Computers in Biology and Medicine 155 (2023) 106671

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Computational drug repurposing by exploiting large-scale gene expression data: Strategy, methods and applications

Hao He [a,b,1], Hongrui Duo [a,1], Youjin Hao [a], Xiaoxi Zhang [a], Xinyi Zhou [a], Yujie Zeng [a], Yinghong Li [c], Bo Li [a,*]

[a] College of Life Sciences, Chongqing Normal University, Chongqing, 400044, PR China
[b] State Key Laboratory of Medical Neurobiology and MOE Frontiers Center for Brain Science, Institutes of Brain Science, Fudan University, Shanghai, 200032, PR China
[c] The Key Laboratory on Big Data for Bio Intelligence, Chongqing University of Posts and Telecommunications, Chongqing, 400065, PR China

ARTICLE INFO

Keywords:
Gene expression profiling
Computational drug repurposing
Drug discovery
RNA-seq
Drug combination

ABSTRACT

De novo drug development is an extremely complex, time-consuming and costly task. Urgent needs for therapies of various diseases have greatly accelerated searches for more effective drug development methods. Luckily, drug repurposing provides a new and effective perspective on disease treatment. Rapidly increased large-scale transcriptome data paints a detailed prospect of gene expression during disease onset and thus has received wide attention in the field of computational drug repurposing. However, how to efficiently mine transcriptome data and identify new indications for old drugs remains a critical challenge. This review discussed the irreplaceable role of transcriptome data in computational drug repurposing and summarized some representative databases, tools and strategies. More importantly, it proposed a practical guideline through establishing the correspondence between three gene expression data types and five strategies, which would facilitate researchers to adopt appropriate strategies to deeply mine large-scale transcriptome data and discover more effective therapies.

1. Introduction

New drug development is characterized by time-consuming, huge investment, low success and high risk [1,2]. New drugs often have unpredictable side effects, so about 90% of experimental drugs fail in the first phase of clinical trial [3–5]. The high failure rate has severely slowed drug development progress, making drugs unable to meet clinical demands. Therefore, there is an urgent need to seek the new more low-cost and efficient drug development strategies.

Drug repurposing (also known as drug repositioning) is a new methodology to reanalyze and re-evaluate drug candidates and/or drugs in clinical use to discover their new indications [6,7]. Potential targets for drug repurposing are approved drugs, because their safeties, targets and mechanism of action, and efficacies were well established [8–10]. Indeed, drug repurposing can greatly speed up drug development process and reduce costs [11–13]. At present, drug repurposing has been successfully carried out in the drug development for various diseases [14,15], including cancer [16–21], neurodegenerative diseases [22], neuropsychiatric disorders [23,24], COVID-19 [25], tuberculosis [26], malaria [27] and autoimmune inflammatory diseases [28]. For example, thalidomide was initially used as an antiemetic drug for pregnant women, and can also be used to treat acute pancreatitis, chronic hepatitis C, and gastritis caused by *helicobacter pylori* [29]. Minoxidil (mainly as an antihypertensive drug) was found to have a good therapeutic effect on hair loss [30].

Rapidly accumulated high-throughput omics data [31–42] and chemical structure information [2,43–48] have opened a window to explore new uses for old drugs. Especially, in-depth mining huge amount of transcriptome data stored in gene expression profiling databases (such as GEO and ArrayExpress) will offer possibilities to solve many problems in the life sciences [49,50]. The transcriptome data not only reflects the detailed gene expression profiles during disease onset and progression [51,52], but also provides the valuable data sources for accurate drug repurposing analysis [53,54]. As shown in Fig. 1, computational drug repurposing began with the construction of the databases of pharmacology, gene mutations, and gene expression profiles. In 2000, Hughes et al. constructed a gene expression signatures database derived from 300 mutational and chemical perturbations on

* Corresponding author.
E-mail address: libcell@cqnu.edu.cn (B. Li).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.compbiomed.2023.106671
Received 25 November 2022; Received in revised form 5 February 2023; Accepted 10 February 2023
Available online 12 February 2023

Saccharomyces cerevisiae, and they proposed that specific gene expression patterns could be applied to characterize unknown drugs [55]. Subsequently, Iconix constructed a toxicogenomic database containing gene expression profiles of rats treated with 600 compounds, which greatly facilitated the identification or prediction of new chemicals by functional analysis [56,57]. More importantly, Lamb and coworkers used the gene chip technique to analyze genome-wide expression profiles of 5 human tumor cell lines treated with 1309 drugs and generated a disease-drug Connectivity Map (CMap) that connects small molecules, gene expression and diseases by gene-expression signatures. CMap is no doubt a milestone in mining transcriptomics for drug reposition [58]. Since then, researchers have focused on impacts of various perturbations on gene expression. Hu et al. systematically analyzed gene expression profiles of disease interference and established a disease-drug interaction network [59]. Xiao et al. comprehensively investigated more than 3000 human and mouse transcriptome data and a perturbation network was defined to predict associations of gene expression and regulation, cancer and drug according to perturbed gene expression features [60]. Based on these valuable explorations, scientists began to focus great attention on expanding the scale of perturbation transcriptome data through integrated design. Later, the library of Integrated Network-based Cellular Signatures (LINCS) L1000 platform, a low-cost and high-throughput transcriptomics technique was applied to profile gene expressions of a collection of human cell lines treated by more than 30,000 compounds, in which 978 landmarks (genes) can cover 82% of the whole transcriptome information [61]. On the basis of this platform, CMap2.0 (https://clue.io/cmap) greatly expanded the scale and abundance of disturbed transcriptome data, in which more than 1.3 million gene expression signatures were collected after gene knockdown, gene over-expression, and small molecule perturbation on 77 cell lines. Through comprehensive analysis of gene expression profiles, CMap2.0 can quickly forecast the small molecules with therapeutic potential and possible mechanisms of action.

