

# A survey of current trends in computational drug repositioning

Jiao Li, Si Zheng, Bin Chen, Atul J. Butte, S. Joshua Swamidass and Zhiyong Lu

Corresponding author. Jiao Li, Institute of Medical Information, Chinese Academy of Medical Sciences, 3rd Yabao Road, Chaoyang District, Beijing 100020, China, Tel.: (86)10-52328740; E-mail: li.jiao@imicams.ac.cn

## Abstract

Computational drug repositioning or repurposing is a promising and efficient tool for discovering new uses from existing drugs and holds the great potential for precision medicine in the age of big data. The explosive growth of large-scale genomic and phenotypic data, as well as data of small molecular compounds with granted regulatory approval, is enabling new developments for computational repositioning. To achieve the shortest path toward new drug indications, advanced data processing and analysis strategies are critical for making sense of these heterogeneous molecular measurements. In this review, we show recent advancements in the critical areas of computational drug repositioning from multiple aspects. First, we summarize available data sources and the corresponding computational repositioning strategies. Second, we characterize the commonly used computational techniques. Third, we discuss validation strategies for repositioning studies, including both computational and experimental methods. Finally, we highlight potential opportunities and use-cases, including a few target areas such as cancers. We conclude with a brief discussion of the remaining challenges in computational drug repositioning.

**Key words:** computational drug repositioning; integrative strategies; genome; phenome; chemical structure; drug combination; prediction validation

## Introduction

Over the past decades, *de novo* drug discovery has grown to be time-consuming and costly, despite the advances in genomics, life sciences and technology. Investments in pharmaceutical R&D have steadily increased, while the number of new drug approvals has stagnated [1]. Indeed, failures are spread throughout the drug development pipeline, and it takes billions of

investment dollars and an average of about 9–12 years to bring a new drug to the market [2]. Improving R&D productivity remains the most important priority for pharmaceutical industry [3]. In light of these challenges, drug repositioning, which concerns the detection and development of new clinical indications for those existing drugs, or for those that are in the development pipeline, has emerged as an increasingly important

**Jiao Li**, PhD, is a division chief of the Institute of Medical Information, Chinese Academy of Medical Sciences. Her research interests include data modeling for drug discovery and translational bioinformatics.

**Si Zheng** is a faculty member of the Institute of Medical Information, Chinese Academy of Medical Sciences. Her research focuses on complex disease molecular network modeling and mining.

**Bin Chen**, PhD, is a postdoctoral scholar at Stanford School of Medicine. He focuses on developing methods to leverage various data in translational drug discovery to help discover new therapeutics faster.

**Atul Butte**, MD, PhD, is an associate professor of Pediatrics and Genetics, and by courtesy, Medicine, Pathology and Computer Science, and is Chief of the Division of Systems Medicine at Stanford University and Lucile Packard Children's Hospital. He has authored nearly 200 publications.

**S. Joshua Swamidass**, MD, PhD. Both a physician and a computational scientist, Dr Swamidass is an assistant professor of Laboratory and Genomic Medicine in the Pathology and Immunology Department of Washington University in St. Louis.

**Zhiyong Lu**, PhD, is Earl Stadtman investigator at the National Center for Biotechnology Information (NCBI), National Institutes of Health (NIH) where he leads the biomedical text mining research group.

**Submitted:** 9 December 2014; **Received (in revised form):** 4 March 2015

Published by Oxford University Press 2015. This work is written by US Government employees and is in the public domain in the US.

strategy for the new drug discovery. It could substantially reduce the risks of development and the costs, and shorten the lag between drug discovery and availability [4]. Among the 84 drug products introduced to market in 2013, new indications of existing drugs accounted for 20% [5]. Drug repositioning has played a key role in drug discovery and precision medicine paradigm [6].

In recent years, drug repositioning is becoming strongly supported by governments, non-trading organizations and academic institutions. For example, both the United States (National Center for Advancing Translational Sciences) and the United Kingdom (Medical Research Council) have launched large-scale funding programs in this area with a goal to extend molecules that already have undergone significant research and development by the pharmaceutical industry to more new indications. Furthermore, the US Food and Drug Administration (FDA) is also enabling drug repositioning, with the creation of several public databases specifically for computational drug repositioning. There are also substantial economic incentives to reposition marketed drugs for the treatment of orphan and rare disorders [7]. All of these efforts significantly promoted drug repositioning research.

Historically, the discovery of new uses of old drugs is mostly through serendipity [8] or resulted from a better understanding of the drugs' mechanism of action. For example, the monoclonal antibody bevacizumab, originally developed to treat patients with metastatic colon cancer and non-small cell lung cancer by inhibiting angiogenesis, is now being used to slow or reverse abnormal vascularization of the retina in exudative (wet) macular degeneration [9]. With the accumulation of the large volumes of omics data, bioinformatics plays an increasingly important role in the discovery of new drug indications [4]. Depending on where the discovery comes from, these newly proposed computational methods can be categorized as either 'drug based' or 'disease based' [10]. Traditional studies mostly focus on exploring the shared characteristics among drug compounds such as chemical structures [11, 12] and side effects [13]. Other methods include rescreening the existing pharmacopeia against new targets to uncover the novel drug indications [14], looking for similarities of molecular activities [15], or exploring the relationships between drugs and diseases.

With the drug-related data growth and open data initiatives, a set of new repositioning strategies and techniques has emerged with integrating data from various sources, like pharmacological, genetic, chemical or clinical data. These methods can accumulate evidence supporting discovery of new uses or indications of existing drugs. In this review, we summarize recent progress in computational drug repositioning as the following four parts (see Figure 1): repositioning strategies (with available data sets), computational approaches, validation methods and application areas.

## Computational repositioning strategies

### Genome

Rapid advances in genomics have led to the generation of large volumes of genomic and transcriptomic data for a diverse set of disease samples, normal tissue samples, animal models and cell lines. Much of these data are publicly available. Together with other phenotypic, and clinical database, these data sets provide a unique opportunity to understand disease mechanism, elucidate drug mechanism of actions and identify new use of old drugs. Among those, transcriptomic profiles, such as gene

expression data are most widely used, while other genomic and genetic profiles have been explored for drug repositioning as well.

One key source of data behind several repurposing efforts is the Connectivity Map (CMap) [16] project and its extended project Library of Integrated Network-Based Cellular Signatures (LINCS) [17], which produced large-scale gene expression profiles from human cancer cell lines treated with different drug compounds under different conditions. CMap aims to construct a detailed map for functional associations among diseases, genetic perturbations and drug actions. By integrating with other functional genomics databases [e.g. NCBI Gene Expression Omnibus (GEO)] [18], its data have been extensively explored in drug repositioning studies. One approach using these data is to look for inverse drug-disease relationships by comparing drug gene expression profiles and disease gene expression profiles. This approach is also referred as 'signature reversion'. For example, by systematically comparing gene expression signatures of inflammatory bowel disease (IBD) derived from GEO against a set of drug gene expression signatures comprising 164 drug compounds from CMap, Dudley *et al.* [19] inferred several new interesting drug-disease pairs and validated one pair in IBD preclinical models. In another case, Jahchan *et al.* [20] used a similar systematic drug-repositioning bioinformatics approach to query a large compendium of gene expression profiles to identify antidepressant drugs for the treatment of small cell lung cancer.

