlingerio, M., Koutra, D., Eliassi-Rad, T., & Faloutsos, C. (2012). NetSimile: a scalable approach to size-independent network similarity. <http://arxiv.org/abs/1209.2684>
- Besnard, J., Ruda, G. F., Setola, V., Abecassis, K., Rodriguez, R. M., Huang, X. P., et al. (2012). Automated design of ligands to polypharmacological profiles. *Nature* 492, 215–220.
- Bhardwaj, N., Abyzov, A., Clarke, D., Shou, C., & Gerstein, M. B. (2011). Integration of protein motions with molecular networks reveals different mechanisms for permanent and transient interactions. *Protein Sci* 20, 1745–1754.
- Bhardwaj, N., Yan, K. K., & Gerstein, M. B. (2010). Analysis of diverse regulatory networks in a hierarchical context shows consistent tendencies for collaboration in the middle levels. *Proc Natl Acad Sci USA* 107, 6841–6846.

- Bhattacharyya, R. P., Remenyi, A., Yeh, B. J., & Lim, W. A. (2006). Domains, motifs, and scaffolds: the role of modular interactions in the evolution and wiring of cell signaling circuits. *Annu Rev Biochem* 75, 655–680.
- Bickerton, G. R., Paolini, G. V., Besnard, J., Muresan, S., & Hopkins, A. L. (2012). Quantifying the chemical beauty of drugs. *Nat Chem* 4, 90–98.
- Bishop, K. J. M., Klajn, L., & Grzybowski, B. A. (2006). The core and most useful molecules in organic chemistry. *Angew Chem Intl Ed* 45, 5348–5354.
- Blackstrom, L., Boldi, P., Rosa, M., Ugander, J., & Vigna, S. (2011). Four degrees of separation. <http://arxiv.org/abs/1111.4570>
- Blank, L. M., Kuepfer, L., & Sauer, U. (2005). Large-scale 13C-flux analysis reveals mechanistic principles of metabolic network robustness to null mutations in yeast. *Genome Biol* 6, R49.
- Bleakley, K., & Yamanishi, Y. (2009). Supervised prediction of drug–target interactions using bipartite local models. *Bioinformatics* 25, 2397–2403.
- Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., & Hwang, D. -U. (2006). Complex networks: structure and dynamics. *Physics Rep* 424, 175–308.
- Böde, C., Kovács, I. A., Szalay, M. S., Palotai, R., Korcsmáros, T., & Csermely, P. (2007). Network analysis of protein dynamics. *FEBS Lett* 581, 2776–2782.
- Bodenmiller, B., Wanka, S., Kraft, C., Urban, J., Campbell, D., Pedrioli, P. G., et al. (2010). Phosphoproteomic analysis reveals interconnected system-wide responses to perturbations of kinases and phosphatases in yeast. *Sci Signal* 3, rs4.
- Boettiger, C., & Hastings, A. (2012). Early warning signals and the prosecutor's fallacy. *Proc Biol Sci* 279, 4734–4739.
- Bogan, A. A., & Thorn, K. S. (1998). Anatomy of hot spots in protein interfaces. *J Mol Biol* 280, 1–9.
- Boginski, V., & Commander, C. W. (2009). Identifying critical nodes in protein–protein interaction networks. In S. Butenko, W. A. Chaovalitwongse, & P. M. Pardalos (Eds.), *Clustering challenges in biological networks* (pp. 153–167). Singapore: World Scientific Publishing.
- Bonchev, D., & Buck, G. A. (2007). From molecular to biological structure and back. *J Chem Inf Model* 47, 909–917.
- Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D., Marlow, C., Settle, J. E., et al. (2012). A 61-million-person experiment in social influence and political mobilization. *Nature* 489, 295–298.
- Bonnet, E., Tatar, M., Joshi, A., Michael, T., Marchal, K., Berx, G., et al. (2010). Module network inference from a cancer gene expression data set identifies microRNA regulated modules. *PLoS One* 5, e10162.
- Bordbar, A., Lewis, N. E., Schellenberger, J., Palsson, B. O., & Jamshidi, N. (2010). Insight into human alveolar macrophage and *M. tuberculosis* interactions via metabolic reconstructions. *Mol Syst Biol* 6, 422.
- Bordbar, A., Feist, A. M., Usaite-Black, R., Woodcock, J., Palsson, B. O., & Famili, I. (2011). A multi-tissue type genome-scale metabolic network for analysis of whole-body systems physiology. *BMC Syst Biol* 5, 180.
- Borisy, A. A., Elliott, P. J., Hurst, N. W., Lee, M. S., Lehar, J., Price, E. R., et al. (2003). Systematic discovery of multicomponent therapeutics. *Proc Natl Acad Sci USA* 100, 7977–7982.
- Borklu Yuçel, E., & Ulgen, K. O. (2011). A network-based approach on elucidating the multi-faceted nature of chronological aging in *S. cerevisiae*. *PLoS One* 6, e29284.
- Borrett, S. R. (2012). Throughflow centrality is a global indicator of the functional importance of species in ecosystems. <http://arxiv.org/abs/1209.0725>
- Bovolenta, L. A., Acencio, M. L., & Lemke, N. (2012). HTRIdb: an open-access database for experimentally verified human transcriptional regulation interactions. *BMC Genomics* 13, 405.
- Bozóky, B., Savchenko, A., Csermely, P., Korcsmáros, T., Dúli, Z., Pontén, F., Székely, L., & Klein, G. (in press). Novel signatures of cancer associated fibroblasts. *Int J Cancer*, 132. <http://dx.doi.org/10.1002/ijc.28035>.
- Bray, D., & Duke, T. (2004). Conformational spread: the propagation of allosteric states in large multiprotein complexes. *Annu Rev Biophys Biomol Struct* 33, 53–73.
- Brede, M. (2010). Coordinated and uncoordinated optimization of networks. *Phys Rev E* 81, 066104.
- Brehme, M., Hantschel, O., Colinge, J., Kaupe, I., Planyavsky, M., Kocher, T., et al. (2009). Charting the molecular network of the drug target Bcr-Abl. *Proc Natl Acad Sci USA* 106, 7414–7419.
- Breitkreutz, B. J., Stark, C., & Tyers, M. (2003). Osprey: a network visualization system. *Genome Biol* 4, R22.
- Breitkreutz, D., Hlatky, L., Rietman, E., & Tuszynski, J. A. (2012). Molecular signaling network complexity is correlated with cancer patient survivability. *Proc Natl Acad Sci USA* 109, 9209–9212.
- Brennan, R. J., Nikolskaya, T., & Bureeva, S. (2009). Network and pathway analysis of compound–protein interactions. *Methods Mol Biol* 575, 225–247.
- Brennan, M. D., Cheong, R., & Levchenko, A. (2012). Systems biology. How information theory handles cell signaling and uncertainty. *Science* 338, 334–335.
- Brinda, K. V., & Vishveshwara, S. (2005). A network representation of protein structures: implications for protein stability. *Biophys J* 89, 4159–4170.
- Broehe, S., Faust, K., Lima-Mendez, G., Vanderstocken, G., & van Helden, J. (2008). Network Analysis Tools: from biological networks to clusters and pathways. *Nat Protoc* 3, 1616–1629.
- Bromberg, K. D., Ma'ayan, A., Neves, S. R., & Iyengar, R. (2008). Design logic of a cannabinoid receptor signaling network that triggers neurite outgrowth. *Science* 320, 903–909.
- Brouwers, L., Iskar, M., Zeller, G., van Noort, V., & Bork, P. (2011). Network neighbors of drug targets contribute to drug side-effect similarity. *PLoS One* 6, e22187.
- Brown, D., & Superti-Furga, G. (2003). Rediscovering the sweet spot in drug discovery. *Drug Discov Today* 8, 1067–1077.
- Brown, K. R., Otasek, D., Ali, M., McGuffin, M. J., Xie, W., Devani, B., et al. (2009). NAViGaTOR: network analysis, visualization and graphing Toronto. *Bioinformatics* 25, 3327–3329.
- Brown, J. R., Magid-Slav, M., Sanseau, P., & Rajpal, D. K. (2011). Computational biology approaches for selecting host–pathogen drug targets. *Drug Discov Today* 16, 229–236.
- Bu, D., Zhao, Y., Cai, L., Xue, H., Zhu, X., Lu, H., et al. (2003). Topological structure analysis of the protein–protein interaction network in budding yeast. *Nucleic Acids Res* 31, 2443–2450.
- Budd, W. T., Weaver, D. E., Anderson, J., & Zehner, Z. E. (2012). microRNA dysregulation in prostate cancer: network analysis reveals preferential regulation of highly connected nodes. *Chem Biodivers* 9, 857–867.
- Budovsky, A., Abramovich, A., Cohen, R., Chalifa-Caspi, V., & Fraifeld, V. (2007). Longevity network: construction and implications. *Mech Ageing Dev* 128, 117–124.
- Buganin, Y., Faddah, D. A., Cheng, A. W., Itskovich, E., Markoulaki, S., Ganz, K., et al. (2012). Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. *Cell* 150, 1209–1222.
- Buldirev, S. V., Parshani, R., Paul, G., Stanley, H. E., & Havlin, S. (2010). Catastrophic cascade of failures in interdependent networks. *Nature* 464, 1025–1028.
- Bunnage, M. E. (2011). Getting pharmaceutical R&D back on target. *Nat Chem Biol* 7, 335–339.
- Burkard, T. R., Planyavsky, M., Kaupe, I., Breitwieser, F. P., Burckstummer, T., Bennett, K. L., et al. (2011). Initial characterization of the human central proteome. *BMC Syst Biol* 5, 17.
- Burt, R. S. (1995). *Structural holes: the Social Structure of Competition*. Cambridge: Harvard University Press.
- Bush, K., Courvalin, P., Dantas, G., Davies, J., Eisenstein, B., Huovinen, P., et al. (2011). Tackling antibiotic resistance. *Nat Rev Microbiol* 9, 894–896.
- Bultinck, J., Lievens, S., & Tavernier, J. (2012). Protein–protein interactions: network analysis and applications in drug discovery. *Curr Pharm Des* 18, 4619–4629.
- Byrne, A. B., Weirauch, M. T., Koeva, M., Dixon, S. J., Stuart, J. M., et al. (2007). A global analysis of genetic interactions in *Caenorhabditis elegans*. *J Biol* 6, 8.
- Bysell, H., Månsson, R., Hansson, P., & Malmsten, M. (2011). Microgels and microcapsules in peptide and protein drug delivery. *Adv Drug Deliv Rev* 63, 1172–1185.
- Calderwood, M. A., Venkatesan, K., Xing, L., Chase, M. R., Vazquez, A., Holthaus, A. M., et al. (2007). Epstein–Barr virus and virus human protein interaction maps. *Proc Natl Acad Sci USA* 104, 7606–7611.
- Calin, G. A., & Croce, C. M. (2006). MicroRNA signatures in human cancers. *Nat Rev Cancer* 6, 857–866.
- Calzolari, D., Bruschi, S., Coquin, L., Schofield, J., Feala, J. D., Reed, J. C., et al. (2008). Search algorithms as a framework for the optimization of drug combinations. *PLoS Comput Biol* 4, e1000249.
- Calzone, L., Fages, F., & Soliman, S. (2006). BIOCHAM: an environment for modeling biological systems and formalizing experimental knowledge. *Bioinformatics* 22, 1805–1807.
- Campillos, M., Kuhn, M., Gavin, A. C., Jensen, L. J., & Bork, P. (2008). Drug target identification using side-effect similarity. *Science* 321, 263–266.
- Cardenas, M. E., Cruz, M. C., Del Poeta, M., Chung, N., Perfect, J. R., & Heitman, J. (1999). Antifungal activities of antineoplastic agents: *Saccharomyces cerevisiae* as a model system to study drug action. *Clin Microbiol Rev* 12, 583–611.
- Care, M. A., Bradford, J. R., Needham, C. J., Bulpitt, A. J., & Westhead, D. R. (2009). Combining the interactome and deleterious SNP predictions to improve disease gene identification. *Hum Mutat* 30, 485–492.
- Caron, E., Ghosh, S., Matsuo, Y., Ashton-Beaucage, D., Therrien, M., Lemieux, S., et al. (2010). A comprehensive map of the mTOR signaling network. *Mol Syst Biol* 6, 453.
- Cascante, M., Boros, L. G., Comin-Anduix, B., de Atauri, P., Centelles, J. J., & Lee, P. W. (2002). Metabolic control analysis in drug discovery and disease. *Nat Biotechnol* 20, 243–249.
- Caspi, R., Altman, T., Dreher, K., Fulcher, C. A., Subhraveti, P., Keseler, I. M., et al. (2012). The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res* 40, D742–D753.
- Castro, M. A., Wang, X., Fletcher, M. N., Meyer, K. B., & Markowitz, F. (2012). RedeR: R/Bioconductor package for representing modular structures, nested networks and multiple levels of hierarchical associations. *Genome Biol* 13, R29.
- Catania, C., Binder, E., & Cota, D. (2011). mTORC1 signaling in energy balance and metabolic disease. *Int J Obes (Lond)* 35, 751–761.
- Chait, R., Craney, A., & Kishony, R. (2007). Antibiotic interactions that select against resistance. *Nature* 446, 668–671.
- Chang, W., Ma, L., Lin, L., Gu, L., Liu, X., Cai, H., et al. (2009). Identification of novel hub genes associated with liver metastasis of gastric cancer. *Int J Cancer* 125, 2844–2853.
- Chang, R. L., Xie, L., Bourne, P. E., & Palsson, B. O. (2010). Drug off-target effects predicted using structural analysis in the context of a metabolic network model. *PLoS Comput Biol* 6, e1000938.
- Chanumolu, S. K., Rout, C., & Chauhan, R. S. (2012). UniDrug-target: a computational tool to identify unique drug targets in pathogenic bacteria. *PLoS One* 7, e32833.
- Chaurasia, G., & Futschik, M. (2012). The integration and annotation of the human interactome in the UniHI Database. *Methods Mol Biol* 812, 175–188.
- Chautard, E., Thierry-Mieg, N., & Ricard-Blum, S. (2009). Interaction networks: from protein functions to drug discovery. A review. *Pathol Biol* 57, 324–333.
- Chautard, E., Thierry-Mieg, N., & Ricard-Blum, S. (2010). Interaction networks as a tool to investigate the mechanisms of aging. *Biogerontology* 11, 463–473.
- Chavali, A. K., D'Auria, K. M., Hewlett, E. L., Pearson, R. D., & Papin, J. A. (2012). A metabolic network approach for the identification and prioritization of antimicrobial drug targets. *Trends Microbiol* 20, 113–123.
- Chen, N., & Karantza, V. (2011). Autophagy as a therapeutic target in cancer. *Cancer Biol Ther* 11, 157–168.
- Chen, Y., & Klionsky, D. J. (2011). The regulation of autophagy—unanswered questions. *J Cell Sci* 124, 161–170.
- Chen, J. Y., Shen, C., & Sivachenko, A. Y. (2006a). Mining Alzheimer disease relevant proteins from integrated protein interactome data. *Pac Symp Biocomput*, 367–378.
- Chen, J. Y., Shen, C., Yan, Z., Brown, D. P., & Wang, M. (2006b). A systems biology case study of ovarian cancer drug resistance. *Comput Syst Bioinformatics Conf*, 389–398.

- Chen, H., Ding, L., Wu, Z., Yu, T., Dhanapalan, L., & Chen, J. Y. (2009a). Semantic web for integrated network analysis in biomedicine. *Brief Bioinform* 10, 177–192.
- Chen, J., Aronow, B. J., & Jegga, A. G. (2009b). Disease candidate gene identification and prioritization using protein interaction networks. *BMC Bioinforma* 10, 73.
- Chen, J. Y., Mamidipalli, S., & Huan, T. (2009c). HAPPI: an online database of comprehensive human annotated and predicted protein interactions. *BMC Genomics* 10, S16.
- Chen, C. C., Lin, C. Y., Lo, Y. S., & Yang, J. M. (2009d). PPIsearch: a web server for searching homologous protein–protein interactions across multiple species. *Nucleic Acids Res* 37, W369–W375.
- Chen, B., Wild, D., & Guha, R. (2009e). PubChem as a source of polypharmacology. *J Chem Inf Model* 49, 2044–2055.
- Chen, B., Dong, X., Jiao, D., Wang, H., Zhu, Q., Ding, Y., et al. (2010a). Chem2Bio2RDF: a semantic framework for linking and data mining chemogenomic and systems chemistry data. *BMC Bioinforma* 11, 255.
- Chen, L. H., Chiou, G. Y., Chen, Y. W., Li, H. Y., & Chiou, S. H. (2010b). MicroRNA and aging: a novel modulator in regulating the aging network. *Ageing Res Rev* 9, S59–S66.
- Chen, X., Xu, J., Huang, B., Li, J., Wu, X., Ma, L., et al. (2011). A sub-pathway-based approach for identifying drug response principal network. *Bioinformatics* 27, 649–654.
- Chen, X., Liu, M. X., & Yan, G. Y. (2012a). Drug–target interaction prediction by random walk on the heterogeneous network. *Mol Biosyst* 8, 1970–1978.
- Chen, L. C., Yeh, H. Y., Yeh, C. Y., Arias, C. R., & Soo, V. W. (2012b). Identifying co-targets to fight drug resistance based on a random walk model. *BMC Syst Biol* 6, 5.
- Chen, X., Liang, H., Zhang, J., Zen, K., & Zhang, C. Y. (2012c). Secreted microRNAs: a new form of intercellular communication. *Trends Cell Biol* 22, 125–132.
- Cheng, Y., LeGall, T., Oldfield, C. J., Mueller, J. P., Van, Y. Y., Romero, P., et al. (2006). Rational drug design via intrinsically disordered protein. *Trends Biotechnol* 24, 435–442.
- Cheng, C. Y., & Hu, Y. J. (2010). Extracting the abstraction pyramid from complex networks. *BMC Bioinforma* 11, 411.
- Cheng, F., Liu, C., Jiang, J., Lu, W., Li, W., Liu, G., et al. (2012a). Prediction of drug–target interactions and drug repositioning via network-based inference. *PLoS Comput Biol* 8, e1002503.
- Cheng, F., Zhou, Y., Li, W., Liu, G., & Tang, Y. (2012b). Prediction of chemical–protein interactions network with weighted network-based inference method. *PLoS One* 7, e41064.
- Cheng, T. M., Gulati, S., Agius, R., & Bates, P. A. (2012c). Understanding cancer mechanisms through network dynamics. *Brief Funct Genomics* 11, 543–560.
- Chennubhotla, C., & Bahar, I. (2006). Markov propagation of allosteric effects in biomolecular systems: application to GroEL–GroES. *Mol Syst Biol* 2, 36.
- Chennubhotla, C., & Bahar, I. (2007). Markov methods for hierarchical coarse-graining of large protein dynamics. *J Comput Biol* 14, 765–776.
- Chennubhotla, C., Yang, Z., & Bahar, I. (2008). Coupling between global dynamics and signal transduction pathways: a mechanism of allostery for chaperonin GroEL. *Mol Biosyst* 4, 287–292.
- Cheong, R., Rhee, A., Wang, C. J., Nemenman, I., & Levchenko, A. (2011). Information transduction capacity of noisy biochemical signaling networks. *Science* 334, 354–358.
- Chettaoui, C., Delaplace, F., Manceny, M., & Malo, M. (2007). Games network and application to PAs system. *Biosystems* 87, 136–141.
- Chin, C. S., & Samanta, M. P. (2003). Global snapshot of a protein interaction network—a percolation based approach. *Bioinformatics* 19, 2413–2419.
- Cho, D. Y., Kim, Y. A., & Przytycka, T. M. (2012). Chapter 5: network biology approach to complex diseases. *PLoS Comput Biol* 8, e1002820.
- Chong, C. R., & Sullivan, D. J., Jr. (2007). New uses for old drugs. *Nature* 448, 645–646.
- Chovanova, E., Pavelka, A., Benes, P., Strnad, O., Brezovsky, J., Kozlikova, B., et al. (2012). CAVER 3.0: a tool for the analysis of transport pathways in dynamic protein structures. *PLoS Comput Biol* 8, e1002708.
- Christakis, N. A., & Fowler, J. H. (2011). *Connected: The Surprising Power of Our Social Networks and How They Shape Our Lives—How Your Friends' Friends' Friends Affect Everything You Feel, Think, and Do*. Boston: Back Bay Books.
- Christiansen, J. A. (1953). The elucidation of reaction mechanisms by the method of intermediates in quasi-stationary concentrations. *Adv Catal* 5, 311–353.
- Christofk, H. R., Vander Heiden, M. G., Harris, M. H., Ramanathan, A., Gerszten, R. E., Wei, R., et al. (2008). The M2 splice isoform of pyruvate kinase is important for cancer metabolism and tumour growth. *Nature* 452, 230–233.
- Chu, L. H., & Chen, B. S. (2008). Construction of a cancer-perturbed protein–protein interaction network for discovery of apoptosis drug targets. *BMC Syst Biol* 2, 56.
- Chua, H. N., & Roth, F. P. (2011). Discovering the targets of drugs via computational systems biology. *J Biol Chem* 286, 23653–23658.
- Chuang, H. Y., Lee, E., Liu, Y. T., Lee, D., & Ideker, T. (2007). Network-based classification of breast cancer metastasis. *Mol Syst Biol* 3, 140.
- Chuang, H. Y., Rassenti, L., Salcedo, M., Licon, K., Kohlmann, A., Haferlach, T., et al. (2012). Subnetwork-based analysis of chronic lymphocytic leukemia identifies pathways that associate with disease progression. *Blood* 120, 2639–2649.
- Chung, H. J., Park, C. H., Han, M. R., Lee, S., Ohn, J. H., Kim, J., et al. (2005). ArrayXPath II: mapping and visualizing microarray gene-expression data with biomedical ontologies and integrated biological pathway resources using Scalable Vector Graphics. *Nucleic Acids Res* 33, W621–W626.
- Ciliberti, S., Martin, O. C., & Wagner, A. (2007). Robustness can evolve gradually in complex regulatory gene networks with varying topology. *PLoS Comput Biol* 3, 2.
- Ciriello, G., Mina, M., Guzzi, P. H., Cannataro, M., & Guerra, C. (2012a). AlignNemo: a local network alignment method to integrate homology and topology. *PLoS One* 7, e38107.
- Ciriello, G., Cerami, E., Sander, C., & Schultz, N. (2012b). Mutual exclusivity analysis identifies oncogenic network modules. *Genome Res* 22, 398–406.
- Clackson, T., & Wells, J. A. (1995). A hot spot of binding energy in a hormone–receptor interface. *Science* 267, 383–386.
- Clarke, D., Bhardwaj, N., & Gerstein, M. B. (2012). Novel insights through the integration of structural and functional genomics data with protein networks. *J Struct Biol* 179, 320–326.
- Clauset, A., Moore, C., & Newman, M. E. (2008). Hierarchical structure and the prediction of missing links in networks. *Nature* 453, 98–101.
- Clemente, J. C., Ursell, L. K., Parfrey, L. W., & Knight, R. (2012). The impact of the gut microbiota on human health: an integrative view. *Cell* 148, 1258–1270.
- Cloutier, M., & Wang, E. (2011). Dynamic modeling and analysis of cancer cellular network motifs. *Integr Biol* 3, 724–732.
- Cohen, I. R. (1992). The cognitive paradigm and the immunological homunculus. *Immunol Today* 13, 490–494.
- Cohen, A. A., Geva-Zatorsky, N., Eden, E., Frenkel-Morgenstern, M., Issaeva, I., Sigal, A., et al. (2008). Dynamic proteomics of individual cancer cells in response to a drug. *Science* 322, 1511–1516.
- Cokol, M., Iossifov, I., Weinreb, C., & Rzhetsky, A. (2005). Emergent behavior of growing knowledge about molecular interactions. *Nat Biotechnol* 23, 1243–1247.
- Coleman, R. G., & Sharp, K. A. (2010). Protein pockets: inventory, shape and comparison. *J Chem Inf Model* 50, 589–603.
- Colizza, V., Flammini, A., Serrano, M. A., & Vespignani, A. (2006). Detecting rich-club ordering in complex networks. *Nat Phys* 2, 110–115.
- Cornelius, S. P., Kath, W. L., & Motter, A. E. (2011). Controlling complex networks with compensatory perturbations. <http://arxiv.org/abs/1105.3726>
- Cosgrove, E. J., Zhou, Y., Gardner, T. S., & Kolaczky, E. D. (2008). Predicting gene targets of perturbations via network-based filtering of mRNA expression compendia. *Bioinformatics* 24, 2482–2490.
- Cosgrove, B. D., Alexopoulos, L. G., Hang, T. C., Hendriks, B. S., Sorger, P. K., Griffith, L. G., et al. (2010). Cytokine-associated drug toxicity in human hepatocytes is associated with signaling network dysregulation. *Mol Biosyst* 6, 1195–1206.
- Costa, P. R., Acencio, M. L., & Lemke, N. (2010). A machine learning approach for genome-wide prediction of morbid and druggable human genes based on systems-level data. *BMC Genomics* 11, S9.
- Costanzo, M., Baryshnikov, A., Bellay, J., Kim, Y., Spear, E. D., Sevier, C. S., et al. (2010). The genetic landscape of a cell. *Science* 327, 425–431.
- Coulombe, B. (2011). Mapping the disease protein interactome: toward a molecular medicine GPS to accelerate drug and biomarker discovery. *J Proteome Res* 10, 120–125.
- Cowan, N. J., Chastain, E. J., Vilhena, D. A., Freudenberg, J. S., & Bergstrom, C. T. (2012). Nodal dynamics, not degree distributions, determine the structural controllability of complex networks. *PLoS One* 7, e38398.
- Cowley, M. J., Pinesse, M., Kassahn, K. S., Waddell, N., Pearson, J. V., Grimmond, S. M., et al. (2012). PINA v2.0: mining interactome modules. *Nucleic Acids Res* 40, D862–D865.
- Cowper-Salari, R., Cole, M. D., Karagas, M. R., Lupien, M., & Moore, J. H. (2011). Layers of epistasis: genome-wide regulatory networks and network approaches to genome-wide association studies. *Wiley Interdiscip Rev Syst Biol Med* 3, 513–526.
- Croft, D., O'Kelly, G., Wu, G., Haw, R., Gillespie, M., Matthews, L., et al. (2011). Reactome: a database of reactions, pathways and biological processes. *Nucleic Acids Res* 39, D691–D697.
- Crombach, A., Wotton, K. R., Cicin-Sain, D., Ashyraliyev, M., & Jaeger, J. (2012). Efficient reverse-engineering of a developmental gene regulatory network. *PLoS Comput Biol* 8, e1002589.
- Crul, T., Toth, N., Piotto, S., Literati-Nagy, P., Tory, K., Haldimann, P., et al. (2013). Hydroxamic acid derivatives: pleiotropic Hsp co-inducers restoring homeostasis and robustness. *Curr Pharm Des* 19, 309–346.
- Csermely, P. (2004). Strong links are important—but weak links stabilize them. *Trends Biochem Sci* 29, 331–334.
- Csermely, P. (2008). Creative elements: network-based predictions of active centres in proteins, cellular and social networks. *Trends Biochem Sci* 33, 569–576.
- Csermely, P. (2009). *Weak Links: The Universal Key to the Stability of Networks and Complex Systems*. Heidelberg, Germany: Springer Verlag.
- Csermely, P. (in press). The appearance and promotion of creativity at various levels of interdependent networks. *Talent Development Excellence*, 5.
- Csermely, P., Ágoston, V., & Pongor, S. (2005). The efficiency of multi-target drugs: the network approach might help drug design. *Trends Pharmacol Sci* 26, 178–182.
- Csermely, P., Palotai, R., & Nussinov, R. (2010). Induced fit, conformational selection and independent dynamic segments: an extended view of binding events. *Trends Biochem Sci* 35, 539–546.
- Csermely, P., Sandhu, K. S., Hazai, E., Hoksza, Z., Kiss, H. J. M., Miozzo, F., et al. (2012). Disordered proteins and network disorder in network representations of protein structure, dynamics and function. Hypotheses and a comprehensive review. *Curr Prot Pept Sci* 13, 19–33.
- Csoka, A. B., & Szyf, M. (2009). Epigenetic side-effects of common pharmaceuticals: a potential new field in medicine and pharmacology. *Med Hypotheses* 73, 770–780.
- Cui, Q. (2010). A network of cancer genes with co-occurring and anti-co-occurring mutations. *PLoS One* 5, e13180.
- Cui, Q., Yu, Z., Purisima, E. O., & Wang, E. (2006). Principles of microRNA regulation of a human cellular signaling network. *Mol Syst Biol* 2, 46.
- Cui, Q., Ma, Y., Jaramillo, M., Bari, H., Awan, A., Yang, S., et al. (2007). A map of human cancer signaling. *Mol Syst Biol* 3, 152.
- Cui, Q., Purisima, E. O., & Wang, E. (2009). Protein evolution on a human signaling network. *BMC Syst Biol* 3, 21.
- Dai, L., Vorselen, D., Korolev, K. S., & Gore, J. (2012). Generic indicators for loss of resilience before a tipping point leading to population collapse. *Science* 336, 1175–1177.
- Daily, M. D., Upadhyaya, T. J., & Gray, J. J. (2008). Contact rearrangements form coupled networks from local motions in allosteric proteins. *Proteins* 71, 455–466.