In the last two decades, drug repurposing based on transcriptome data has made great progress. Here, we summarized several representative and well-known databases (Table 1), and corresponding tools (Table 2). For researchers, the availability of multiple databases and tools is not only an opportunity, but also a great challenge. For those who lack knowledge of bioinformatics and computational biology, it is still very challenging to choose and adopt appropriate strategies to deeply mine large-scale transcriptome datasets for drug repurposing. In order to simplify and clarify the procedure, this review aims at proposing a practical guideline by establishing the correspondence between gene expression data types and analysis strategies. To facilitate discussion, transcriptome datasets were categorized into three groups based on data resources: (i) from healthy and patients, (ii) from patients with and without drug intervention, as well as (iii) from patient-only (Fig. 2). In addition, five popular strategies were summarized in detail based on main research objectives (drug, disease and target) in drug repurposing: drug-disease connection, drug combination, drug-drug connection, drug-target connection (association), and disease-disease connection. What's more, this paper also summarizes the latest progress in transcriptome data-based drug repurposing and illustrates some ideas on the tomorrow of drug repurposing combined with scRNA-seq.

2. Computational strategies for drug repurposing based on gene expression profiling

Gene expression profiling reflects alterations of various transcripts during disease progression, and thus can provide important clues to slow down disease progression or cure them by interfering with or reversing gene expression processes. Currently, there are two main strategies for drug repurposing using transcriptome data: identifying the specific states and assessing correlations between data of patients or individuals. Gene signatures can be any subsets of genes correlated with a specific biological state [102,103]. A gene signature consisting of a series of up- and down-regulated DEGs can adequately reflect the changes in gene expression perturbed by stimulus [104,105]. Generally, differentially expressed genes (DEGs) or sets of representative genes are used to characterize the biological status of a patient or assess the correlation between individuals. Their popularity in drug repurposing is mainly due to two reasons: (i) global gene expression profiles obscure some important expression patterns, while gene signatures identified by feature selection can simplify the analysis process and better discriminate different biological states [106]; and (ii) gene signatures are relatively independent and can be compared across data platforms [72,107]. It is worth noting that, when the expression profiles from control groups (healthy tissue or individuals) are lacking, the whole genome-wide expression level derived from patients could also be used to explore relationships based on the correlation of higher dimensional vectors. In most cases, the relationship between two states can be determined by the correlation of gene expression vectors for a given signature (or gene set), rather than all genes in the genome.

2.1. Analysis strategies based on gene expression profiles of healthy people vs. patients

2.1.1. Drug-disease connection

High-throughput DNA microarray and RNA sequencing techniques allow to analyze gene expression changes at the whole genome level, which is helpful to discover drug-disease connection [108]. Studies have shown that restoring abnormal gene expression levels of patients to normal levels may have therapeutic potential [52,58]. That is, if the gene expression signature stimulated by a drug is significantly and

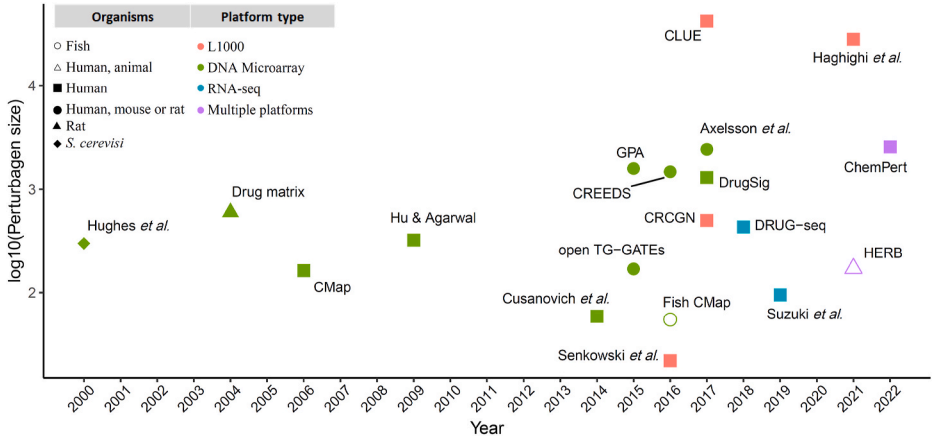

Fig. 1. The typical approaches or tools used in computational drug repurposing over the past twenty years.

Table 1
Most representative gene expression profile databases used for drug repurposing.

Database	Statistics Information	Category	URL
NCBI GEO [62]	(1) 177,273 series (5,107,078 samples); (2) 5959 organisms.	CGED	http://www.ncbi.nlm.nih.gov/geo/
ArrayExpress [63]	(1) 75,391 experiments (2,589,677 assays); (2) > 3300 organisms.	CGED	https://www.ebi.ac.uk/arrayexpress/
Expression Atlas [64]	(1) 4315 experiments (153,212 assays); (2) 65 species.	CGED	https://www.ebi.ac.uk/gxa/home
GTEx [65]	(1) 17,382 RNA-seq samples; (2) 54 tissue types and 2 cell lines on human.	CGED	https://commonfund.nih.gov/gtex
CCLE [66]	(1) 1457 cell lines; (2) 136,488 unique datasets.	CGED	https://portals.broadinstitute.org/ccle
TCGA [67]	(1) 85,552 cases (67 primary sites); (2) Human (33 cancer types).	TGED	https://portal.gdc.cancer.gov/
ICGC [68]	(1) 24,289 donors (445 tissues); (2) Human (22 cancer primary sites).	TGED	https://dcc.icgc.org/
MGI-GXD [69]	(1) 107,436 expression assays; (2) 16,562 genes from mouse, including data from numerous strains of wild-type mice and from >4900 mouse mutants.	CGED	http://www.informatics.jax.org/expression.shtml
GENT [70]	(1) More than 68,000 samples from human; (2) 72 different normal and tumor tissues.	CGED; TGED	http://gent2.appex.kr/gent2/
MTD [71]	(1) 254 datasets, 102 tissues/cell lines; (2) 4 organisms (human, mouse, rat and pig).	CGED	http://mtd.cbi.ac.cn/
DP14 & DP92 [72]	DP14/92: (1) 276/857 profiles; (2) 14/92 compounds, 1/3 cell line, 3/3 time points.	CPED	http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE60408
open TG-GATEs [73]	(1) 20,118 assays; (2) 170 compounds perturbation from *in vivo* and *in vitro*.	CPED	https://toxico.nibiohn.go.jp/english/
Fish CMap [74]	(1) 3516 microarrays (55 experimental conditions); (2) zebrafish and fathead minnow.	CPED	GSE38070, GSE60202, GSE70807, and GSE70936
CREEDS [75]	(1) 4205 signatures; (2) mammalian cells.	CPED	http://amp.pharm.mssm.edu/creeds
CRCGN	(1) 5996 signatures; (2) 500 perturbagens, multiple cell lines.	CPED	https://carcinogenome.org/
DrugSig [76]	(1) 7000 microarrays; (2) 1300 drugs, human	CPED	http://biotechlab.fudan.edu.cn/database/drugsig/
DRUG-seq [77]	(1) 433 compounds; (2) one cell line	CPED	http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE120222
Drug Repurposing Hub [78]	(1) 16,826 samples; (2) 7934 compounds	CPED	http://www.broadinstitute.org/repurposing
Haghighi et al. [79]	(1) 28,000 Genetic and Chemical Perturbations; (2) cell lines	CPED	http://broad.io/rosetta
HERB [80]	(1) 1037 high-throughput experiments, 6164 profiles; (2) 20 herbs and 152 ingredients	CPED	http://herb.ac.cn/
TMNP [81]	(1) 21,466 signatures; (2) multiple cell lines	CPED; DGED; GPED	http://www.bcxnfz.top/TMNP
DrugMatrix [82]	(1) 12,921 samples; (2) 8 different tissues of rats (perturbated by 600 compounds)	CPED	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE59927
DA Cusanovich et al. [83]	(1) 201 samples; (2) perturbations of knocking down 57 transcription factors	GPED	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE50588
Gene perturbation atlas (GPA) [60]	(1) 3072 transcriptome profiles of single-gene perturbations; (2) 1170 different cell lines/tissues of human or mouse	GPED	http://biocc.hrbmu.edu.cn/GPA/
Senkowski et al. [84]	(1) 1065 profiles; (2) 22 compounds, in three distinct cultured cell models	CPED	http://data.genometry.com/
Suzuki et al. [85]	(1) 3240 RNA-seq and 3393 ATAC-seq libraries; (2) 23 cell lines treated with 95 compounds	CPED	https://kero.hgc.jp/
ChemPert [86]	(1) 82,270 transcriptional signatures, 2566 unique perturbagens; (2) 167 cell types	CPED	https://chempert.uni.lu/