Another approach, 'guilt-by-association', looks for drugs that provoke similar transcriptional responses, positing that they could share similar mode of action (MoA) [21]. The availability of many public repositories, such as the Drug versus Disease (DvD) [22], the database for Annotation, Visualization and Integrated Discovery (DAVID) [23] and the Gene Set Enrichment Analysis (GSEA) [24], provides such an opportunity for the comparison of drug and disease signatures from gene expression profiles.

Recently, noncoding RNAs, especially the microRNA (miRNAs), have been shown in regulating kinds of cell activities [25, 26], thus becoming promising therapeutic targets for drug repositioning [27]. For example, Liu *et al.* [28] developed an in silico drug repositioning strategy based on miRNA-TF feed-forward loops (FFLs). miRNAs and transcript factors (TFs) were found to be significantly enriched in cystic fibrosis (CF) associated gene regulations from public available data sources. Then they constructed FFLs in CF by defining specific TFs and miRNAs as two regulatory elements. Forty-eight existing drugs showing ability to influence the expression of miRNA that are part of FFLs were repurposed for the treatment of CF patients. Jiang *et al.* [29] predicted new indications for existing drugs by constructing small molecule-miRNA network for each cancer. Rukov *et al.* [30] developed a web server that links miRNA expression and drug function by combining data on miRNA targeting and protein-drug interactions. SM2miR is a database containing manually curated relationships between experimentally validated molecules and miRNA [31].

In addition to transcriptomic data, other genomic profiles (e.g. genetic mutations) can be applied to drug repositioning. For instance, Garnett *et al.* [32] carried out a large-scale screening of human cancer cell lines with 130 clinical/preclinical drugs. A multivariate analysis of genetic and gene expression profiles of cancer cell lines showed that a few mutated cancer genes that are associated with drug sensitivity may serve as potential biomarkers of drug response. To some extent, these mutations reflect the molecular activity of drugs, and can be regarded as drug signatures during the repositioning process. In another

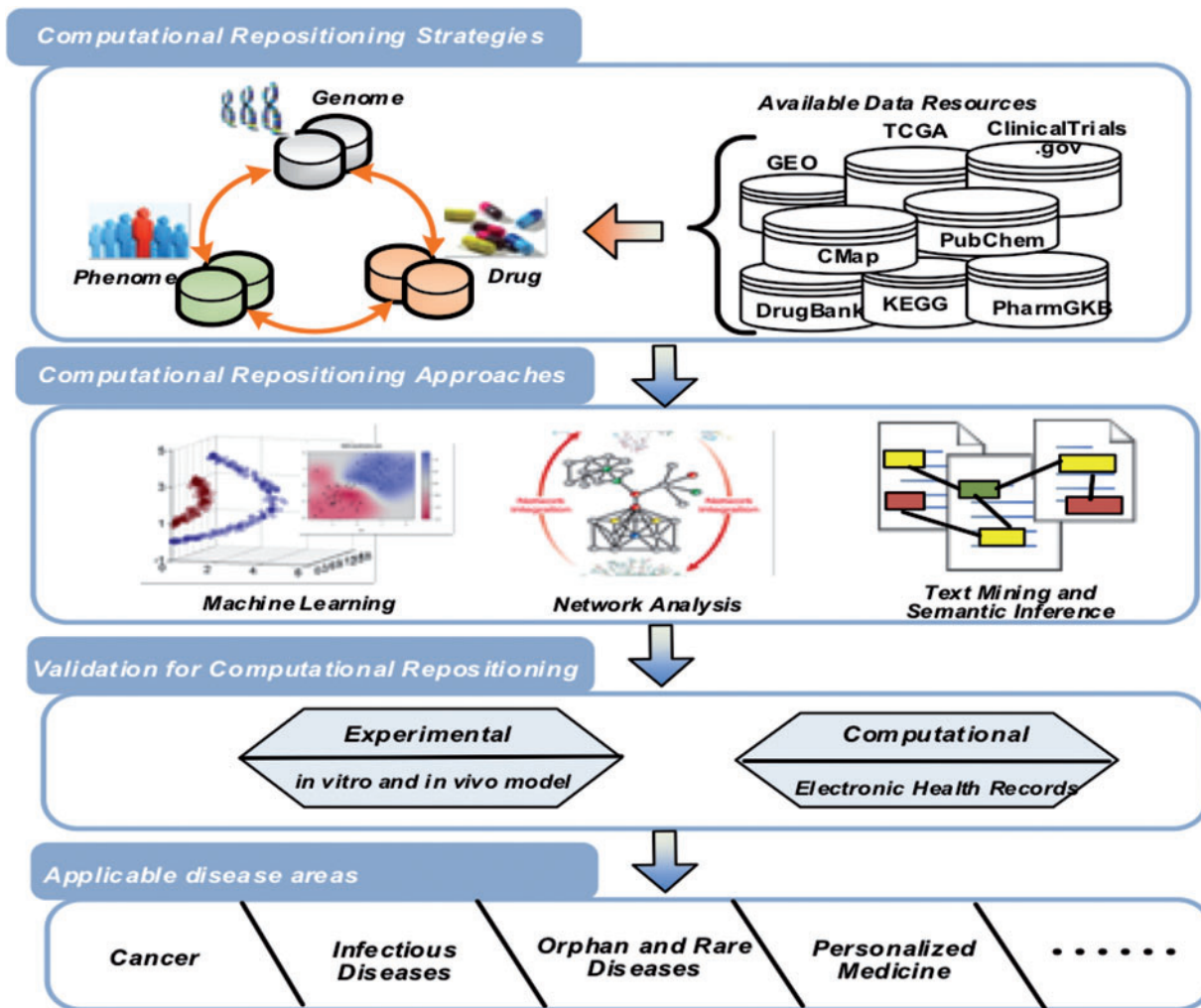


Figure 1. Overview of recent progress in computational drug repositioning.

case, Okada et al. [33] performed a three-stage genome-wide association study (GWAS) meta-analysis of rheumatoid arthritis (RA) patients and linked the risk loci to known RA drug targets. In their study, logistic regression models assuming additive effects of the allele dosages were used to assess the relationship of the single-nucleotide polymorphisms (SNPs) and RA. In total, 101 RA risk loci were identified (e.g. 42 are novel), and they showed significant overlapping with approved RA drug target genes. Furthermore, several drugs approved for other diseases were connected to RA risk genes, indicating they could be repositioned for RA. In another GWAS [34], a catalog of disease-associated genes from published genome-wide associations studies were further integrated with targets of drugs from pharmaceutical projects. In this way, the drugs with targets mapped to the disease-associated genes from GWAS data may be repositioned.

### Phenome

The phenome, defined as the comprehensive collection of phenotypic information, was emerged as a new source for drug repositioning. In recent years, the phenome-wide association study (PheWAS) has become increasingly popular as a systematic approach to identify important genetic associations with human diseases [35]. For instance, Denny et al. [36] performed a

large-scale application of the PheWAS using electronic medical records (EMRs), and demonstrated that PheWAS is a useful tool to enhance the analysis of the genomic basis and to detect novel associations between genetic markers and human diseases.