- Dall'Astra, L., Barrat, A., Barthélemy, M., & Vespignani, A. (2006). Vulnerability of weighted networks. *J Stat Mech* 2006, P04006.
- Daminelli, S., Haupt, V. J., Reimann, M., & Schroeder, M. (2012). Drug repositioning through incomplete bi-cliques in an integrated drug–target–disease network. *Integr Biol* 4, 778–788.
- Dancey, J. E., & Chen, H. X. (2006). Strategies for optimizing combinations of molecularly targeted anticancer agents. *Nat Rev Drug Discov* 5, 649–659.
- Dancik, V., Seiler, K. P., Young, D. W., Schreiber, S. L., & Clemons, P. A. (2010). Distinct biological network properties between the targets of natural products and disease genes. *J Am Chem Soc* 132, 9259–9261.
- D'Angelo, M. A., Raices, M., Panowski, S. H., & Hetzer, M. W. (2009). Age-dependent deterioration of nuclear pore complexes causes a loss of nuclear integrity in postmitotic cells. *Cell* 136, 284–295.
- D'Antonio, M., Pendino, V., Sinha, S., & Ciccarelli, F. D. (2012). Network of Cancer Genes (NCG 3.0): integration and analysis of genetic and network properties of cancer genes. *Nucleic Acids Res* 40, D978–D983.
- Das, J., & Yu, H. (2012). HINT: high-quality protein interactomes and their applications in understanding human disease. *BMC Syst Biol* 6, 92.
- Dasika, M. S., Burgard, A., & Maranas, C. D. (2006). A computational framework for the topological analysis and targeted disruption of signal transduction networks. *Biophys J* 91, 382–398.
- Davis, M. J., Shin, C. J., Jing, N., & Ragan, M. A. (2012). Rewiring the dynamic interactome. *Mol Biosyst* 8, 2054–2066.
- de Chasse, B., Navratil, V., Tafforeau, L., Hiet, M. S., Aublin-Gex, A., Agaogue, S., et al. (2008). Hepatitis C virus infection protein network. *Mol Syst Biol* 4, 230.
- De Las Rivas, J., & Fontanillo, C. (2010). Protein–protein interactions essentials: key concepts to building and analyzing interactome networks. *PLoS Comput Biol* 6, e1000807.
- De Las Rivas, J., & Fontanillo, C. (2012). Protein–protein interaction networks: unraveling the wiring of molecular machines within the cell. *Brief Funct Genomics* 11, 489–496.
- De Las Rivas, J., & Prieto, C. (2012). Protein interactions: mapping interactome networks to support drug target discovery and selection. *Methods Mol Biol* 910, 279–296.
- de Leon, J. (2011). Highlights of drug package inserts and the website DailyMed: the need for further improvement in package inserts to help busy prescribers. *J Clin Psychopharmacol* 31, 263–265.
- de Magalhaes, J. P., Wuttke, D., Wood, S. H., Plank, M., & Vora, C. (2012). Genome–environment interactions that modulate aging: powerful targets for drug discovery. *Pharmacol Rev* 64, 88–101.
- de Cavalcante, S., H. L. D., Oirá, M., Sornette, D., & Gauthier, D. J. (2013). Predictability and control of extreme events in complex systems. <http://arxiv.org/abs/1301.0244>
- DeDecker, B. S. (2000). Allosteric drugs: thinking outside the active-site box. *Chem Biol* 7, R103–R107.
- Dehmer, M., Barbarini, N., Varmuza, K., & Graber, A. (2009). A large scale analysis of information-theoretic network complexity measures using chemical structures. *PLoS One* 4, e8057.
- DeJongh, M., Bockstege, B., Frybarger, P., Hazekamp, N., Kammeraad, J., & McGeehan, T. (2012). CytosEED: a Cytoscape plugin for viewing, manipulating and analyzing metabolic models created by the Model SEED. *Bioinformatics* 28, 891–892.
- del Sol, A., & O'Meara, P. (2005). Small-world network approach to identify key residues in protein–protein interaction. *Proteins* 58, 672–682.
- del Sol, A., Fujihashi, H., Amorós, D., & Nussinov, R. (2006). Residues crucial for maintaining short paths in network communication mediate signaling in proteins. *Mol Syst Biol* 2, 19.
- del Sol, A., Araúzo-Bravo, M. J., Amorós, D., & Nussinov, R. (2007). Modular architecture of protein structures and allosteric communications: potential implications for signaling proteins and regulatory linkages. *Genome Biol* 8, R92.
- del Sol, A., Balling, R., Hood, L., & Galas, D. (2010). Diseases as network perturbations. *Curr Opin Biotechnol* 21, 566–571.
- Delmotte, A., Tate, E. W., Yaliraki, S. N., & Barahona, M. (2011). Protein multi-scale organization through graph partitioning and robustness analysis: application to the myosin–myosin light chain interaction. *Phys Biol* 8, 055010.
- Delvenne, J. -C., Yaliraki, S. N., & Barahona, M. (2010). Stability of graph communities across time scales. *Proc Natl Acad Sci USA* 107, 12755–12760.
- Deng, M., Mehta, S., Sun, F., & Chen, T. (2002). Inferring domain–domain interactions from protein–protein interactions. *Genome Res* 12, 1540–1548.
- Derényi, I., Farkas, I., Palla, G., & Vicsek, T. (2004). Topological phase transitions of random networks. *Phys A* 334, 583–590.
- di Bernardo, D., Thompson, M. J., Gardner, T. S., Chobot, S. E., Eastwood, E. L., Wojtovich, A. P., et al. (2005). Chemogenomic profiling on a genome-wide scale using reverse-engineered gene networks. *Nat Biotechnol* 23, 377–383.
- Di Paola, L., De Ruvo, M., Paci, P., Santoni, D., & Giuliani, A. (in press). Protein contact networks: An emerging paradigm in chemistry. *Chem Rev*. <http://dx.doi.org/10.1021/cr3002356>.
- Ding, Y., Chen, M., Liu, Z., Ding, D., Ye, Y., Zhang, M., et al. (2012). atBioNet—an integrated network analysis tool for genomics and biomarker discovery. *BMC Genomics* 13, 325.
- Dixit, A., & Verkhrivker, G. M. (2012). Probing molecular mechanisms of the Hsp90 chaperone: biophysical modeling identifies key regulators of functional dynamics. *PLoS One* 7, e37605.
- Dixon, S. J., Fedyshyn, Y., Koh, J. L., Prasad, T. S., Chahwan, C., Chua, G., et al. (2008). Significant conservation of synthetic lethal genetic interaction networks between distantly related eukaryotes. *Proc Natl Acad Sci USA* 105, 16653–16658.
- Dixon, S. J., Costanzo, M., Baryshnikov, A., Andrews, B., & Boone, C. (2009). Systematic mapping of genetic interaction networks. *Annu Rev Genet* 43, 601–625.
- Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., et al. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485, 376–380.
- Doench, J. G., & Sharp, P. A. (2004). Specificity of microRNA target selection in translational repression. *Genes Dev* 18, 504–511.
- Dogrusoz, U., Erson, E. Z., Giral, E., Demir, E., Babur, O., Cetintas, A., et al. (2006). PATIKAweb: a Web interface for analyzing biological pathways through advanced querying and visualization. *Bioinformatics* 22, 374–375.
- Doncheva, N. T., Klein, K., Domingues, F. S., & Albrecht, M. (2011). Analyzing and visualizing residue networks of protein structures. *Trends Biochem Sci* 36, 179–182.
- Doncheva, N. T., Assenov, Y., Domingues, F. S., & Albrecht, M. (2012a). Topological analysis and interactive visualization of biological networks and protein structures. *Nat Protoc* 7, 670–685.
- Doncheva, N. T., Kacprowski, T., & Albrecht, M. (2012b). Recent approaches to the prioritization of candidate disease genes. *Wiley Interdiscip Rev Syst Biol Med* 4, 429–442.
- Donges, J. F., Donner, R. V., & Kurths, J. (2012). Testing time series reversibility using complex network methods. <http://arxiv.org/abs/1211.1162>
- Donner, R. V., Small, M., Donges, J. F., Marwan, N., Zou, Y., Xiang, R., et al. (2011). Recurrence-based time series analysis by means of complex network methods. *Int J Bifurcation Chaos* 14, 1019–1046.
- Dorogovtsev, S. N., & Mendes, J. F. F. (2002). Evolution of networks. *Adv Phys* 51, 1079–1187.
- Doye, J. P. K. (2002). The network topology of a potential energy landscape: a static scale-free network. *Phys Rev Lett* 88, 238701.
- Drew, K. L. M., Baiman, H., Khwaounjoo, P., Yu, B., & Reynisson, J. (2012). Size estimation of chemical space: how big is it? *J Pharm Pharmacol* 64, 490–495.
- Drews, J. (2003). Strategic trends in the drug industry. *Drug Discov Today* 8, 411–420.
- Dreze, M., Charlotiaux, B., Milstein, S., Vidalain, P. O., Yildirim, M. A., Zhong, Q., et al. (2009). 'Edgetic' perturbation of a *C. elegans* BCL2 ortholog. *Nat Methods* 6, 843–849.
- Du, D., Lee, C. F., & Li, X. -Q. (2012). Systematic differences in signal emitting and receiving revealed by PageRank analysis of a human protein interactome. *PLoS One* 7, e44872.
- Duarte, N. C., Becker, S. A., Jamshidi, N., Thiele, I., Mo, M. L., Vo, T. D., et al. (2007). Global reconstruction of the human metabolic network based on genomic and bibliomic data. *Proc Natl Acad Sci USA* 104, 1777–1782.
- Dudley, J. T., Sirota, M., Shenoy, M., Pai, R. K., Roedder, S., Chiang, A. P., et al. (2011). Computational repositioning of the anticonvulsant topiramate for inflammatory bowel disease. *Sci Transl Med* 3, 96ra76.
- Dunkel, P., Chai, C. L., Sperlagh, B., Huleatt, P. B., & Matyus, P. (2012). Clinical utility of neuroprotective agents in neurodegenerative diseases: current status of drug development for Alzheimer's, Parkinson's and Huntington's diseases, and amyotrophic lateral sclerosis. *Expert Opin Investig Drugs* 21, 1267–1308.
- Durmus Tekir, S., Umit, P., Eren Toklu, A., & Ulgen, K. O. (2010). Reconstruction of protein–protein interaction network of insulin signaling in *Homo sapiens*. *J Biomed Biotechnol* 2010, 690925.
- Eckert, H., & Bajorath, J. (2007). Molecular similarity analysis in virtual screening: foundations, limitations and novel approaches. *Drug Discov Today* 12, 225–233.
- Edberg, A., Soeria-Atmadja, D., Bergman Laurila, J., Johansson, F., Gustafsson, M. G., & Hammerling, U. (2012). Assessing relative bioactivity of chemical substances using quantitative molecular network topology analysis. *J Chem Inf Model* 52, 1238–1249.
- Eduati, F., Corradin, A., Di Camillo, B., & Toffolo, G. (2010). A Boolean approach to linear prediction for signaling network modeling. *PLoS One* 5, e12789.
- Edwards, J. S., & Palsson, B. O. (2000a). The *Escherichia coli* MG1655 in silico metabolic genotype: its definition, characteristics, and capabilities. *Proc Natl Acad Sci USA* 97, 5528–5533.
- Edwards, J. S., & Palsson, B. O. (2000b). Robustness analysis of the *Escherichia coli* metabolic network. *Biotechnol Prog* 16, 927–939.
- Edwards, A. M., Isserlin, R., Bader, G. D., Frye, S. V., Willson, T. M., & Yu, F. H. (2011). Too many roads not taken. *Nature* 470, 163–165.
- Edwards, S. A., Wagner, J., & Grater, F. (2012). Dynamic prestress in a globular protein. *PLoS Comput Biol* 8, e1002509.
- Ehrlich, P. (1908). Experimental researches on specific therapy. *On immunity with special relationship between distribution and action of antigens. Harben Lecture*. London: Lewis.
- Einstein, A. (1934). On the method of theoretical physics. The Herbert Spencer Lecture. *Philos Sci* 1, 163–169.
- Ekins, S., Bugrim, A., Brovold, L., Kirillov, E., Nikolsky, Y., Rakhmatulin, E., et al. (2006). Algorithms for network analysis in systems-ADME/Tox using the MetaCore and MetaDrug platforms. *Xenobiotica* 36, 877–901.
- Emig, D., Cline, M. S., Lengauer, T., & Albrecht, M. (2008). Integrating expression data with domain interaction networks. *Bioinformatics* 24, 2546–2548.
- ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57–74.
- Engin, B. H., Keskin, O., Nussinov, R., & Gursoy, A. (2012). A strategy based on protein–protein interface motifs may help in identifying drug off-targets. *J Chem Inf Model* 52, 2273–2286.
- Ergun, A., Lawrence, C. A., Kohanski, M. A., Brennan, T. A., & Collins, J. J. (2007). A network biology approach to prostate cancer. *Mol Syst Biol* 3, 82.
- Erler, J. T., & Linding, R. (2010). Network-based drugs and biomarkers. *J Pathol* 220, 290–296.
- Erler, J. T., & Linding, R. (2012). Network medicine strikes a blow against breast cancer. *Cell* 149, 731–733.
- Erman, B. (2011). Relationships between ligand binding sites, protein architecture and correlated paths of energy and conformational fluctuations. *Phys Biol* 8, 056003.

- Eronen, L., & Toivonen, H. (2012). Biomine: predicting links between biological entities using network models of heterogeneous databases. *BMC Bioinforma* 13, 119.
- Erten, S., Bebek, G., Ewing, R. M., & Koyuturk, M. (2011a). DADA: degree-aware algorithms for network-based disease gene prioritization. *BioData Min* 4, 19.
- Erten, S., Bebek, G., & Koyuturk, M. (2011b). Vavien: an algorithm for prioritizing candidate disease genes based on topological similarity of proteins in interaction networks. *J Comput Biol* 18, 1561–1574.
- Escribá, P. V., González-Ros, J. M., Goñi, F. M., Kinnunen, P. K., Vigh, L., Sánchez-Magraner, L., et al. (2008). Membranes: a meeting point for lipids, proteins and therapies. *J Cell Mol Med* 12, 829–875.
- Estrada, E. (2006). Virtual identification of essential proteins within the protein interaction network of yeast. *Proteomics* 6, 35–40.
- Estrada, E. (2010). Universality in protein residue networks. *Biophys J* 98, 890–900.
- Estrada, E., & Rodríguez-Velázquez, J. A. (2005). Subgraph centrality in complex networks. *Phys Rev E* 71, 056103.
- Estrada, E., & Uriarte, E. (2001). Recent advances on the role of topological indices in drug discovery research. *Curr Med Chem* 8, 1573–1588.
- Estrada, E., Uriarte, E., Molina, E., Simon-Manso, Y., & Milne, G. W. (2006). An integrated *in silico* analysis of drug-binding to human serum albumin. *J Chem Inf Model* 46, 2709–2724.
- Falkowski, T., Bartelheimer, J., & Spiliopoulou, M. (2006). Mining and visualizing the evolution of subgroups in social networks. *Proc IEEE/WIC/ACM Intl Conf Web Intelligence* (pp. 52–58).
- Fan, T. W., Lorkiewicz, P. K., Sellers, K., Moseley, H. N., Higashi, R. M., & Lane, A. N. (2012). Stable isotope-resolved metabolomics and applications for drug development. *Pharmacol Ther* 133, 366–391.
- Fang, G., Wang, W., Paunic, V., Oatley, B., Haznadar, M., Steinbach, M., et al. (2011). Construction and functional analysis of human genetic interaction networks with genome-wide association data. <http://arxiv.org/abs/1101.3343>
- Farkas, I. J., Korcsmáros, T., Kovács, I. A., Mihálik, Á., Palotai, R., Simkó, G. I., et al. (2011). Network-based tools in the identification of novel drug-targets. *Sci Signal* 4, pt3.
- Farkas, I. J., Szanto-Varnagy, A., & Korcsmáros, T. (2012). Linking proteins to signaling pathways for experiment design and evaluation. *PLoS One* 7, e36202.
- Fatumo, S., Plaimas, K., Mallm, J. P., Schramm, G., Adebijoyi, E., Oswald, M., et al. (2009). Estimating novel potential drug targets of *Plasmodium falciparum* by analysing the metabolic network of knock-out strains *in silico*. *Infect Genet Evol* 9, 351–358.
- Fatumo, S., Plaimas, K., Adebijoyi, E., & Konig, R. (2011). Comparing metabolic network models based on genomic and automatically inferred enzyme information from *Plasmodium* and its human host to define drug targets *in silico*. *Infect Genet Evol* 11, 708–715.
- Faulon, J. -L., & Bender, A. (2010). *Handbook of Chemoinformatics Algorithms*. Boca Raton: CRC Press.
- Fayos, J., & Fayos, C. (2007). Wind data mining by Kohonen neural networks. *PLoS One* 2, e210.
- Fazekas, D., Koltai, M., Türei, D., Módos, D., Pálfi, M., Dúl, Z., et al. (2013). Signalink 2—a signaling pathway resource with multi-layered regulatory networks. *BMC Syst Biol* 7, 7.
- Fearnley, L. G., & Nielsen, L. K. (2012). PATHLOGIC-S: a scalable Boolean framework for modelling cellular signalling. *PLoS One* 7, e41977.
- Feldman, I., Rzhetsky, A., & Vitkup, D. (2008). Network properties of genes harboring inherited disease mutations. *Proc Natl Acad Sci USA* 105, 4323–4328.
- Fell, D. A. (1998). Increasing the flux in metabolic pathways: a metabolic control analysis perspective. *Biotechnol Bioeng* 58, 121–124.
- Fernandez, M., Caballero, J., Fernandez, L., & Sarai, A. (2011). Genetic algorithm optimization in drug design QSAR: Bayesian-regularized genetic neural networks (BRGNN) and genetic algorithm-optimized support vectors machines (GA-SVM). *Mol Divers* 15, 269–289.
- Ferrarini, L., Bertelli, L., Feala, J., McCulloch, A. D., & Paternostro, G. (2005). A more efficient search strategy for aging genes based on connectivity. *Bioinformatics* 21, 338–348.
- Ferro, A., Giugno, R., Pigola, G., Pulvirenti, A., Skripin, D., Bader, G. D., et al. (2007). NetMatch: a Cytoscape plugin for searching biological networks. *Bioinformatics* 23, 910–912.
- Fialkowski, M., Bishop, K. J. M., Chubukov, V. A., Campbell, C. J., & Grzybowski, B. A. (2005). Architecture and evolution of organic chemistry. *Angew Chem Int Ed* 44, 7263–7269.
- Fingar, D. C., & Inoki, K. (2012). Deconvolution of mTORC2 “in silico”. *Sci Signal* 5, pe12.
- Fischer, E. (1894). Einfluss der Configuration auf die Wirkung der Enzyme. *Ber Dtsch Chem Ges* 27, 2984–2993.
- Fliri, A. F., Loging, W. T., Thadeio, P. F., & Volkmann, R. A. (2005). Analysis of drug-induced effect patterns to link structure and side effects of medicines. *Nat Chem Biol* 1, 389–397.
- Fliri, A. F., Loging, W. T., & Volkmann, R. A. (2009). Drug effects viewed from a signal transduction network perspective. *J Med Chem* 52, 8038–8046.
- Fliri, A. F., Loging, W. T., & Volkmann, R. A. (2010). Cause–effect relationships in medicine: a protein network perspective. *Trends Pharmacol Sci* 31, 547–555.
- Florez, A. F., Park, D., Bhak, J., Kim, B. C., Kuchinsky, A., Morris, J. H., et al. (2010). Protein network prediction and topological analysis in *Leishmania major* as a tool for drug target selection. *BMC Bioinforma* 11, 484.
- Folger, O., Jerby, L., Frezza, C., Gottlieb, E., Ruppin, E., & Shlomi, T. (2011). Predicting selective drug targets in cancer through metabolic networks. *Mol Syst Biol* 7, 501.
- Fonseca, S. G., Lipson, K. L., & Urano, F. (2007). Endoplasmic reticulum stress signaling in pancreatic beta-cells. *Antioxid Redox Signal* 9, 2335–2344.
- Forbes, S. A., Bindal, N., Bamford, S., Cole, C., Kok, C. Y., Beare, D., et al. (2011). COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res* 39, D945–D950.
- Forster, J., Famili, I., Fu, P., Palsson, B. O., & Nielsen, J. (2003). Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. *Genome Res* 13, 244–253.
- Fortunato, S. (2010). Community detection in graphs. *Phys Rep* 486, 75–174.
- Foster, K. R. (2011). The sociobiology of molecular systems. *Nat Rev Genet* 12, 193–203.
- Fox, A. D., Hescott, B. J., Blumer, A. C., & Slonim, D. K. (2011). Connectedness of PPI network neighborhoods identifies regulatory hub proteins. *Bioinformatics* 27, 1135–1142.
- Franke, L., van Bakel, H., Fokkens, L., de Jong, E. D., Egmont-Petersen, M., & Wijmenga, C. (2006). Reconstruction of a functional human gene network, with an application for prioritizing positional candidate genes. *Am J Hum Genet* 78, 1011–1025.
- Fraser, I. D., & Germain, R. N. (2009). Navigating the network: signaling cross-talk in hematopoietic cells. *Nat Immunol* 10, 327–331.
- Freeman, L. C. (1978). Centrality in social networks I: conceptual clarification. *Social Networks* 1, 215–239.
- Freeman, M. (2000). Feedback control of intercellular signalling in development. *Nature* 408, 313–319.
- Freeman, T. C., Goldovsky, L., Brosch, M., van Dongen, S., Maziere, P., Grocock, R. J., et al. (2007). Construction, visualisation, and clustering of transcription networks from microarray expression data. *PLoS Comput Biol* 3, 2032–2042.
- Friedman, N. (2004). Inferring cellular networks using probabilistic graphical models. *Science* 303, 799–805.
- Frolkis, A., Knox, C., Lim, E., Jewison, T., Law, V., Hau, D. D., et al. (2010). SMPDB: the small molecule pathway database. *Nucleic Acids Res* 38, D480–D487.
- Fudenberg, G., Getz, G., Meyerson, M., & Mirny, L. A. (2011). High order chromatin architecture shapes the landscape of chromosomal alterations in cancer. *Nat Biotechnol* 29, 1109–1113.
- Fujimori, S., Hirai, N., Masuoka, K., Oshikubo, T., Yamashita, T., Washio, T., et al. (2012). IRView: a database and viewer for protein interaction regions. *Bioinformatics* 28, 1949–1950.
- Fullwood, M. J., Liu, M. H., Pan, Y. F., Liu, J., Xu, H., Mohamed, Y. B., et al. (2009). An oestrogen-receptor-alpha-bound human chromatin interactome. *Nature* 462, 58–64.
- Fung, D. C., Li, S. S., Goel, A., Hong, S. H., & Wilkins, M. R. (2012). Visualization of the interactome: what are we looking at? *Proteomics* 12, 1669–1686.
- Funke, C., Schneider, S. A., Berg, D., & Kell, D. B. (in press). Genetics and iron in the systems biology of Parkinson's disease and some related disorders. *Neurochem Int*. <http://dx.doi.org/10.1016/j.neuint.2012.11.015>.
- Gambari, R., Fabbri, E., Borgatti, M., Lampronti, I., Finotti, A., Brognara, E., et al. (2011). Targeting microRNAs involved in human diseases: a novel approach for modification of gene expression and drug development. *Biochem Pharmacol* 82, 1416–1429.
- Gamermann, D., Montagud, A., Jaime Infante, R. A., Triana, J., de Córdoba, P. F., & Urchueguía, J. F. (2012). PyNetMet: Python tools for efficient work with networks and metabolic models. <http://arxiv.org/abs/1211.7196>.
- Gandhi, T. K., Zhong, J., Mathivanan, S., Karthick, L., Chandrika, K. N., Mohan, S. S., et al. (2006). Analysis of the human protein interactome and comparison with yeast, worm and fly interaction datasets. *Nat Genet* 38, 285–293.
- Ganesan, A. (2008). The impact of natural products upon modern drug discovery. *Curr Opin Chem Biol* 12, 306–317.
- Gansner, E. R., & North, S. C. (2000). An open graph visualization system and its applications to software engineering. *Softwar Pract Experience* 30, 1203–1233.
- Gao, Z., Li, H., Zhang, H., Liu, X., Kang, L., Luo, X., et al. (2008). PDTD: a web-accessible protein database for drug target identification. *BMC Bioinforma* 9, 104.
- Gao, J., Ade, A. S., Tarcea, V. G., Weymouth, T. E., Mirel, B. R., Jagadish, H. V., et al. (2009). Integrating and annotating the interactome using the MiMI plugin for cytoscape. *Bioinformatics* 25, 137–138.
- Gao, X., Wang, H., Yang, J. J., Liu, X., & Liu, Z. R. (2012). Pyruvate kinase M2 regulates gene transcription by acting as a protein kinase. *Mol Cell* 45, 598–609.
- García, I., Munteanu, C. R., Fall, Y., Gomez, G., Uriarte, E., & Gonzalez-Diaz, H. (2009). QSAR and complex network study of the chiral HMGR inhibitor structural diversity. *Bioorg Med Chem* 17, 165–175.
- García-Domenech, R., Galvez, J., de Julian-Ortiz, J. V., & Pogliani, L. (2008). Some new trends in chemical graph theory. *Chem Rev* 108, 1127–1169.
- Gardino, A. K., & Yaffe, M. B. (2011). 14–3–3 proteins as signaling integration points for cell cycle control and apoptosis. *Semin Cell Dev Biol* 22, 688–695.
- Gardner, T. S., di Bernardo, D., Lorenz, D., & Collins, J. J. (2003). Inferring genetic networks and identifying compound mode of action via expression profiling. *Science* 301, 102–105.
- Garg, A., Mohanram, K., De Micheli, G., & Xenarios, I. (2012). Implicit methods for qualitative modeling of gene regulatory networks. *Methods Mol Biol* 786, 397–443.
- Gáspár, E. M., & Csermely, P. (2012). Rigidity and flexibility of biological networks. *Briefings Funct Genomics* 11, 443–456.
- Gaspar, P. M., Fuglestad, B., Komives, E. A., Markwick, P. R., & McCammon, J. A. (2012). Allosteric networks in thrombin distinguish procoagulant vs. anticoagulant activities. *Proc Natl Acad Sci USA* 109, 21216–21222.
- Garten, Y., Tatonetti, N. P., & Altman, R. B. (2010). Improving the prediction of pharmacogenes using text-derived drug–gene relationships. *Pac Symp Biocomput*, 305–314.
- Gaulton, A., Bellis, L. J., Bento, A. P., Chambers, J., Davies, M., Hersey, A., et al. (2012). ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res* 40, D1100–D1107.
- Geenen, S., Taylor, P. N., Snoep, J. L., Wilson, I. D., Kenna, J. G., & Westerhoff, H. V. (2012). Systems biology tools for toxicology. *Arch Toxicol* 8, 1251–1271.
- Gehlenborg, N., O'Donoghue, S. I., Baliga, N. S., Goesmann, A., Hibbs, M. A., Kitano, H., et al. (2010). Visualization of omics data for systems biology. *Nat Methods* 7, S56–S68.
- Gehring, R., Schumm, P., Youssef, M., & Scoglio, C. (2010). A network-based approach for resistance transmission in bacterial populations. *J Theor Biol* 262, 97–106.
- Gerber, S., Assmus, H., Bakker, B., & Klipp, E. (2008). Drug-efficacy depends on the inhibitor type and the target position in a metabolic network—a systematic study. *J Theor Biol* 252, 442–455.