Note: CGED - Comprehensive gene expression database; TGED - Tumor gene expression database; CPED - Compounds-perturbation gene expression database; DGED - Disease gene expression database; GPED - Genetic perturbations gene expression database.

negatively correlated with that in disease state, indicating that this drug may have the potency to cure that disease. Presently, a large number of computational methods have been developed to predict drug-disease connections, such as the Connective Map (CMap) [58,61]. CMap is one of the most comprehensive and systematic resources for drug-disease association analyses, which describes the disease, physiology, and biology states stimulated by drugs. CMap stored a huge number of expression profiling induced by small molecules, in which GSEA algorithm can be used to compare differences in gene signatures between the treatment and the control groups, bridging the gap between biological states and similar (same) or opposite responses [58].

As shown in Fig. 3(A), when drug-induced signatures are matched to that in the disease state, CMap analysis yields four results: strong negative correlation, no correlation, weak positive correlation, and strong positive correlation. The drug may be considered as a candidate if the perturbed gene expression profile is negatively correlated with that in the disease state. Conversely, if the gene expression profile after drug perturbation is positively correlated with that of the patient, the drug may induce or aggravate the disease.

For instance, Noort and co-workers used 127 identified DEGs as gene signatures and predicted three drugs (citalopram, enilconazole, and troglitazone) with the potential to treat colorectal cancer [109]. These three drugs showed significant inhibitory effects on mouse cancer and human tumor cells *in vitro*. A similar approach was used to predict possible drugs for asthma and papillary renal cell carcinoma [52,110].

2.1.2. Drug-drug combinations

Combination drug therapy (CDT) has been highly concerned in biomedicine community nowadays [111–114]. Owing to the complex interactions between cell components as well as simultaneous dysregulations of pathways or networks in diseases [115], monotherapy has

Table 2
Representative state-of-art tools for drug repurposing based on large-scale transcriptome data.

Tool	Description	Relationship	Category	URL
CLUE (CMap2.0) [58, 61]	Cloud platform for analyzing perturbed L1000 transcriptome, P100 and GCP proteome data.	Drug-drug Drug-disease Drug-target	Web tool	https://clue.io
LINCS Canvas Browser [87]	Web server for gene signature analysis.	Drug-drug Drug-disease Drug-target Disease-disease	Web tool	http://www.maayanlab.net/LINCS/LCB/
L1000CDS2 [88]	A web-based tool for analyzing small molecule signatures, pairwise combinations, and predicting drug targets based on L1000 data.	Drug-disease Drug-drug Drug-target Drug combination	Web tool	https://maayanlab.cloud/L1000CDS2
PharmacoGx [89]	An R package for calculating the correlation between drug dose and molecular features in cancer cell lines.	Drug-disease Drug-drug Drug-target	R package	https://github.com/bhklab/PharmacoGx
DrugComboRanker [90]	A computational tool for analyzing synergistic drug combinations and their mechanisms of action.	Drug combination	Java app	https://github.com/methodistmab/DrugComboRanker
SEP-L1000 [91]	A platform for ADRs ranking of approved drugs and pre-clinical small-molecule chemicals based compound structure and gene expression features.	Drug-disease Drug-drug	Web tool	http://maayanlab.net/SEP-L1000/
Mantra [92]	An integrated data platform for analyzing the Mode of Action (MoA) of novel drugs and repositioning approved drugs.	Drug-drug	Web tool	http://mantra.tigem.it
DvD [93]	An R package for comparing gene expression profiles and visualizing of interaction networks.	Drug-disease Drug-drug Drug-target	R package	https://saezlab.github.io/DrugVsDisease/
DSEA [94]	A tool to analyze transcriptional responses of different, phenotypically-related drugs to find shared pathways.	Drug-target	Web tool	https://dsea.tigem.it/
L1000FWD [95]	Web tool to visualize gene expression signatures induced by small molecules, predict function of small molecules, and explore the models of action of drugs.	Drug-drug Drug-disease Drug combination	Web tool	http://amp.pharm.mssm.edu/L1000FWD
ksRepo [96]	A platform for analyzing perturbation effects of approved drugs and other chemical.	Drug-disease Drug-drug Drug-target	R package	https://github.com/adam-sam-brown/ksRepo
gene2drug [97]	A computational framework for pathway-based drug repurposing and gene expression perturbated by small compounds.	Drug-target	Web tool	https://gene2drug.tigem.it/
iLINCS [98]	A web platform for gene signature-driven drug repurposing based on LINCS data and Omics signatures.	Drug-disease Drug-drug Drug-target	Web tool	http://www.ilincs.org
DeSigN [99]	A web tool for analyzing DEGs perturbed by 140 drugs and predicting potential applications of drugs.	Drug-disease Drug-drug	Web tool	http://design.cancerresearch.my/
Cogena [100]	An effective framework for gene co-expression and enrichment analysis.	Drug-disease Drug-drug	R package	https://github.com/zhilongjia/cogena
Toppgenes Suite [101]	A portal for gene function enrichment and candidate gene ranking.	Drug-disease Drug-drug Disease-disease	Web tool	https://toppgene.cchmc.org/

intrinsic limitations in their maximum therapeutic benefits [116,117], which has accelerated the development of CDT. Existing studies have shown that a balanced multi-drug combination may be more effective than monotherapy [116,118–120], and has an advantage in avoiding drug resistance [117,121–128].