Meanwhile, clinical side effects are shown to be capable of profiling drug-related human phenotypic information and can subsequently help discover new therapeutic uses. For example, Yang et al. [37] used drug side effects as features to predict its indications. Ye et al. [38] identified novel indications based on the hypothesis that similar side-effect profiles may share similar therapeutic properties. Bisgin et al. [39] developed a Latent Dirichlet Allocation model for drug repositioning that adopted the phenome information from the Side Effect Resource (SIDER) [40]. Using drug side-effect profiles to suggest its novel indications has shown to be attractive but its practical use would require deep understanding of the underlying molecular/pathological mechanisms.

Finally, the phenome can be incorporated with other kinds of data for drug repositioning. As an example, Hoehndorf et al. [41] developed an integrated system to predict novel drug-disease associations by linking genotype-disease associations with drug-gene associations. In this model, beginning with integrating phenotype ontologies for disease and gene or genotype, they derived a semantic similarity-based score to measure genotype-disease associations. With this approach, most of the

known drug–disease associations have been retrieved and the new associations may indicate a new repositioning opportunity. Although some researchers have demonstrated the potential correlations between genome and phenome [42], there is still an urgent need to understand these correlations better and turn them into disease treatment or personalized health care. For example, the BRCA mutation (mutation in either BRCA1 or BRCA2 genes) was found to be associated with the risk of getting breast cancer for ovarian cancer patients [43]. Because BRCA mutations are clinically significant [44], a deeper understanding of the relationship between BRCA mutation status and cancer phenotype will be important for making precise treatment decisions for patients.

## Drug chemical structures

The drug chemical structures can also point toward repositioning opportunities. Moreover, publicly available databases of chemical structures, high-throughput screening data and literature-derived biochemical data containing massive amounts of information useful for repositioning [45–48]. The key insight behind these approaches is that the molecules with similar chemical structures often affect proteins and biological systems in similar ways. Similarity may be measured in many different methods using different structural features, including 2D topological fingerprints or 3D conformations, and is an active area of research. The way to incorporating similarity between chemicals into repositioning inferences is also an active area of research.

For example, Swamidass *et al.* [46] proposed to use chemical structure to infer which targets modulate disease-relevant phenotype. Knowing which targets modulate disease-relevant phenotypes is a signal that can indicate what other drugs might work to treat the disease.

Recently, chemical structure information turned to be integrated with other types of data for computational repositioning. Wang *et al.* [49] proposed an integrated repositioning model that incorporated drug chemical structure, molecular activity and side effect. All three different types of data were integrated to define a kernel function used by a support vector machine (SVM) classifier. This method was further compared with other methods, and showed high efficiency. Similarly, Tan *et al.* [50] incorporated drug chemical structure similarity and gene semantic similarity to construct a drug similarity network, which was further used to extract novel indications. Ng *et al.* [51] proposed a novel algorithm called 'ligENTS' to define novel drug–target associations by mapping the drug to its global pharmacological space according to its chemical information. Then structural systems biology platform was integrated to reposition approved drugs for malaria. A full review of chemical structure-based approaches to repositioning is beyond our scope, but these studies show a trend toward including chemical structures alongside other types of data.

## Drug combinations

As many diseases are driven by complex molecular and environmental interactions, targeting a single component may not be sufficient to disrupt those mechanisms, and interest in early drug discovery stages has increasingly evolved to target multiple molecules using combined drugs or multi-target inhibitors. For example, the activated B-cell-like (ABC) subtype of diffuse large B-cell lymphoma (DLBCL) is a malignant cancer with poor prognosis. Constitutive activation of the NF- $\kappa$ B by I $\kappa$ B kinase

(IKK) has shown to be a key pathogenic factor. Ceribelli *et al.* [52] screened 466 drugs that have been approved or in early stage for cancer therapy and found that ibrutinib, a kinase inhibitor that can block B-cell receptor signaling pathway to activate IKK, shows a significant synergistic effect with JQ1 in killing ABC DLBCL cells both *in vitro* and *in vivo*, suggesting that the combination of JQ1 and ibrutinib might be a new effective therapy.

Current combination strategies rely mainly on clinical and empirical experience, computational prediction is thus highly demanded [53]. Huang *et al.* [54] used side effects as features for drug–drug combinations (DDCs), and further classify the safe DDCs from the unsafe ones. Sun *et al.* [55] developed a model to predict effective drug combinations by integrating gene expression profiles of multiple drugs. Existing drug combinations were extracted from the Drug Combination Database (DCDB) [56]. They started to identify important features related to drug combinations by using statistical methods that look for side effects, genes or disease pathways that could be affected by drugs in the combination. Those features were further used to construct a machine-learning classifier for predicting effective drug combinations.

Other than *in silico* methods, experimental characterizations of drug efficiency (e.g. library screening and cell viability assays) were also adopted to identify new drug combinations. Kang *et al.* [57] identified antileukemic drugs that could be combined with imatinib to overcome drug resistance in BCR-ABL + leukemia. They first used library screening, literature search and correlation analysis to select 11 candidate drugs that might be combined with imatinib. Dose responses for these candidates with/without imatinib were applied in an iterative search algorithm to identify effective combinations that can overcome drug resistance. These predicted combinations were further confirmed in preclinical models.

## Available data sources

Currently, many different kinds of data sources (e.g. genetic, pharmacogenomics, clinical, chemical agent) are available for supporting computational drug repositioning, and to some extent they have promoted the development of various repositioning strategies. In Table 1, we summarize a brief list of those frequently used data sources.

## Computational repositioning approaches

### Machine learning

For the various data resources that support the exploration of repositioning opportunities, machine learning-based models can leverage the data to study the underlying systems for the prediction of novel associations between drugs and diseases. In recent years, an increasing number of machine learning methods have been proposed when coupled with the elucidation of various features.

Menden *et al.* [58] developed machine-learning models to predict the response of cancer cell lines to drug treatment, quantified through IC<sub>50</sub> values. In the model, cancer genomic features of the cell lines (e.g. mutation status of 77 oncogenes and microsatellite status) and chemical properties (e.g. structural fingerprint) were used to build a feed-forward perceptron neural network model and a random forest regression model. The predicted IC<sub>50</sub> values were further validated by a cross-validation and an independent blind test. Napolitano *et al.* [59] focused on a drug-centered approach to predict drug