- Gerstein, M. B., Kundaje, A., Hariharan, M., Landt, S. G., Yan, K. K., Cheng, C., et al. (2012). Architecture of the human regulatory network derived from ENCODE data. *Nature* 489, 91–100.
- Gertsbak, I. B., & Shprungin, Y. (2010). Models of network reliability. *Analysis, combinatorics and Monte Carlo*. Boca Raton: CRC Press.
- Getoor, L., & Diehl, C. P. (2005). Link mining: a survey. *ACM SIGKDD Explor Newslett* 7, 3–12.
- Geva-Zatorsky, N., Dekel, E., Cohen, A. A., Danon, T., Cohen, L., & Alon, U. (2010). Protein dynamics in drug combinations: a linear superposition of individual drug responses. *Cell* 140, 643–651.
- Ghazalpour, A., Doss, S., Yang, X., Aten, J., Toomey, E. M., Van Nas, A., et al. (2004). Thematic review series: the pathogenesis of atherosclerosis. Toward a biological network for atherosclerosis. *J Lipid Res* 45, 1793–1805.
- Ghosh, R., & Lerman, K. (2012). Rethinking centrality: the role of dynamical processes in social network analysis. <http://arxiv.org/abs/1209.4616>
- Ghosh, A., & Vishveshwara, S. (2007). A study of communication pathways in methionyl-tRNA synthetase by molecular dynamics simulations and structure network analysis. *Proc Natl Acad Sci USA* 104, 15711–15716.
- Ghosh, A., & Vishveshwara, S. (2008). Variations in clique and community patterns in protein structures during allosteric communication: investigation of dynamically equilibrated structures of methionyl tRNA synthetase complexes. *Biochemistry* 47, 11398–11407.
- Ginsburg, I. (1999). Multi-drug strategies are necessary to inhibit the synergistic mechanism causing tissue damage and organ failure in post infectious sequelae. *Inflammopharmacology* 7, 207–217.
- Girvan, M., & Newman, M. E. (2002). Community structure in social and biological networks. *Proc Natl Acad Sci USA* 99, 7821–7826.
- Glaser, B. (2010). Genetic analysis of complex disease—a roadmap to understanding or a colossal waste of money. *Pediatr Endocrinol Rev* 7, 258–265.
- Goehler, H., Lalowski, M., Stelzl, U., Waelter, S., Stroedicke, M., Worm, U., et al. (2004). A protein interaction network links G1T1, an enhancer of huntingtin aggregation, to Huntington's disease. *Mol Cell* 15, 853–865.
- Goel, R., Harsha, H. C., Pandey, A., & Prasad, T. S. (2012). Human Protein Reference Database and Human Proteinpedia as resources for phosphoproteome analysis. *Mol Biosyst* 8, 453–463.
- Goh, K. I., & Choi, I. G. (2012). Exploring the human diseaseome: the human disease network. *Brief Funct Genomics* 11, 533–542.
- Goh, K. I., Cusick, M. E., Valle, D., Childs, B., Vidal, M., & Barabasi, A. L. (2007). The human disease network. *Proc Natl Acad Sci USA* 104, 8685–8690.
- Gombs, I., Crul, T., Piotto, S., Gøngör, B., Török, Z., Balogh, G., et al. (2011). Membrane-lipid therapy in operation: the HSP co-inducer BGP-15 activates stress signal transduction pathways by remodeling plasma membrane rafts. *PLoS One* 6, e28818.
- Goncalves, J. P., Graos, M., & Valente, A. X. (2009). POLAR MAPPER: a computational tool for integrated visualization of protein interaction networks and mRNA expression data. *J R Soc Interface* 6, 881–896.
- Gong, Y., & Zhang, Z. (2007). Cell Frame: a data structure for abstraction of cell biology experiments and construction of perturbation networks. *Ann NY Acad Sci* 1115, 249–266.
- Gonzalez, M. W., & Kahn, M. G. (2012). Chapter 4: protein interactions and disease. *PLoS Comput Biol* 8, e1002819.
- Gonzalez-Diaz, H., & Prado-Prado, F. J. (2008). Unified QSAR and network-based computational chemistry approach to antimicrobials, part 1: multispecies activity models for antifungals. *J Comp Chem* 29, 656–667.
- Gonzalez-Diaz, H., Romaris, F., Duardo-Sanchez, A., Perez-Montoto, L. G., Prado-Prado, F., Patlewicz, G., et al. (2010a). Predicting drugs and proteins in parasite infections with topological indices of complex networks: theoretical backgrounds, applications and legal issues. *Curr Pharm Design* 16, 2737–2764.
- Gonzalez-Diaz, H., Duardo-Sanchez, A., Ubeira, F. M., Prado-Prado, F., Perez-Montoto, L. G., Concu, R., et al. (2010b). Review of MARCH-INSIDE & complex networks prediction of drugs: ADMET, anti-parasite activity, metabolizing enzymes and cardiotoxicity proteome biomarkers. *Curr Drug Metab* 11, 379–406.
- Goodey, N. M., & Benkovic, S. J. (2008). Allosteric regulation and catalysis emerge via a common route. *Nat Chem Biol* 4, 474–482.
- Gordo, S., & Giral, E. (2009). Knitting and untying the protein network: modulation of protein ensembles as a therapeutic strategy. *Protein Sci* 18, 481–493.
- Gosline, S. J., Spencer, S. J., Ursu, O., & Fraenkel, E. (2012). SAMNet: a network-based approach to integrate multi-dimensional high throughput datasets. *Integr Biol* 4, 1415–1427.
- Gothard, C. M., Soh, S., Gothard, N. A., Kowalczyk, B., Wei, Y., Baytekin, B., et al. (2012). Rewiring chemistry: algorithmic discovery and experimental validation of one-pot reactions in the network of organic chemistry. *Angew Chem Int Ed* 51, 7922–7927.
- Gottlieb, A., Stein, G. Y., Ruppin, E., & Sharan, R. (2011a). PREDICT: a method for inferring novel drug indications with application to personalized medicine. *Mol Syst Biol* 7, 496.
- Gottlieb, A., Magger, O., Berman, I., Ruppin, E., & Sharan, R. (2011b). PRINCIPLE: a tool for associating genes with diseases via network propagation. *Bioinformatics* 27, 3325–3326.
- Grady, D., Thiemann, C., & Brockmann, D. (2012). Robust classification of salient links in complex networks. *Nat Commun* 3, 864.
- Grassler, J., Koschutski, D., & Schreiber, F. (2012). CentiLib: comprehensive analysis and exploration of network centralities. *Bioinformatics* 28, 1178–1179.
- Graudenzi, A., Serra, R., Villani, M., Colacci, A., & Kauffman, S. A. (2011a). Robustness analysis of a Boolean model of gene regulatory network with memory. *J Comput Biol* 18, 559–577.
- Graudenzi, A., Serra, R., Villani, M., Damiani, C., Colacci, A., & Kauffman, S. A. (2011b). Dynamical properties of a Boolean model of gene regulatory network with memory. *J Comput Biol* 18, 1291–1303.
- Greene, L. H., & Higan, V. A. (2003). Uncovering network systems within protein structures. *J Mol Biol* 334, 781–791.
- Greer, E. L., & Brunet, A. (2008). Signaling networks in aging. *J Cell Sci* 121, 407–412.
- Griffith, O. L., Montgomery, S. B., Bernier, B., Chu, B., Kasaian, K., Aerts, S., et al. (2008). ORegAnno: an open-access community-driven resource for regulatory annotation. *Nucleic Acids Res* 36, D107–D113.
- Gros, C. (2012). Pushing the complexity barrier: diminishing returns in the sciences. *Complex Systems* 21, 183.
- Grosdidier, S., & Fernández-Recio, J. (2012). Protein-protein docking and hot-spot prediction for drug discovery. *Curr Pharm Des* 18, 4607–4618.
- Grzybowski, B. A., Bishop, K. J. M., Kowalczyk, B., & Wilmer, C. E. (2009). The 'wired' universe of organic chemistry. *Nat Chem* 1, 31–36.
- Guarente, L. (1993). Synthetic enhancement in gene interaction: a genetic tool come of age. *Trends Genet* 9, 362–366.
- Gudivada, R. C., Qu, X. A., Chen, J., Jegga, A. G., Neumann, E. K., & Aronow, B. J. (2008). Identifying disease-causal genes using Semantic Web-based representation of integrated genomic and phenomic knowledge. *J Biomed Inform* 41, 717–729.
- Guimera, R., & Amaral, L. A. (2005). Functional cartography of complex metabolic networks. *Nature* 433, 895–900.
- Guimera, R., & Sales-Pardo, M. (2009). Missing and spurious interactions and the reconstruction of complex networks. *Proc Natl Acad Sci USA* 106, 22073–22078.
- Guimera, R., Danon, L., Diaz-Guilera, A., Giral, F., & Arenas, A. (2003). Self-similar community structure in a network of human interactions. *Phys Rev E* 68, 065103.
- Guimera, R., Sales-Pardo, M., & Amaral, L. A. (2007a). Classes of complex networks defined by role-to-role connectivity profiles. *Nat Phys* 3, 63–69.
- Guimera, R., Sales-Pardo, M., & Amaral, L. A. (2007b). A network-based method for target selection in metabolic networks. *Bioinformatics* 23, 1616–1622.
- Gulmann, C., Sheehan, K. M., Kay, E. W., Liotta, L. A., & Petricoin, E. F., III (2006). Array-based proteomics: mapping of protein circuitries for diagnostics, prognostics, and therapy guidance in cancer. *J Pathol* 208, 595–606.
- Gulsoy, G., Gandhi, B., & Kahveci, T. (2012). Topac: alignment of gene regulatory networks using topology-aware coloring. *J Bioinform Comput Biol* 10, 1240001.
- Günther, S., Kuhn, M., Dunkel, M., Campillos, M., Senger, C., Petsalaki, E., et al. (2008). SuperTarget and Matador: resources for exploring drug–target relationships. *Nucleic Acids Res* 36, D919–D922.
- Guo, J., -T., Xu, D., Kim, D., & Xu, Y. (2003). Improving the performance of DomainParser for structural domain partition using neural network. *Nucleic Acids Res* 31, 944–952.
- Guo, H., Ingolia, N. T., Weissman, J. S., & Bartel, D. P. (2010). Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature* 466, 835–840.
- Guo, X., Gao, L., Wei, C., Yang, X., Zhao, Y., & Dong, A. (2011). A computational method based on the integration of heterogeneous networks for predicting disease–gene associations. *PLoS One* 6, e24171.
- Gupta, E. K., & Ito, M. K. (2002). Lovastatin and extended-release niacin combination product: the first drug combination for the management of hyperlipidemia. *Heart Dis* 4, 124–137.
- Gupta, G. P., Nguyen, D. X., Chiang, A. C., Bos, P. D., Kim, J. Y., Nadal, C., et al. (2007). Mediators of vascular remodeling co-opted for sequential steps in lung metastasis. *Nature* 446, 765–770.
- Gupta, R., Bhattacharyya, A., Agosto-Perez, F. J., Wickramasinghe, P., & Davuluri, R. V. (2011). MPromDb update 2010: an integrated resource for annotation and visualization of mammalian gene promoters and CHIP-seq experimental data. *Nucleic Acids Res* 39, D92–D97.
- Gutfraind, A., Meyers, L. A., & Safro, I. (2012). Multiscale network generation. <http://arxiv.org/abs/1207.4266>
- Gyurkó, D., Söti, C., Steták, A., & Csermely, P. (in press). System level mechanisms of adaptation, learning, memory formation and evolvability: the role of chaperone and other networks. *Curr Prot Pept Sci*, 14. <http://arxiv.org/abs/1206.0094>.
- Hakes, L., Pinney, J. W., Robertson, D. L., & Lovell, S. C. (2008). Protein–protein interaction networks and biology—what's the connection? *Nat Biotechnol* 26, 69–72.
- Halabi, N., Rivoire, O., Leibler, S., & Ranganathan, R. (2009). Protein sectors: evolutionary units of three-dimensional structure. *Cell* 138, 774–786.
- Hallén, K., Björkregren, J., & Tegnér, J. (2006). Detection of compound mode of action by computational integration of whole-genome measurements and genetic perturbations. *BMC Bioinforma* 7, 51.
- Hallock, P., & Thomas, M. A. (2012). Integrating the Alzheimer's disease proteome and transcriptome: a comprehensive network model of a complex disease. *OMICS* 16, 37–49.
- Hamp, T., & Rost, B. (2012). Alternative protein–protein interfaces are frequent exceptions. *PLoS Comput Biol* 8, e1002623.
- Han, J. D., Bertin, N., Hao, T., Goldberg, D. S., Berriz, G. F., Zhang, L. V., et al. (2004a). Evidence for dynamically organized modularity in the yeast protein–protein interaction network. *Nature* 430, 88–93.
- Han, K., Ju, B. H., & Jung, H. (2004b). WebInterViewer: visualizing and analyzing molecular interaction networks. *Nucleic Acids Res* 32, W89–W95.
- Han, J. D., Dupuy, D., Bertin, N., Cusick, M. E., & Vidal, M. (2005). Effect of sampling on topology predictions of protein–protein interaction networks. *Nat Biotechnol* 23, 839–844.
- Hanahan, D., & Weinberg, R. A. (2000). The hallmarks of cancer. *Cell* 100, 57–70.
- Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell* 144, 646–674.
- Haney, S., Bardwell, L., & Nie, Q. (2010). Ultrasensitive responses and specificity in cell signaling. *BMC Syst Biol* 4, 119.
- Hansen, N. T., Brunak, S., & Altman, R. B. (2009). Generating genome-scale candidate gene lists for pharmacogenomics. *Clin Pharmacol Ther* 86, 183–189.
- Harrison, R., Papp, B., Pál, C., Oliver, S. G., & Delneri, D. (2007). Plasticity of genetic interactions in metabolic networks of yeast. *Proc Natl Acad Sci U S A* 104, 2307–2312.



- Hartman, J. L. T., Garvik, B., & Hartwell, L. (2001). Principles for the buffering of genetic variation. *Science* 291, 1001–1004.
- Hartsperger, M. L., Strache, R., & Stumpflen, V. (2010). HiNO: an approach for inferring hierarchical organization from regulatory networks. *PLoS One* 5, e13698.
- Hartwell, L. H., Szankasi, P., Roberts, C. J., Murray, A. W., & Friend, S. H. (1997). Integrating genetic approaches into the discovery of anticancer drugs. *Science* 278, 1064–1068.
- Hasan, S., Bonde, B. K., Buchan, N. S., & Hall, M. D. (2012). Network analysis has diverse roles in drug discovery. *Drug Discov Today* 17, 869–874.
- Hase, T., Tanaka, H., Suzuki, Y., Nakagawa, S., & Kitano, H. (2009). Structure of protein interaction networks and their implications on drug design. *PLoS Comput Biol* 5, e1000550.
- Hattori, M., Tanaka, N., Kanehisa, M., & Goto, S. (2010). SIMCOMP/SUBCOMP: chemical structure search servers for network analyses. *Nucleic Acids Res* 38, W652–W656.
- Havugimana, P. C., Hart, G. T., Nepusz, T., Yang, H., Turinsky, A. L., Li, Z., et al. (2012). A census of human soluble protein complexes. *Cell* 150, 1068–1081.
- Hayes, K. R., Vollrath, A. L., Zastrow, G. M., McMillan, B. J., Craven, M., Jovanovich, S., et al. (2005). EDGE: a centralized resource for the comparison, analysis, and distribution of toxicogenomic information. *Mol Pharmacol* 67, 1360–1368.
- He, Z., Zhang, J., Shi, X. H., Hu, L. L., Kong, X., Cai, Y. D., et al. (2010). Predicting drug–target interaction networks based on functional groups and biological features. *PLoS One* 5, e9603.
- He, Y., Zhang, M., Ju, Y., Yu, Z., Lv, D., Sun, H., et al. (2012). dbDEPC 2.0: updated database of differentially expressed proteins in human cancers. *Nucleic Acids Res* 40, D964–D971.
- Hebb, D. O. (1949). *The Organization of Behavior*. New York: Wiley & Sons.
- Hecker, N., Ahmed, J., von Eichborn, J., Dunkel, M., Macha, K., Eckert, A., et al. (2012). SuperTarget goes quantitative: update on drug–target interactions. *Nucleic Acids Res* 40, D1113–D1117.
- Heemskerck, J., Farkas, R., & Kaufmann, P. (2012). Neuroscience networking: linking discovery to drugs. *Neurospychopharmacology* 37, 287–289.
- Hegreness, M., Shores, N., Damian, D., Hartl, D., & Kishony, R. (2008). Accelerated evolution of resistance in multidrug environments. *Proc Natl Acad Sci USA* 105, 13977–13981.
- Henney, A., & Superti-Furga, G. (2008). A network solution. *Nature* 455, 730–731.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behav Brain Sci* 33, 61–83.
- Henry, C. S., DeJongh, M., Best, A. A., Frybarger, P. M., Linsay, B., & Stevens, R. L. (2010). High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nat Biotechnol* 28, 977–982.
- Herzog, F., Kahraman, A., Boehringer, D., Mak, R., Bracher, A., Walzthoeni, T., et al. (2012). Structural probing of a protein phosphatase 2A network by chemical cross-linking and mass spectrometry. *Science* 337, 1348–1352.
- Hernandez, P., Huerta-Cepas, J., Montaner, D., Al-Shahrour, F., Valls, J., Gomez, L., et al. (2007). Evidence for systems-level molecular mechanisms of tumorigenesis. *BMC Genomics* 8, 185.
- Herrgard, M. J., Swainston, N., Dobson, P., Dunn, W. B., Arga, K. Y., Arvas, M., et al. (2008). A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology. *Nat Biotechnol* 26, 1155–1160.
- Hert, J., Keiser, M. J., Irwin, J. J., Oprea, T. I., & Shoichet, B. K. (2008). Quantifying the relationships among drug classes. *J Chem Inf Model* 48, 755–765.
- Hidalgo, C. A., & Rodriguez-Sickert, C. (2008). The dynamics of a mobile phone network. *Phys A* 387, 3017–3024.
- Hidalgo, C. A., Blumm, N., Barabasi, A. L., & Christakis, N. A. (2009). A dynamic network approach for the study of human phenotypes. *PLoS Comput Biol* 5, e1000353.
- Higuero, A. P., Schreyer, A., Bickerton, G. R., Pitt, W. R., Groom, C. R., & Blundell, T. L. (2009). Atomic interactions and profile of small molecules disrupting protein–protein interfaces: the TIMBAL database. *Chem Biol Drug Des* 74, 457–467.
- Hillenmeyer, M. E., Fung, E., Wildenhain, J., Pierce, S. E., Hoon, S., Lee, W., et al. (2008). The chemical genomic portrait of yeast: uncovering a phenotype for all genes. *Science* 320, 362–365.
- Hintze, A., & Adami, C. (2008). Evolution of complex modular biological networks. *PLoS Comput Biol* 4, e23.
- Holford, M., Li, N., Nadkarni, P., & Zhao, H. (2005). VitaPad: visualization tools for the analysis of pathway data. *Bioinformatics* 21, 1596–1602.
- Holme, P. (2011). Metabolic robustness and network modularity: a model study. *PLoS One* 6, e16605.
- Holme, P., & Saramäki (2011). Temporal networks. *Physics Rep* 519, 97–125.
- Hoogboom, C., Theocharis, G., & Kevrekidis, P. G. (2010). Discrete breathers at the interface between a diatomic and a monoatomic granular chain. *Phys Rev E* 82, 061303.
- Hooper, S. D., & Bork, P. (2005). Medusa: a simple tool for interaction graph analysis. *Bioinformatics* 21, 4432–4433.
- Hopkins, A. L. (2008). Network pharmacology: the next paradigm in drug discovery. *Nat Chem Biol* 4, 682–690.
- Hopkins, A. L., Mason, J. S., & Overington, J. P. (2006). Can we rationally design promiscuous drugs? *Curr Opin Struct Biol* 16, 127–136.
- Hormozdiari, F., Salari, R., Bafna, V., & Sahinalp, S. C. (2010). Protein–protein interaction network evaluation for identifying potential drug targets. *J Comput Biol* 17, 669–684.
- Hornbeck, P. V., Kornhauser, J. M., Tkachev, S., Zhang, B., Skrzypek, E., Murray, B., et al. (2012). PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res* 40, D261–D270.
- Hornberg, J. J., Bruggeman, F. J., Westerhoff, H. V., & Lankelma, J. (2006). Cancer: a systems biology disease. *Biosystems* 83, 81–90.
- Hsing, M., Byler, K. G., & Cherkasov, A. (2008). The use of Gene Ontology terms for predicting highly-connected ‘hub’ nodes in protein–protein interaction networks. *BMC Syst Biol* 2, 80.
- Hsu, C. W., Juan, H. F., & Huang, H. C. (2008). Characterization of microRNA-regulated protein–protein interaction network. *Proteomics* 8, 1975–1979.
- Hu, G., & Agarwal, P. (2009). Human disease–drug network based on genomic expression profiles. *PLoS One* 4, e6536.
- Hu, Y., & Bajorath, J. (2010). Polypharmacology directed compound data mining: identification of promiscuous chemotypes with different activity profiles and comparison to approved drugs. *J Chem Inf Model* 50, 2112–2118.
- Hu, Y., & Bajorath, J. (2011). Target family-directed exploration of scaffolds with different SAR profiles. *J Chem Inf Model* 51, 3138–3148.
- Hu, T. M., & Hayton, W. L. (2011). Architecture of the drug–drug interaction network. *J Clin Pharm Ther* 36, 135–143.
- Hu, Z., Hung, J. H., Wang, Y., Chang, Y. C., Huang, C. L., Huyck, M., et al. (2009). VisANT 3.5: multi-scale network visualization, analysis and inference based on the gene ontology. *Nucleic Acids Res* 37, W115–W121.
- Hu, T., Sinnott-Armstrong, N. A., Kiralis, J. W., Andrew, A. S., Karagas, M. R., & Moore, J. H. (2011). Characterizing genetic interactions in human disease association studies using statistical epistasis networks. *BMC Bioinforma* 12, 364.
- Hu, K., Xiang, J., Yang, W., Xu, X., & Tang, Y. (2012). Link prediction in complex networks by multi degree preferential-attachment indices. <http://arxiv.org/abs/1211.1790>
- Huan, T., Sivachenko, A. Y., Harrison, S. H., & Chen, J. Y. (2008). ProteoLens: a visual analytic tool for multi-scale database-driven biological network data mining. *BMC Bioinforma* 9, S5.
- Huang, S. (2001). Genomics, complexity and drug discovery: insights from Boolean network models of cellular regulation. *Pharmacogenomics* 2, 203–222.
- Huang, J., Zhu, H., Haggarty, S. J., Spring, D. R., Hwang, H., Jin, F., et al. (2004). Finding new components of the target of rapamycin (TOR) signaling network through chemical genetics and proteome chips. *Proc Natl Acad Sci USA* 101, 16594–16599.
- Huang, S., Ernberg, I., & Kauffman, S. (2009). Cancer attractors: a systems view of tumors from a gene network dynamics and developmental perspective. *Semin Cell Dev Biol* 20, 869–876.
- Huang, H., Liu, C. C., & Zhou, X. J. (2010a). Bayesian approach to transforming public gene expression repositories into disease diagnosis databases. *Proc Natl Acad Sci USA* 107, 6823–6828.
- Huang, D. Z., Zhou, T., Lafleur, K., Nevado, C., & Cafilisch, A. (2010b). Kinase selectivity potential for inhibitors targeting the ATP binding site: a network analysis. *Bioinformatics* 26, 198–204.
- Huang, D., Zhou, X., Lyon, C. J., Hsueh, W. A., & Wong, S. T. (2010c). MicroRNA-integrated and network-embedded gene selection with diffusion distance. *PLoS One* 5, e13748.
- Huang, Z., Zhu, L., Cao, Y., Wu, G., Liu, X., Chen, Y., et al. (2011a). ASD: a comprehensive database of allosteric proteins and modulators. *Nucleic Acids Res* 39, D663–D669.
- Huang, R., Southall, N., Wang, Y., Yasgar, A., Shinn, P., Jadhav, A., et al. (2011b). The NCGC pharmaceutical collection: a comprehensive resource of clinically approved drugs enabling repurposing and chemical genomics. *Sci Transl Med* 3, 80ps16.
- Huang, T., Cai, Y. D., Chen, L., Hu, L. L., Kong, X. Y., Li, Y. X., et al. (2012a). Selection of reprogramming factors of induced pluripotent stem cells based on the protein interaction network and functional profiles. *Protein Pept Lett* 19, 113–119.
- Huang, H., Wu, X., Sonachalam, M., Mandape, S. N., Pandey, R., MacDorman, K. F., et al. (2012b). PAGED: a pathway and gene-set enrichment database to enable molecular phenotype discoveries. *BMC Bioinforma* 13, S2.
- Hue, M., Riffle, M., Vert, J. P., & Noble, W. S. (2010). Large-scale prediction of protein–protein interactions from structures. *BMC Bioinforma* 11, 144.
- Hughes, T. R. (2002). Yeast and drug discovery. *Funct Integr Genomics* 2, 199–211.
- Hughes, T. R., Marton, M. J., Jones, A. R., Roberts, C. J., Stoughton, R., Armour, C. D., et al. (2000). Functional discovery via a compendium of expression profiles. *Cell* 102, 109–126.
- Huthmacher, C., Hoppe, A., Bulik, S., & Holzthutter, H. G. (2010). Antimalarial drug targets in *Plasmodium falciparum* predicted by stage-specific metabolic network analysis. *BMC Syst Biol* 4, 120.
- Hwang, W. C., Zhang, A., & Ramanathan, M. (2008). Identification of information flow-modulating drug targets: a novel bridging paradigm for drug discovery. *Clin Pharmacol Ther* 84, 563–572.
- Hwang, D., Lee, I. Y., Yoo, H., Gehlenborg, N., Cho, J. H., Petritis, B., et al. (2009). A systems approach to prion disease. *Mol Syst Biol* 5, 252.
- Hwang, T., Zhang, W., Xie, M., Liu, J., & Kuang, R. (2011). Inferring disease and gene set associations with rank coherence in networks. *Bioinformatics* 27, 2692–2699.
- Ideker, T., & Krogan, N. J. (2012). Differential network biology. *Mol Syst Biol* 8, 565.
- Ideker, T., & Lauffenburger, D. (2003). Building with a scaffold: emerging strategies for high- to low-level cellular modeling. *Trends Biotechnol* 21, 255–262.
- Ideker, T. E., Thorsson, V., & Karp, R. M. (2000). Discovery of regulatory interactions through perturbation: inference and experimental design. *Pac Symp Biocomput*, 305–316.
- Iguchi, H., Kosaka, N., & Ochiya, T. (2010). Versatile applications of microRNA in anti-cancer drug discovery: from therapeutics to biomarkers. *Curr Drug Discov Technol* 7, 95–105.
- Inoue, K., Shimozone, S., Yoshida, H., & Kurata, H. (2012). Application of approximate pattern matching in two dimensional spaces to grid layout for biochemical network maps. *PLoS One* 7, e37739.
- International Human Genome Sequencing Consortium (2004). Finishing the euchromatic sequence of the human genome. *Nature* 431, 931–945.
- Iorio, F., Tagliaferri, R., & di Bernardo, D. (2009). Identifying network of drug mode of action by gene expression profiling. *J Comput Biol* 16, 241–251.
- Iorio, F., Bosotti, R., Scacheri, E., Belcastro, V., Mithbaokar, P., Ferriero, R., et al. (2010). Discovery of drug mode of action and drug repositioning from transcriptional responses. *Proc Natl Acad Sci USA* 107, 14621–14626.