Fig. 3(B) illustrates the process of predicting effective drug combinations using transcriptome data. Through gene signature enrichment analysis, bioinformatics tools such as gene2drug could be used to predict inhibitors or activators for specific disordered signaling pathways according to gene expression profiles [97]. In general, an effective drug combination (or synergistic drug pair) ordinarily inhibits two targets in a linear pathway [129,130] or two proteins in parallel pathways crucial to the disease [131,132], or reverses two major disordered signaling pathways associated with the disease [106,133,134], or inhibits two entry pathways of infected viruses such as SARS-CoV-2 [135]. An alternative approach is to score predicted drug combinations based on the ability to reversing disease inordinate signatures. Theoretically, the optimal drug combination should have a strong negative correlation with the disease state.

Currently, there are several well-known and popular databases for drug combination screening, such as SYNERGxDB [136], DrugCombDB [137], CellMinerCDB [138], DCDB [139] and SynergyFinder [140]. Of which, SYNERGxDB is suitable for determining effective therapeutic combinations and cancer biomarkers. DrugCombDB and DCDB are comprehensive databases for drug combination prediction. CellMinerCDB performs integrated molecular and pharmacological analysis within or across cancer cell line datasets. SynergyFinder 2.0 can be used to visually analyze the synergistic effects of multiple drug combination [140].

Varied approaches (including machine learning algorithms) have been applied to predict drug combination [116,141]. For instance, DrugComboRanker approach could predict the synergistic effects of drug combination acting on different signaling network modules, which has been used to screen drug combinations that can overcome drug resistance in cancer cells [90]. Lee et al. developed the Combinatorial Drug Assembler (CDA), a genomics and bioinformatics system, which detected the best pattern matching combinatorial drug pairs. Two combination drug pairs (alsterpaullone + scriptaid, irinotecan + semustine) predicted by CDA significantly induced apoptosis in non-small cell lung cancer and trip-negative breast cancer cells and showed good synergistic effects [124]. In addition, using genomic information, drug targets and pharmacological information of various types of cancer cell lines, a personalized drug combination synergy prediction pipeline based on machine learning strategy was proposed by

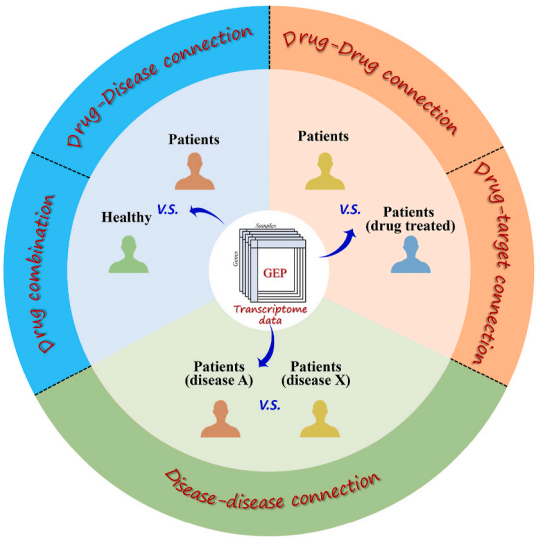

Fig. 2. Different data types in transcriptomics and employed drug repurposing schemes.

Jeon et al. [142]. Similarly, Pulkkinen et al. developed a mechanism-agnostic optimization method for identifying pairwise or higher-order combinations with maximal cancer-selectivity [143].

2.2. Analysis strategies toward gene expression profiles before and after drug intervention

2.2.1. Drug-drug connection

Gene expression profiles and previous studies indicated that similar drugs may have common clinical indications [144–147]. In view of this, two different drugs are considered functionally similar if they can induce similar gene expression profiles in cells. They may be involved in the same model of action and have the potential to treat the same or similar diseases.

As shown in Fig. 4(A), if a gene signature induced by a compound is positively correlated with that of a drug, the compound is regarded as a potential substitute for the drug. Conversely, a negative correlation implies that they may mutually attenuate the effect. This strategy of finding drug substitutes will facilitate overcoming drug resistance.

To date, this strategy has been widely applied to predict novel targets or functions of various chemicals. For instance, Lim et al. identified a new target of urotensin-II receptor antagonist (KR-37524) based on its similar expression patterns with other drugs [148]. Further experiments verified this. Moreover, ÖZDEMİR et al. found that Rho-kinase is a new target of two antidiabetic drugs (metformin and tolbutamide) by CMap [149].

2.2.2. Drug-target connection

As mentioned above, when whole-genome expression profile is used to study the action model of a drug, the real signal may be masked by other responses of the same pathway, which hinders the detection of the real signal [150]. Therefore, the combination of robust targets and well-qualified disease-related biomarkers enhances understanding of the mechanism of action and also facilitates drug development [151]. Fortunately, gene editing technology can successfully solve this problem, because it simplifies the question by interfering with only one gene.