Table 1. Data resources supporting computational repositioning strategies

Repositioning strategies	Database	URL	Brief description
Genome	ArrayExpress	<a href="http://www.ebi.ac.uk/arrayexpress">http://www.ebi.ac.uk/arrayexpress</a>	Public repositories of functional genomics data.
	Gene Expression Omnibus (GEO)	<a href="http://www.ncbi.nlm.nih.gov/geo">http://www.ncbi.nlm.nih.gov/geo</a>	Gene expression patterns under different biological conditions.
	Gene Expression Atlas	<a href="http://www.ebi.ac.uk/gxa">http://www.ebi.ac.uk/gxa</a>	Genomic data (e.g. Exome, SNP, Methylation, mRNA, miRNA, Clinical) of >10 000 patient tissue samples across > 30 common cancers.
	The Cancer Genome Atlas (TCGA)	<a href="http://cancergenome.nih.gov">http://cancergenome.nih.gov</a>	Genomic data (e.g. DNA copy number, mRNA expression, mutation data) of > 1000 cancer cell lines.
	Cancer Cell Line Encyclopedia (CCLE)	<a href="http://www.broadinstitute.org/ccle">http://www.broadinstitute.org/ccle</a>	A comprehensive description of genomic, transcriptomic and epigenomic changes.
	International Cancer Genome Consortium	<a href="https://icgc.org">https://icgc.org</a>	Gene expression profiles of > 1000 drugs across three primary cell lines.
	The Connectivity Map (CMap)	<a href="http://www.broadinstitute.org/cmap">http://www.broadinstitute.org/cmap</a>	Aim to produce > 1 million gene expression profiles of drugs and genetic perturbagens across > 15 cell lines.
	Library of Integrated Network-based Cellular Signatures (LINCS)	<a href="http://www.lincsproject.org">http://www.lincsproject.org</a>	Collections of annotated gene signatures from different sources.
	Molecular Signature Database (MsigDB)	<a href="http://www.broadinstitute.org/gsea/msigdb">http://www.broadinstitute.org/gsea/msigdb</a>	Functional annotation tools.
	Gene Signature Database (GeneSigDB)	<a href="http://compbio.dfci.harvard.edu/genesigdb">http://compbio.dfci.harvard.edu/genesigdb</a>	Resource for understanding high-level functions and utilities of the biological system from molecular-level information.
Genome/ Phenome	Database for Annotation, Visualization and Integrated Discovery (DAVID)	<a href="http://david.abcc.ncifcrf.gov">http://david.abcc.ncifcrf.gov</a>	Tool to determine if an a priori defined gene signature shows statistically significant, concordant differences between two biological states.
	Kyoto Encyclopedia of Genes and Genomes (KEGG)	<a href="http://www.genome.jp/kegg">http://www.genome.jp/kegg</a>	Computational pipeline for comparing disease and drug-response gene expression signatures from publicly available resources.
	Gene Set Enrichment Analysis (GSEA)	<a href="http://www.broadinstitute.org/gsea">http://www.broadinstitute.org/gsea</a>	Genetic variation on drug response.
	Drug versus Disease (DvD)	<a href="http://www.ebi.ac.uk/saezrodriguez/dvd">www.ebi.ac.uk/saezrodriguez/dvd</a>	Relationship between genes and genetic phenotypes, particularly disorders.
	The Pharmacogenetics and Pharmacogenomics Knowledge Base (PharmGKB)	<a href="http://www.pharmgkb.org">http://www.pharmgkb.org</a>	Adverse drug reactions of > 900 drugs.
	Online Mendelian Inheritance in Man (OMIM)	<a href="http://www.omim.org">http://www.omim.org</a>	A registry and result database of publicly and privately supported clinical studies.
	Side Effect Resource (SIDER)	<a href="http://sideeffects.embl.de">http://sideeffects.embl.de</a>	Information about FDA approved drugs, such as brand name, therapeutic products.
	ClinicalTrials.gov	<a href="http://www.clinicaltrials.gov">http://www.clinicaltrials.gov</a>	Known examples of drug combinations, models for drug combinations.
	Drugs@FDA Database	<a href="http://www.fda.gov/Drugs/InformationOnDrugs/ucm135821.htm">http://www.fda.gov/Drugs/InformationOnDrugs/ucm135821.htm</a>	A comprehensive, publicly accessible collection of approved and investigational drugs.
	Drug Combination Database (DCDB)	<a href="http://www.cls.zju.edu.cn/dcdb">http://www.cls.zju.edu.cn/dcdb</a>	3D structure of proteins, nucleic acids.
Phenome/ Drug	The NCGC Pharmaceutical Collection (NPC)	<a href="http://tripod.nih.gov/npc/">http://tripod.nih.gov/npc/</a>	A database containing chemical structures representing approved drugs, chemical isolates from traditional medicinal herbs and regulated chemicals.
	Protein Data Bank (PDB)	<a href="http://www.rcsb.org/pdb/home/home.do">http://www.rcsb.org/pdb/home/home.do</a>	Detailed drug (i.e. chemical, pharmacological and pharmaceutical) data with comprehensive drug target (i.e. sequence, structure, and pathway) information.
	SWEETLEAD	<a href="https://simtk.org/home/sweetlead">https://simtk.org/home/sweetlead</a>	A large literature-derived database of molecule structures and molecule-protein interactions. This includes a catalog of approved drugs.
	DrugBank	<a href="http://www.drugbank.ca/">http://www.drugbank.ca/</a>	A repository of biological assay results for hundreds of thousands of molecules.
	ChEMBL	<a href="https://www.ebi.ac.uk/chembl/">https://www.ebi.ac.uk/chembl/</a>	
	PubChem	<a href="https://pubchem.ncbi.nlm.nih.gov/">https://pubchem.ncbi.nlm.nih.gov/</a>	

therapeutic class by using drug-related features (e.g. drug chemical structure similarity, drug molecular target similarity and drug gene expression similarity). They merged these features into a single drug similarity matrix, which was used as a kernel for SVM classification. Other than drug-related features, Gottlieb et al. [60] integrated various disease-related features (e.g. phenotypic and genetic features). Based on these features, drug–drug and disease–disease similarity measures were computed to construct classification features. They further used a logistic regression classifier to predict novel drug indications.

Moreover, there exist machine learning algorithms that use collaborative filtering techniques to predict unknown drug–disease associations. For instance, Zhang et al. [61] proposed a unified computational framework for integrating multiple aspects of drug similarity and disease similarity. Briefly, genome (e.g. drug target protein, disease gene), phenotype (e.g. disease phenotype, drug side effect) and chemical structure (e.g. drug chemical structure) were integrated to extract drug similarity matrix and disease similarity matrix. Based on all this information, the authors turned the drug–disease network analysis into an optimization problem. This computational framework shows high efficiency in exploring novel drug indications. Yang et al. [62] used a causal inference–probabilistic matrix factorization approach to infer drug–disease associations. In this model, they integrated multilevel relations to construct causal networks connecting drug–target–pathway–gene–disease, PMF modes were learned based on known interactions. This approach can predict novel drug–disease associations, thus providing potential values for drug repositioning.

## Network analysis

Network-based analysis is another widely used strategy for computational drug repositioning. With the advances of high-throughput technology and bioinformatics methods, molecular interactions in the biological systems can be modeled by networks. Previous studies have suggested that drug–target network, drug–drug network, drug–disease network, protein interaction network, transcriptional networks and signaling networks are useful in the identification of therapeutic targets or characteristics of drug targets [63–65], thus providing new opportunities for drug discovery or repositioning.

Li et al. [66] developed a bipartite drug–target network method to identify potential new indications of an existing drug through its relation to similar drugs. In the bipartite network model, drug pair similarity integrated drug chemical structure similarity, common drug targets and their interactions. Built on their past success, the same authors recently built a causal network (CauseNet) [67]—a multilayered pathway of gene, disease and drug target—to identify new therapeutic uses of existing drugs. In the causal network, the transition likelihood of each chain is estimated on the basis of known drug–disease treatment association. Wu et al. [68] applied network clustering to a drug–disease heterogeneous network to identify closely connected modules of disease and drugs, which can be used for extracting possible drug–disease pairs for drug repositioning. In the network, two nodes (one drug or one disease) with shared genes/targets and enriched features (biological process, pathway and phenotypes) were connected and the connection was weighted by a Jaccard score. Jin et al. [69] developed a novel method to repurpose drugs for cancer therapeutics by leveraging off-target effects (OTEs) that may affect important cancer cell signaling pathways. The OTEs of drugs on signaling proteins were recognized by a hybrid model composed by a network

component called cancer-signaling bridges and a regression model called Bayesian Factor Regression Model.