- Iossifov, I., Zheng, T., Baron, M., Gilliam, T. C., & Rzhetsky, A. (2008). Genetic-linkage mapping of complex hereditary disorders to a whole-genome molecular-interaction network. *Genome Res* 18, 1150–1162.
- Isin, B., Estiu, G., Wiest, O., & Oltvai, Z. N. (2012). Identifying ligand binding conformations of the  $\beta_2$ -adrenergic receptor by using its agonists as computational probes. *PLoS One* 7, e50186.
- Iskar, M., Zeller, G., Zhao, X. M., van Noort, V., & Bork, P. (2012). Drug discovery in the age of systems biology: the rise of computational approaches for data integration. *Curr Opin Biotechnol* 23, 609–616.
- Ispolatov, I., & Maslov, S. (2008). Detection of the dominant direction of information flow and feedback links in densely interconnected regulatory networks. *BMC Bioinforma* 9, 424.
- Iyer, P., Hu, Y., & Bajorath, J. (2011a). SAR monitoring of evolving compound data sets using activity landscapes. *J Chem Inf Model* 51, 532–540.
- Iyer, P., Stumpfe, D., & Bajorath, J. (2011b). Molecular mechanism-based network-like similarity graphs reveal relationships between different types of receptor ligands and structural changes that determine agonistic, inverse-agonistic, and antagonistic effects. *J Chem Inf Model* 51, 1281–1286.
- Iyer, P., Wawer, M., & Bajorath, J. (2011c). Comparison of two- and three-dimensional activity landscape representations for different compound data sets. *Med Chem Comm* 2, 113–118.
- Jacobs, D. J., Dallakyan, S., Wood, G. G., & Heckathorne, A. (2003). Network rigidity at finite temperature: relationships between thermodynamic stability, the non-additivity of entropy, and cooperativity in molecular systems. *Phys Rev E* 68, 061109.
- Jacobs, D. J., Rader, A. J., Kuhn, L. A., & Thorpe, M. F. (2001). Protein flexibility predictions using graph theory. *Proteins* 44, 150–165.
- Jayawardhana, B., Kell, D. B., & Rattray, M. (2008). Bayesian inference of the sites of perturbations in metabolic pathways via Markov chain Monte Carlo. *Bioinformatics* 24, 1191–1197.
- Jamshidi, N., & Palsson, B. O. (2007). Investigating the metabolic capabilities of *Mycobacterium tuberculosis* H37Rv using the in silico strain iNJ661 and proposing alternative drug targets. *BMC Syst Biol* 1, 26.
- Jamshidi, N., & Palsson, B. O. (2008). Top-down analysis of temporal hierarchy in biochemical reaction networks. *PLoS Comput Biol* 4, e1000177.
- Janjic, V., & Przulj, N. (2012). Biological function through network topology: a survey of the human diseasome. *Brief Funct Genomics* 11, 522–532.
- Jeon, J., Nam, H.-J., Choi, Y. S., Yang, J.-S., Hwang, J., & Kim, S. (2011). Molecular evolution of protein conformational changes revealed by a network of evolutionarily coupled residues. *Mol Biol Evol* 28, 2675–2685.
- Jeong, H., Mason, S. P., Barabasi, A. L., & Oltvai, Z. N. (2001). Lethality and centrality in protein networks. *Nature* 411, 41–42.
- Jerne, N. K. (1974). Towards a network theory of the immune system. *Ann Immunol* 125C, 373–389.
- Jerne, N. K. (1984). Idiotypic networks and other preconceived ideas. *Immunol Rev* 79, 5–23.
- Jessulat, M., Pitre, S., Gui, Y., Hooshyar, M., Omidi, K., Samanfar, B., et al. (2011). Recent advances in protein–protein interaction prediction: experimental and computational methods. *Expert Opin Drug Discov* 6, 921–935.
- Jia, J., Zhu, F., Ma, X., Cao, Z., Li, Y., & Chen, Y. Z. (2009). Mechanisms of drug combinations: interaction and network perspectives. *Nat Rev Drug Discov* 8, 111–128.
- Jiang, X., Liu, B., Jiang, J., Zhao, H., Fan, M., Zhang, J., et al. (2008). Modularity in the genetic disease-phenotype network. *FEBS Lett* 582, 2549–2554.
- Jiang, R., Gan, M., & He, P. (2011). Constructing a gene semantic similarity network for the inference of disease genes. *BMC Systems Biol* 5, S2.
- Jianu, R., Yu, K., Cao, L., Nguyen, V., Salomon, A. R., & Laidlaw, D. H. (2010). Visual integration of quantitative proteomic data, pathways, and protein interactions. *IEEE Trans Vis Comput Graph* 16, 609–620.
- Jin, G., Zhou, X., Wang, H., Zhao, H., Cui, K., Zhang, X. S., et al. (2008). The knowledge-integrated network biomarkers discovery for major adverse cardiac events. *J Proteome Res* 7, 4013–4021.
- Jin, G., Fu, C., Zhao, H., Cui, K., Chang, J., & Wong, S. T. (2012). A novel method of transcriptional response analysis to facilitate drug repositioning for cancer therapy. *Cancer Res* 72, 33–44.
- Jonsson, P. F., & Bates, P. A. (2006). Global topological features of cancer proteins in the human interactome. *Bioinformatics* 22, 2291–2297.
- Johnson, M., & Maggiora, G. M. (Eds.). (1990). *Concepts and applications of molecular similarity*. New York: John Wiley & Sons.
- Jonsson, P. F., Cavanna, T., Zicha, D., & Bates, P. A. (2006). Cluster analysis of networks generated through homology: automatic identification of important protein communities involved in cancer metastasis. *BMC Bioinforma* 7, 2.
- Joseph, A., & Chen, G. (2012). Composite centrality: a natural scale for complex evolving networks. <http://arxiv.org/abs/1211.3951>
- Joseph, R. E., Xie, Q., & Andreotti, A. H. (2010). Identification of an allosteric signaling network within Tec family kinases. *J Mol Biol* 403, 231–242.
- Jothi, R., Balaji, S., Wuster, A., Grochow, J. A., Gsponer, J., Przytycka, T. M., et al. (2009). Genomic analysis reveals a tight link between transcription factor dynamics and regulatory network architecture. *Mol Syst Biol* 5, 294.
- Judson, R. S., Martin, M. T., Egeghy, P., Gangwal, S., Reif, D. M., Kothiyi, P., et al. (2012). Aggregating data for computational toxicology applications: the U.S. Environmental Protection Agency (EPA) Aggregated Computational Toxicology Resource (ACToR) system. *Int J Mol Sci* 13, 1805–1831.
- Jung, J. P., Moyano, J. V., & Collier, J. H. (2011). Multifactorial optimization of endothelial cell growth using modular synthetic extracellular matrices. *Integr Biol* 3, 185–196.
- Jurman, G., Filosi, M., Visintainer, R., Riccadonna, S., & Furlanello, C. (2012a). Stability indicators in network reconstruction. <http://arxiv.org/abs/1209.1654>
- Jurman, G., Visintainer, R., Riccadonna, S., Filosi, M., & Furlanello, C. (2012b). A global distance for network comparison. <http://arxiv.org/abs/1201.2931>
- Kahle, J. J., Gulbahce, N., Shaw, C. A., Lim, J., Hill, D. E., Barabasi, A. L., et al. (2011). Comparison of an expanded ataxia interactome with patient medical records reveals a relationship between macular degeneration and ataxia. *Hum Mol Genet* 20, 510–527.
- Kaltenbach, L. S., Romero, E., Becklin, R. R., Chettier, R., Bell, R., Phansalkar, A., et al. (2007). Huntingtin interacting proteins are genetic modifiers of neurodegeneration. *PLoS Genet* 3, e82.
- Kandasamy, K., Mohan, S. S., Raju, R., Keerthikumar, S., Kumar, G. S., Venugopal, A. K., et al. (2010). NetPath: a public resource of curated signal transduction pathways. *Genome Biol* 11, R3.
- Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., & Tanabe, M. (2012). KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res* 40, D109–D114.
- Kannan, N., & Vishveshwara, S. (1999). Identification of side-chain clusters in protein structures by a graph spectral method. *J Mol Biol* 292, 441–464.
- Kanté, M. M., Limouzy, V., Mary, A., & Nourine, L. (2011). On the enumeration of minimal dominating sets and related notions. *Lect Notes Comput Sci* 6914, 298–309.
- Kaplowitz, N. (2001). Drug-induced liver disorders: implications for drug development and regulation. *Drug Saf* 24, 483–490.
- Kar, G., Gursoy, A., & Keskin, O. (2009). Human cancer protein–protein interaction network: a structural perspective. *PLoS Comput Biol* 5, e1000601.
- Karlebach, G., & Shamir, R. (2010). Minimally perturbing a gene regulatory network to avoid a disease phenotype: the glioma network as a test case. *BMC Systems Biol* 4, 15.
- Karni, S., Soreq, H., & Sharan, R. (2009). A network-based method for predicting disease-causing genes. *J Comput Biol* 16, 181–189.
- Karp, P. D., Paley, S. M., Krummenacker, M., Latendresse, M., Dale, J. M., Lee, T. J., et al. (2010). Pathway Tools version 13.0: integrated software for pathway/genome informatics and systems biology. *Brief Bioinform* 11, 40–79.
- Kashtan, N., Itzkovitz, S., Milo, R., & Alon, U. (2004). Efficient sampling algorithm for estimating subgraph concentrations and detecting network motifs. *Bioinformatics* 20, 1746–1758.
- Kauffman, S., Peterson, C., Samuelsson, B., & Troein, C. (2003). Random Boolean network models and the yeast transcriptional network. *Proc Natl Acad Sci USA* 100, 14796–14799.
- Keiser, M. J., Roth, B. L., Armbruster, B. N., Ernsberger, P., Irwin, J. J., & Shoichet, B. K. (2007). Relating protein pharmacology by ligand chemistry. *Nat Biotechnol* 25, 197–206.
- Keiser, M. J., Setola, V., Irwin, J. J., Laggner, C., Abbas, A. I., Hufeisen, S. J., et al. (2009). Predicting new molecular targets for known drugs. *Nature* 462, 175–181.
- Keiser, M. J., Irwin, J. J., & Shoichet, B. K. (2010). The chemical basis of pharmacology. *Biochemistry* 49, 10267–10276.
- Keith, C. T., & Zimmermann, G. R. (2004). Multi-target lead discovery for networked systems. *Curr Drug Discov* 2004, 19–23.
- Keith, C. T., Borisy, A. A., & Stockwell, B. R. (2005). Multicomponent therapeutics for networked systems. *Nat Rev Drug Discov* 4, 71–78.
- Kell, D. B. (2006). Systems biology, metabolic modelling and metabolomics in drug discovery and development. *Drug Discov Today* 11, 1085–1092.
- Kell, D. B., Dobson, P. D., Bilisland, E., & Oliver, S. G. (2013). The promiscuous binding of pharmaceutical drugs and their transporter-mediated uptake into cells: what we (need to) know and how we can do so. *Drug Discov Today* 18, 218–239.
- Kellenberger, E., Schalon, C., & Rognan, D. (2008). How to measure the similarity between protein ligand-binding sites? *Curr Comput Aided Drug Des* 4, 209–220.
- Kelley, B. P., Yuan, B., Lewitter, F., Sharan, R., Stockwell, B. R., & Ideker, T. (2004). PathBLAST: a tool for alignment of protein interaction networks. *Nucleic Acids Res* 32, W83–W88.
- Kelley, R., & Ideker, T. (2005). Systematic interpretation of genetic interactions using protein networks. *Nat Biotechnol* 23, 561–566.
- Kenific, C. M., Thorburn, A., & Debnath, J. (2010). Autophagy and metastasis: another double-edged sword. *Curr Opin Cell Biol* 22, 241–245.
- Kerrien, S., Aranda, B., Breuza, L., Bridge, A., Broackes-Carter, F., Chen, C., et al. (2012). The IntAct molecular interaction database in 2012. *Nucleic Acids Res* 40, D841–D846.
- Keskin, O., Ma, B., & Nussinov, R. (2005). Hot regions in protein–protein interactions: the organization and contribution of structurally conserved hot spot residues. *J Mol Biol* 345, 1281–1294.
- Keskin, O., Gursoy, A., Ma, B., & Nussinov, R. (2007). Towards drugs targeting multiple proteins in a systems biology approach. *Curr Top Med Chem* 7, 943–951.
- Khazaei, T., McGuigan, A., & Mahadevan, R. (2012). Ensemble modeling of cancer metabolism. *Front Physiol* 3, 135.
- Kholodenko, B. N. (2006). Cell-signalling dynamics in time and space. *Nat Rev Mol Cell Biol* 7, 165–176.
- Kholodenko, B. N., Kiyatkin, A., Bruggeman, F. J., Sontag, E., Westerhoff, H. V., & Hoek, J. B. (2002). Untangling the wires: a strategy to trace functional interactions in signaling and gene networks. *Proc Natl Acad Sci USA* 99, 12841–12846.
- Kier, L. B., & Hall, L. H. (2005). The prediction of ADMET properties using structure information representations. *Chem Biodivers* 2, 1428–1437.
- Kim, M., & Leskovec, J. (2011). Multiplicative attribute graph model of real-world networks. *Internet Math* 8, 113–160.
- Kim, P. M., Lu, L. J., Xia, Y., & Gerstein, M. B. (2006). Relating three-dimensional structures to protein networks provides evolutionary insights. *Science* 314, 1938–1941.
- Kim, D., Rath, O., Kolch, W., & Cho, K. H. (2007). A hidden oncogenic positive feedback loop caused by crosstalk between Wnt and ERK pathways. *Oncogene* 26, 4571–4579.

- Kim, J. M., Jung, Y. S., Sungur, E. A., Han, K. H., Park, C., & Sohn, I. (2008). A copula method for modeling directional dependence of genes. *BMC Bioinforma* 9, 225.
- Kim, H. U., Kim, T. Y., & Lee, S. Y. (2010). Genome-scale metabolic network analysis and drug targeting of multi-drug resistant pathogen *Acinetobacter baumannii* AYE. *Mol Biosyst* 6, 339–348.
- Kim, H. U., Kim, S. Y., Jeong, H., Kim, T. Y., Kim, J. J., Choy, H. E., et al. (2011a). Integrative genome-scale metabolic analysis of *Vibrio vulnificus* for drug targeting and discovery. *Mol Syst Biol* 7, 460.
- Kim, Y., Kim, T. K., Yoo, J., You, S., Lee, I., Carlson, G., et al. (2011b). Principal network analysis: identification of subnetworks representing major dynamics using gene expression data. *Bioinformatics* 27, 391–398.
- Kim, H. U., Sohn, S. B., & Lee, S. Y. (2012). Metabolic network modeling and simulation for drug targeting and discovery. *Biotechnol J* 7, 330–342.
- Kinnings, S. L., Xie, L., Fung, K. H., Jackson, R. M., & Bourne, P. E. (2010). The *Mycobacterium tuberculosis* drugome and its polypharmacological implications. *PLoS Comput Biol* 6, e1000976.
- Kirkpatrick, P., & Ellis, C. (2004). Chemical space. *Nature* 432, 823.
- Kirkwood, T. B., & Kowald, A. (1997). Network theory of aging. *Exp Gerontol* 32, 395–399.
- Kirouac, D. C., Saez-Rodriguez, J., Swantek, J., Burke, J. M., Lauffenburger, D. A., & Sorger, P. K. (2012). Creating and analyzing pathway and protein interaction compendia for modelling signal transduction networks. *BMC Syst Biol* 6, 29.
- Kiss, H. J., Mihalik, A., Nanasi, T., Ory, B., Spiro, Z., Soti, C., et al. (2009). Ageing as a price of cooperation and complexity: self-organization of complex systems causes the gradual deterioration of constituent networks. *Bioessays* 31, 651–664.
- Kitano, H. H. (2004a). Biological robustness. *Nature Rev Genetics* 5, 826–837.
- Kitano, H. H. (2004b). Cancer as a robust system: implications to anticancer therapy. *Nat Rev Cancer* 4, 227–235.
- Kitano, H. H. (2007). A robustness-based approach to systems-oriented drug design. *Nat Rev Drug Discov* 6, 202–210.
- Kitano, H., Funahashi, A., Matsuoka, Y., & Oda, K. (2005). Using process diagrams for the graphical representation of biological networks. *Nat Biotechnol* 23, 961–966.
- Kitsak, M., Gallos, L. K., Havlin, S., Liljeros, F., Muchnik, L., Stanley, H. E., et al. (2010). Identifying influential spreaders in complex networks. *Nat Phys* 6, 888–893.
- Kiyosawa, N., Manabe, S., Sanbuissho, A., & Yamoto, T. (2010). Gene set-level network analysis using a toxicogenomics database. *Genomics* 96, 39–49.
- Klamt, S., & Gilles, E. D. (2004). Minimal cut sets in biochemical reaction networks. *Bioinformatics* 20, 226–234.
- Klinke, D. J., II (2010). Signal transduction networks in cancer: quantitative parameters influence network topology. *Cancer Res* 70, 1773–1782.
- Klinke, D. J., II, & Finley, S. D. (2012). Timescale analysis of rule-based biochemical reaction networks. *Biotechnol Prog* 28, 33–44.
- Klukas, C., & Schreiber, F. (2007). Dynamic exploration and editing of KEGG pathway diagrams. *Bioinformatics* 23, 344–350.
- Klussmann, E., & Scott, J. (2008). *Protein-Protein Interactions as New Drug Targets*. Heidelberg: Springer.
- Knox, C., Law, V., Jewison, T., Liu, P., Ly, S., Frolkis, A., et al. (2011). DrugBank 3.0: a comprehensive resource for 'omics' research on drugs. *Nucleic Acids Res* 39, D1035–D1041.
- Koch, C. (2012). Modular biological complexity. *Science* 337, 531–532.
- Kohler, J., Baumbach, J., Taubert, J., Specht, M., Skusa, A., Rugg, A., et al. (2006). Graph-based analysis and visualization of experimental results with ONDEX. *Bioinformatics* 22, 1383–1390.
- Kohler, S., Bauer, S., Horn, D., & Robinson, P. N. (2008). Walking the interactome for prioritization of candidate disease genes. *Am J Hum Genet* 82, 949–958.
- Kola, I., & Bell, J. (2011). A call to reform the taxonomy of human disease. *Nat Rev Drug Discov* 10, 641–642.
- Kolb, P., Ferreira, R. S., Irwin, J. J., & Shoichet, B. K. (2009). Docking and cheminformatic screens for new ligands and targets. *Curr Opin Biotechnol* 20, 429–436.
- Kolchinsky, A., Lourenco, A., Li, L., & Rocha, L. M. (2013). Evaluation of linear classifiers on articles containing pharmacokinetic evidence of drug–drug interactions. *Pac Symp Biocomput* 2013, 409–420.
- Kolodkin, A., Boogerd, F. C., Plant, N., Bruggeman, F. J., Goncharuk, V., Lunshof, J., et al. (2012). Emergence of the silicon human and network targeting drugs. *Eur J Pharm Sci* 46, 190–197.
- Komurov, K., & White, M. (2007). Revealing static and dynamic modular architecture of the eukaryotic protein interaction network. *Mol Syst Biol* 3, 110.
- Konrat, R. (2009). The protein meta-structure: a novel concept for chemical and molecular biology. *Cell Mol Life Sci* 66, 3625–3639.
- Korcsmáros, T., Szalay, M. S., Bóde, C., Kovács, I. A., & Csermely, P. (2007). How to design multi-target drugs: target-search options in cellular networks. *Expert Op Drug Discov* 2, 799–808.
- Korcsmáros, T., Farkas, I. J., Szalay, M. S., Rovó, P., Fazekas, D., Spiró, Z., et al. (2010). Uniformly curated signaling pathways reveal tissue-specific cross-talks, novel pathway components, and drug target candidates. *Bioinformatics* 26, 2042–2050.
- Korcsmáros, T., Szalay, M. S., Rovó, P., Palotai, R., Fazekas, D., Lenti, K., et al. (2011). Signalogs: orthology-based identification of novel signaling pathway components in three metazoans. *PLoS One* 6, e19240.
- Koshland, D. E. (1958). Application of a theory of enzyme specificity to protein synthesis. *Proc Natl Acad Sci USA* 44, 98–104.
- Kotelnikova, E., Yuryev, A., Mazo, I., & Daraselia, N. (2010). Computational approaches for drug repositioning and combination therapy design. *J Bioinform Comput Biol* 8, 593–606.
- Kotera, M., Yamanishi, Y., Moriya, Y., Kanehisa, M., & Goto, S. (2012). GENIES: gene network inference engine based on supervised analysis. *Nucleic Acids Res* 40, W162–W167.
- Kotlyar, M., Fortney, K., & Jurisica, I. (2012). Network-based characterization of drug-regulated genes, drug targets, and toxicity. *Methods* 57, 499–507.
- Kovács, I. A., Szalay, M. S., & Csermely, P. (2005). Water and molecular chaperones act as weak links of protein folding networks: energy landscape and punctuated equilibrium changes point towards a game theory of proteins. *FEBS Lett* 579, 2254–2260.
- Kovács, I. A., Palotai, R., Szalay, M. S., & Csermely, P. (2010). Community landscapes: a novel, integrative approach for the determination of overlapping network modules. *PLoS One* 7, e12528.
- Kowalik, M., Gothard, C. M., Drews, A. M., Gothard, N. A., Wieckiewicz, A., Fuller, P. E., et al. (2012). Parallel optimization of synthetic pathways within the network of organic chemistry. *Angew Chem Int Ed* 51, 7928–7932.
- Kozakov, D., Hall, D. R., Chuang, G. Y., Cencic, R., Brenke, R., Grove, L. E., et al. (2011). Structural conservation of druggable hot spots in protein–protein interfaces. *Proc Natl Acad Sci USA* 108, 13528–13533.
- Kozhenkov, S., & Baitaluk, M. (2012). Mining and integration of pathway diagrams from imaging data. *Bioinformatics* 28, 739–742.
- Krauthammer, M., Kaufmann, C. A., Gilliam, T. C., & Rzhetsky, A. (2004). Molecular triangulation: bridging linkage and molecular-network information for identifying candidate genes in Alzheimer's disease. *Proc Natl Acad Sci USA* 101, 15148–15153.
- Krein, M. P., & Sukumar, N. (2011). Exploration of the topology of chemical spaces with network measures. *J Phys Chem A* 115, 12905–12918.
- Krek, A., Grun, D., Poy, M. N., Wolf, R., Rosenberg, L., Epstein, E. J., et al. (2005). Combinatorial microRNA target predictions. *Nat Genet* 37, 495–500.
- Krings, G., Karsai, M., Bernharsson, S., Blondel, V. D., & Saramäki, J. (2012). Effects of time window size and placement on the structure of aggregated networks. *EPJ Data Sci* 1, 4.
- Krishnan, A., Zbilut, J. P., Tomita, M., & Giuliani, A. (2008). Proteins as networks: usefulness of graph theory in protein science. *Curr Protein Pept Sci* 9, 28–38.
- Krzywinski, M., Birol, I., Jones, S. J., & Marra, M. A. (2012). Hive plots—rational approach to visualizing networks. *Brief Bioinform* 13, 627–644.
- Kuchaiev, O., & Przulj, N. (2011). Integrative network alignment reveals large regions of global network similarity in yeast and human. *Bioinformatics* 27, 1390–1396.
- Kuchaiev, O., Stevanovic, A., Hayes, W., & Przulj, N. (2011). GraphCrunch 2: software tool for network modeling, alignment and clustering. *BMC Bioinforma* 12, 24.
- Kuhn, M., Campillos, M., Letunic, I., Jensen, L. J., & Bork, P. (2010). A side effect resource to capture phenotypic effects of drugs. *Mol Syst Biol* 6, 343.
- Kuhn, B., Fuchs, J. E., Reutlinger, M., Stahl, M., & Taylor, N. R. (2011). Rationalizing tight ligand binding through cooperative interaction networks. *J Chem Inf Model* 51, 3180–3198.
- Kuhn, M., Szklarczyk, D., Franceschini, A., von Mering, C., Jensen, L. J., & Bork, P. (2012). STITCH 3: zooming in on protein–chemical interactions. *Nucleic Acids Res* 40, D876–D880.
- Kumar, N., Afeyan, R., Kim, H. D., & Lauffenburger, D. A. (2008). Multipathway model enables prediction of kinase inhibitor cross-talk effects on migration of Her2-overexpressing mammary epithelial cells. *Mol Pharmacol* 73, 1668–1678.
- Kung, C., Hixon, J., Choe, S., Marks, K., Gross, S., Murphy, E., et al. (2012). Small molecule activation of PKM2 in cancer cells induces serine auxotrophy. *Chem Biol* 19, 1187–1198.
- Kushwaha, S. K., & Shakya, M. (2009). PINAT1.0: protein interaction network analysis tool. *Bioinformatics* 3, 419–421.
- Kushwaha, S. K., & Shakya, M. (2010). Protein interaction network analysis—approach for potential drug target identification in *Mycobacterium tuberculosis*. *J Theor Biol* 262, 284–294.
- Lage, K., Karlberg, E. O., Storling, Z. M., Olason, P. I., Pedersen, A. G., Rigina, O., et al. (2007). A human phenome–interactome network of protein complexes implicated in genetic disorders. *Nat Biotechnol* 25, 309–316.
- Lage, K., Hansen, N. T., Karlberg, E. O., Eklund, A. C., Roque, F. S., Donahoe, P. K., et al. (2008). A large-scale analysis of tissue-specific pathology and gene expression of human disease genes and complexes. *Proc Natl Acad Sci USA* 105, 20870–20875.
- Lai, Y. H., Li, Z. C., Chen, L. L., Dai, Z., & Zou, X. Y. (2012). Identification of potential host proteins for influenza A virus based on topological and biological characteristics by proteome-wide network approach. *J Proteomics* 75, 2500–2513.
- Lamb, J., Crawford, E. D., Peck, D., Modell, J. W., Blat, I. C., Wrobel, M. J., et al. (2006). The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* 313, 1929–1935.
- Landis, S. C., Amara, S. G., Asadullah, K., Austin, C. P., Blumenstein, R., Bradley, E. W., et al. (2012). A call for transparent reporting to optimize the predictive value of preclinical research. *Nature* 490, 187–191.
- Langfelder, P., Luo, R., Oldham, M. C., & Horvath, S. (2011). Is my network module preserved and reproducible? *PLoS Comput Biol* 7, e1001057.
- Laskowski, R. A., Luscombe, N. M., Swindells, M. B., & Thornton, J. M. (1996). Protein clefts in molecular recognition and function. *Protein Sci* 5, 2438–2452.
- Latora, V., & Marchiori, M. (2001). Efficient behaviour of small-world networks. *Phys Rev Lett* 87, 198701.
- Le, D. H., & Kwon, Y. K. (2012). GPEC: a Cytoscape plug-in for random walk-based gene prioritization and biomedical evidence collection. *Comput Biol Chem* 37, 17–23.
- Ledford, H. (2012). Drug candidates derailed in case of mistaken identity. *Nature* 483, 519.
- Lee, G. M., & Craik, C. S. (2009). Trapping moving targets with small molecules. *Science* 324, 213–215.
- Lee, D. S., Park, J., Kay, K. A., Christakis, N. A., Oltvai, Z. N., & Barabasi, A. L. (2008a). The implications of human metabolic network topology for disease comorbidity. *Proc Natl Acad Sci USA* 105, 9880–9885.
- Lee, I., Lehner, B., Crombie, C., Wong, W., Fraser, A. G., & Marcotte, E. M. (2008b). A single gene network accurately predicts phenotypic effects of gene perturbation in *Caenorhabditis elegans*. *Nat Genet* 40, 181–188.