To comprehensively analyze the relationship between gene expression signatures, drugs, and diseases, Subramanian et al. obtained the expression profiles of more than 20,000 through gene knockdown or gene overexpressing [61]. According to the correlations between gene signatures on gene expression levels, the new connections were predicted and updated *via* CLUE. Fig. 4(B) showed the comparison of gene signatures perturbed by small molecules. If there is a strong negative correlationship between two gene signatures, it indicates that the small molecule may be able to inhibit gene expression. If there is a positive

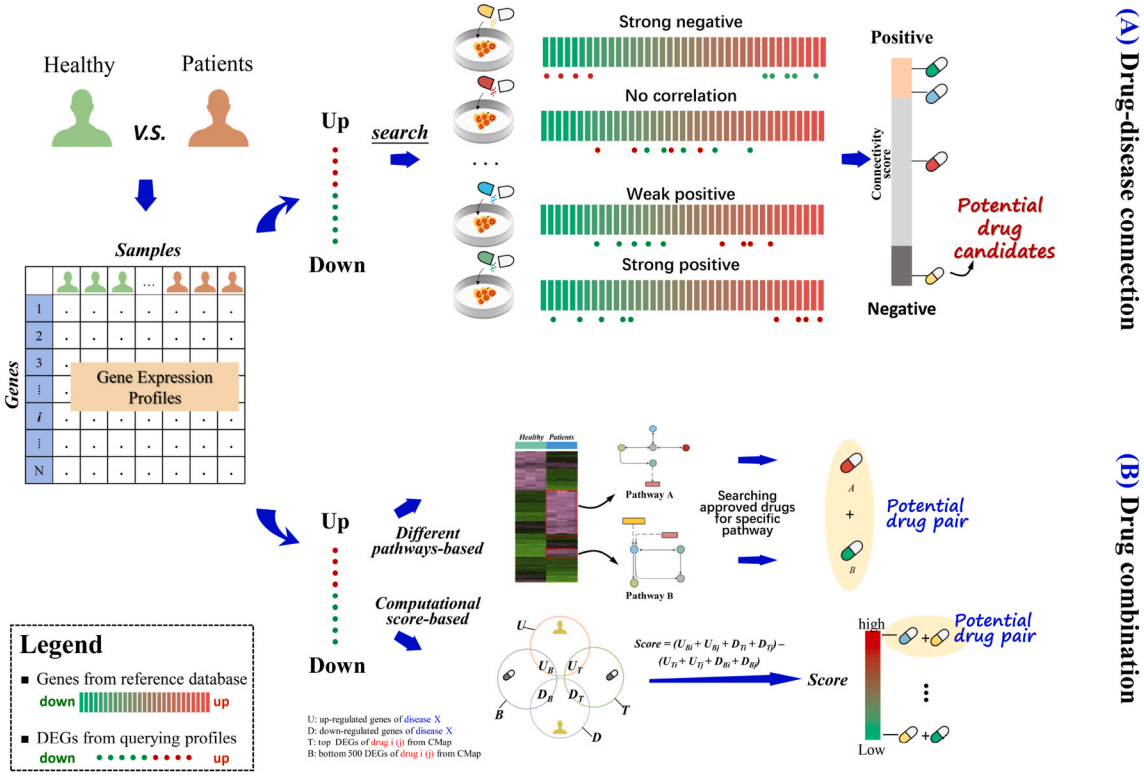

Fig. 3. Two reusing scenarios of the transcriptome profiles from patients compared to healthy subjects. (A) Identifying new indications of existent drugs. (B) Discovering drug-drug combinations by the cell signaling/metabolic pathways or the specific computational scores.

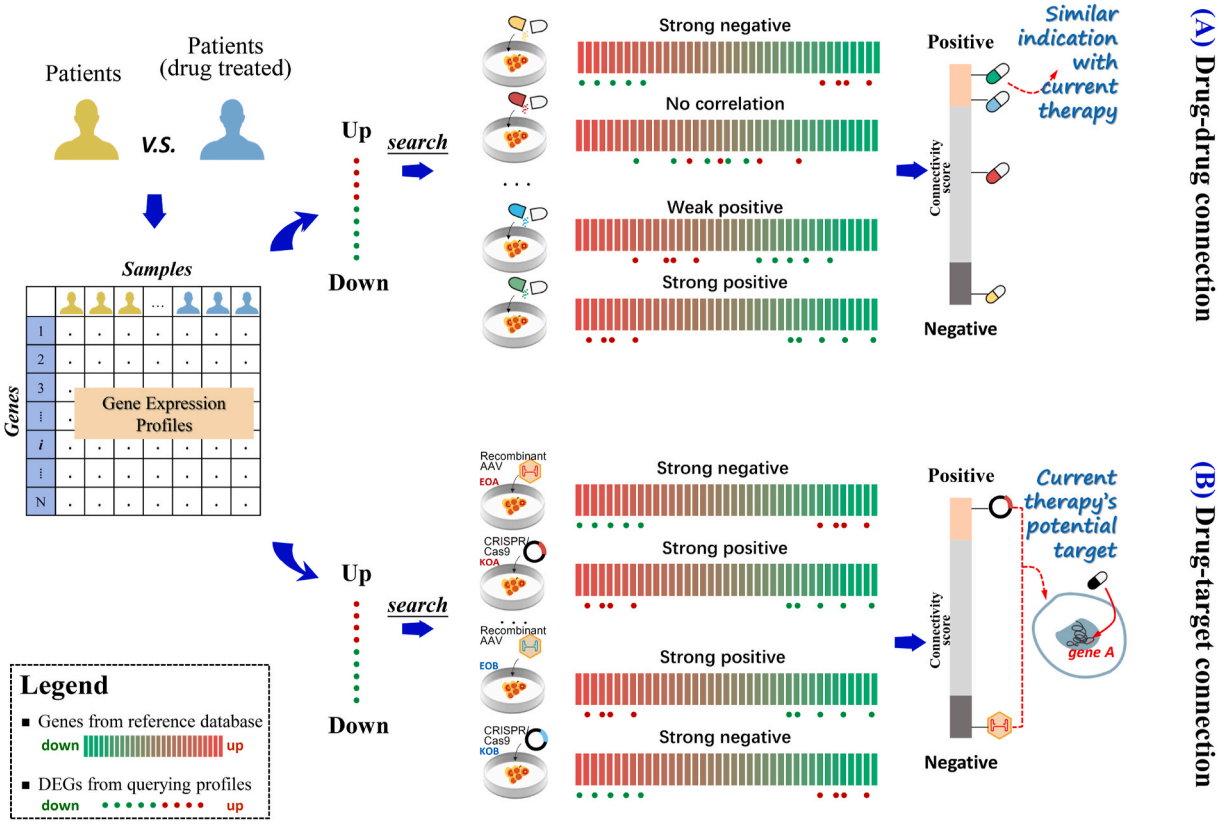

Fig. 4. Two reusing scenarios of the transcriptome profiles of patients with and without drug intervention. (A) Identifying new indications in terms of drug-drug connection. (B) Discovering new indications in terms of drug-target connection.

correlation, it can promote gene expression. If a gene is not a known target of a small molecule, it can be considered as a new candidate target. Therefore, new drug-target interactions can be established.