## Text mining and semantic inference

The biomedical and pharmaceutical knowledge available in literature [70, 71] or databases contains vast amount of information for drugs and diseases, which can be automatically mined [72–74] and retrieved [75], thanks to the recent advances in text mining research [76]. Therefore, it is possible to detect novel indications for existing drugs by finding relevant knowledge through text mining approaches [77]. One important basis rooted in this method is the occurrence of biological ontology, which makes it possible for the comparison and analysis of biological information from different sources.

A recent study by Andronis et al. [78] summarized several literature mining approaches and sources for drug repurposing. For example, if one study finds that disease A was caused by the lack of nutrition B while another study reports that drug C used for another disease was an activator of nutrition B, then drug C might be repurposed for disease A through literature mining.

Moreover, semantic technologies have facilitated the integration of various data sources and the discovery of new drug indications. For instance, Zhu et al. [79] developed an ontology to model FDA-approved breast cancer drugs and their relations with pathways, drugs, genes, SNPs and diseases. New drug–disease pairs were inferred from the ontology-based knowledge base. Chen et al. [80] developed a statistical model to assess drug–target associations from a semantic linked network comprised by drugs, chemical compounds, protein targets, diseases, side effects and pathways and their relations. The model considered the topology and semantics of the subgraph between a drug and a target. The similar drug–drug pair from different disease areas may indicate a potential repositioning opportunity.

## Validation for computational drug repositioning

Computational models often predict a handful interesting hits; however, the ultimate goal of drug repositioning is to move only one or two hits into clinical application to benefit patients. Therefore, experimental validation of these computer-generated hits becomes important. Despite some known limitations, *in vitro* and *in vivo* models (e.g. cell-based targeted assays and mouse models) have been increasingly used to validate the candidate hits for preclinical drug evaluation. For example, Zerbini et al. [81] identified several FDA-approved drugs showed sensitivity for clear cell renal cell carcinoma (ccRCC). To confirm the drug efficacy, they further used apoptosis assays and xenograft mouse models and demonstrated pentamidine as a potential therapeutic agent for ccRCC, as it can significantly induce apoptosis in tumor cells and slow tumor growth. Kang et al. [57] used cell viability assays to validate the drug combinations generated from a computational search algorithm. Subsequently, drug combinations with significant efficacy in cancer cell killing were confirmed. Végner et al. [82] performed an experimental validation study for a previous reported computational drug repositioning strategy where they confirmed the efficacy of top-predicted drugs through ACE inhibition assay, COX inhibition assay, dopaminergic agonist and antagonist assay.

In addition to *in vitro* and *in vivo* models, electronic health/medical records may be helpful to validate the predictions. For example, Khatri et al. [83] identified atorvastatin as a new therapeutic for organ transplantation using meta-analysis of public microarray data sets and validated the beneficial effect of

Table 2. Current and future target areas for computational drug repositioning

Target area	Key findings	Main data sources	Main methods
Cancer	Repositioning of auranofin for the treatment of gastrointestinal stromal tumors [88].	The FDA-approved drug library	High-throughput screening
	Repositioning of irinotecan for the treatment of breast cancer [89].	Genome (expression, structure, activity), clinicaltrials.gov	Biomarker-guided repurposing
	Repositioning of the anthelmintic drug mebendazole for the treatment for colon cancer [90].	The Pharmacok 1600 library, NCI 60 data, CMap	In silico drug screening
	Repositioning of metformin for various types of cancer including breast, colorectal, endometrial, esophageal, etc [91].	Clinical trials for cancer treatment with metformin	Clinical studies
Infectious disease	Repositioning of an existing phenothiazine-like antipsychotic drug, trifluoperazine, as a potential anti-CSC (cancer stem cell) agent [92].	CMap	In silico drug screening, in vitro and in vivo models
Rare and orphan diseases	Identified novel <i>Plasmodium falciparum</i> targets of drug-like active compounds [51].	DrugBank	Drug chemical structure analysis, network analysis
	Drug repositioning for orphan genetic diseases [93], such as proposing HDAC1 and TSPO as two significant targets for epileptic syndromes.	Gene Expression Omnibus, OMIM phenotypes, DrugBank, CMap, FDA	Systematic analysis of gene coexpression, text mining
Personalized medicine	Establishment of a new resource, the Rare Disease Repurposing Database (RDRD) [94].	FDA orphan designation database, FDA-approved drugs	Database analysis
	ISM strategy for personalized AML therapy defined dasatinib and sunitinib could be repositioned for subsets of AML [95].	Drug-response profiles, clinical responses information	Ex vivo drug sensitivity and resistance testing (DSRT)
	Generation of a novel method, PREDICT, for inferring novel drug indications with application to personalized medicine [60].	DrugBank, OMIM, ArrayExpress, DCDB, KEGG DRUG databases, SIDER, UniProt, clinicaltrials.gov	Text mining, machine learning, network analysis
	Current opportunities and existing challenges of drug repositioning for personalized treatment [96].	Genome (expression, structure), clinical trials, Gene Expression Omnibus, DrugBank, etc.	Experimental approaches (cell-based screening, in vivo mouse studies), computational approaches (network analysis)

atorvastatin on graft survival by retrospective analysis of EMRs of a single-center cohort of 2515 renal transplant patients followed for up to 22 years. Xu *et al.* [84] used electronic health records to validate metformin, a drug for type 2 diabetes mellitus, associated with cancer mortality.

Understanding and selecting the appropriate validation model are critical for the success of the prediction, as the contexts of the validation models may be different from those used to make the prediction or some validation models are not reliable per se. For example, it is debatable that genomic responses in mouse models mimic human inflammatory diseases [85, 86]. Half of the hepatocellular carcinoma (HCC) cell lines are not significantly correlated to the HCC tumors from TCGA, while a few rarely used ovarian cancer cell lines more closely resemble ovarian tumor profiles than common used cell lines [87]. In addition to the selection of the right model, the selection of the right hits for validation is also critical. Some drugs may not be favored by physicians or biologists owing to various reasons such as high toxicity, high cost and low bioavailability. The early involvement of all stakeholders would likely be to increase the rate of success in translational research.

### Current and future target areas: cancers, infectious and orphan diseases and personalized medicine

There are several high potential applications of drug repositioning in disease or related therapeutic areas. One important application is the anticancer drug discovery. Motivated by the benefit of drug repositioning and the high demand for anticancer drugs, searching novel anticancer therapies from existing drugs has become increasingly popular. Drug repositioning also turns out to be a promising strategy for the discovery of anti-infectious drugs that can overcome drug resistance, as the emergence of drug resistance is awful to human beings and can largely reduce the drug efficacy. Alternative therapeutics for orphan and rare diseases can also be identified from approved drugs, and this strategy has gained various grants supporting research in this field. Finally, drug repositioning provides a new accessibility to personalized treatment. Table 2 shows examples of drug repositioning applied in cancer, infectious, orphan diseases and personalized medicine.