- Lee, E., Jung, H., Radivojac, P., Kim, J. W., & Lee, D. (2009). Analysis of AML genes in dysregulated molecular networks. *BMC Bioinforma* 10, S2.
- Lee, S., Lee, K. H., Song, M., & Lee, D. (2011). Building the process-drug-side effect network to discover the relationship between biological processes and side effects. *BMC Bioinforma* 12, S2.
- Lee, P., Lakshmanan, L. V. S., & Yu, J. X. (2012a). On top-k structural similarity search. *IEEE 28th Intl Conf Data Engin* (pp. 774–785).
- Lee, H. S., Bae, T., Lee, J. H., Kim, D. G., Oh, Y. S., Jang, Y., et al. (2012b). Rational drug repositioning guided by an integrated pharmacological network of protein, disease and drug. *BMC Syst Biol* 6, 80.
- Lee, J. H., Kim, D. G., Bae, T. J., Rho, K., Kim, J. T., Lee, J. J., et al. (2012c). CDA: combinatorial drug discovery using transcriptional response modules. *PLoS One* 7, e42573.
- Lee, M. J., Ye, A. S., Gardino, A. K., Heijink, A. M., Sorger, P. K., MacBeath, G., et al. (2012d). Sequential application of anticancer drugs enhances cell death by rewiring apoptotic signaling networks. *Cell* 149, 780–794.
- Leeson, P. D., & Springthorpe, B. (2007). The influence of drug-like concepts on decision-making in medicinal chemistry. *Nat Rev Drug Discov* 6, 881–890.
- Lehár, J., Zimmermann, G. R., Krueger, A. S., Molnar, R. A., Ledell, J. T., Heilbut, A. M., et al. (2007). Chemical combination effects predict connectivity in biological systems. *Mol Syst Biol* 3, 80.
- Leicht, E. A., Holme, P., & Newman, M. E. (2006). Vertex similarity in networks. *Phys Rev E* 73, 026120.
- Lemke, N., Heredia, F., Barcellos, C. K., Dos Reis, A. N., & Mombach, J. C. (2004). Essentiality and damage in metabolic networks. *Bioinformatics* 20, 115–119.
- Lentz, H. H. K., Selhorts, T., & Sokolov, I. M. (2012). Unfolding accessibility provides a macroscopic approach to temporal networks. <http://arxiv.org/abs/1210.2283>
- Lepoivre, C., Bergon, A., Lopez, F., Perumal, N. B., Nguyen, C., Imbert, J., et al. (2012). TranscriptionBrowser 3.0: introducing a new compendium of molecular interactions and a new visualization tool for the study of gene regulatory networks. *BMC Bioinforma* 13, 19.
- Lewis, B. P., Shih, I. H., Jones-Rhoades, M. W., Bartel, D. P., & Burge, C. B. (2003). Prediction of mammalian microRNA targets. *Cell* 115, 787–798.
- Lewis, B. P., Burge, C. B., & Bartel, D. P. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* 120, 15–20.
- Lewis, A. C., Jones, N. S., Porter, M. A., & Deane, C. M. (2012). What evidence is there for the homology of protein–protein interactions? *PLoS Comput Biol* 8, e1002645.
- Lewitzky, M., Simister, P. C., & Feller, S. M. (2012). Beyond ‘furballs’ and ‘dumpling soups’—towards a molecular architecture of signaling complexes and networks. *FEBS Lett* 586, 2740–2750.
- Li, W., & Kurata, H. (2005). A grid layout algorithm for automatic drawing of biochemical networks. *Bioinformatics* 21, 2036–2042.
- Li, Q., & Lai, L. (2007). Prediction of potential drug targets based on simple sequence properties. *BMC Bioinforma* 8, 353.
- Li, J., & Li, A. (2013). Characters and patterns of communities in networks. <http://arxiv.org/abs/1301.2957>
- Li, Y., & Patra, J. C. (2010). Genome-wide inferring gene–phenotype relationship by walking on the heterogeneous network. *Bioinformatics* 26, 1219–1224.
- Li, Z., Zhang, S., Wang, Y., Zhang, X. S., & Chen, L. (2007). Alignment of molecular networks by integer quadratic programming. *Bioinformatics* 23, 1631–1639.
- Li, J., Zhu, X., & Chen, J. Y. (2009a). Building disease-specific drug–protein connectivity maps from molecular interaction networks and PubMed abstracts. *PLoS Comput Biol* 5, e1000450.
- Li, L., Zhang, K., Lee, J., Cordes, S., Davis, D. P., & Tang, Z. (2009b). Discovering cancer genes by integrating network and functional properties. *BMC Med Genomics* 2, 61.
- Li, F., Li, P., Xu, W., Peng, Y., Bo, X., & Wang, S. (2010a). PerturbationAnalyzer: a tool for investigating the effects of concentration perturbation on protein interaction networks. *Bioinformatics* 26, 275–277.
- Li, X., Gianoulis, T. A., Yip, K. Y., Gerstein, M., & Snyder, M. (2010b). Extensive in vivo metabolite–protein interactions revealed by large-scale systematic analyses. *Cell* 143, 639–650.
- Li, L., Bum-Erdene, K., Baenziger, P. H., Rosen, J. J., Hemmert, J. R., Nellis, J. A., et al. (2010c). BioDrugScreen: a computational drug design resource for ranking molecules docked to the human proteome. *Nucleic Acids Res* 38, D765–D773.
- Li, J., Lenferink, A. E., Deng, Y., Collins, C., Cui, Q., Purisima, E. O., et al. (2010d). Identification of high-quality cancer prognostic markers and metastasis network modules. *Nat Commun* 1, 34.
- Li, L., Zhou, X., Ching, W. K., & Wang, P. (2010e). Predicting enzyme targets for cancer drugs by profiling human metabolic reactions in NCI-60 cell lines. *BMC Bioinforma* 11, 501.
- Li, M., Wang, J., Chen, X., Wang, H., & Pan, Y. (2011a). A local average connectivity-based method for identifying essential proteins from the network level. *Comput Biol Chem* 35, 143–150.
- Li, Y., Wen, Z., Xiao, J., Yin, H., Yu, L., Yang, L., et al. (2011b). Predicting disease-associated substitution of a single amino acid by analyzing residue interactions. *BMC Bioinforma* 12, 14.
- Li, Q., Li, X., Li, C., Chen, L., Song, J., Tang, Y., et al. (2011c). A network-based multi-target computational estimation scheme for anticoagulant activities of compounds. *PLoS One* 6, e14774.
- Li, S., Zhang, B., & Zhang, N. (2011d). Network target for screening synergistic drug combinations with application to traditional Chinese medicine. *BMC Syst Biol* 5, S10.
- Li, X. L., Qian, L., Bittner, M. L., & Dougherty, E. R. (2011e). Characterization of drug efficacy regions based on dosage and frequency schedules. *IEEE Trans Biomed Eng* 58, 488–498.
- Li, L., Tibeche, C., Fu, C., Kaneko, T., Moran, M. F., Schiller, M. R., et al. (2012a). The human phosphotyrosine signaling network: evolution and hotspots of hijacking in cancer. *Genome Res* 22, 1222–1230.
- Li, H., Lee, Y., Chen, J. L., Rebman, E., Li, J., & Lussier, Y. A. (2012b). Complex-disease networks of trait-associated single-nucleotide polymorphisms (SNPs) unveiled by information theory. *J Am Med Inform Assoc* 19, 295–305.
- Li, G., Ruan, X., Auerbach, R. K., Sandhu, K. S., Zheng, M., Wang, P., et al. (2012c). Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell* 148, 84–98.
- Liang, H., & Li, W. H. (2007). MicroRNA regulation of human protein protein interaction network. *RNA* 13, 1402–1408.
- Liang, S., Fuhrman, S., & Somogyi, R. (1998a). Reveal, a general reverse engineering algorithm for inference of genetic network architectures. *Pac Symp Biocomput* 1998, 18–29.
- Liang, J., Edelsbrunner, H., & Woodward, C. (1998b). Anatomy of protein pockets and cavities: measurement of binding site geometry and implications for ligand design. *Protein Sci* 7, 1884–1897.
- Liang, Z., Xu, M., Teng, M., & Niu, L. (2006). NetAlign: a web-based tool for comparison of protein interaction networks. *Bioinformatics* 22, 2175–2177.
- Liang, D., Han, G., Feng, X., Sun, J., Duan, Y., & Lei, H. (2012). Concerted perturbation observed in a hub network in Alzheimer's disease. *PLoS One* 7, e40498.
- Liao, C. S., Lu, K., Baym, M., Singh, R., & Berger, B. (2009). IsoRankN: spectral methods for global alignment of multiple protein networks. *Bioinformatics* 25, i253–i258.
- Liao, X., Xia, Q., Qian, Y., Zhang, L., Hu, G., & Mi, Y. (2011). Pattern formation in oscillatory complex networks consisting of excitable nodes. *Phys Rev E* 83, 056204.
- Liben-Nowell, D., & Kleinberg, J. (2007). The link prediction problem for social networks. *J Am Soc Inf Sci Technol* 58, 1019–1031.
- Licata, L., Briganti, L., Peluso, D., Perfetto, L., Iannucelli, M., Galeota, E., et al. (2012). MINT, the molecular interaction database: 2012 update. *Nucleic Acids Res* 40, D857–D861.
- Lieberman-Aiden, E., van Berkum, N. L., Williams, L., Imakaev, M., Ragozcy, T., Telling, A., et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326, 289–293.
- Lim, J., Hao, T., Shaw, C., Patel, A. J., Szabo, G., Rual, J. F., et al. (2006). A protein–protein interaction network for human inherited ataxias and disorders of Purkinje cell degeneration. *Cell* 125, 801–814.
- Lin, C. Y., Chin, C. H., Wu, H. H., Chen, S. H., Ho, C. W., & Ko, M. T. (2008). Hubba: hub objects analyzer—a framework of interactome hubs identification for network biology. *Nucleic Acids Res* 36, W438–W443.
- Lin, C. C., Chen, Y. J., Chen, C. Y., Oyang, Y. J., Juan, H. F., & Huang, H. C. (2012). Crosstalk between transcription factors and microRNAs in human protein interaction network. *BMC Syst Biol* 6, 18.
- Linding, R., Jensen, L. J., Ostheimer, G. J., van Vugt, M. A., Jorgensen, C., Miron, I. M., et al. (2007). Systematic discovery of in vivo phosphorylation networks. *Cell* 129, 1415–1426.
- Linding, R., Jensen, L. J., Pasculescu, A., Olhovskiy, M., Colwill, K., Bork, P., et al. (2008). NetworkKIN: a resource for exploring cellular phosphorylation networks. *Nucleic Acids Res* 36, D695–D699.
- Lindow, M., & Kauppinen, S. (2012). Discovering the first microRNA-targeted drug. *J Cell Biol* 199, 407–412.
- Lindsay, M. A. (2005). Finding new drug targets in the 21st century. *Drug Discov Today* 10, 1683–1687.
- Linghu, B., Snitkin, E. S., Hu, Z., Xia, Y., & Delisi, C. (2009). Genome-wide prioritization of disease genes and identification of disease–disease associations from an integrated human functional linkage network. *Genome Biol* 10, R91.
- Lipinski, C., & Hopkins, A. (2004). Navigating chemical space for biology and medicine. *Nature* 432, 855–861.
- Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* 46, 3–26.
- Lipton, S. A. (2004). Turning down, but not off. Neuroprotection requires a paradigm shift in drug development. *Nature* 428, 473.
- Liu, Y., & Bahar, I. (2010). Toward understanding allosteric signaling mechanisms in the ATPase domain of molecular chaperones. *Pac Symp Biocomput* 2010, 269–280.
- Liu, R., & Hu, J. (2011). Computational prediction of heme-binding residues by exploiting residue interaction network. *PLoS One* 6, e25560.
- Liu, J., & Nussinov, R. (2008). Allosteric effects in the marginally stable von Hippel-Lindau tumor suppressor protein and allostery-based rescue mutant design. *Proc Natl Acad Sci USA* 105, 901–906.
- Liu, T., Lin, Y., Wen, X., Jorissen, R. N., & Gilson, M. K. (2007a). BindingDB: a web-accessible database of experimentally determined protein–ligand binding affinities. *Nucleic Acids Res* 35, D198–D201.
- Liu, M., Liberzon, A., Kong, S. W., Lai, W. R., Park, P. J., Kohane, I. S., et al. (2007b). Network-based analysis of affected biological processes in type 2 diabetes models. *PLoS Genet* 3, e96.
- Liu, Z. P., Wu, L. Y., Wang, Y., Zhang, X. S., & Chen, L. N. (2008a). Analysis of protein surface patterns by pocket similarity network. *Protein Pept Lett* 15, 448–455.
- Liu, Z. P., Wu, L. Y., Wang, Y., & Zhang, X. S. (2008b). Protein cavity clustering based on community structure of pocket similarity network. *Int J Bioinform Res Appl* 4, 445–460.
- Liu, Y. L., Wise, P. H., & Butte, A. J. (2009). The “etiome”: identification and clustering of human disease etiological factors. *BMC Bioinforma* 10, S14.
- Liu, Y., Gierasch, L. M., & Bahar, I. (2010a). Role of Hsp70 ATPase domain intrinsic dynamics and sequence evolution in enabling its functional interactions with NEFs. *PLoS Comput Biol* 6, e1000931.
- Liu, Y., Hu, B., Fu, C., & Chen, X. (2010b). DCDB: drug combination database. *Bioinformatics* 26, 587–588.
- Liu, Y. Y., Slotine, J. J., & Barabasi, A. L. (2011). Controllability of complex networks. *Nature* 473, 167–173.
- Liu, Y. Y., Slotine, J. J., & Barabasi, A. L. (2012). Control centrality and hierarchical structure in complex networks. *PLoS One* 7, e44459.
- Liu, Z., He, J., -L., & Srivastava, J. (2013). Cliques in complex networks reveal link formation and community evolution. <http://arxiv.org/abs/1301.0803>

- Lo, K., Raftery, A. E., Dombek, K. M., Zhu, J., Schadt, E. E., Bumgarner, R. E., et al. (2012). Integrating external biological knowledge in the construction of regulatory networks from time-series expression data. *BMC Syst Biol* 6, 101.
- Logue, J. S., & Morrison, D. K. (2012). Complexity in the signaling network: insights from the use of targeted inhibitors in cancer therapy. *Genes Dev* 26, 641–650.
- Longabaugh, W. J. (2012). BioTapestry: a tool to visualize the dynamic properties of gene regulatory networks. *Methods Mol Biol* 786, 359–394.
- Lopez-Bigas, N., & Ouzounis, C. A. (2004). Genome-wide identification of genes likely to be involved in human genetic disease. *Nucleic Acids Res* 32, 3108–3114.
- Lopez-Bigas, N., Audit, B., Ouzounis, C., Parra, G., & Guigo, R. (2005). Are splicing mutations the most frequent cause of hereditary disease? *FEBS Lett* 579, 1900–1903.
- Lorenz, D. R., Cantor, C. R., & Collins, J. J. (2009). A network biology approach to aging in yeast. *Proc Natl Acad Sci USA* 106, 1145–1150.
- Loscalzo, J., & Barabasi, A. L. (2011). Systems biology and the future of medicine. *Wiley Interdiscip Rev Syst Biol Med* 3, 619–627.
- Lounkine, E., Wawer, M., Wassermann, A. M., & Bajorath, J. (2010). SARANEA: a freely available program to mine structure–activity and structure–selectivity relationship information in compound data sets. *J Chem Inf Model* 50, 68–78.
- Lounkine, E., Keiser, M. J., Whitebread, S., Mikhailov, D., Hamon, J., Jenkins, J. L., et al. (2012). Large-scale prediction and testing of drug activity on side-effect targets. *Nature* 486, 361–367.
- Lovász, L. (2009). Very large graphs. *Curr Dev Math* 2008, 67–128.
- Lovász, L. (2012). *Large networks and graph limits*. American Mathematical Society Colloquium Publications vol. 60, Providence, Rhode Island, USA: American Mathematical Society.
- Lowe, J. A., Jones, P., & Wilson, D. M. (2010). Network biology as a new approach to drug discovery. *Curr Opin Drug Discov Devel* 13, 524–526.
- Lü, L., & Zhou, T. (2011). Link prediction in complex networks: a survey. *Phys A* 390, 1150–1170.
- Lu, M., Zhang, Q., Deng, M., Miao, J., Guo, Y., Gao, W., et al. (2008). An analysis of human microRNA and disease associations. *PLoS One* 3, e3420.
- Lu, L., Jin, C. H., & Zhou, T. (2009). Similarity index based on local paths for link prediction of complex networks. *Phys Rev E* 80, 046122.
- Lu, J. J., Pan, W., Hu, Y. J., & Wang, Y. T. (2012). Multi-target drugs: the trend of drug research and development. *PLoS One* 7, e40262.
- Ludemann, A., Weicht, D., Selbig, J., & Kopka, J. (2004). PaVESy: pathway visualization and editing system. *Bioinformatics* 20, 2841–2844.
- Lum, P. Y., Derry, J. M., & Schadt, E. E. (2009). Integrative genomics and drug development. *Pharmacogenomics* 10, 203–212.
- Luni, C., Shoemaker, J. E., Sanft, K. R., Petzold, L. R., & Doyle, F. J., III (2010). Confidence from uncertainty—a multi-target drug screening method from robust control theory. *BMC Syst Biol* 4, 161.
- Luo, H., Chen, J., Shi, L., Mikailov, M., Zhu, H., Wang, K., et al. (2011). DRAR-CPI: a server for identifying drug repositioning potential and adverse drug reactions via the chemical–protein interactome. *Nucleic Acids Res* 39, W492–W498.
- Luppi, B., Bigucci, F., Cerchiara, T., & Zecchi, V. (2010). Chitosan-based hydrogels for nasal drug delivery: from inserts to nanoparticles. *Expert Opin Drug Deliv* 7, 811–828.
- Lusis, A. J., & Weiss, J. N. (2010). Cardiovascular networks: systems-based approaches to cardiovascular disease. *Circulation* 121, 157–170.
- Ma, X., & Gao, L. (2012). Biological network analysis: insights into structure and functions. *Brief Funct Genomics* 11, 434–442.
- Ma, H., & Goryanin, I. (2008). Human metabolic network reconstruction and its impact on drug discovery and development. *Drug Discov Today* 13, 402–408.
- Ma, H., & Zeng, A. P. (2003). Reconstruction of metabolic networks from genome data and analysis of their global structure for various organisms. *Bioinformatics* 19, 270–277.
- Ma, H. W., Zhao, X. M., Yuan, Y. J., & Zeng, A. P. (2004). Decomposition of metabolic network into functional modules based on the global connectivity structure of reaction graph. *Bioinformatics* 20, 1870–1876.
- Ma, H., Sorokin, A., Mazein, A., Selkov, A., Selkov, E., Demin, O., et al. (2007). The Edinburgh human metabolic network reconstruction and its functional analysis. *Mol Syst Biol* 3, 135.
- Ma, C. W., Xiu, Z. L., & Zeng, A. P. (2012a). Discovery of intramolecular signal transduction network based on a new protein dynamics model of energy dissipation. *PLoS One* 7, e31529.
- Ma, J., Zhang, X., Ung, C. Y., Chen, Y. Z., & Li, B. (2012b). Metabolic network analysis revealed distinct routes of deletion effects between essential and non-essential genes. *Mol Biosyst* 8, 1179–1186.
- Ma'ayan, A. (2008). Network integration and graph analysis in mammalian molecular systems biology. *IET Syst Biol* 2, 206–221.
- Ma'ayan, A., Jenkins, S. L., Goldfarb, J., & Iyengar, R. (2007). Network analysis of FDA approved drugs and their targets. *Mt Sinai J Med* 74, 27–32.
- Macpherson, J. I., Pinney, J. W., & Robertson, D. L. (2009). JNets: exploring networks by integrating annotation. *BMC Bioinforma* 10, 95.
- Madhamshekar, P. B., Maetschke, S. R., Davis, M. J., Reverter, A., & Ragan, M. A. (2012). Gene regulatory network inference: evaluation and application to ovarian cancer allows the prioritization of drug targets. *Genome Med* 4, 41.
- Maeno, Y., & Ohsawa, Y. (2008). Discovering covert node in networked organization. <http://arxiv.org/abs/0803.3363>
- Mandl, J., Meszaros, T., Banhegyi, G., Hunyady, L., & Csala, M. (2009). Endoplasmic reticulum: nutrient sensor in physiology and pathology. *Trends Endocrinol Metab* 20, 194–201.
- Mani, K. M., Lefebvre, C., Wang, K., Lim, W. K., Dalla-Favera, R., et al. (2008). A systems biology approach to prediction of oncogenes and molecular perturbation targets in B-cell lymphomas. *Mol Syst Biol* 4, 169.
- Mantzaris, A. V., Bassett, D. S., Wymbs, N. F., Estrada, E., Porter, M. A., Mucha, P. J., Grafton, S. T., & Higham, D. J. (in press). Dynamic network centrality summarizes learning in the human brain. *J Compl Netw*. <http://arxiv.org/abs/1207.5047>.
- Mar, J. C., & Quackenbush, J. (2009). Decomposition of gene expression state space trajectories. *PLoS Comput Biol* 5, e1000626.
- Marbach, D., Prill, R. J., Schaffter, T., Mattiussi, C., Floreano, D., & Stolovitzky, G. (2010). Revealing strengths and weaknesses of methods for gene network inference. *Proc Natl Acad Sci USA* 107, 6286–6291.
- Marbach, D., Costello, J. C., Kuffner, R., Vega, N. M., Prill, R. J., Camacho, D. M., et al. (2012). Wisdom of crowds for robust gene network inference. *Nat Methods* 9, 796–804.
- Margineanu, D. G. (2012). Systems biology impact on antiepileptic drug discovery. *Epilepsy Res* 98, 104–115.
- Martin, Y. C., Kofron, J. L., & Traphagen, L. M. (2002). Do structurally similar molecules have similar biological activity? *J Med Chem* 45, 4350–4358.
- Martin, A., Ochagavia, M. E., Rabasa, L. C., Miranda, J., Fernandez-de-Cossio, J., & Bringas, R. (2010). BisoGenet: a new tool for gene network building, visualization and analysis. *BMC Bioinforma* 11, 91.
- Martin, A. J., Vidotto, M., Boscaroli, F., Di Domenico, T., Walsh, I., & Tosatto, S. C. (2011). RING: networking interacting residues, evolutionary information and energetics in protein structures. *Bioinformatics* 27, 2003–2005.
- Martin, F., Thomson, T. M., Sewer, A., Drubin, D. A., Mathis, C., Weisensee, D., et al. (2012). Assessment of network perturbation amplitude by applying high-throughput data to causal biological networks. *BMC Syst Biol* 6, 54.
- Martinez-Romero, M., Vazquez-Naya, J. M., Rabunal, J. R., Pita-Fernandez, S., Macenlle, R., Castro-Alvarino, J., et al. (2010). Artificial intelligence techniques for colorectal cancer drug metabolism: ontology and complex network. *Curr Drug Metab* 11, 347–368.
- Marton, M. J., DeRisi, J. L., Bennett, H. A., Iyer, V. R., Meyer, M. R., Roberts, C. J., et al. (1998). Drug target validation and identification of secondary drug target effects using DNA microarrays. *Nat Med* 4, 1293–1301.
- Maslov, S., & Ispolatov, I. (2007). Propagation of large concentration changes in reversible protein-binding networks. *Proc Natl Acad Sci USA* 104, 13655–13660.
- Maslov, S., & Sneppen, K. (2002). Specificity and stability in topology of protein networks. *Science* 296, 910–913.
- Mathur, S., & Dinakarpanian, D. (2011). Drug repositioning using disease associated biological processes and network analysis of drug targets. *AMIA Annu Symp Proc* 2011, 305–311.
- Matsuura, M., Nakazawa, H., Hashimoto, T., & Mitsuhashi, S. (1980). Combined antibacterial activity of amoxicillin with clavulanic acid against ampicillin-resistant strains. *Antimicrob Agents Chemother* 17, 908–911.
- McDermott, A. M., Heneghan, H. M., Miller, N., & Kerin, M. J. (2011). The therapeutic potential of microRNAs: disease modulators and drug targets. *Pharm Res* 28, 3016–3029.
- McDowall, M. D., Scott, M. S., & Barton, G. J. (2009). PIPs: human protein–protein interaction prediction database. *Nucleic Acids Res* 37, D651–D656.
- McGary, K. L., Park, T. J., Woods, J. O., Cha, H. J., Wallingford, J. B., & Marcotte, E. M. (2010). Systematic discovery of nonobvious human disease models through orthologous phenotypes. *Proc Natl Acad Sci USA* 107, 6544–6549.
- McLaughlin, R. N., Jr., Poelwijk, F. J., Raman, A., Gosal, W. S., & Ranganathan, R. (2012). The spatial architecture of protein function and adaptation. *Nature* 491, 138–142.
- McManus, K. J., Barrett, I. J., Nouhi, Y., & Hieter, P. (2009). Specific synthetic lethal killing of RAD54B-deficient human colorectal cancer cells by FEN1 silencing. *Proc Natl Acad Sci USA* 106, 3276–3281.
- Mehlhorn, K., & Näher, S. (1999). *The LEDA Platform of Combinatorial and Geometric Computing*. Cambridge, UK: Cambridge University Press.
- Mehraban, S., Shirazi, A. H., Zamani, M., & Jafari, G. R. (2013). Coupling between time series: a network view. <http://arxiv.org/abs/1301.1010>
- Mei, H., Xia, T., Feng, G., Zhu, J., Lin, S. M., & Qiu, Y. (2012). Opportunities in systems biology to discover mechanisms and repurpose drugs for CNS diseases. *Drug Discov Today* 17, 1208–1216.
- Mei, J. P., Kwok, C. K., Yang, P., Li, X. L., & Zheng, J. (2013). Drug–target interaction prediction by learning from local information and neighbors. *Bioinformatics* 29, 238–245.
- Meil, A., Durand, P., & Wojcik, J. (2005). PIMWalker: visualising protein interaction networks using the HUPO PSI molecular interaction format. *Appl Bioinformatics* 4, 137–139.
- Memisevic, V., & Przulj, N. (2012). C-GRAAL: common-neighbors-based global GRAph Alignment of biological networks. *Integr Biol* 4, 734–743.
- Mencher, S. K., & Wang, L. G. (2005). Promiscuous drugs compared to selective drugs (promiscuity can be a virtue). *BMC Clin Pharmacol* 5, 3.
- Mészáros, B., Tóth, J., Vértessy, B. G., Dosztányi, Z., & Simon, I. (2011). Proteins with complex architecture as potential targets for drug design: a case study of Mycobacterium tuberculosis. *PLoS Comput Biol* 7, e1002118.
- Michaelis, M. L., Seyb, K. I., & Ansar, S. (2005). Cytoskeletal integrity as a drug target. *Curr Alzheimer Res* 2, 227–229.
- Mihalik, A., & Csermely, P. (2011). Heat shock partially dissociates the overlapping modules of the yeast protein–protein interaction network: a systems level model of adaptation. *PLoS Comput Biol* 7, e1002187.
- Milenkovic, T., Memisevic, V., Bonato, A., & Przulj, N. (2011). Dominating biological networks. *PLoS One* 6, e23016.
- Millan, M. J. (2006). 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.
- Miller, M. L., Jensen, L. J., Diella, F., Jorgensen, C., Tinti, M., Li, L., et al. (2008). Linear motif atlas for phosphorylation-dependent signaling. *Sci Signal* 1, ra2.
- Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., & Alon, U. (2002). Network motifs: simple building blocks of complex networks. *Science* 298, 824–827.