On basis of this strategy, Pabon et al. propounded a new pipeline that can investigate the correlation between small molecules and protein targets, which may accelerate the identification and development of novel chemical classes by screening compound-target interactions. The authors then validated 152 FAD approved drugs and 3104 potential targets [150]. Their correlation analysis provided new insights on cellular responses to disrupting protein interactions and highlighted the complex genetic phenotypes of drug treatments.

2.3. Analysis strategies toward the gene expression profiles for single or multiple patients

To predict the therapeutic effects of new potential drugs or repositioned drugs, researchers rationalized the use of gene expression data from single or multiple patients based on disease-disease similarities. Chiang et al. suggested that if two diseases have similar gene expression profiles, a drug used for one disease may also be effective for the other disease [152]. Actually, this is also the process of exploiting disease-disease connection to discover new repurposed drugs.

Although there is a large amount of gene expression data from pathological samples from patients, healthy control samples are usually lacking. This means that researchers cannot directly obtain gene signatures of diseases, but only rely on gene expression data of patients. To make better use of the existing transcriptome data, the full gene lists from patients or pathological states only could also be used to explore new correlations between different diseases. Alternatively, the analysis can be simplified by determining the specific order of the entire gene list, or by using the validated 978 representative genes from the L1000 database. Through ranking each gene based on average expression level across multiple samples, gene lists were generated and high-dimensional vectors were applied to evaluate their correlations. As shown in Fig. 5, if the global gene expression profile of disease-A shows a strongly positive correlation with that of disease-X, the drug for disease-A may be a candidate for the treatment of disease-X.

Predicting new repositioned drugs using disease-disease connection has two advantages. Firstly, it significantly facilitated the analysis of mechanisms for some complex diseases, such as cancer, diabetes and cardiovascular diseases [153,154]. Besides, it provides an economic way for the drug markets of rare diseases. Due to the small number of people affected by rare diseases, research institutions and funds for drug development are scarce [155], and an *in silico* way with low cost could help break through its dilemma. For instance, Suthram et al. constructed a framework by integrating disease-related transcriptome data and protein interaction networks. On the basis of this framework, 138 connections between multiple diseases and fourteen drug-shared pairs were identified [156].

3. Opportunities and challenges

In recent years, transcriptome data has been widely used for drug repurposing. Its popularity may be due to: (i) compared with other data types, the generation of transcriptome data requires fewer cells and lower cost [58]; (ii) current methods for analyzing longer reads and higher throughout transcriptome data have better performance in capturing complex genomic and environmental impacts [107]; (iii) numerous studies have shown that gene-expression signature analysis is considered to be an important approach for elucidating the action model of drugs and mechanisms of disease [58,157,158]. It is clear that huge amount of transcriptome data not only accelerates the discovery of drug candidates for multiple diseases, but also plays a crucial role in revealing new relationships between drugs, genes, and diseases.

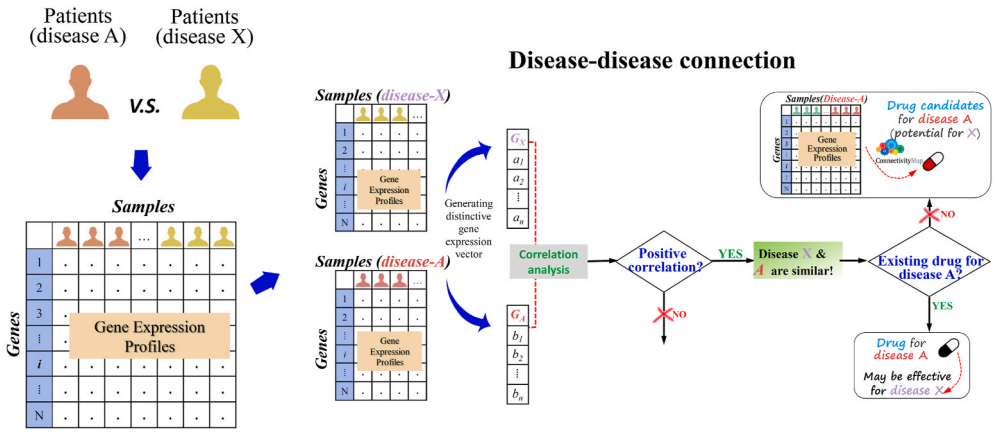

Fig. 5. Discovering new indications in terms of disease-disease connection *via* transcriptome profiles from two groups of patients with different diseases.

Everything has two sides. The application of transcriptome data in drug repurposing also has some limitations and challenges. For instance, the majority the perturbed gene expression profiles for reference are derived from a limited number of cell lines, thus there is a long way to go from prediction to application. The predictive power of computational models is affected by some unavoidable factors (e.g., data missing, data bias, and limited types of perturbed experiments). In addition, limitations of gold standard datasets for drug repurposing also affect the performances of these models. Although several models are available, effective methods to conquer these problems are still lacking.

4. Perspectives

With the availability of huge amount of transcriptome data, researchers have realized the unique significance of transcriptome data in characterizing gene perturbations, exploring molecular interactions and precision medicine. The increasing attention has inevitably made more transcriptome data available for drug repurposing, which requires stringent requirements for data storage and efficient searching.

Through the review of transcriptome data in the area of drug repurposing, we have witnessed the advances in measuring methods and various dimensions of transcriptome data. Taking the milestone project CMap as an example, the scale of perturbed gene expression data was expanded from 564 to 1.3 million in ten years (year 2006–2017). As a result, the rapidly growing databases require faster development of analysis tools to meet the expansion of data. In addition, superior statistic models are becoming more increasingly popular in computational drug repurposing, especially the application of machine learning-based, deep learning-based [159,160] and artificial intelligence methods [161–163]. To improve the accuracy of repurposed drugs, the combination of different computational methods will be the future trend.