### Discussion

Each of the aforementioned computational repositioning strategies and approaches has their methodological advantages and limitations. A combination of these methods is often desired for achieving better results. For example, Wang *et al.* [49] integrated drug chemical structures, target protein sequences and phenotype data and applied machine learning approach (i.e. SVM) to identify drug-disease relationships, and further made the drug-disease network analysis. Gottlieb *et al.* [60] integrated chemical structures, drug side effects, drug target protein sequences and target protein interactions and phenotype data and applied network analysis for target protein distance calculation, applied text mining to identify disease phenotypes and used machine learning algorithms to classify true and false drug-disease associations. These integrative methods showed better performance in both sensitivity and specificity when comparing with individual methods.

Despite several successful use cases of computational drug repositioning, challenges remain. First, the transformation of theoretical computational models into practical use is far from straightforward, due to some inevitable factors like missing data, data bias and technical limitations of computational methods. For example, many *in silico* repositioning approaches search

potential drug-target interactions through chemical structure information. Thus, the lack of high-resolution structural data for targets makes such methods inadequate. Large scale of drug-induced transcriptomic profiles (e.g. gene expression data in CMap) might have experimental variation across batches. Additional gene expression normalization techniques are needed to ensure the systematical genome-wide analysis in an unbiased way [97]. Second, the lack of structured gold standard for drug repositioning made it hard to compare and evaluate the performance of computational methods. In response, several recent studies focus on curating a comprehensive and public catalog of existing drug indications [98, 99]. Although common metrics (e.g. sensitivity, specificity, and area under ROC curve (AUC)) are adopted, previous studies performed evaluation on their own data sets rather than a shared gold standard for various reasons. Third, although computational drug repositioning may merely shorten the process of drug discovery in preclinical and Phase I trials, challenges may still exist after Phase II trials for the repositioned drugs. For instance, as the test population becomes larger for Phase II trials, more resources will be needed for the further drug validation [100]. Recently, some researchers have proposed the idea of repositioning the approved drugs for the treatment of Ebola, and the debate has been surprisingly intense [101]. Antiviral drugs or drugs with immune system's modulation drugs were repositioned to treat Ebola, but World Health Organization has ignored these proposed FDA-approved drugs, owing to the deficiency of experimental tests, as well as the potential drug toxicity.

### Conclusions

In summary, we believe computational drug repositioning research is of great significance to improve human health through discovering new uses for existing drugs. In fact, a number of studies have already been carried out with various degrees of success. It has great potentials to accelerate drug discovery with interesting opportunities in several particular disease areas (e.g. cancer). As both public and private sectors from around the world are supporting drug-repositioning research with various funding opportunities and special programs, now is the time for the research community to further develop techniques and methods toward new discoveries and breakthroughs.

#### Key Points

- Various types of big data (e.g. genomic, phenotypic, clinical data, chemical structure) are publicly available for computational drug repositioning research.
- Integrative repositioning strategies for heterogeneous data can help find new uses of old drugs.
- Computational drug repositioning shows promises in accelerating drug discoveries for cancer and infectious and orphan diseases.

### Acknowledgment

The authors would like to thank Dr Pankaj Agarwal for his helpful discussion.

### Funding

NIH Intramural Research Program, National Library of Medicine (to Z.L.), the National Key Technology Research



and Development Program of China (Grant No. 2013BAI06B01), the National Population and Health Scientific Data Sharing Program of China, the Fundamental Research Funds for the Central Universities (Grant No. 13R0101).

## References

- Booth B, Zimmel R. Prospects for productivity. *Nat Rev Drug Discov* 2004;**3**(5):451–6.
- Dickson M, Gagnon JP. The cost of new drug discovery and development. *Discov Med* 2004;**4**(22):172–9.
- Paul SM, Mytelka DS, Dunwiddie CT, et al. How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nat Rev Drug Discov* 2010;**9**(3):203–14.
- Shaughnessy AF. Old drugs, new tricks. *BMJ* 2011;**342**:d741.
- Graul AI, Cruces E, Stringer M. The year's new drugs & biologics, 2013: Part I. *Drugs Today (Barc)* 2014;**50**(1):51–100.
- Shameer K, Readhead B, Dudley JT. Computational and experimental advances in drug repositioning for accelerated therapeutic stratification. *Curr Top Med Chem* 2015;**15**(1):5–20.
- Power A, Berger AC, Ginsburg GS. Genomics-enabled drug repositioning and repurposing: insights from an IOM Roundtable activity. *JAMA* 2014;**311**(20):2063–4.
- Bolgar B, Arany A, Temesi G, et al., Drug repositioning for treatment of movement disorders: from serendipity to rational discovery strategies. *Curr Top Med Chem* 2013;**13**(18):2337–63.
- Rich RM, Rosenfeld PJ, Puliafito CA, et al. Short-term safety and efficacy of intravitreal bevacizumab (Avastin) for neovascular age-related macular degeneration. *Retina* 2006;**26**(5):495–511.
- Dudley JT, Deshpande T, Butte AJ. Exploiting drug-disease relationships for computational drug repositioning. *Brief Bioinform* 2011;**12**(4):303–11.
- Keiser MJ, Setola V, Irwin JJ, et al. Predicting new molecular targets for known drugs. *Nature* 2009;**462**(7270):175–81.
- Ha S, Seo YJ, Kwon MS, et al. IDMap: facilitating the detection of potential leads with therapeutic targets. *Bioinformatics* 2008;**24**(11):1413–15.
- von Eichborn J, Murgueitio MS, Dunkel M, et al. PROMISCUOUS: a database for network-based drug-repositioning. *Nucleic Acids Res* 2011;**39**:D1060–6.
- Gloeckner C, Garner AL, Mersha F, et al. Repositioning of an existing drug for the neglected tropical disease Onchocerciasis. *Proc Natl Acad Sci USA* 2010;**107**(8):3424–9.
- Iorio F, Bosotti R, Scacheri E, et al. Discovery of drug mode of action and drug repositioning from transcriptional responses. *Proc Natl Acad Sci USA* 2010;**107**(33):14621–6.
- Lamb J, Crawford ED, Peck D, et al. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* 2006;**313**(5795):1929–35.
- Vidovic D, Koleti A, Schurer SC. Large-scale integration of small molecule-induced genome-wide transcriptional responses, Kinome-wide binding affinities and cell-growth inhibition profiles reveal global trends characterizing systems-level drug action. *Front Genet* 2014;**5**:342.
- Barrett T, Suzek TO, Troup DB, et al. NCBI GEO: mining millions of expression profiles—database and tools. *Nucleic Acids Res*, 2005;**33**:D562–6.
- Dudley JT, Sirota M, Shenoy M, et al. Computational repositioning of the anticonvulsant topiramate for inflammatory bowel disease. *Sci Transl Med* 2011;**3**(96):96ra76.
- Jahchan NS, Dudley JT, Mazur PK, et al. A drug repositioning approach identifies tricyclic antidepressants as inhibitors of small cell lung cancer and other neuroendocrine tumors. *Cancer Discov* 2013;**3**(12):1364–77.
- Iorio F, Rittman T, Ge H, et al. Transcriptional data: a new gateway to drug repositioning? *Drug Discov Today* 2013;**18**(7–8):350–7.
- Pacini C, Iorio F, Goncalves E, et al. DvD: An R/Cytoscape pipeline for drug repurposing using public repositories of gene expression data. *Bioinformatics* 2013;**29**(1):132–4.
- Dennis G, Jr., Sherman BT, Hosack DA, et al., DAVID: database for annotation, visualization, and integrated discovery. *Genome Biol*, 2003;**4**(5):P3.
- Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* 2005;**102**(43):15545–50.
- Ding XM. MicroRNAs: regulators of cancer metastasis and epithelial-mesenchymal transition (EMT). *Chin J Cancer* 2014;**33**(3):140–7.
- Wen X, Deng FM, Wang J. MicroRNAs as predictive biomarkers and therapeutic targets in prostate cancer. *Am J Clin Exp Urol* 2014;**2**(3):219–30.
- Xing Z, Li D, Yang L, et al. MicroRNAs and anticancer drugs. *Acta Biochim Biophys Sin (Shanghai)* 2014;**46**(3):233–9.
- Liu Z, Borlak J, Tong W. Deciphering miRNA transcription factor feed-forward loops to identify drug repurposing candidates for cystic fibrosis. *Genome Med* 2014;**6**(12):94.
- Jiang W, Chen X, Liao M, et al., Identification of links between small molecules and miRNAs in human cancers based on transcriptional responses. *Sci Rep* 2012;**2**:282.
- Rukov JL, Wilentzik R, Jaffe I, et al. PharmacomiR: linking microRNAs and drug effects. *Brief Bioinform* 2014;**15**(4):648–59.
- Liu X, Wang S, Meng F, et al. SM2miR: a database of the experimentally validated small molecules' effects on microRNA expression. *Bioinformatics* 2013;**29**(3):409–411.
- Garnett MJ, Edelman EJ, Heidorn SJ, et al. Systematic identification of genomic markers of drug sensitivity in cancer cells. *Nature* 2012;**483**(7391):570–5.
- Okada Y, Wu D, Trynka G, et al. Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature* 2014;**506**(7488):376–81.
- Sanseau P, Agarwal P, Barnes MR, et al. Use of genome-wide association studies for drug repositioning. *Nat Biotechnol* 2012;**30**(4):317–20.
- Hebbring SJ. The challenges, advantages and future of phenotype-wide association studies. *Immunology* 2014;**141**(2):157–65.
- Denny JC, Bastarache L, Ritchie MD, et al. Systematic comparison of phenotype-wide association study of electronic medical record data and genome-wide association study data. *Nat Biotechnol* 2013;**31**(12):1102–10.
- Yang L, Agarwal P. Systematic drug repositioning based on clinical side-effects. *PLoS One* 2011;**6**(12):e28025.
- Ye H, Liu Q, Wei J. Construction of drug network based on side effects and its application for drug repositioning. *PLoS One* 2014;**9**(2):e87864.
- Bisgin H, Liu Z, Fang H, et al. A phenotype-guided drug repositioning through a latent variable model. *BMC Bioinformatics* 2014;**15**:267.