- Mimeault, M., & Batra, S. K. (2010). Frequent deregulations in the hedgehog signaling network and cross-talks with the epidermal growth factor receptor pathway involved in cancer progression and targeted therapies. *Pharmacol Rev* 62, 497–524.
- Mirshahvalad, A., Beauchesne, O. H., Archambault, E., & Rosvall, M. (2013). Effect of resampling schemes on significance analysis of clustering and ranking. *PLoS One* 8, e53943.
- Missiuro, P. V., Liu, K., Zou, L., Ross, B. C., Zhao, G., Liu, J. S., et al. (2009). Information flow analysis of interactome networks. *PLoS Comput Biol* 5, e1000350.
- Mithani, A., Preston, G. M., & Hein, J. (2009). Rahnuma: hypergraph-based tool for metabolic pathway prediction and network comparison. *Bioinformatics* 25, 1831–1832.
- Mizutani, S., Pauwels, E., Stoven, V., Goto, S., & Yamanishi, Y. (2012). Relating drug–protein interaction network with drug side effects. *Bioinformatics* 28, i522–i528.
- Moazed, D. (2011). Mechanisms for the inheritance of chromatin states. *Cell* 146, 510–518.
- Möbius, A., Neklioudov, A., Díaz-Sánchez, A., Hoffmann, K. H., Fachat, A., & Schreiber, M. (1997). Optimization by thermal cycling. *Phys Rev Lett* 79, 4297–4301.
- Mones, E., Vicsek, L., & Vicsek, T. (2012). Hierarchy measure for complex networks. *PLoS One* 7, e33799.
- Moon, H. S., Bhak, J., Lee, K. H., & Lee, D. (2005). Architecture of basic building blocks in protein and domain structural interaction networks. *Bioinformatics* 21, 1479–1486.
- Moran, L. B., & Graeber, M. B. (2008). Towards a pathway definition of Parkinson's disease: a complex disorder with links to cancer, diabetes and inflammation. *Neurogenetics* 9, 1–13.
- Moreno-Sanchez, R., Saavedra, E., Rodriguez-Enriquez, S., Gallardo-Perez, J. C., Quezada, H., & Westerhoff, H. V. (2010). Metabolic control analysis indicates a change of strategy in the treatment of cancer. *Mitochondrion* 10, 626–639.
- Mori, M. A., Liu, M., Bezy, O., Almind, K., Shapiro, H., Kasif, S., et al. (2010). A systems biology approach identifies inflammatory abnormalities between mouse strains prior to development of metabolic disease. *Diabetes* 59, 2960–2971.
- Morita, H., & Takano, M. (2009). Residue network in protein native structure belongs to the universality class of a three-dimensional critical percolation cluster. *Phys Rev E* 79, 020901.
- Moriya, H., Shimizu-Yoshida, Y., & Kitano, H. (2006). In vivo robustness analysis of cell division cycle genes in *Saccharomyces cerevisiae*. *PLoS Genet* 2, 111.
- Morris, R. G., & Barthelemy, M. (2012). Transport on coupled spatial networks. *Phys Rev Lett* 109, 128703.
- Morris, J. H., Huang, C. C., Babbitt, P. C., & Ferrin, T. E. (2007). structureViz: linking Cytoscape and UCSF Chimera. *Bioinformatics* 23, 2345–2347.
- Morris, M. K., Melas, I., & Saez-Rodriguez, J. (2013). Construction of cell type-specific logic models of signaling networks using CellNOpt. *Methods Mol Biol* 930, 179–214.
- Morselli, E., Galluzzi, L., Kepp, O., Vicencio, J. M., Criollo, A., Maiuri, M. C., et al. (2009). Anti- and pro-tumor functions of autophagy. *Biochim Biophys Acta* 1793, 1524–1532.
- Moschopoulos, C. N., Pavlopoulos, G. A., Likothanassis, S., & Kossida, S. (2011). Analyzing protein–protein interaction networks with web tools. *Curr Bioinformatics* 6, 389–397.
- Motter, A. E. (2010). Improved network performance via antagonism: from synthetic rescues to multi-drug combinations. *Bioessays* 32, 236–245.
- Motter, A. E., Gulbahce, N., Almaas, E., & Barabasi, A. L. (2008). Predicting synthetic rescues in metabolic networks. *Mol Syst Biol* 4, 168.
- Mucha, P. J., Richardson, T., Macon, K., Porter, M. A., & Onnela, J.-P. (2010). Community structure in time-dependent multiscale, and multiplex networks. *Science* 328, 876–878.
- Mueller, K., Ash, C., Pennisi, E., & Smith, O. (2012). The gut microbiota. *Science* 336, 1245.
- Murabito, E., Smallbone, K., Swinton, J., Westerhoff, H. V., & Steuer, R. (2011). A probabilistic approach to identify putative drug targets in biochemical networks. *J R Soc Interface* 8, 880–895.
- Murrell, P. (2012). *Hyperdraw: Visualizing Hypergraphs*. (R package version 1.8.0. <http://www.bioconductor.org/packages/release/bioc/html/hyperdraw.html>)
- Musmeci, N., Battison, S., Caldarelli, G., Puliga, M., & Gabrielli, A. (in press). Bootstrapping topological properties and systemic risk of complex networks using the fitness model. *J Stat Phys*. <http://dx.doi.org/10.1007/s10955-013-0720-1>.
- Nacher, J. C., & Schwartz, J. M. (2008). A global view of drug–therapy interactions. *BMC Pharmacol* 8, 5.
- Nacher, J. C., & Schwartz, J. M. (2012). Modularity in protein complex and drug interactions reveals new polypharmacological properties. *PLoS One* 7, e30028.
- Nagasaki, M., Saito, A., Jeong, E., Li, C., Kojima, K., Ikeda, E., et al. (2011). Cell illustrator 4.0: a computational platform for systems biology. *Stud Health Technol Inform* 162, 160–181.
- Nam, H., Lewis, N. E., Lerman, J. A., Lee, D. H., Chang, R. L., Kim, D., et al. (2012). Network context and selection in the evolution to enzyme specificity. *Science* 337, 1101–1104.
- Navlakha, S., & Kingsford, C. (2010). The power of protein interaction networks for associating genes with diseases. *Bioinformatics* 26, 1057–1063.
- Navlakha, S., & Kingsford, C. (2011). Network archeology uncovering ancient networks from present-day interactions. *PLoS Comput Biol* 7, e1001119.
- Navratil, V., de Chasse, B., Meyniel, L., Delmotte, S., Gautier, C., Andre, P., et al. (2009). VirHostNet: a knowledge base for the management and the analysis of proteome-wide virus–host interaction networks. *Nucleic Acids Res* 37, D661–D668.
- Navratil, V., de Chasse, B., Combe, C. R., & Lotteau, V. (2011). When the human viral infectome and diseaseome networks collide: towards a systems biology platform for the aetiology of human diseases. *BMC Syst Biol* 5, 13.
- Nayal, M., & Honig, B. (2006). On the nature of cavities on protein surfaces: application to the identification of drug-binding sites. *Proteins* 63, 892–906.
- Nelander, S., Wang, W., Nilsson, B., She, Q. B., Pratilas, C., Rosen, N., et al. (2008). Models from experiments: combinatorial drug perturbations of cancer cells. *Mol Syst Biol* 4, 216.
- Nelson, M. R., Wegmann, D., Ehm, M. G., Kessner, D., St Jean, P., Verzilli, C., et al. (2012). An abundance of rare functional variants in 202 drug target genes sequenced in 14,002 people. *Science* 337, 100–104.
- Nemenman, I., Escala, G. S., Hlavacek, W. S., Unkefer, P. J., Unkefer, C. J., & Wall, M. E. (2007). Reconstruction of metabolic networks from high-throughput metabolite profiling data: in silico analysis of red blood cell metabolism. *Ann N Y Acad Sci* 1115, 102–115.
- Neph, S., Stergachis, A. B., Reynolds, A., Sandstrom, R., Borenstein, E., & Stamatoyannopoulos, J. A. (2012). Circuitry and dynamics of human transcription factor regulatory networks. *Cell* 150, 1274–1286.
- Nepusz, T., & Vicsek, T. (2012). Controlling edge dynamics in complex networks. *Nat Phys* 8, 568–573.
- Nepusz, T., Petroczi, A., Negyessy, L., & Bazzo, F. (2008). Fuzzy communities and the concept of bridgeness in complex networks. *Phys Rev E* 77, 016107.
- Newman, M. E. J. (2011). Complex systems: a survey. *Am J Phys* 79, 800–810.
- Ng, S. K., Zhang, Z., Tan, S. H., & Lin, K. (2003). InterDom: a database of putative interacting protein domains for validating predicted protein interactions and complexes. *Nucleic Acids Res* 31, 251–254.
- Nguyen, T. P., & Jordan, F. (2010). A quantitative approach to study indirect effects among disease proteins in the human protein interaction network. *BMC Syst Biol* 4, 103.
- Nguyen, T. P., Liu, W. C., & Jordan, F. (2011). Inferring pleiotropy by network analysis: linked diseases in the human PPI network. *BMC Syst Biol* 5, 179.
- Nguyen, L. K., Matallanas, D., Croucher, D. R., von Kriegsheim, A., & Kholodenko, B. N. (2013). Signalling by protein phosphatases and drug development: a systems-centred view. *FEBS J* 280, 751–765.63.
- Nibbe, R. K., Koyuturk, M., & Chance, M. R. (2010). An integrative -omics approach to identify functional sub-networks in human colorectal cancer. *PLoS Comput Biol* 6, e1000639.
- Nicosia, V., Criado, R., Romance, M., Russo, G., & Latora, V. (2012). Controlling centrality in complex networks. *Sci Rep* 2, 218.
- Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science* 314, 1560–1563.
- Nussinov, R., & Tsai, C. J. (2012). The different ways through which specificity works in orthosteric and allosteric drugs. *Curr Pharm Des* 18, 1311–1316.
- Nussinov, R., Tsai, C.-J., & Csermely, P. (2011). Allo-network drugs: harnessing allostery in cellular networks. *Trends Pharmacol Sci* 32, 686–693.
- NWB Team (2006). *Network Workbench Tool*. Indiana University, Northeastern University, and University of Michigan. (<http://nwb.slis.indiana.edu>)
- Oberhardt, M. A., Goldberg, J. B., Hogardt, M., & Papin, J. A. (2010). Metabolic network analysis of *Pseudomonas aeruginosa* during chronic cystic fibrosis lung infection. *J Bacteriol* 192, 5534–5548.
- Ohlson, S. (2008). Designing transient binding drugs: a new concept for drug discovery. *Drug Discov Today* 13, 433–439.
- Oprea, T. I., Nielsen, S. K., Ursu, O., Yang, J. J., Taboureaux, O., Mathias, S. L., et al. (2011). Associating drugs, targets and clinical outcomes into an integrated network affords a new platform for computer-aided drug repurposing. *Mol Inform* 30, 100–111.
- Orlev, N., Shamir, R., & Shiloh, Y. (2004). PIVOT: protein interactions visualization tool. *Bioinformatics* 20, 424–425.
- Oti, M., & Brunner, H. G. (2007). The modular nature of genetic diseases. *Clin Genet* 71, 1–11.
- Oti, M., Snel, B., Huynen, M. A., & Brunner, H. G. (2006). Predicting disease genes using protein–protein interactions. *J Med Genet* 43, 691–698.
- Overington, J. P., Al-Lazikani, B., & Hopkins, A. L. (2006). How many drug targets are there? *Nat Rev Drug Discov* 5, 993–996.
- Ozbabacan, S. E. A., Gursoy, A., Keskin, O., & Nussinov, R. (2010). Conformational ensembles, signal transduction and residue hot spots: application to drug discovery. *Curr Op Drug Discov Dev* 13, 527–537.
- Pabuwal, V., & Li, Z. (2009). Comparative analysis of the packing topology of structurally important residues in helical membrane and soluble proteins. *Protein Eng Des Sel* 22, 67–73.
- Pache, R. A., & Aloy, P. (2012). A novel framework for the comparative analysis of biological networks. *PLoS One* 7, e31220.
- Pache, R. A., Ceol, A., & Aloy, P. (2012). NetAligner—a network alignment server to compare complexes, pathways and whole interactomes. *Nucleic Acids Res* 40, W157–W161.
- Pacifico, S., Liu, G., Guest, S., Parrish, J. R., Fotouhi, F., & Finley, R. L., Jr. (2006). A database and tool, IM Browser, for exploring and integrating emerging gene and protein interaction data for *Drosophila*. *BMC Bioinforma* 7, 195.
- Pacini, C., Iorio, F., Gonçalves, E., Iskar, M., Klabunde, T., Bork, P., et al. (2013). DvD: an R/Cytoscape pipeline for drug repurposing using public repositories of gene expression data. *Bioinformatics* 29, 132–134.
- Padiadpu, J., Vashisht, R., & Chandra, N. (2010). Protein–protein interaction networks suggest different targets have different propensities for triggering drug resistance. *Syst Synth Biol* 4, 311–322.
- Paek, E., Park, J., & Lee, K. J. (2004). Multi-layered representation for cell signaling pathways. *Mol Cell Proteomics* 3, 1009–1022.
- Pál, C., Papp, B., Lercher, M. J., Csermely, P., Oliver, S. G., & Hurst, L. D. (2006). Chance and necessity in the evolution of minimal metabolic networks. *Nature* 440, 667–670.
- Pálffy, M., Remenyi, A., & Korcsmaros, T. (2012). Endosomal crosstalk: meeting points for signaling pathways. *Trends Cell Biol* 22, 447–456.
- Palla, G., Derenyi, I., Farkas, I., & Vicsek, T. (2005). Uncovering the overlapping community structure of complex networks in nature and society. *Nature* 435, 814–818.
- Palla, G., Barabasi, A.-L., & Vicsek, T. (2007). Quantifying social group evolution. *Nature* 446, 664–667.

- Palumbo, M. C., Colosimo, A., Giuliani, A., & Farina, L. (2007). Essentiality is an emergent property of metabolic network wiring. *FEBS Lett* 581, 2485–2489.
- Pan, H., Lee, J. C., & Hilser, V. J. (2000). Binding sites in *Escherichia coli* dihydrofolate reductase communicate by modulating the conformational ensemble. *Proc Natl Acad Sci USA* 97, 12020–12025.
- Pandey, G., Zhang, B., Chang, A. N., Myers, C. L., Zhu, J., Kumar, V., et al. (2010). An integrative multi-network and multi-classifier approach to predict genetic interactions. *PLoS Comput Biol* 6, e1000928.
- Pandini, A., Fornili, A., Fraternali, F., & Kleinjung, J. (2012). Detection of allosteric signal transmission by information-theoretic analysis of protein dynamics. *FASEB J* 26, 868–881.
- Paolini, G. V., Shapland, R. H. B., van Hoorn, W. P., Mason, J. S., & Hopkins, A. L. (2006). Global mapping of pharmacological space. *Nature Biotech* 24, 805–815.
- Papatsoris, A. G., Karamouzis, M. V., & Papavassiliou, A. G. (2007). The power and promise of “rewiring” the mitogen-activated protein kinase network in prostate cancer therapeutics. *Mol Cancer Ther* 6, 811–819.
- Papin, J. A., Reed, J. L., & Palsson, B. O. (2004). Hierarchical thinking in network biology: the unbiased modularization of biochemical networks. *Trends Biochem Sci* 29, 641–647.
- Papin, J. A., Hunter, T., Palsson, B. O., & Subramaniam, S. (2005). Reconstruction of cellular signalling networks and analysis of their properties. *Nat Rev Mol Cell Biol* 6, 99–111.
- Papp, E., & Csermely, P. (2006). Chemical chaperones: mechanisms of action and potential use. *Handb Exp Pharmacol* 172, 405–416.
- Papp, B., Pal, C., & Hurst, L. D. (2004). Metabolic network analysis of the causes and evolution of enzyme dispensability in yeast. *Nature* 429, 661–664.
- Park, K., & Kim, D. (2008). Binding similarity network of ligand. *Proteins* 71, 960–971.
- Park, K., & Kim, D. (2011). Modeling allosteric signal propagation using protein structure networks. *BMC Bioinforma* 12, S23.
- Parsons, A. B., Lopez, A., Givoni, I. E., Williams, D. E., Gray, C. A., Porter, J., et al. (2006). Exploring the mode-of-action of bioactive compounds by chemical-genetic profiling in yeast. *Cell* 126, 611–625.
- Pasi, M., Tiberti, M., Arrigoni, A., & Papaleo, E. (2012). xPyder: a PyMOL plugin to analyze coupled residues and their networks in protein structures. *J Chem Inf Model* 52, 1865–1874.
- Pavlopoulos, G. A., Wegener, A. L., & Schneider, R. (2008). A survey of visualization tools for biological network analysis. *BioData Min* 1, 12.
- Pawson, T., & Linding, R. (2008). Network medicine. *FEBS Lett* 582, 1266–1270.
- Pe'er, D., & Hacohen, N. (2011). Principles and strategies for developing network models in cancer. *Cell* 144, 864–873.
- Peltason, L., Iyer, P., & Bajorath, J. (2010). Rationalizing three-dimensional activity landscapes and the influence of molecular representations on landscape topology and the formation of activity cliffs. *J Chem Inf Model* 50, 1021–1033.
- Penrod, N. M., Cowper-Salari, R., & Moore, J. H. (2011). Systems genetics for drug target discovery. *Trends Pharmacol Sci* 32, 623–630.
- Perra, N., Gonçalves, B., Pastor-Satorras, R., & Vespignani, A. (2012). Activity driven modeling of time varying networks. *Scientific Reports* 2, 469.
- Perumal, D., Lim, C. S., & Sakharkar, M. K. (2009). A comparative study of metabolic network topology between a pathogenic and a non-pathogenic bacterium for potential drug target identification. *Summit Transl Bioinformatics* 2009, 100–104.
- Pfützner, R., Scholtes, I., Garas, A., Tessone, C. J., & Schweitzer, F. (2012). Betweenness preference: quantifying correlations in the topological dynamics of temporal networks. <http://arxiv.org/abs/1208.0588>
- Pham, L., Christadore, L., Schaus, S., & Kolaczyk, E. D. (2011). Network-based prediction for sources of transcriptional dysregulation using latent pathway identification analysis. *Proc Natl Acad Sci USA* 108, 13347–13352.
- Phan, H. T., & Sternberg, M. J. (2012). PINALOG: a novel approach to align protein interaction networks—implications for complex detection and function prediction. *Bioinformatics* 28, 1239–1245.
- Piazza, F., & Sanejouand, Y. -H. (2008). Discrete breathers in protein structures. *Phys Biol* 5, 026001.
- Piazza, F., & Sanejouand, Y. -H. (2009). Long-range energy transfer in proteins. *Phys Biol* 6, 046014.
- Pinter, R. Y., Rokhlenko, O., Yeger-Lotem, E., & Ziv-Ukelson, M. (2005). Alignment of metabolic pathways. *Bioinformatics* 21, 3401–3408.
- Platzer, A., Perco, P., Lukas, A., & Mayer, B. (2007). Characterization of protein-interaction networks in tumors. *BMC Bioinforma* 8, 224.
- Pommier, Y., & Cherfils, J. (2005). Interfacial inhibition of macromolecular interactions: nature's paradigm for drug discovery. *Trends Pharmacol Sci* 26, 138–145.
- Pons, C., Glaser, F., & Fernandez-Recio, J. (2011). Prediction of protein-binding areas by small-world residue networks and application to docking. *BMC Bioinforma* 12, 378.
- Portales-Casamar, E., Kirov, S., Lim, J., Lithwick, S., Swanson, M. L., Ticoll, A., et al. (2007). PAZAR: a framework for collection and dissemination of cis-regulatory sequence annotation. *Genome Biol* 8, R207.
- Portales-Casamar, E., Thongjuea, S., Kwon, A. T., Arenillas, D., Zhao, X., Valen, E., et al. (2010). JASPAR 2010: the greatly expanded open-access database of transcription factor binding profiles. *Nucleic Acids Res* 38, D105–D110.
- Pósfai, M., Liu, Y. -Y., Slotine, J. -J., & Barabási, A. -L. (2013). Effect of correlations on network controllability. *Sci Rep* 3, 1067.
- Prado-Prado, F. J., Gonzalez-Diaz, H., de la Vega, O. M., Ubeira, F. M., & Chou, K. C. (2008). Unified QSAR approach to antimicrobials. Part 3: first multi-tasking QSAR model for Input-Coded prediction, structural back-projection, and complex networks clustering of antiprotozoal compounds. *Bioorg Med Chem* 16, 5871–5880.
- Prado-Prado, F. J., de la Vega, O. M., Uriarte, E., Ubeira, F. M., Chou, K. C., & Gonzalez-Diaz, H. (2009). Unified QSAR approach to antimicrobials. 4. Multi-target QSAR modeling and comparative multi-distance study of the giant components of antiviral drug–drug complex networks. *Bioorg Med Chem* 17, 569–575.
- Prado-Prado, F. J., Ubeira, F. M., Borges, F., & Gonzalez-Diaz, H. (2010). Unified QSAR & network-based computational chemistry approach to antimicrobials. II. Multiple distance and triadic census analysis of antiparasitic drugs complex networks. *J Comp Chem* 31, 164–173.
- Prado-Prado, F. J., Garcia, I., Garcia-Mera, X., & Gonzalez-Diaz, H. (2011). Entropy multi-target QSAR model for prediction of antiviral drug complex networks. *Chemomet Intell Lab Syst* 107, 227–233.
- Prieto, C., & De Las Rivas, J. (2006). APID: Agile Protein Interaction DataAnalyzer. *Nucleic Acids Res* 34, W298–W302.
- Prieto, C., & De Las Rivas, J. (2010). Structural domain–domain interactions: assessment and comparison with protein–protein interaction data to improve the interactome. *Proteins* 78, 109–117.
- Prill, R. J., Saez-Rodriguez, J., Alexopoulos, L. G., Sorger, P. K., & Stolovitzky, G. (2011). Crowdsourcing network inference: the DREAM predictive signaling network challenge. *Sci Signal* 4, mr7.
- Prinz, F., Schlange, T., & Asadullah, K. (2011). Believe it or not: how much can we rely on published data on potential drug targets? *Nat Rev Drug Discov* 10, 712.
- Pritchard, J. R., Bruno, P. M., Gilbert, L. A., Capron, K. L., Lauffenburger, D. A., & Hemann, M. T. (2013). Defining principles of combination drug mechanisms of action. *Proc Natl Acad Sci USA* 110, E170–E179.
- Promislow, D. E. (2004). Protein networks, pleiotropy and the evolution of senescence. *Proc Biol Sci* 271, 1225–1234.
- Prussia, A., Thepchatri, P., Snyder, J. P., & Plemper, R. K. (2011). Systematic approaches towards the development of host-directed antiviral therapeutics. *Int J Mol Sci* 12, 4027–4052.
- Przulj, N., Corneil, D. G., & Jurisica, I. (2006). Efficient estimation of graphlet frequency distributions in protein–protein interaction networks. *Bioinformatics* 22, 974–980.
- Pujadas, E., & Feinberg, A. P. (2012). Regulated noise in the epigenetic landscape of development and disease. *Cell* 148, 1123–1131.
- Pujana, M. A., Han, J. D., Starita, L. M., Stevens, K. N., Tewari, M., Ahn, J. S., et al. (2007). Network modeling links breast cancer susceptibility and centrosome dysfunction. *Nat Genet* 39, 1338–1349.
- Pujol, A., Mosca, R., Farrés, J., & Aloy, P. (2010). Unveiling the role of network and systems biology in drug discovery. *Trends Pharmacol Sci* 31, 115–123.
- Qu, X. A., Gudivada, R. C., Jegga, A. G., Neumann, E. K., & Aronow, B. J. (2009). Inferring novel disease indications for known drugs by semantically linking drug action and disease mechanism relationships. *BMC Bioinforma* 10, S4.
- Rader, A. J., & Brown, S. M. (2011). Correlating allostery with rigidity. *Mol Biosyst* 7, 464–471.
- Radicchi, F., Ramasco, J. J., & Fortunato, S. (2011). Information filtering in complex weighted networks. *Phys Rev E* 83, 046101.
- Radiojac, P., Peng, K., Clark, W. T., Peters, B. J., Mohan, A., Boyle, S. M., et al. (2008). An integrated approach to inferring gene–disease associations in humans. *Proteins* 72, 1030–1037.
- Raj, T., Shulman, J. M., Keenan, B. T., Chibnik, L. B., Evans, D. A., Bennett, D. A., et al. (2012). Alzheimer disease susceptibility loci: evidence for a protein network under natural selection. *Am J Hum Genet* 90, 720–726.
- Rajasethupathy, P., Vaytaden, S. J., & Bhalla, U. S. (2005). Systems modeling: a pathway to drug discovery. *Curr Opin Chem Biol* 9, 400–406.
- Raman, K., & Chandra, N. (2008). Mycobacterium tuberculosis interactome analysis unravels potential pathways to drug resistance. *BMC Microbiol* 8, 234.
- Raman, K., Yeturu, K., & Chandra, N. (2008). targetTB: a target identification pipeline for *Mycobacterium tuberculosis* through an interactome, reactome and genome-scale structural analysis. *BMC Syst Biol* 2, 109.
- Raman, K., Vashisht, R., & Chandra, N. (2009). Strategies for efficient disruption of metabolism in *Mycobacterium tuberculosis* from network analysis. *Mol Biosyst* 5, 1740–1751.
- Raman, M. P., Singh, S., Devi, P. R., & Velmurugan, D. (2012). Uncovering potential drug targets for tuberculosis using protein networks. *Bioinformation* 8, 403–406.
- Rao, F., & Cafllisi, A. (2004). The protein folding network. *J Mol Biol* 342, 299–306.
- Ravasz, E., Somera, A. L., Mongru, D. A., Oltvai, Z. N., & Barabási, A. L. (2002). Hierarchical organization of modularity in metabolic networks. *Science* 297, 1551–1555.
- Ravikumar, B., Sarkar, S., Davies, J. E., Futter, M., Garcia-Arencibia, M., Green-Thompson, Z. W., et al. (2010). Regulation of mammalian autophagy in physiology and pathophysiology. *Physiol Rev* 90, 1383–1435.
- Ray, M., Ruan, J., & Zhang, W. (2008). Variations in the transcriptome of Alzheimer's disease reveal molecular networks involved in cardiovascular diseases. *Genome Biol* 9, R148.
- Real, E., Rain, J. C., Battaglia, V., Jallet, C., Perrin, P., Tordo, N., et al. (2004). Antiviral drug discovery strategy using combinatorial libraries of structurally constrained peptides. *J Virol* 78, 7410–7417.
- Rees, S., Morrow, D., & Kenakin, T. (2002). GPCR drug discovery through the exploitation of allosteric drug binding sites. *Receptors Channels* 8, 261–268.
- Reisen, F., Weisel, M., Kriegl, J. M., & Schneider, G. (2010). Self-organizing fuzzy graphs for structure-based comparison of protein pockets. *J Proteome Res* 9, 6498–6510.
- Reja, R., Venkatakrishnan, A. J., Lee, J., Kim, B. C., Ryu, J. W., Gong, S., et al. (2009). Mitochondriome: mitochondrial protein interactome database, and its application in ‘aging network’ analysis. *BMC Genomics* 10, S20.
- Remenyi, A., Good, M. C., & Lim, W. A. (2006). Docking interactions in protein kinase and phosphatase networks. *Curr Opin Struct Biol* 16, 676–685.
- Ren, J., Xie, L., Li, W. W., & Bourne, P. E. (2010). SMAP-WS: a parallel web service for structural proteome-wide ligand-binding site comparison. *Nucleic Acids Res* 38, W441–W444.
- Reppas, A. I., & Lawyer, G. (2012). Low k-shells identify bridge elements critical to disease flow in small-world networks. *AIP Conf Proc* 1479, 1426–1429.

- Resendis-Antonio, O. (2009). Filling kinetic gaps: dynamic modeling of metabolism where detailed kinetic information is lacking. *PLoS One* 4, e4967.
- Resendis-Antonio, O., Checa, A., & Encarnacion, S. (2010). Modeling core metabolism in cancer cells: surveying the topology underlying the Warburg effect. *PLoS One* 5, e12383.
- Reynolds, K. A., McLaughlin, R. N., & Ranganathan, R. (2011). Hot spots for allosteric regulation on protein surfaces. *Cell* 147, 1564–1575.
- Rhodes, D. R., & Chinnaiyan, A. M. (2005). Integrative analysis of the cancer transcriptome. *Nat Genet* 37, S31–S37.
- Rhodes, D. R., Kalyana-Sundaram, S., Mahavisno, V., Varambally, R., Yu, J., Briggs, B. B., et al. (2007b). OncoPrint 3.0: genes, pathways, and networks in a collection of 18,000 cancer gene expression profiles. *Neoplasia* 9, 166–180.
- Rhodes, D. R., Kalyana-Sundaram, S., Tomlins, S. A., Mahavisno, V., Kasper, N., Varambally, R., et al. (2007a). Molecular concepts analysis links tumors, pathways, mechanisms, and drugs. *Neoplasia* 9, 443–454.
- Rickman, D. S., Soong, T. D., Moss, B., Mosquera, J. M., Dlabal, J., Terry, S., et al. (2012). Oncogene-mediated alterations in chromatin conformation. *Proc Natl Acad Sci USA* 109, 9083–9088.
- Riera-Fernandez, P., Munteanu, C. R., Escobar, M., Prado-Prado, F., Martin-Romalde, R., Pereira, D., et al. (2012). New Markov–Shannon Entropy models to assess connectivity quality in complex networks: from molecular to cellular pathway, parasite–host, neural, industry, and legal–social networks. *J Theor Biol* 293, 174–188.
- Rito, T., Wang, Z., Deane, C. M., & Reinert, G. (2010). How threshold behaviour affects the use of subgraphs for network comparison. *Bioinformatics* 26, i611–i617.
- Rivera, C. G., Chu, L.-H., Bader, J. S., & Popel, A. S. (2012). Applications of network bioinformatics to cancer angiogenesis. In A. S. Azmi (Ed.), *Systems Biology in Cancer Research and Drug Discovery* (pp. 229–244). Dordrecht: Springer Science+Business Media.
- Rocha, G. Z., Dias, M. M., Ropelle, E. R., Osorio-Costa, F., Rossato, F. A., Vercesi, A. E., et al. (2011). Metformin amplifies chemotherapy-induced AMPK activation and antitumoral growth. *Clin Cancer Res* 17, 3993–4005.
- Rogers, D. J., & Tanimoto, T. T. (1960). A computer program for classifying plants. *Science* 132, 1115–1118.
- Roguev, A., Bandyopadhyay, S., Zofall, M., Zhang, K., Fischer, T., Collins, S. R., et al. (2008). Conservation and rewiring of functional modules revealed by an epistasis map in fission yeast. *Science* 322, 405–410.
- Rohn, H., Hartmann, A., Junker, A., Junker, B. H., & Schreiber, F. (2012). FluxMap: a VANTED add-on for the visual exploration of flux distributions in biological networks. *BMC Syst Biol* 6, 33.
- Romero, P., Wagg, J., Green, M. L., Kaiser, D., Krummenacker, M., & Karp, P. D. (2005). Computational prediction of human metabolic pathways from the complete human genome. *Genome Biol* 6, R2.
- Rosado, J. O., Henriques, J. P., & Bonatto, D. (2011). A systems pharmacology analysis of major chemotherapy combination regimens used in gastric cancer treatment: predicting potential new protein targets and drugs. *Curr Cancer Drug Targets* 11, 849–869.
- Rosen, Y., & Elman, N. M. (2009). Carbon nanotubes in drug delivery: focus on infectious diseases. *Expert Opin Drug Deliv* 6, 517–530.
- Rosvall, M., & Bergstrom, C. T. (2010). Mapping change in large networks. *PLoS One* 5, e8694.
- Rosvall, M., & Bergstrom, C. T. (2011). Multilevel compression of random walks on networks reveals hierarchical organization in large integrated systems. *PLoS One* 6, e18209.
- Rotem, E., Loinger, A., Ronin, I., Levin-Reisman, I., Gabay, C., Shores, N., et al. (2010). Regulation of phenotypic variability by a threshold-based mechanism underlies bacterial persistence. *Proc Natl Acad Sci USA* 107, 12541–12546.
- Rothkegel, A., & Lehnertz, K. (2012). Conedy: a scientific tool to investigate Complex Network Dynamics. <http://arxiv.org/abs/1202.3074>
- Roy, J., & Cyert, M. S. (2009). Cracking the phosphatase code: docking interactions determine substrate specificity. *Sci Signal* 2, re9.
- Rozenblatt-Rosen, O., Deo, R. C., Padi, M., Adelman, G., Calderwood, M. A., Rolland, T., et al. (2012). Interpreting cancer genomes using systematic host network perturbations by tumour virus proteins. *Nature* 487, 491–495.
- Rual, J. F., Venkatesan, K., Hao, T., Hirozane-Kishikawa, T., Dricot, A., Li, N., et al. (2005). Towards a proteome-scale map of the human protein–protein interaction network. *Nature* 437, 1173–1178.
- Ruan, W., Buerkle, T., & Dudeck, J. W. (2004). Mapping various information sources to a semantic network. *Stud Health Technol Inform* 107, 430–433.
- Ruffner, H., Bauer, A., & Bouwmeester, T. (2007). Human protein–protein interaction networks and the value for drug discovery. *Drug Discov Today* 12, 709–716.
- Russo, F., Di Bella, S., Nigita, G., Macca, V., Lagana, A., Giugno, R., et al. (2012). miRanda: extracellular circulating microRNAs database. *PLoS One* 7, e47786.
- Ruths, D. A., Nakhleh, L., Iyengar, M. S., Reddy, S. A., & Ram, P. T. (2006). Hypothesis generation in signaling networks. *J Comput Biol* 13, 1546–1557.
- Ruths, D., Muller, M., Tseng, J. T., Nakhleh, L., & Ram, P. T. (2008a). The signaling petri net-based simulator: a non-parametric strategy for characterizing the dynamics of cell-specific signaling networks. *PLoS Comput Biol* 4, e1000005.
- Ruths, D., Nakhleh, L., & Ram, P. T. (2008b). Rapidly exploring structural and dynamic properties of signaling networks using PathwayOracle. *BMC Syst Biol* 2, 76.
- Rzhetsky, A., Iossifov, I., Koike, T., Krauthammer, M., Kra, P., Morris, M., et al. (2004). GeneWays: a system for extracting, analyzing, visualizing, and integrating molecular pathway data. *J Biomed Inform* 37, 43–53.
- Rzhetsky, A., Wajngurt, D., Park, N., & Zheng, T. (2007). Probing genetic overlap among complex human phenotypes. *Proc Natl Acad Sci USA* 104, 11694–11699.
- Saadatpour, A., Wang, R. S., Liao, A., Liu, X., Loughran, T. P., Albert, I., et al. (2011). Dynamical and structural analysis of a T cell survival network identifies novel candidate therapeutic targets for large granular lymphocyte leukemia. *PLoS Comput Biol* 7, e1002267.
- Saavedra, S., Reed-Tsochas, F., & Uzzi, B. (2011). Common organizing mechanisms in ecological and socio-economic networks. <http://arxiv.org/abs/1110.0376>
- Sacco, F., Perfetto, L., Castagnoli, L., & Cesareni, G. (2012). The human phosphatase interactome: an intricate family portrait. *FEBS Lett* 586, 2732–2739.
- Sachs, K., Perez, O., Pe'er, D., Lauffenburger, D. A., & Nolan, G. P. (2005). Causal protein–signaling networks derived from multiparameter single-cell data. *Science* 308, 523–529.
- Salwinski, L., Miller, C. S., Smith, A. J., Pettit, F. K., Bowie, J. U., & Eisenberg, D. (2004). The Database of Interacting Proteins: 2004 update. *Nucleic Acids Res* 32, D449–D451.
- San Miguel, M., Johnson, J. H., Kertesz, J., Kaski, K., Diaz-Guilera, A., MacKay, R. S., et al. (2012). Challenges in complex systems science. *Eur Phys J Spec Top* 214, 245–271.
- Sanchez Claros, C., & Tramontano, A. (2012). Detecting mutually exclusive interactions in protein–protein interaction maps. *PLoS One* 7, e38765.
- Sandhu, K. S., Li, G., Poh, H. M., Quek, Y. L. K., Sia, Y. Y., Peh, S. Q., et al. (2012). Large scale functional organization of long-range chromatin interaction networks. *Cell Reports* 2, 1207–1219.
- Sanseau, P., Agarwal, P., Barnes, M. R., Pastinen, T., Richards, J. B., Cardon, L. R., et al. (2012). Use of genome-wide association studies for drug repositioning. *Nat Biotechnol* 30, 317–320.
- Santonico, E., Castagnoli, L., & Cesareni, G. (2005). Methods to reveal domain networks. *Drug Discov Today* 10, 1111–1117.
- Sanz-Pamplona, R., Berenguer, A., Sole, X., Cordero, D., Crous-Bou, M., Serra-Musach, J., et al. (2012). Tools for protein–protein interaction network analysis in cancer research. *Clin Transl Oncol* 14, 3–14.
- Sardiu, M. E., & Washburn, M. P. (2011). Building protein–protein interaction networks with proteomics and informatics tools. *J Biol Chem* 286, 23645–23651.
- Sariyüce, A. E., Saule, E., Kaya, K., & Catalyürek, Ü. V. (2012). Shattering and compressing networks for centrality analysis. <http://arxiv.org/abs/1209.6007>
- Sarkar, F. H., Li, Y., Wang, Z., Kong, D., & Ali, S. (2010). Implication of microRNAs in drug resistance for designing novel cancer therapy. *Drug Resist Updat* 13, 57–66.
- Saslis-Lagoudakis, C. H., Savolainen, V., Williamson, E. M., Forest, F., Wagstaff, S. J., Baral, S. R., et al. (2012). Phylogenies reveal predictive power of traditional medicine in bioprospecting. *Proc Natl Acad Sci USA* 109, 15835–15840.
- Satoh, J. (2012). Molecular network of microRNA targets in Alzheimer's disease brains. *Exp Neurol* 235, 436–446.
- Satoh, J., Tabunoki, H., & Arima, K. (2009). Molecular network analysis suggests aberrant CREB-mediated gene regulation in the Alzheimer disease hippocampus. *Dis Markers* 27, 239–252.
- Savino, R., Paduano, S., Preiano, M., & Terracciano, R. (2012). The proteomics big challenge for biomarkers and new drug-targets discovery. *Int J Mol Sci* 13, 13926–13948.
- Scannell, J. W., Blanckley, A., Boldon, H., & Warrington, B. (2012). Diagnosing the decline in pharmaceutical R&D efficiency. *Nat Rev Drug Discov* 11, 191–200.
- Schadt, E. E., Friend, S. H., & Shaywitz, D. A. (2009). A network view of disease and compound screening. *Nat Rev Drug Discov* 8, 286–295.
- Schaefer, C. F., Anthony, K., Krupa, S., Buchoff, J., Day, M., Hannay, T., et al. (2009). PID: The Pathway Interaction Database. *Nucleic Acids Res* 37, D674–D679.
- Schaffter, T., Marbach, D., & Floreano, D. (2011). GeneNetWeaver: in silico benchmark generation and performance profiling of network inference methods. *Bioinformatics* 27, 2263–2270.
- Scheer, M., Grote, A., Chang, A., Schomburg, I., Munaretto, C., Rother, M., et al. (2011). BRENDA, the enzyme information system in 2011. *Nucleic Acids Res* 39, D670–D676.
- Scheffer, M., Bascompte, J., Brock, W. A., Brovkin, V., Carpenter, S. R., Dakos, V., et al. (2009). Early-warning signals for critical transitions. *Nature* 461, 53–59.
- Schlecht, U., Miranda, M., Suresh, S., Davis, R. W., & St Onge, R. P. (2012). Multiplex assay for condition-dependent changes in protein–protein interactions. *Proc Natl Acad Sci USA* 109, 9213–9218.
- Schleker, S., Sun, J., Raghavan, B., Srnec, M., Muller, N., Koepfinger, M., et al. (2012). The current Salmonella–host interactome. *Proteomics Clin Appl* 6, 117–133.
- Schmelzle, K., Kane, S., Gridley, S., Lienhard, G. E., & White, F. M. (2006). Temporal dynamics of tyrosine phosphorylation in insulin signaling. *Diabetes* 55, 2171–2179.
- Schneider, C. M., Moreira, A. A., Andrade, J. S., Jr., Havlin, S., & Herrmann, H. J. (2011). Mitigation of malicious attacks on networks. *Proc Natl Acad Sci USA* 108, 3838–3841.
- Schreiber, S. L., & Bernstein, B. E. (2002). Signaling network model of chromatin. *Cell* 111, 771–778.
- Schreyer, A., & Blundell, T. (2009). CREDO: a protein–ligand interaction database for drug discovery. *Chem Biol Drug Des* 73, 157–167.
- Schulz, M., Bakker, B. M., & Klipp, E. (2009). Tide: a software for the systematic scanning of drug targets in kinetic network models. *BMC Bioinforma* 10, 344.
- Schumm, P., Scoglio, C. M., & van der Merwe, D. (2010). A network model of successive partitioning-limited solute diffusion through the stratum corneum. *J Theor Biol* 262, 471–477.
- Schuster, S., Krefl, J.-U., Schroeter, A., & Pfeiffer, T. (2008). Use of game-theoretical methods in biochemistry and biophysics. *J Biol Phys* 34, 1–17.
- Schwobbermeyer, H., & Wunschi, R. (2012). MAVisto: a tool for biological network motif analysis. *Methods Mol Biol* 804, 263–280.
- Searls, D. B. (2003). Pharmacophylogenomics: genes, evolution and drug targets. *Nat Rev Drug Discov* 2, 613–623.
- Secrier, M., Pavlopoulos, G. A., Aerts, J., & Schneider, R. (2012). Arena3D: visualizing time-driven phenotypic differences in biological systems. *BMC Bioinforma* 13, 45.
- Seebacher, J., & Gavin, A. C. (2011). SnapShot: protein–protein interaction networks. *Cell* 144 1000–1000e1.
- Segal, E., Shapira, M., Regev, A., Pe'er, D., Botstein, D., Koller, D., et al. (2003). Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data. *Nat Genet* 34, 166–176.