Additionally, although the disease microenvironment is gaining attention in drug development and targeting therapy, how to use transcriptome data for microenvironment (e.g. tumor microenvironment) evaluation is an urgent barrier to be broken. Fortunately, the emerging single-cell RNA-sequencing technology can provide a new perspective for disease microenvironment research, which will certainly promote the development of drug repurposing. For instance, the information obtained from different cell subsets makes it possible to distinguish the tumor and paracancerous tissues. This not only gives a real picture of the microenvironment *in vivo* but also provides the opportunity to predict the side effects of new therapies. More significantly, it also allows for personalized and precise treatment.

Conflicts of interest

None Declared.

Author contributions

Conceptualization & Data Curation: B. Li, H. He and H. R. Duo; Writing-Original Draft: H. He; Writing-Reviewing and Editing: Y. J. Hao, X. X. Zhang, X. Y. Zhou and Y. J. Zeng; Resources & Visualization: H. He and Y. H. Li; Supervision & Funding Acquisition: B. Li, Y. J. Hao and Y. H. Li. All authors reviewed the manuscript.

Declaration of competing interest

The authors declare that there are no conflicts of interest in this work.

Acknowledgments

This project was supported by Science and Technology Research Program of Chongqing Municipal Education Commission (KJQN202100538, KJQN202100642), National Natural Science Foundation of China (62101087), China Postdoctoral Science Foundation (2021MD703942), Chongqing Primary and Secondary School Innovative Talent Training Project (CY210709), and Open Fund of Yunnan Key Laboratory of Plant Reproductive Adaptation and Evolutionary Ecology, Yunnan University.

References

[1] N. Nosengo, Can you teach old drugs new tricks? Nature 534 (2016) 314–316.
[2] Y. Zhou, Y. Zhang, X. Lian, F. Li, C. Wang, F. Zhu, Y. Qiu, Y. Chen, Therapeutic target database update 2022: facilitating drug discovery with enriched comparative data of targeted agents, Nucleic Acids Res. 50 (2022) D1398–D1407.
[3] N. Fleming, How artificial intelligence is changing drug discovery, Nature 557 (2018) S55–S57.
[4] H. Dowden, J. Munro, Trends in clinical success rates and therapeutic focus, Nat. Rev. Drug Discov. 18 (2019) 495–496.
[5] F. Zhu, X.X. Li, S.Y. Yang, Y.Z. Chen, Clinical success of drug targets prospectively predicted by in silico study, Trends Pharmacol. Sci. 39 (2018) 229–231.
[6] M.A. Farha, E.D. Brown, Drug repurposing for antimicrobial discovery, Nat. Microbiol. 4 (2019) 565–577.
[7] T.T. Ashburn, K.B. Thor, Drug repositioning: identifying and developing new uses for existing drugs, Nat. Rev. Drug Discov. 3 (2004) 673–683.
[8] W. Xue, F. Yang, P. Wang, G. Zheng, Y. Chen, X. Yao, F. Zhu, What contributes to serotonin-norepinephrine reuptake inhibitors' dual-targeting mechanism? The key role of transmembrane domain 6 in human serotonin and norepinephrine transporters revealed by molecular dynamics simulation, ACS Chem. Neurosci. 9 (2018) 1128–1140.
[9] Y.H. Li, X.X. Li, J.J. Hong, Y.X. Wang, J.B. Fu, H. Yang, C.Y. Yu, F.C. Li, J. Hu, W. W. Xue, Y.Y. Jiang, Y.Z. Chen, F. Zhu, Clinical trials, progression-speed differentiating features and swiftness rule of the innovative targets of first-in-class drugs, Brief, Bioinformatics 21 (2020) 649–662.
[10] K. Amahong, W. Zhang, Y. Zhou, S. Zhang, J. Yin, F. Li, H. Xu, T. Yan, Z. Yue, Y. Liu, T. Hou, Y. Qiu, L. Tao, L. Han, F. Zhu, CovInter: interaction data between coronavirus RNAs and host proteins, Nucleic Acids Res. 51 (2023) gkac834.
[11] J.P. Jourdan, R. Bureau, C. Rochais, P. Dallemagne, Drug repositioning: a brief overview, J. Pharm. Pharmacol. 72 (2020) 1145–1151.