40. Kuhn M, Campillos M, Letunic I, et al. A side effect resource to capture phenotypic effects of drugs. *Mol Syst Biol* 2010;6:343.
41. Hoehndorf R, Oellrich A, Rebholz-Schuhmann D, et al. Linking PharmGKB to phenotype studies and animal models of disease for drug repurposing. *Pac Symp Biocomput* 2012:388–399.
42. Flintoft L. Disease genetics: phenome-wide association studies go large. *Nat Rev Genet* 2014;15(1):2.
43. Vencken PM, Kriege M, Hoening M, et al. The risk of primary and contralateral breast cancer after ovarian cancer in BRCA1/BRCA2 mutation carriers: implications for counseling. *Cancer* 2013;119(5):955–62.
44. Hyman DM, Zhou Q, Iasonos A, et al. Improved survival for BRCA2-associated serous ovarian cancer compared with both BRCA-negative and BRCA1-associated serous ovarian cancer. *Cancer* 2012;118(15):3703–9.
45. Novick PA, Ortiz OF, Poelman J, et al. SWEETLEAD: an in silico database of approved drugs, regulated chemicals, and herbal isolates for computer-aided drug discovery. *PLoS One* 2013;8(11):e79568.
46. Swamidass SJ. Mining small-molecule screens to repurpose drugs. *Brief Bioinform* 2011;12(4):327–35.
47. Berman HM, Westbrook J, Feng Z, et al. The protein data bank. *Nucleic Acids Res* 2000;28(1):235–42.
48. Pihan E, Colliandre L, Guichou JF, et al. e-Drug3D: 3D structure collections dedicated to drug repurposing and fragment-based drug design. *Bioinformatics* 2012;28(11):1540–4.
49. Wang Y, Chen S, Deng N, et al. Drug repositioning by kernel-based integration of molecular structure, molecular activity, and phenotype data. *PLoS One* 2013;8(11):e78518.
50. Tan FJ, Yang RZ, Xu XX, et al. Drug repositioning by applying 'expression profiles' generated by integrating chemical structure similarity and gene semantic similarity. *Mol Biosyst* 2014;10(5):1126–38.
51. Ng C, Hauptman R, Zhang Y, et al. Anti-infectious drug repurposing using an integrated chemical genomics and structural systems biology approach. *Pac Symp Biocomput* 2014:136–47.
52. Ceribelli M, Kelly PN, Shaffer AL, et al. Blockade of oncogenic I $\kappa$ B kinase activity in diffuse large B-cell lymphoma by bromodomain and extraterminal domain protein inhibitors. *Proc Natl Acad Sci USA* 2014;111(31):11365–70.
53. Wu M, Sirota M, Butte AJ, Chen B. Characteristics of drug combination therapy in oncology by analyzing clinical trial data on clinicaltrials.gov. *Pac Symp Biocomput* 2015;20:68–79.
54. Huang H, Zhang P, Qu XA, et al. Systematic prediction of drug combinations based on clinical side-effects. *Sci Rep* 2014;4:7160.
55. Sun Y, Xiong Y, Xu Q, et al. A hadoop-based method to predict potential effective drug combination. *Biomed Res Int* 2014:196858.
56. Liu Y, Hu B, Fu C, et al. DCDB: drug combination database. *Bioinformatics* 2010;26(4):587–8.
57. Kang Y, Hodges A, Ong E, et al. Identification of drug combinations containing imatinib for treatment of BCR-ABL+ Leukemias. *PLoS One* 2014;9(7):e102221.
58. Menden MP, Iorio F, Garnett M, et al. Machine learning prediction of cancer cell sensitivity to drugs based on genomic and chemical properties. *PLoS One* 2013;8(4):e61318.
59. Napolitano F, Zhao Y, Moreira VM, et al. Drug repositioning: a machine-learning approach through data integration. *J Cheminform* 2013;5:30.
60. Gottlieb A, Stein GY, Ruppin E, et al. PREDICT: a method for inferring novel drug indications with application to personalized medicine. *Mol Syst Biol* 2011;7:496.
61. Zhang P, Wang F, Hu J. Towards drug repositioning: a unified computational framework for integrating multiple aspects of drug similarity and disease similarity. *AMIA Annu Symp Proc* 2014:1258–67.
62. Yang J, Li Z, Fan X, et al. Drug-disease association and drug-repositioning predictions in complex diseases using causal inference-probabilistic matrix factorization. *J Chem Inf Model* 2014;54(9):2562–9.
63. Yildirim MA, Goh KI, Cusick ME, et al. Drug-target network. *Nat Biotechnol* 2007;25(10):1119–26.
64. Schadt EE. Molecular networks as sensors and drivers of common human diseases. *Nature* 2009;461(7261):218–23.
65. Iorio F, Saez-Rodriguez J, di Bernardo D. Network based elucidation of drug response: from modulators to targets. *BMC Syst Biol* 2013;7:139.
66. Li J, Lu Z. A new method for computational drug repositioning using drug pairwise similarity. *Proc IEEE Int Conf Bioinformatics Biomed* 2012;2012:1–4.
67. Li J, Lu Z. Pathway-based drug repositioning using causal inference. *BMC Bioinformatics* 2013;14 (Suppl 16): S3.
68. Wu C, Gudivada RC, Aronow BJ, et al. Computational drug repositioning through heterogeneous network clustering. *BMC Syst Biol* 2013;7:S6.
69. Jin G, Fu C, Zhao H, et al. A novel method of transcriptional response analysis to facilitate drug repositioning for cancer therapy. *Cancer Res* 2012;72(1):33–44.
70. Lu Z. PubMed and beyond: a survey of web tools for searching biomedical literature. *Database (Oxford)* 2011;2011:baq036.
71. Islamaj Dogan R, Murray GC, Neveol A, et al. Understanding PubMed user search behavior through log analysis. *Database (Oxford)* 2009;2009:bap018.
72. Leaman R, Islamaj Dogan R, Lu Z. DNorm: disease name normalization with pairwise learning to rank. *Bioinformatics* 2013;29(22):2909–17.
73. Li J, Lu Z. Systematic identification of pharmacogenomics information from clinical trials. *J Biomed Inform* 2012;45(5):870–8.
74. Leaman R, Wei CH, Lu Z. tmChem: a high performance approach for chemical named entity recognition and normalization. *J Cheminform* 2015;7 (Suppl 1): S3.
75. Wei CH, Kao HY, Lu Z. PubTator: a web-based text mining tool for assisting biocuration. *Nucleic Acids Res* 2013;41 (Web Server issue):W518–22.
76. Arighi CN, Wu CH, Cohen KB, et al. BioCreative-IV virtual issue. *Database (Oxford)* 2014;2014:bau-039.
77. Tari LB, Patel JH. Systematic drug repurposing through text mining. *Methods Mol Biol* 2014;1159:253–67.
78. Andronis C, Sharma A, Virvilis V, et al. Literature mining, ontologies and information visualization for drug repurposing. *Brief Bioinform* 2011;12(4):357–68.
79. Zhu Q, Tao C, Shen F, et al. Exploring the pharmacogenomics knowledge base (PharmGKB) for repositioning breast cancer drugs by leveraging Web ontology language (OWL) and cheminformatics approaches. *Pac Symp Biocomput* 2014:172–82.
80. Chen B, Ding Y, Wild DJ. Assessing drug target association using semantic linked data. *PLoS Comput Biol* 2012;8(7):e1002574.
81. Zerbini LF, Bhasin MK, de Vasconcellos JF, et al. Computational repositioning and preclinical validation of