- Sengupta, U., Ukil, S., Dimitrova, N., & Agrawal, S. (2009). Expression-based network biology identifies alteration in key regulatory pathways of type 2 diabetes and associated risk/complications. *PLoS One* 4, e8110.
- Sergina, N. V., Rausch, M., Wang, D., Blair, J., Hann, B., Shokat, K. M., et al. (2007). Escape from HER-family tyrosine kinase inhibitor therapy by the kinase-inactive HER3. *Nature* 445, 437–441.
- Serra-Musach, J., Aguilar, H., Iorio, F., Comellas, F., Berenguer, A., Brunet, J., et al. (2012). Cancer develops, progresses and responds to therapies through restricted perturbation of the protein–protein interaction network. *Integr Biol* 4, 1038–1048.
- Sethi, A., Eargle, J., Blacka, A. A., & Luthey-Schulten, Z. (2009). Dynamical networks in tRNA:protein complexes. *Proc Natl Acad Sci USA* 106, 6620–6625.
- Shakarian, P., & Paulo, D. (2012). Large social networks can be targeted for viral marketing with small seed sets. <http://arxiv.org/abs/1205.4431>
- Shalgi, R., Lieber, D., Oren, M., & Pilpel, Y. (2007). Global and local architecture of the mammalian microRNA-transcription factor regulatory network. *PLoS Comput Biol* 3, e131.
- Sharan, R., & Ideker, T. (2006). Modeling cellular machinery through biological network comparison. *Nat Biotechnol* 24, 427–433.
- Sharan, R., Suthram, S., Kelley, R. M., Kuhn, T., McCuine, S., Uetz, P., et al. (2005). Conserved patterns of protein interaction in multiple species. *Proc Natl Acad Sci USA* 102, 1974–1979.
- Sharma, A., Chavali, S., Tabassum, R., Tandon, N., & Bharadwaj, D. (2010a). Gene prioritization in type 2 diabetes using domain interactions and network analysis. *BMC Genomics* 11, 84.
- Sharma, S. V., Haber, D. A., & Settleman, J. (2010b). Cell line-based platforms to evaluate the therapeutic efficacy of candidate anticancer agents. *Nat Rev Cancer* 10, 241–253.
- Shen, J., Zhang, J., Luo, X., Zhu, W., Yu, K., Chen, K., et al. (2007). Predicting protein–protein interactions based only on sequences information. *Proc Natl Acad Sci USA* 104, 4337–4341.
- Shen, Y., Liu, J., Estiu, G., Isin, B., Ahn, Y. Y., Lee, D. S., et al. (2010). Blueprint for antimicrobial hit discovery targeting metabolic networks. *Proc Natl Acad Sci USA* 107, 1082–1087.
- Shi, Y. (2009). Serine/threonine phosphatases: mechanism through structure. *Cell* 139, 468–484.
- Shiraishi, T., Matsuyama, S., & Kitano, H. (2010). Large-scale analysis of network bistability for human cancers. *PLoS Comput Biol* 6, e1000851.
- Shlomi, T., Cabili, M. N., Herrgard, M. J., Palsson, B. O., & Ruppin, E. (2008). Network-based prediction of human tissue-specific metabolism. *Nat Biotechnol* 26, 1003–1010.
- Shlomi, T., Cabili, M. N., & Ruppin, E. (2009). Predicting metabolic biomarkers of human inborn errors of metabolism. *Mol Syst Biol* 5, 263.
- Shmulevich, I., & Kauffman, S. A. (2004). Activities and sensitivities in Boolean network models. *Phys Rev Lett* 93, 048701.
- Shmulevich, I., Dougherty, E. R., & Zhang, W. (2002). Gene perturbation and intervention in probabilistic Boolean networks. *Bioinformatics* 18, 1319–1331.
- Simkó, G. I., Gyurkó, D., Veres, D. V., Nánási, T., & Csermely, P. (2009). Network strategies to understand the aging process and help age-related drug design. *Genome Med* 1, 90.
- Simonis, N., Rual, J. F., Lemmens, I., Boxus, M., Hirozane-Kishikawa, T., Gatot, J. S., et al. (2012). Host–pathogen interactome mapping for HTLV-1 and -2 retroviruses. *Retrovirology* 9, 26.
- Singer, T. (2007). Extrapolation of preclinical data into clinical reality translational science. *Ernst Schering Res Found Workshop* 2007, 1–5.
- Singh, S., Malik, B. K., & Sharma, D. K. (2007). Choke point analysis of metabolic pathways in *E. histolytica*: a computational approach for drug target identification. *Bioinformatics* 23, 68–72.
- Sirota, M., Dudley, J. T., Kim, J., Chiang, A. P., Morgan, A. A., Sweet-Cordero, A., et al. (2011). Discovery and preclinical validation of drug indications using compendia of public gene expression data. *Sci Transl Med* 3, 96ra77.
- Small, D. H. (2007). Neural network dysfunction in Alzheimer's disease: a drug development perspective. *Drug News Perspect* 20, 557–563.
- Small, B. G., McColl, B. W., Allmendinger, R., Pahle, J., Lopez-Castejon, G., Rothwell, N. J., et al. (2011). Efficient discovery of anti-inflammatory small-molecule combinations using evolutionary computing. *Nat Chem Biol* 7, 902–908.
- Smith, G. R., & Sternberg, M. J. (2002). Prediction of protein–protein interactions by docking methods. *Curr Opin Struct Biol* 12, 28–35.
- Smoot, M. E., Ono, K., Ruschinski, J., Wang, P. L., & Ideker, T. (2011). Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* 27, 431–432.
- Son, S.-W., Christensen, C., Bizhani, G., Foster, D. V., Grassberger, P., & Paczuski, M. (2012). Sampling properties of directed networks. *Phys Rev E* 86, 046104.
- Song, B., & Lee, H. (2012). Prioritizing disease genes by integrating domain interactions and disease mutations in a protein–protein interaction network. *Int J Innov Comput Inf Contr* 8, 1327–1338.
- Song, J., & Singh, M. (2013). From hub proteins to hub modules: the relationship between essentiality and centrality in the yeast interactome at different scales of organization. *PLoS Comput Biol* 9, e1002910.
- Song, B., Sridhar, P., Kahveci, T., & Ranka, S. (2009). Double iterative optimisation for metabolic network-based drug target identification. *Int J Data Min Bioinform* 3, 124–144.
- Sornette, D., & Osorio, I. (2011). Prediction. In I. Osorio, H. P. Zaveri, M. G. Frei, & S. Arthurs (Eds.), *Epilepsy: The Intersection of Neurosciences, Biology, Mathematics, Physics and Engineering* (pp. 203–240). London, UK: CRC Press, Taylor & Francis Group.
- Sóti, C., & Csermely, P. (2007). Aging cellular networks: chaperones as major participants. *Exp Gerontol* 42, 113–119.
- Sóti, C., Nagy, E., Gircz, Z., Vigh, L., Csermely, P., & Ferdinandy, P. (2005). Heat shock proteins as emerging therapeutic targets. *Br J Pharmacol* 146, 769–780.
- Spiró, Z., Kovács, I. A., & Csermely, P. (2008). Drug–therapy networks and the prediction of novel drug targets. *J Biol* 7, 20.
- Spizzo, R., Nicoloso, M. S., Croce, C. M., & Calin, G. A. (2009). SnapShot: microRNAs in cancer. *Cell* 137, 586–586e1.
- Squartini, T., Picciolo, F., Ruzzenenti, F., & Garlaschelli, D. (2012). Reciprocity of weighted networks. <http://arxiv.org/abs/1208.4208>
- Sreenivasiah, P. K., Rani, S., Cayetano, J., Arul, N., & Kim do, H. (2012). IPAVS: Integrated Pathway Resources, Analysis and Visualization System. *Nucleic Acids Res* 40, D803–D808.
- Sridhar, P., Kahveci, T., & Ranka, S. (2007). An iterative algorithm for metabolic network-based drug target identification. *Pac Symp Biocomput* 2007, 88–99.
- Sridhar, P., Song, B., Kahveci, T., & Ranka, S. (2008). Mining metabolic networks for optimal drug targets. *Pac Symp Biocomput* 2008, 291–302.
- Srihari, S., & Leong, H. W. (2012). Employing functional interactions for characterization and detection of sparse complexes from yeast PPI networks. *Int J Bioinform Res Appl* 8, 286–304.
- Srihari, S., & Leong, H. W. (2013). A survey of computational methods for protein complex prediction from protein interaction networks. *J Bioinform Comput Biol* 11, 123002.
- Srivas, R., Hannum, G., Ruschinski, J., Ono, K., Wang, P. L., Smoot, M., et al. (2011). Assembling global maps of cellular function through integrative analysis of physical and genetic networks. *Nat Protoc* 6, 1308–1323.
- Stark, C., Breitkreutz, B. J., Chatr-Aryamontri, A., Boucher, L., Oughtred, R., Livstone, M. S., et al. (2011). The BioGRID Interaction Database: 2011 update. *Nucleic Acids Res* 39, D698–D704.
- Stegmaier, P., Krull, M., Voss, N., Kel, A. E., & Wingender, E. (2010). Molecular mechanistic associations of human diseases. *BMC Syst Biol* 4, 124.
- Stein, A., Ceol, A., & Aloy, P. (2011). 3did: identification and classification of domain-based interactions of known three-dimensional structure. *Nucleic Acids Res* 39, D718–D723.
- Stelzl, U., Worm, U., Lalowski, M., Haenig, C., Brembeck, F. H., Goehler, H., et al. (2005). A human protein–protein interaction network: a resource for annotating the proteome. *Cell* 122, 957–968.
- Steták, A., Veress, R., Ovádi, J., Csermely, P., Kéri, G., & Ullrich, A. (2007). Nuclear translocation of the tumor marker pyruvate kinase M2 induces programmed cell death. *Cancer Res* 67, 1602–1608.
- Stites, E. C., Trampont, P. C., Ma, Z., & Ravichandran, K. S. (2007). Network analysis of oncogenic Ras activation in cancer. *Science* 318, 463–467.
- Stojimirović, A., & Yu, Y.-K. (2009). ITM Probe: analyzing information flow in protein networks. *Bioinformatics* 25, 2447–2449.
- Stokic, D., Hanel, R., & Thurner, S. (2009). A fast and efficient gene-network reconstruction method from multiple over-expression experiments. *BMC Bioinforma* 10, 253.
- Straub, F. B., & Szabolcsi, G. (1964). O dinamicaszékiz aspektah sztkurú fermentov. (On the dynamic aspects of protein structure). In A. E. Braunstein (Ed.), *Molecular Biology, Problems and Perspectives* (pp. 182–187). Moscow: Izdat. Nauka.
- Stumpf, M. P., & Wiuf, C. (2010). Incomplete and noisy network data as a percolation process. *J R Soc Interface* 7, 1411–1419.
- Stumpf, M. P., Wiuf, C., & May, R. M. (2005). Subnets of scale-free networks are not scale-free: sampling properties of networks. *Proc Natl Acad Sci USA* 102, 4221–4224.
- Stumpf, M. P., Thorne, T., de Silva, E., Stewart, R., An, H. J., Lappe, M., et al. (2008). Estimating the size of the human interactome. *Proc Natl Acad Sci USA* 105, 6959–6964.
- Su, J. G., Xu, X. J., Li, C. H., Chen, W. Z., & Wang, C. X. (2011). Identification of key residues for protein conformational transition using elastic network model. *J Chem Phys* 135, 174101.
- Suderman, M., & Hallett, M. (2007). Tools for visually exploring biological networks. *Bioinformatics* 23, 2651–2659.
- Sugaya, N., & Furuya, T. (2011). Dr. PIAS: an integrative system for assessing the druggability of protein–protein interactions. *BMC Bioinforma* 12, 50.
- Sugaya, N., Ikeda, K., Tashiro, T., Takeda, S., Otomo, J., Ishida, Y., et al. (2007). An integrative in silico approach for discovering candidates for drug-targetable protein–protein interactions in interactome data. *BMC Pharmacol* 7, 10.
- Sugaya, N., Kanai, S., & Furuya, T. (2012). Dr. PIAS 2.0: an update of a database of predicted druggable protein–protein interactions. *Database*, bas034.
- Sun, W., & He, J. (2011). From isotropic to anisotropic side chain representations: comparison of three models for residue contact estimation. *PLoS One* 6, e19238.
- Sun, J., & Zhao, Z. (2010). A comparative study of cancer proteins in the human protein–protein interaction network. *BMC Genomics* 11, S5.
- Sun, J., Faloutsos, C., Papadimitriou, S., & Yu, P. S. (2007). Graphscope: parameter-free mining of large, time-evolving graphs. *Proc 13th ACM SIGKDD Intl Conf Knowledge Discovery Data Mining* (pp. 687–696).
- Sun, J., Wum, Y., Xu, H., & Zhao, Z. (2012). DTome: a web-based tool for drug–target interaction construction. *BMC Bioinforma* 13, S7.
- Sun, Y., Zhu, R., Ye, H., Tang, K., Zhao, J., Chen, Y., Liu, Q., Cao, Z. (in press). Towards a bioinformatics analysis of anti-Alzheimer's herbal medicines from a target network perspective. *Brief Bioinform*. <http://dx.doi.org/10.1093/bib/bbs025>.
- Suthram, S., Dudley, J. T., Chiang, A. P., Chen, R., Hastie, T. J., & Butte, A. J. (2010). Network-based elucidation of human disease similarities reveals common functional modules enriched for pluripotent drug targets. *PLoS Comput Biol* 6, e1000662.
- Szalay-Bekó, M., Palotai, R., Szappanos, B., Kovács, I. A., Papp, B., & Csermely, P. (2012). ModuLand plug-in for Cytoscape: determination of hierarchical layers of overlapping modules and community centrality. *Bioinformatics* 28, 2202–2204.
- Szilágyi, A., Csermely, P., Nussinov, R. (in press). Allo-network drugs: extension of the allosteric drug concept to protein–protein interaction and signaling networks. *Curr Top Med Chem*, 13.
- Szklarczyk, D., Franceschini, A., Kuhn, M., Simonovic, M., Roth, A., Minguez, P., et al. (2011). The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res* 39, D561–D568.

- Tabei, Y., Pauwels, E., Stoven, V., Takemoto, K., & Yamanishi, Y. (2012). Identification of chemogenomic features from drug–target interaction networks using interpretable classifiers. *Bioinformatics* 28, 1487–1494.
- Taboureau, O., Nielsen, S. K., Audouze, K., Weinhold, N., Edsgard, D., Roque, F. S., et al. (2011). ChemProt: a disease chemical biology database. *Nucleic Acids Res* 39, D367–D372.
- Tajer, A., & Poor, H. V. (2012). Quick search for rare events. <http://arxiv.org/abs/1210.2406>
- Takarabe, M., Okuda, S., Itoh, M., Tokimatsu, T., Goto, S., & Kanehisa, M. (2008). Network analysis of adverse drug interactions. *Genome Inform* 20, 252–259.
- Takarabe, M., Shigemizu, D., Kotera, M., Goto, S., & Kanehisa, M. (2011). Network-based analysis and characterization of adverse drug–drug interactions. *J Chem Inf Model* 51, 2977–2985.
- Takarabe, M., Kotera, M., Nishimura, Y., Goto, S., & Yamanishi, Y. (2012). Drug target prediction using adverse event report systems: a pharmacogenomic approach. *Bioinformatics* 28, i611–i618.
- Takigawa, I., Tsuda, K., & Mamitsuka, H. (2011). Mining significant substructure pairs for interpreting polypharmacology in drug–target network. *PLoS One* 6, e16999.
- Talchai, C., Xuan, S., Lin, H. V., Sussel, L., & Accili, D. (2012). Pancreatic beta cell dedifferentiation as a mechanism of diabetic beta cell failure. *Cell* 150, 1223–1234.
- Tanaka, R., Yi, T. M., & Doyle, J. (2005). Some protein interaction data do not exhibit power law statistics. *FEBS Lett* 579, 5140–5144.
- Tanaka, N., Ohno, K., Niimi, T., Moritomo, A., Mori, K., & Orita, M. (2009). Small-world phenomena in chemical library networks: application to fragment-based drug discovery. *J Chem Inf Model* 49, 2677–2686.
- Tang, S., Liao, J. C., Dunn, A. R., Altman, R. B., Spudis, J. A., & Schmidt, J. P. (2007). Predicting allosteric communication in myosin via a pathway of conserved residues. *J Mol Biol* 373, 1361–1373.
- Tang, J., Scellato, S., Musolesi, M., Mascolo, C., & Latora, V. (2010). Small-world behavior in time-varying graphs. *Phys Rev E* 81, 055101.
- Taniguchi, C. M., Emanuelli, B., & Kahn, C. R. (2006). Critical nodes in signalling pathways: insights into insulin action. *Nat Rev Mol Cell Biol* 7, 85–96.
- Taylor, I. W., Linding, R., Warde-Farley, D., Liu, Y., Pesquita, C., Faria, D., et al. (2009). Dynamic modularity in protein interaction networks predicts breast cancer outcome. *Nat Biotechnol* 27, 199–204.
- Tegnér, J., & Björkegren, J. (2007). Perturbations to uncover gene networks. *Trends Genet* 23, 34–41.
- Tegnér, J., Yeung, M. K., Hasty, J., & Collins, J. J. (2003). Reverse engineering gene networks: integrating genetic perturbations with dynamical modeling. *Proc Natl Acad Sci USA* 100, 5944–5949.
- Tehver, R., Chen, J., & Thirumalai, D. (2009). Allosteric wiring diagrams in the transitions that drive the GroEL reaction cycle. *J Mol Biol* 387, 390–406.
- Temkin, O. N., & Bonchev, D. G. (1992). Application of graph theory to chemical kinetics. *J Chem Educ* 69, 544–550.
- Tentner, A. R., Lee, M. J., Ostheimer, G. J., Samson, L. D., Lauffenburger, D. A., & Yaffe, M. B. (2012). Combined experimental and computational analysis of DNA damage signaling reveals context-dependent roles for Erk in apoptosis and G1/S arrest after genotoxic stress. *Mol Syst Biol* 8, 568.
- Terfve, C. D., Cokelaer, T., Henriques, D., Macnamara, A., Gonçalves, E., Morris, M. K., et al. (2012). CellNOptR: a flexible toolkit to train protein signaling networks to data using multiple logic formalisms. *BMC Syst Biol* 6, 133.
- Theocharidis, A., van Dongen, S., Enright, A. J., & Freeman, T. C. (2009). Network visualization and analysis of gene expression data using BioLayout Express(3D). *Nat Protoc* 4, 1535–1550.
- Thorn, C. F., Klein, T. E., & Altman, R. B. (2010). Pharmacogenomics and bioinformatics: PharmGKB. *Pharmacogenomics* 11, 501–505.
- Tiligada, E. (2006). Chemotherapy: induction of stress responses. *Endocr Relat Cancer* 13, S115–S124.
- Tomida, A., & Tsuruo, T. (1999). Drug resistance mediated by cellular stress response to the microenvironment of solid tumors. *Anticancer Drug Des* 14, 169–177.
- Tomlinson, I. P., Novelli, M. R., & Bodmer, W. F. (1996). The mutation rate and cancer. *Proc Natl Acad Sci USA* 93, 14800–14803.
- Tomba, P. (2012). On the supertertiary structure of proteins. *Nat Chem Biol* 8, 597–600.
- Tong, A. H., Lesage, G., Bader, G. D., Ding, H., Xu, H., Xin, X., et al. (2004). Global mapping of the yeast genetic interaction network. *Science* 303, 808–813.
- Torkamani, A., & Schork, N. J. (2009). Identification of rare cancer driver mutations by network reconstruction. *Genome Res* 19, 1570–1578.
- Tranchevent, L. C., Barriot, R., Yu, S., Van Vooren, S., Van Loo, P., Coessens, B., et al. (2008). ENDEAVOUR update: a web resource for gene prioritization in multiple species. *Nucleic Acids Res* 36, W377–W384.
- Tsai, C. J., Kumar, S., Ma, B., & Nussinov, R. (1999). Folding funnels, binding funnels, and protein function. *Protein Sci* 8, 1181–1190.
- Tsai, C. J., Ma, B., & Nussinov, R. (2009). Protein–protein interaction networks: how can a hub protein bind so many different partners? *Trends Biochem Sci* 34, 594–600.
- Tu, Z., Argmann, C., Wong, K. K., Mitnaul, L. J., Edwards, S., Sach, I. C., et al. (2009). Integrating siRNA and protein–protein interaction data to identify an expanded insulin signaling network. *Genome Res* 19, 1057–1067.
- Tun, K., Menghini, M., D'Andrea, L., Dhar, P., Tanaka, H., & Giuliani, A. (2011). Why so few drug targets: a mathematical explanation? *Curr Comput Aided Drug Des* 7, 206–213.
- Tuikkala, J., Vahamaa, H., Salmela, P., Nevalainen, O. S., & Aittokallio, T. (2012). A multilevel layout algorithm for visualizing physical and genetic interaction networks, with emphasis on their modular organization. *BioData Min* 5, 2.
- Tunçbag, N., McCallum, S., Huang, S. S., & Fraenkel, E. (2012). SteinerNet: a web server for integrating 'omic' data to discover hidden components of response pathways. *Nucleic Acids Res* 40, W505–W509.
- Tuske, S., Sarafianos, S. G., Clark, A. D., Jr., Ding, J., Naeger, L. K., White, K. L., et al. (2004). Structures of HIV-1 RT-DNA complexes before and after incorporation of the anti-AIDS drug tenofovir. *Nat Struct Mol Biol* 11, 469–474.
- Uetz, P., Dong, Y. A., Zeretzke, C., Atzler, C., Baiker, A., Berger, B., et al. (2006). Herpesviral protein networks and their interaction with the human proteome. *Science* 311, 239–242.
- Ulitsky, I., Maron-Katz, A., Shavit, S., Sagiv, D., Linhart, C., Elkon, R., et al. (2010). Expander: from expression microarrays to networks and functions. *Nat Protoc* 5, 303–322.
- Ummanni, R., Mundt, F., Pospisil, H., Venz, S., Scharf, C., Barrett, C., et al. (2011). Identification of clinically relevant protein targets in prostate cancer with 2D-DIGE coupled mass spectrometry and systems biology network platform. *PLoS One* 6, e16833.
- Valente, A. X. C. N. (2010). Prediction in the hypothesis-rich regime. <http://arxiv.org/abs/1003.3551>
- Valente, T. W. (2012). Network interventions. *Science* 337, 49–53.
- van Laarhoven, T., Nabuurs, S. B., & Marchiori, E. (2011). Gaussian interaction profile kernels for predicting drug–target interaction. *Bioinformatics* 27, 3036–3043.
- Vandin, F., Clay, P., Upfal, E., & Raphael, B. J. (2012). Discovery of mutated subnetworks associated with clinical data in cancer. *Pac Symp Biocomput* 2012, 55–66.
- Vanunu, O., Magger, O., Ruppim, E., Shlomi, T., & Sharan, R. (2010). Associating genes and protein complexes with disease via network propagation. *PLoS Comput Biol* 6, e1000641.
- Vaquerizas, J. M., Kummerfeld, S. K., Teichmann, S. A., & Luscombe, N. M. (2009). A census of human transcription factors: function, expression and evolution. *Nat Rev Genet* 10, 252–263.
- Varnek, A., & Baskin, I. I. (2011). Chemoinformatics as a theoretical chemistry discipline. *Mol Inf* 30, 20–32.
- Vashisht, R., Mondal, A. K., Jain, A., Shah, A., Vishnoi, P., Priyadarshini, P., et al. (2012). Crowd sourcing a new paradigm for interactome driven drug target identification in *Mycobacterium tuberculosis*. *PLoS One* 7, e39808.
- Vassilev, L. T., Vu, B. T., Graves, B., Carvajal, D., Podlaski, F., Filipovic, Z., et al. (2004). In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. *Science* 303, 844–848.
- Vazquez, A. (2009). Optimal drug combinations and minimal hitting sets. *BMC Syst Biol* 3, 81.
- Vazquez, A., Markert, E. K., & Oltvai, Z. N. (2011). Serine biosynthesis with one carbon catabolism and the glycine cleavage system represents a novel pathway for ATP generation. *PLoS One* 6, e25881.
- Vergoulis, T., Vlachos, I. S., Alexiou, P., Georgakilas, G., Maragkakias, M., Reczko, M., et al. (2012). TarBase 6.0: capturing the exponential growth of miRNA targets with experimental support. *Nucleic Acids Res* 40, D222–D229.
- Videla, S., Guziolowski, C., Eduati, F., Thiele, S., Grabe, N., Saez-Rodriguez, J., et al. (2012). Revisiting the training of logic models of protein signaling networks with a formal approach based on answer set programming. *Lect Notes Comput Sci* 7605, 342–361.
- Vigh, L., Literati, P. N., Horvath, I., Torok, Z., Balogh, G., Glatz, A., et al. (1997). Bimoclozolol: a nontoxic, hydroxylamine derivative with stress protein-inducing activity and cytoprotective effects. *Nat Med* 3, 1150–1154.
- Vilar, S., Gonzalez-Diaz, H., Santana, L., & Uriarte, E. (2009). A network-QSAR model for prediction of genetic-component biomarkers in human colorectal cancer. *J Theor Biol* 261, 449–458.
- Vina, D., Uriarte, E., Orallo, F., & Gonzalez-Diaz, H. (2009). Alignment-free prediction of a drug–target complex network based on parameters of drug connectivity and protein sequence of receptors. *Mol Pharm* 6, 825–835.
- Vishveshwara, S., Ghosh, A., & Hansia, P. (2009). Intra and inter-molecular communications through protein structure network. *Curr Protein Pept Sci* 10, 146–160.
- Vlasblom, J., Wu, S., Pu, S., Superina, M., Liu, G., Orsi, C., et al. (2006). GenePro: a Cytoscape plug-in for advanced visualization and analysis of interaction networks. *Bioinformatics* 22, 2178–2179.
- Volinia, S., Galasso, M., Costinean, S., Tagliavini, L., Gamberoni, G., Drusco, A., et al. (2010). Reprogramming of miRNA networks in cancer and leukemia. *Genome Res* 20, 589–599.
- von Eichborn, J., Murgueitio, M. S., Dunkel, M., Koerner, S., Bourne, P. E., & Preissner, R. (2011). PROMISCUOUS: a database for network-based drug-repositioning. *Nucleic Acids Res* 39, D1060–D1066.
- Wachi, S., Yoneda, K., & Wu, R. (2005). Interactome–transcriptome analysis reveals the high centrality of genes differentially expressed in lung cancer tissues. *Bioinformatics* 21, 4205–4208.
- Wagner, A., & Fell, D. A. (2001). The small world inside large metabolic networks. *Proc Biol Sci* 268, 1803–1810.
- Wang, B. (2012). On sampling social networking services. <http://arxiv.org/abs/1209.2486>
- Wang, R. S., & Albert, R. (2011). Elementary signaling modes predict the essentiality of signal transduction network components. *BMC Syst Biol* 5, 44.
- Wang, J., Zhang, S., Wang, Y., Chen, L., & Zhang, X. S. (2009). Disease-aging network reveals significant roles of aging genes in connecting genetic diseases. *PLoS Comput Biol* 5, e1000521.
- Wang, J., Lu, M., Qiu, C., & Cui, Q. (2010a). TransmiR: a transcription factor-microRNA regulation database. *Nucleic Acids Res* 38, D119–D122.
- Wang, Y., Bolton, E., Dracheva, S., Karapetyan, K., Shoemaker, B. A., Suzek, T. O., et al. (2010b). An overview of the PubChem BioAssay resource. *Nucleic Acids Res* 38, D255–D266.
- Wang, X., Gulbahce, N., & Yu, H. (2011a). Network-based methods for human disease gene prediction. *Brief Funct Genomics* 10, 280–293.
- Wang, L., Khankhanian, P., Baranzini, S. E., & Mousavi, P. (2011b). iCTNet: a Cytoscape plugin to produce and analyze integrative complex traits networks. *BMC Bioinform* 12, 380.

