Perspectives in Science (2016) 9, 49—52

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/pisc

Computational approaches for the study of the role of small molecules in diseases☆

Monica Campillos[a,b,*]

[a] *Institute of Bioinformatics and Systems Biology, Helmholtz Zentrum München, 85764 Neuherberg, Germany*
[b] *German Center for Diabetes Research, Helmholtz Zentrum München, 85764 Neuherberg, Germany*

Received 21 March 2016; accepted 16 May 2016
Available online 26 August 2016

KEYWORDS
Drug discovery;
Drug target prediction;
Drug—disease connections;
Computational approaches;
Systems biology;
Bioinformatics

Summary An enormous amount of molecular and phenotypic information of drugs as well as diseases is now available in public repositories. Computational analysis of these datasets is facilitating the acquisition of a systems view of how drugs act on our human organism and interfere with diseases. Here, I highlight recent approaches integrating large-scale information of drugs and diseases that are contributing to change our current view on how drugs interfere with human diseases.

Introduction

Small molecules are the substances most often used as therapeutic agents. However, despite the huge investment of pharmaceutical companies in the development of new drugs, only few novel compounds are approved annually for medical treatment (Emanuel, 2015). The high drug attrition rate is due to a lack of efficacy and unexpected toxicity of drugs (Waring et al., 2015), indicating that our understanding on how compounds affect human biological circuits and interfere with diseases is far from complete.

The recent explosion of biological information of drugs in the public domain is facilitating the study of drug action on the human organism in an unprecedented scale. Over the last two decades, several databases storing molecular and phenotypic information of drugs have appeared on the public domain. Examples of drug target databases are DrugBank (Wishart et al., 2006), ChEMBL (Gaulton et al., 2011) and Matador (Gunther et al., 2008). Repositories of *in vivo* and *in vivo* phenotypic effects of drugs include SIDER (Kuhn et al., 2010, 2016), a database of side effects of marketed drugs, warehouses of high-throughput chemical

* Correspondence to: Institute of Bioinformatics and Systems Biology, Helmholtz Zentrum München, 85764 Neuherberg, Germany.
E-mail address: monica.campillos@helmholtz-muenchen.de