- pentamidine for renal cell cancer. *Mol Cancer Ther* 2014;**13**(7): 1929–41.
82. Vegner L, Peragovics A, Tombor L, et al. Experimental confirmation of new drug-target interactions predicted by drug profile matching. *J Med Chem* 2013;**56**(21):8377–88.
  83. Khatri P, Roedder S, Kimura N, et al. A common rejection module (CRM) for acute rejection across multiple organs identifies novel therapeutics for organ transplantation. *J Exp Med* 2013;**210**(11):2205–21.
  84. Xu H, Aldrich MC, Chen Q, et al. Validating drug repurposing signals using electronic health records: a case study of metformin associated with reduced cancer mortality. *J Am Med Inform Assoc* 2015;**22**:179–91.
  85. Seok J, Warren HS, Cuenca AG, et al. Genomic responses in mouse models poorly mimic human inflammatory diseases. *Proc Natl Acad Sci USA* 2013;**110**(9):3507–12.
  86. Takao K, Miyakawa T. Genomic responses in mouse models greatly mimic human inflammatory diseases. *Proc Natl Acad Sci USA* 2015;**112**(4):1167–72.
  87. Domcke S, Sinha R, Levine DA, et al. Evaluating cell lines as tumour models by comparison of genomic profiles. *Nat Commun* 2013;**4**:2126.
  88. Pessetto ZY, Weir SJ, Sethi G, et al. Drug repurposing for gastrointestinal stromal tumor. *Mol Cancer Ther* 2013;**12**(7): 1299–309.
  89. Stenvang J, Kumler I, Nygard SB, et al. Biomarker-guided repurposing of chemotherapeutic drugs for cancer therapy: a novel strategy in drug development. *Front Oncol* 2013;**3**:313.
  90. Nygren P, Fryknas M, Agerup B, et al., Repositioning of the anthelmintic drug mebendazole for the treatment for colon cancer. *J Cancer Res Clin Oncol* 2013;**139**(12):2133–40.
  91. Quinn BJ, Kitagawa H, Memmott RM, et al. Repositioning metformin for cancer prevention and treatment. *Trends Endocrinol Metab* 2013;**24**(9):469–80.
  92. Yeh CT, Wu AT, Chang PM, et al. Trifluoperazine, an antipsychotic agent, inhibits cancer stem cell growth and overcomes drug resistance of lung cancer. *Am J Respir Crit Care Med* 2012;**186**(11):1180–8.
  93. Molineris I, Ala U, Provero P, et al. Drug repositioning for orphan genetic diseases through Conserved Anticoexpressed Gene Clusters (CAGCs). *BMC Bioinformatics* 2013;**14**:288.
  94. Xu K, Cote TR. Database identifies FDA-approved drugs with potential to be repurposed for treatment of orphan diseases. *Brief in Bioinform* 2011;**12**(4):341–5.
  95. Pemovska T, Kontro M, Yadav B, et al. Individualized systems medicine strategy to tailor treatments for patients with chemorefractory acute myeloid leukemia. *Cancer Discov* 2013;**3**(12):1416–29.
  96. Li YY, Jones SJ. Drug repositioning for personalized medicine. *Genome Med* 2012;**4**(3):27.
  97. Iskar M, Campillos M, Kuhn M, et al. Drug-induced regulation of target expression. *PLoS Comput Biol* 2010;**6**(9).
  98. Khare R, Li J, Lu Z. LabeledIn: cataloging labeled indications for human drugs. *J Biomed Inform* 2014;**52**:448–56.
  99. Khare R, Burger JD, Aberdeen JS, et al. Scaling drug indication curation through crowdsourcing. *Database* 2015; **2015**:bav016. doi:10.1093/database/bav016
  100. Shim JS, Liu JO. Recent advances in drug repositioning for the discovery of new anticancer drugs. *Int J Biol Sci* 2014;**10**(7): 654–63.
  101. Enserink M. Infectious diseases. Debate erupts on ‘repurposed’ drugs for Ebola. *Science* 2014;**345**(6198):718–19.