- Wang, C., Jiang, W., Li, W., Lian, B., Chen, X., Hua, L., et al. (2011c). Topological properties of the drug targets regulated by microRNA in human protein–protein interaction network. *J Drug Target* 19, 354–364.
- Wang, W. X., Ni, X., Lai, Y. C., & Grebogi, C. (2012a). Optimizing controllability of complex networks by minimum structural perturbations. *Phys Rev E* 85, 026115.
- Wang, X., Wei, X., Thijssen, B., Das, J., Lipkin, S. M., & Yu, H. (2012b). Three-dimensional reconstruction of protein networks provides insight into human genetic disease. *Nat Biotechnol* 30, 159–164.
- Wang, J., Li, Z. X., Qiu, C. X., Wang, D., & Cui, Q. H. (2012c). The relationship between rational drug design and drug side effects. *Brief Bioinform* 13, 377–382.
- Wang, Z., Liu, J., Yu, Y., Chen, Y., Wang, Y., Lindow, M., et al. (2012d). Modular pharmacology: the next paradigm in drug discovery. *Expert Opin Drug Discov* 7, 667–677.
- Wang, Y. Y., Xu, K. J., Song, J., & Zhao, X. M. (2012e). Exploring drug combinations in genetic interaction network. *BMC Bioinforma* 13, S7.
- Wang, I. M., Zhang, B., Yang, X., Zhu, J., Stepaniants, S., Zhang, C., et al. (2012f). Systems analysis of eleven rodent disease models reveals an inflammatory signature and key drivers. *Mol Syst Biol* 8, 594.
- Warburg, O. (1956). On the origin of cancer cells. *Science* 123, 309–314.
- Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature* 393, 440–442.
- Wawer, M., & Bajorath, J. (2011a). Local structural changes, global data views: graphical substructure–activity relationship trailing. *J Med Chem* 54, 2944–2951.
- Wawer, M., & Bajorath, J. (2011b). Extracting SAR information from a large collection of anti-malarial screening hits by NSG-SPT analysis. *ACS Med Chem Lett* 2, 201–206.
- Wawer, M., Peltason, L., Weskamp, N., Teckentrup, A., & Bajorath, J. (2008). Structure–activity relationship anatomy by network-like similarity graphs and local structure–activity relationship indices. *J Med Chem* 51, 6075–6084.
- Wawer, M., Lounkine, E., Wassermann, A. M., & Bajorath, J. (2010). Data structures and computational tools for the extraction of SAR information from large compound sets. *Drug Discov Today* 15, 630–639.
- Weisel, M., Kriegl, J. M., & Schneider, G. (2010). Architectural repertoire of ligand-binding pockets on protein surfaces. *ChemBioChem* 11, 556–563.
- Wells, J. A., & McClendon, C. L. (2007). Reaching for high-hanging fruit in drug discovery at protein–protein interfaces. *Nature* 450, 1001–1009.
- Wermuth, C. G. (2006). Selective optimization of side activities: the SOSA approach. *Drug Discov Today* 11, 160–164.
- Werner, E. (2003). In silico multicellular systems biology and minimal genomes. *Drug Discov Today* 8, 1121–1127.
- Werner, E. (2005). Genome semantics, in silico multicellular systems and the Central Dogma. *FEBS Lett* 579, 1779–1782.
- Werner, E. (2011). Cancer networks. A general theoretical and computational framework for understanding cancer. <http://arxiv.org/abs/1110.5865>
- Weskamp, N., Huellermeier, E., & Klebe, G. (2009). Merging chemical and biological space: structural mapping of enzyme binding pocket space. *Proteins* 76, 317–330.
- West, J., Bianconi, G., Severini, S., & Teschendorff, A. E. (2012). Differential network entropy reveals cancer system hallmarks. *Sci Rep* 2, 802.
- Westerhoff, H. V., Mosekilde, E., Noe, C. R., & Clemensen, A. M. (2008). Integrating systems approaches into pharmaceutical sciences. *Eur J Pharm Sci* 35, 1–4.
- White, F. M. (2008). Quantitative phosphoproteomic analysis of signaling network dynamics. *Curr Opin Biotechnol* 19, 404–409.
- White, E., & DiPaola, R. S. (2009). The double-edged sword of autophagy modulation in cancer. *Clin Cancer Res* 15, 5308–5316.
- White, J. C., & Mikulecky, D. C. (1981). Application of network thermodynamics to the computer modeling of the pharmacology of anticancer agents: a network model for methotrexate action as a comprehensive example. *Pharmacol Ther* 15, 251–291.
- Wiedermann, M., Donges, J. F., Heitzig, J., & Kurths, J. (2013). Node-weighted interacting network measures improve the representation of real-world complex systems. <http://arxiv.org/abs/1301.0805>.
- Wilson, T. R., Johnston, P. G., & Longley, D. B. (2009). Anti-apoptotic mechanisms of drug resistance in cancer. *Curr Cancer Drug Targets* 9, 307–319.
- Winkler, D. A. (2004). Neural networks as robust tools in drug lead discovery and development. *Mol Biotechnol* 27, 139–168.
- Wishart, D. S., Knox, C., Guo, A. C., Eisner, R., Young, N., Gautam, B., et al. (2009). HMDB: a knowledgebase for the human metabolome. *Nucleic Acids Res* 37, D603–D610.
- Wuif, C., Brameier, M., Hagberg, O., & Stumpf, M. P. (2006). A likelihood approach to analysis of network data. *Proc Natl Acad Sci USA* 103, 7566–7570.
- Wolfson, M., Budovsky, A., Tacutu, R., & Fraifeld, V. (2009). The signaling hubs at the crossroad of longevity and age-related disease networks. *Int J Biochem Cell Biol* 41, 516–520.
- Wong, P., & Frishman, D. (2006). Fold designability, distribution, and disease. *PLoS Comput Biol* 2, e40.
- Wong, P. K., Yu, F., Shahangian, A., Cheng, G., Sun, R., & Ho, C. M. (2008). Closed-loop control of cellular functions using combinatory drugs guided by a stochastic search algorithm. *Proc Natl Acad Sci USA* 105, 5105–5110.
- Wu, Z. X., & Holme, P. (2011). Onion structure and network robustness. *Phys Rev E* 84, 026106.
- Wu, X., Jiang, R., Zhang, M. Q., & Li, S. (2008). Network-based global inference of human disease genes. *Mol Syst Biol* 4, 189.
- Wu, X., Liu, Q., & Jiang, R. (2009a). Align human interactome with phenome to identify causative genes and networks underlying disease families. *Bioinformatics* 25, 98–104.
- Wu, J., Vallenius, T., Ovaska, K., Westermarck, J., Makela, T. P., & Hautaniemi, S. (2009b). Integrated network analysis platform for protein–protein interactions. *Nat Methods* 6, 75–77.
- Wu, G., Feng, X., & Stein, L. (2010). A human functional protein interaction network and its application to cancer data analysis. *Genome Biol* 11, R53.
- Wu, C., Zhu, J., & Zhang, X. (2012). Integrating gene expression and protein–protein interaction network to prioritize cancer-associated genes. *BMC Bioinforma* 13, 182.
- Xi, Y., Chen, Y. P., Qian, C., & Wang, F. (2011). Comparative study of computational methods to detect the correlated reaction sets in biochemical networks. *Brief Bioinform* 12, 132–150.
- Xia, K., Dong, D., & Han, J. D. (2006). IntNetDB v1.0: an integrated protein–protein interaction network database generated by a probabilistic model. *BMC Bioinforma* 7, 508.
- Xia, J., Sun, J., Jia, P., & Zhao, Z. (2011). Do cancer proteins really interact strongly in the human protein–protein interaction network? *Comput Biol Chem* 35, 121–125.
- Xiao, F., Zuo, Z., Cai, G., Kang, S., Gao, X., & Li, T. (2009). miRecords: an integrated resource for microRNA–target interactions. *Nucleic Acids Res* 37, D105–D110.
- Xie, L., Xie, L., & Bourne, P. E. (2009a). A unified statistical model to support local sequence order independent similarity searching for ligand-binding sites and its application to genome-based drug discovery. *Bioinformatics* 25, I305–I312.
- Xie, L., Li, J., & Bourne, P. E. (2009b). Drug discovery using chemical systems biology: identification of the protein–ligand binding network to explain the side effects of CETP inhibitors. *PLoS Comput Biol* 5, e1000387.
- Xie, L., & Bourne, P. E. (2011). Structure-based systems biology for analyzing off-target binding. *Curr Opin Struct Biol* 21, 189–199.
- Xie, Z., -R., & Hwang, M. (2010). An interaction-motif-based scoring function for protein–ligand docking. *BMC Bioinforma* 11, 298.
- Xing, H., & Gardner, T. S. (2006). The mode-of-action by network identification (MNI) algorithm: a network biology approach for molecular target identification. *Nat Protoc* 1, 2551–2554.
- Xiong, H., & Choe, Y. (2008). Dynamical pathway analysis. *BMC Syst Biol* 2, 9.
- Xiong, J., Liu, J., Rayner, S., Tian, Z., Li, Y., & Chen, S. (2010). Pre-clinical drug prioritization via prognosis-guided genetic interaction networks. *PLoS One* 5, e13937.
- Xu, J., & Li, Y. (2006). Discovering disease-genes by topological features in human protein–protein interaction network. *Bioinformatics* 22, 2800–2805.
- Xu, Y., Xu, D., Gabow, H. N., & Gabow, H. (2000). Protein domain decomposition using a graph-theoretic approach. *Bioinformatics* 16, 1091–1104.
- Xu, Q., Xiang, E. W., & Yang, Q. (2011a). Transferring network topological knowledge for predicting protein–protein interactions. *Proteomics* 11, 3818–3825.
- Xu, F., Zhao, C., Li, Y., Li, J., Deng, Y., & Shi, T. (2011b). Exploring virus relationships based on virus–host protein–protein interaction network. *BMC Syst Biol* 5, S11.
- Xue, H., Xian, B., Dong, D., Xia, K., Zhu, S., Zhang, Z., et al. (2007). A modular network model of aging. *Mol Syst Biol* 3, 147.
- Yabuuchi, H., Nijima, S., Takematsu, H., Ida, T., Hirokawa, T., Hara, T., et al. (2011). Analysis of multiple compound–protein interactions reveals novel bioactive molecules. *Mol Syst Biol* 7, 472.
- Yamada, T., Letunic, I., Okuda, S., Kanehisa, M., & Bork, P. (2011). iPath2.0: interactive pathway explorer. *Nucleic Acids Res* 39, W412–W415.
- Yamanishi, Y., Araki, M., Gutteridge, A., Honda, W., & Kanehisa, M. (2008). Prediction of drug–target interaction networks from the integration of chemical and genomic spaces. *Bioinformatics* 24, i232–i240.
- Yamanishi, Y., Kotera, M., Kanehisa, M., & Goto, S. (2010). Drug–target interaction prediction from chemical, genomic and pharmacological data in an integrated framework. *Bioinformatics* 26, i246–i254.
- Yamanishi, Y., Pauwels, E., Saigo, H., & Stoven, V. (2011). Extracting sets of chemical substructures and protein domains governing drug–target interactions. *J Chem Inf Model* 51, 1183–1194.
- Yan, B., & Gregory, S. (2012). Finding missing edges in networks based on their community structure. *Phys Rev E* 85, 056112.
- Yan, K. K., Mazo, I., Yuryev, A., & Maslov, S. (2007b). Prediction and verification of indirect interactions in densely interconnected regulatory networks. <http://arxiv.org/abs/0710.0892>
- Yan, X., Mehan, M. R., Huang, Y., Waterman, M. S., Yu, P. S., & Zhou, X. J. (2007a). A graph-based approach to systematically reconstruct human transcriptional regulatory modules. *Bioinformatics* 23, i577–i586.
- Yang, J., & Jiang, X. F. (2010). A novel approach to predict protein–protein interactions related to Alzheimer's disease based on complex network. *Protein Pept Lett* 17, 356–366.
- Yang, Y., & Leskovec, J. (2012). Structure and overlaps of communities in networks. <http://arxiv.org/abs/1205.6228>.
- Yang, K., Bai, H., Ouyang, Q., Lai, L., & Tang, C. (2008). Finding multiple target optimal intervention in disease-related molecular network. *Mol Syst Biol* 4, 228.
- Yang, L., Luo, H., Chen, J., Xing, Q., & He, L. (2009a). SePreSA: a server for the prediction of populations susceptible to serious adverse drug reactions implementing the methodology of a chemical–protein interactome. *Nucleic Acids Res* 37, W406–W412.
- Yang, L., Chen, J., & He, L. (2009b). Harvesting candidate genes responsible for serious adverse drug reactions from a chemical–protein interactome. *PLoS Comput Biol* 5, e1000441.
- Yang, L., Xu, L., & He, L. (2009c). A CitationRank algorithm inheriting Google technology designed to highlight genes responsible for serious adverse drug reaction. *Bioinformatics* 25, 2244–2250.
- Yang, L., Chen, J., Shi, L., Hudock, M. P., Wang, K., & He, L. (2010). Identifying unexpected therapeutic targets via chemical–protein interaction. *PLoS One* 5, e9568.
- Yang, L., Wang, K. J., Wang, L. S., Jegga, A. G., Qin, S. Y., He, G., et al. (2011). Chemical–protein interactome and its application in off-target identification. *Interdiscip Sci* 3, 22–30.
- Yang, X., Zhang, B., & Zhu, J. (2012). Functional genomics- and network-driven systems biology approaches for pharmacogenomics and toxicogenomics. *Curr Drug Metab* 13, 952–967.
- Yazicioglu, A. Y., Abbas, W., & Egerstedt, M. (2012). A tight lower bound on the controllability of networks with multiple leaders. <http://arxiv.org/abs/1205.3058>

- Ye, H., Yang, L., Cao, Z., Tang, K., & Li, Y. (2012). A pathway profile-based method for drug repositioning. *Chin Sci Bull* 57, 2016–2112.
- Yeh, I., Hanekamp, T., Tsoka, S., Karp, P. D., & Altman, R. B. (2004). Computational analysis of *Plasmodium falciparum* metabolism: organizing genomic information to facilitate drug discovery. *Genome Res* 14, 917–924.
- Yeh, P., Tschumi, A. I., & Kishony, R. (2006). Functional classification of drugs by properties of their pairwise interactions. *Nat Genet* 38, 489–494.
- Yeh, S. H., Yeh, H. Y., & Soo, V. W. (2012). A network flow approach to predict drug targets from microarray data, disease genes and interactome network—case study on prostate cancer. *J Clin Bioinforma* 2, 1.
- Yellaboina, S., Tasneem, A., Zaykin, D. V., Raghavachari, B., & Jothi, R. (2011). DOMINE: a comprehensive collection of known and predicted domain–domain interactions. *Nucleic Acids Res* 39, D730–D735.
- Yeturu, K., & Chandra, N. (2008). PocketMatch: a new algorithm to compare binding sites in protein structures. *BMC Bioinforma* 9, 543.
- Yeturu, K., & Chandra, N. (2011). PocketAlign a novel algorithm for aligning binding sites in protein structures. *J Chem Inf Model* 51, 1725–1736.
- Yeung, M. K., Tegnér, J., & Collins, J. J. (2002). Reverse engineering gene networks using singular value decomposition and robust regression. *Proc Natl Acad Sci USA* 99, 6163–6168.
- Yeung, K. Y., Dombek, K. M., Lo, K., Mittler, J. E., Zhu, J., Schadt, E. E., et al. (2011). Construction of regulatory networks using expression time-series data of a genotyped population. *Proc Natl Acad Sci USA* 108, 19436–19441.
- Yildirim, M. A., Goh, K. I., Cusick, M. E., Barabási, A. L., & Vidal, M. (2007). Drug–target network. *Nat Biotechnol* 25, 1119–1126.
- Yip, K. Y., Alexander, R. P., Yan, K. K., & Gerstein, M. (2010). Improved reconstruction of *in silico* gene regulatory networks by integrating knockout and perturbation data. *PLoS One* 5, e8121.
- Yoon, B. J. (2011). Enhanced stochastic optimization algorithm for finding effective multi-target therapeutics. *BMC Bioinforma* 12, S18.
- Yu, Q., & Huang, J. -F. (2012). The analysis of the druggable families based on topological features in the protein–protein interaction network. *Lett Drug Des Discov* 9, 426–430.
- Yu, H., Zhu, X., Greenbaum, D., Karro, J., & Gerstein, M. (2004a). TopNet: a tool for comparing biological sub-networks, correlating protein properties with topological statistics. *Nucleic Acids Res* 32, 328–337.
- Yu, H., Luscombe, N. M., Lu, H. X., Zhu, X., Xia, Y., Han, J. D., et al. (2004b). Annotation transfer between genomes: protein–protein interologs and protein–DNA regulos. *Genome Res* 14, 1107–1118.
- Yu, H., Greenbaum, D., Xin Lu, H., Zhu, X., & Gerstein, M. (2004c). Genomic analysis of essentiality within protein networks. *Trends Genet* 20, 227–231.
- Yu, L. R., Issaq, H. J., & Veenstra, T. D. (2007a). Phosphoproteomics for the discovery of kinases as cancer biomarkers and drug targets. *Proteomics Clin Appl* 1, 1042–1057.
- Yu, H., Kim, P. M., Sprecher, E., Trifonov, V., & Gerstein, M. (2007b). The importance of bottlenecks in protein networks: correlation with gene essentiality and expression dynamics. *PLoS Comput Biol* 3, e59.
- Yu, H., Chen, J., Xu, X., Li, Y., Zhao, H., Fang, Y., et al. (2012). A systematic prediction of multiple drug–target interactions from chemical, genomic, and pharmacological data. *PLoS One* 7, e37608.
- Zamir, E., & Bastiaens, P. I. (2008). Reverse engineering intracellular biochemical networks. *Nat Chem Biol* 4, 643–647.
- Zampetaki, A., Kiechl, S., Drozdov, I., Willeit, P., Mayr, U., Prokopi, M., et al. (2010). Plasma microRNA profiling reveals loss of endothelial miR-126 and other microRNAs in type 2 diabetes. *Circ Res* 107, 810–817.
- Zanzoni, A., Soler-Lopez, M., & Aloy, P. (2009). A network medicine approach to human disease. *FEBS Lett* 583, 1759–1765.
- Závodszy, P., Abatur, L. V., & Varshavsky, Y. M. (1966). Structure of glyceraldehyde-3-phosphate dehydrogenase and its alteration by coenzyme binding. *Acta Biochim Biophys Acad Sci Hung* 1, 389–403.
- Zelezniak, A., Pers, T. H., Soares, S., Patti, M. E., & Patil, K. R. (2010). Metabolic network topology reveals transcriptional regulatory signatures of type 2 diabetes. *PLoS Comput Biol* 6, e1000729.
- Zhang, Z. D., & Grigorov, M. G. (2006). Similarity networks of protein binding sites. *Proteins* 62, 470–478.
- Zhang, Y., Bo, X. C., Yang, J., & Wang, S. Q. (2005). HBVPathDB: a database of HBV infection-related molecular interaction network. *World J Gastroenterol* 11, 1690–1692.
- Zhang, J. X., Huang, W. J., Zeng, J. H., Huang, W. H., Wang, Y., Zhao, R., et al. (2007). DITOP: drug-induced toxicity related protein database. *Bioinformatics* 23, 1710–1712.
- Zhang, S., Zhang, X. S., & Chen, L. (2008). Biomolecular network querying: a promising approach in systems biology. *BMC Syst Biol* 2, 5.
- Zhang, B., Li, H., Riggins, R. B., Zhan, M., Xuan, J., Zhang, Z., et al. (2009). Differential dependency network analysis to identify condition-specific topological changes in biological networks. *Bioinformatics* 25, 526–532.
- Zhang, J., & Huan, J. (2010). Analysis of network topological features for identifying potential drug targets. *Proc 9th Intl Workshop Data Mining Bioinformatics (BioKDD'10)*, Washington D. C., July 2010.
- Zhang, M., Zhu, C., Jacomy, A., Lu, L. J., & Jegga, A. G. (2011a). The orphan disease networks. *Am J Hum Genet* 88, 755–766.
- Zhang, J., Lushington, G. H., & Huan, J. (2011b). The BioAssay network and its implications to future therapeutic discovery. *BMC Bioinforma* 12, S1.
- Zhang, M., Su, S., Bhatnagar, R. K., Hassett, D. J., & Lu, L. J. (2012). Prediction and analysis of the protein interactome in *Pseudomonas aeruginosa* to enable network-based drug target selection. *PLoS One* 7, e41202.
- Zhao, S., & Iyengar, R. (2012). Systems pharmacology: network analysis to identify multiscale mechanisms of drug action. *Annu Rev Pharmacol Toxicol* 52, 505–521.
- Zhao, S., & Li, S. (2010). Network-based relating pharmacological and genomic spaces for drug target identification. *PLoS One* 5, e11764.
- Zhao, J., Yu, H., Luo, J. H., Cao, Z. W., & Li, Y. X. (2006). Hierarchical modularity of nested bow-ties in metabolic networks. *BMC Bioinforma* 7, 386.
- Zhao, X. M., Iskar, M., Zeller, G., Kuhn, M., van Noort, V., & Bork, P. (2011a). Prediction of drug combinations by integrating molecular and pharmacological data. *PLoS Comput Biol* 7, e1002323.
- Zhao, J., Yang, T. H., Huang, Y., & Holme, P. (2011b). Ranking candidate disease genes from gene expression and protein interaction: a Katz-centrality based approach. *PLoS One* 6, e24306.
- Zheng, W., Brooks, B. R., & Thirumalai, D. (2007). Allosteric transitions in the chaperonin GroEL are captured by a dominant normal mode that is most robust to sequence variations. *Biophys J* 93, 2289–2299.
- Zheng, W., Schafer, N. P., Davtyan, A., Papoian, G. A., & Wolynes, P. G. (2012a). Predictive energy landscapes for protein–protein association. *Proc Natl Acad Sci USA* 109, 19244–19249.
- Zheng, C. S., Xu, X. J., & Ye, H. Z. (2012b). [Computational simulation of multi-target research on the material basis of *Caulis sinomenii* in treating osteoarthritis] (in Chinese). *Zhongguo Zhong Xi Yi Jie He Za Zhi* 32, 375–379.
- Zhong, Q., Simonis, N., Li, Q. R., Charlotiaux, B., Heuze, F., Klitgord, N., et al. (2009). Edgetic perturbation models of human inherited disorders. *Mol Syst Biol* 5, 321.
- Zhou, T., Lü, L., & Zhang, Y. -C. (2009). Predicting missing links via local information. *Eur Phys J* 71, 623–630.
- Zhu, X., Gerstein, M., & Snyder, M. (2007). Getting connected: analysis and principles of biological networks. *Genes Dev* 21, 1010–1024.
- Zhu, J., Zhang, B., & Schadt, E. E. (2008). A systems biology approach to drug discovery. *Adv Genet* 60, 603–635.
- Zhu, M., Gao, L., Li, X., Liu, Z., Xu, C., Yan, Y., et al. (2009). The analysis of the drug-targets based on the topological properties in the human protein–protein interaction network. *J Drug Target* 17, 524–532.
- Zhu, Y. -X., Lü, L., Zhang, Q. -M., & Zhou, T. (2012a). Uncovering missing links with cold ends. *Phys A* 391, 5769–5778.
- Zhu, F., Shi, Z., Qin, C., Tao, L., Liu, X., Xu, F., et al. (2012b). Therapeutic target database update 2012: a resource for facilitating target-oriented drug discovery. *Nucleic Acids Res* 40, D1128–D1136.
- Zhuravlev, P. I., & Papoian, G. A. (2010). Protein functional landscapes, dynamics, allostery: a tortuous path towards a universal theoretical framework. *Q Rev Biophys* 43, 295–332.
- Zimmermann, G. R., Lehar, J., & Keith, C. T. (2007). Multi-target therapeutics: when the whole is greater than the sum of the parts. *Drug Discov Today* 12, 34–42.
- Zlatic, V., Ghoshal, G., & Caldarelli, G. (2009). Hypergraph topological quantities for tagged social networks. *Phys Rev E* 80, 036118.
- Zoncu, R., Efeyan, A., & Sabatini, D. M. (2011). mTOR: from growth signal integration to cancer, diabetes and ageing. *Nat Rev Mol Cell Biol* 12, 21–35.
- Zur, H., Ruppini, E., & Shlomi, T. (2010). iMAT: an integrative metabolic analysis tool. *Bioinformatics* 26, 3140–3142.