http://dx.doi.org/10.1016/j.pisc.2016.08.001

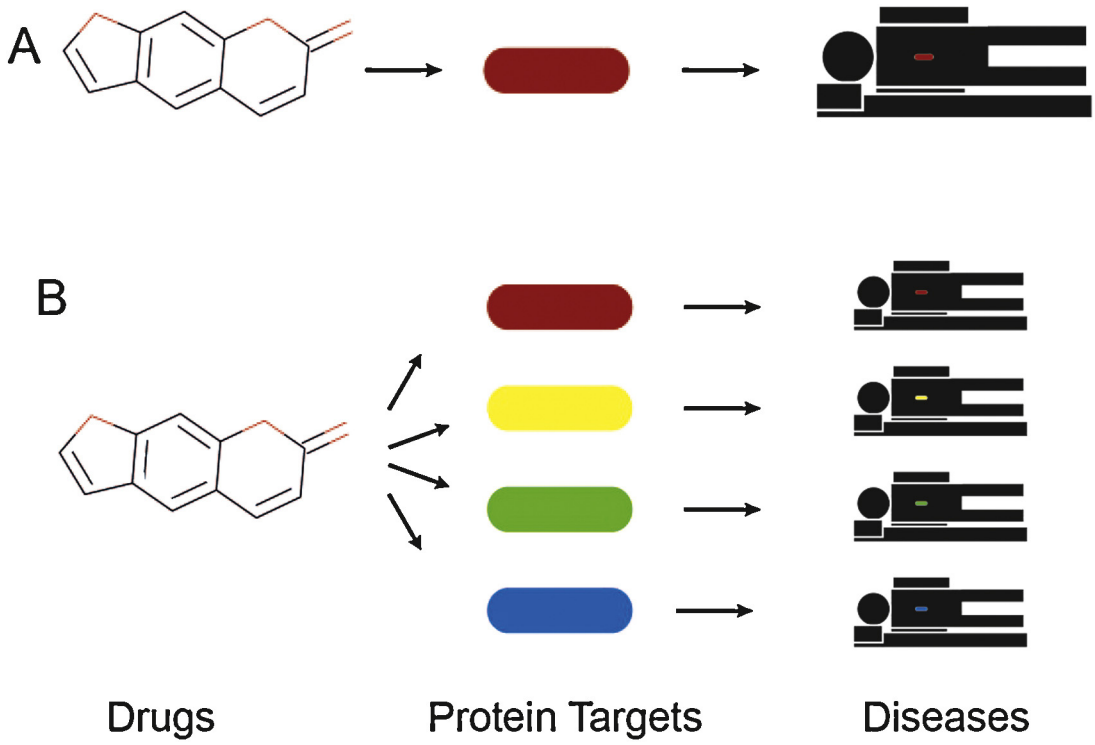

Figure 1 The classical pharmacology view of ''one drug, one target, one disease'' (A) is changing to a more complex scenario of ''many drugs, many targets and many diseases'' (B).

genetics experiments such as ChemBank (Seiler et al., 2008) and PubChem Biassay (Wang et al., 2009) and repositories of gene expression profiles after drug perturbation in cancer cell lines (Lamb et al., 2006).

Resources containing large-scale information of diseases have existed since more than three decades. The first database collecting clinical as well as molecular information of inherited diseases was the 'Online Mendelian Inheritance in Man (OMIM)' (http://omim.org/). More recently, dedicated databases storing genome-wide association disease studies (GWAS), such as the NHGRI GWAS Catalog (Welter et al., 2014) and molecular information of diseases (Pinero et al., 2015) as well as resources offering clinical phenotypes of more than 5000 common and rare diseases (Kohler et al., 2014; Vogt et al., 2014a,b) such as Orphanet and Decipher (Firth et al., 2009) have been released in the public domain.

The integrative analysis of chemical and disease information is changing our view on drug mechanisms of action as well as how drugs interfere with disease mechanisms. The analysis of large-scale drug target information soon evidenced the polypharmacological activity of drugs, that is, the property of drugs to interfere with many protein targets (Anighoro et al., 2014; Jalencas and Mestres, 2013; Peters, 2013). The classical view of ''one drug, one target, one disease'' (Imming et al., 2006) is evolving to a more complex scenario of ''many drugs, many targets and many diseases'' (Mestres et al., 2008; Yildirim et al., 2007) (Fig. 1). Here, I will highlight recent computational efforts that have contributed to enhance our knowledge of drug modes of action and disease relationships.

Results

Elucidation of drug targets

Due to the medical and biological relevance of the discovery of novel drug targets, uncovering new targets of drugs has been an active area on drug discovery research in the last years. Diverse chemo and bio-informatics approaches have been developed to predict drug targets. Chemo-informatics approaches exploit similarities on two and three dimensional structural features of compounds to assign novel targets to compounds (Keiser et al., 2007; Liu et al., 2013; Paolini et al., 2006; Xia et al., 2004), while bioinformatics approaches rely on the analysis of biological properties of drugs. These properties include side effects (Campillos et al., 2008), gene expression profiles after drug perturbation (Lamb et al., 2006; Xia et al., 2004), cytotoxicity profiles of chemicals across a panel of cancer cell lines (Shoemaker, 2006) and bioactivity profiles of chemicals on chemical genetics screens (Petrone et al., 2012).

Biological and chemical properties of compounds have also been exploited in combination to uncover molecular information of compounds, for example in docking approaches where the interaction between compounds and proteins are modeled based on the compound and protein structures (Laird and Blake, 2004) and machine-learning methods that incorporate chemical structure and protein target information (Li et al., 2015).