[12] S. Pushpakom, F. Iorio, P.A. Eyers, K.J. Escott, S. Hopper, A. Wells, A. Doig, T. Guilliams, J. Latimer, C. McNamee, A. Norris, P. Sanseau, D. Cavalla, M. Pirmohamed, Drug repurposing: progress, challenges and recommendations, Nat. Rev. Drug Discov. 18 (2019) 41–58.
[13] M. Lotfi Shahreza, N. Ghadiri, S.R. Mousavi, J. Varshosaz, J.R. Green, A review of network-based approaches to drug repositioning, Briefings Bioinf. 19 (2018) 878–892.
[14] N. Novac, Challenges and opportunities of drug repositioning, Trends Pharmacol. Sci. 34 (2013) 267–272.
[15] T.N. Jarada, J.G. Rokne, R. Alhajj, A review of computational drug repositioning: strategies, approaches, opportunities, challenges, and directions, J. Cheminf. 12 (2020) 46.
[16] W. Sun, P.E. Sanderson, W. Zheng, Drug combination therapy increases successful drug repositioning, Drug Discov. Today 21 (2016) 1189–1195.
[17] C. Xu, D. Ai, D. Shi, S. Suo, X. Chen, Y. Yan, Y. Cao, R. Zhang, N. Sun, W. Chen, J. McDermott, S. Zhang, Y. Zeng, J.J. Han, Accurate drug repositioning through non-tissue-specific core signatures from cancer transcriptomes, Cell Rep. 29 (2019) 1055.
[18] F. Lo-Coco, G. Avvisati, M. Vignetti, C. Thiede, S.M. Orlando, S. Iacobelli, F. Ferrara, P. Fazi, L. Cicconi, E. Di Bona, G. Specchia, S. Sica, M. Divona, A. Levis, W. Fiedler, E. Cerqui, M. Breccia, G. Fioritoni, H.R. Salih, M. Cazzola, L. Melillo, A.M. Carella, C.H. Brandts, E. Morra, M. von Lilienfeld-Toal, B. Hertenstein, M. Wattad, M. Lübbert, M. Hänel, N. Schmitz, H. Link, M.G. Kropp, A. Rambaldi, G. La Nasa, M. Luppi, F. Ciceri, O. Finizio, A. Venditti, F. Fabbiano, K. Döhner, M. Sauer, A. Ganser, S. Amadori, F. Mandelli, H. Döhner, G. Ehninger, R. F. Schlenk, U. Platzbecker, Retinoic acid and arsenic trioxide for acute promyelocytic leukemia, N. Engl. J. Med. 369 (2013) 111–121.
[19] U. Bharadwaj, T.K. Eckols, M. Kolosov, M.M. Kasembeli, A. Adam, D. Torres, X. Zhang, L.E. Dobrolecki, W. Wei, M.T. Lewis, B. Dave, J.C. Chang, M.D. Landis, C.J. Creighton, M.A. Mancini, D.J. Tweardy, Drug-repositioning screening identified piperlongumine as a direct STAT3 inhibitor with potent activity against breast cancer, Oncogene 34 (2015) 1341–1353.
[20] O.S. Kwon, H. Lee, H.J. Kong, E.J. Kwon, J.E. Park, W. Lee, S. Kang, M. Kim, W. Kim, H.J. Cha, Connectivity map-based drug repositioning of bortezomib to reverse the metastatic effect of GALNT14 in lung cancer, Oncogene 39 (2020) 4567–4580.
[21] L. Huang, S. Garrett Injac, K. Cui, F. Braun, Q. Lin, Y. Du, H. Zhang, M. Kogiso, H. Lindsay, S. Zhao, P. Baxter, A. Adekunle, T.K. Man, H. Zhao, X.N. Li, C.C. Lau, S.T.C. Wong, Systems biology-based drug repositioning identifies digoxin as a potential therapy for groups 3 and 4 medulloblastoma, Sci. Transl. Med. 10 (2018), eaat0150.
[22] J.C. Siavelis, M.M. Bourdakou, E.I. Athanasiadis, G.M. Spyrou, K.S. Nikita, Bioinformatics methods in drug repurposing for Alzheimer's disease, Briefings Bioinf. 17 (2016) 322–335.
[23] H.C. So, C.K. Chau, W.T. Chiu, K.S. Ho, C.P. Lo, S.H. Yim, P.C. Sham, Analysis of genome-wide association data highlights candidates for drug repositioning in psychiatry, Nat. Neurosci. 20 (2017) 1342–1349.
[24] H.C. So, C.K. Chau, A. Lau, S.Y. Wong, K. Zhao, Translating GWAS findings into therapies for depression and anxiety disorders: gene-set analyses reveal enrichment of psychiatric drug classes and implications for drug repositioning, Psychol. Med. 49 (2019) 2692–2708.
[25] M.M. Ghahremanpour, J. Tirado-Rives, M. Deshmukh, J.A. Ippolito, C.H. Zhang, I. Cabeza de Vaca, M.E. Liosi, K.S. Anderson, W.L. Jorgensen, Identification of 14 known drugs as inhibitors of the main protease of SARS-CoV-2, ACS Med. Chem. Lett. 11 (2020) 2526–2533.
[26] Y.Y. Wang, C. Cui, L. Qi, H. Yan, X.M. Zhao, DrPOCS: drug repositioning based on projection onto convex sets, IEEE ACM Trans. Comput. Biol. Bioinf 16 (2019) 154–162.
[27] A. Nzila, Z. Ma, K. Chibale, Drug repositioning in the treatment of malaria and TB, Future Med. Chem. 3 (2011) 1413–1426.
[28] K.M. Kingsmore, A.C. Grammer, P.E. Lipsky, Drug repurposing to improve treatment of rheumatic autoimmune inflammatory diseases, Nat. Rev. Rheumatol. 16 (2020) 32–52.
[29] M. Weeber, R. Vos, H. Klein, L.T. De Jong-Van Den Berg, A.R. Aronson, G. Molema, Generating hypotheses by discovering implicit associations in the literature: a case report of a search for new potential therapeutic uses for thalidomide, J. Am. Med. Inf. Assoc. 10 (2003) 252–259.
[30] M. Bagherian, E. Sabeti, K. Wang, M.A. Sartor, Z. Nikolovska-Coleska, K. Najarian, Machine learning approaches and databases for prediction of drug-target interaction: a survey paper, Briefings Bioinf. 22 (2021) 247–269.
[31] B. Li, J. Tang, Q. Yang, S. Li, X. Cui, Y. Li, Y. Chen, W. Xue, X. Li, F. Zhu, NOREVA: normalization and evaluation of MS-based metabolomics data, Nucleic Acids Res. 45 (2017) W162–W170.
[32] Q. Yang, Y. Wang, Y. Zhang, F. Li, W. Xia, Y. Zhou, Y. Qiu, H. Li, F. Zhu, NOREVA: enhanced normalization and evaluation of time-course and multi-class metabolomic data, Nucleic Acids Res. 48 (2020) W436–W448.
[33] J. Fu, Y. Zhang, Y. Wang, H. Zhang, J. Liu, J. Tang, Q. Yang, H. Sun, W. Qiu, Y. Ma, Z. Li, M. Zheng, F. Zhu, Optimization of metabolomic data processing using NOREVA, Nat. Protoc. 17 (2022) 129–151.
[34] J. Tang, Y. Wang, J. Fu, Y. Zhou, Y. Luo, Y. Zhang, B. Li, Q. Yang, W. Xue, Y. Lou, Y. Qiu, F. Zhu, A critical assessment of the feature selection methods used for biomarker discovery in current metaproteomics studies, Briefings Bioinf. 21 (2020) 1378–1390.
[35] J. Tang, J. Fu, Y. Wang, B. Li, Y. Li, Q. Yang, X. Cui, J. Hong, X. Li, Y. Chen, W. Xue, F. Zhu, ANPELA: analysis and performance assessment of the label-free quantification workflow for metaproteomic studies, Briefings Bioinf. 21 (2020) 621–636.
[36] J. Fu, Y. Zhang, J. Liu, X. Lian, J. Tang, F. Zhu, Pharmacometabonomics: data processing and statistical analysis, Briefings Bioinf. 22 (2021) bbab138.
[37] M. Mou, Z. Pan, M. Lu, H. Sun, Y. Wang, Y. Luo, F. Zhu, Application of machine learning in spatial proteomics, J. Chem. Inf. Model. (2022) 5875–5895.
[38] S. Zhang, X. Sun, M. Mou, K. Amahong, H. Sun, W. Zhang, S. Shi, Z. Li, J. Gao, F. Zhu, REGLIV: molecular regulation data of diverse living systems facilitating current multiomics research, Comput. Biol. Med. 148 (2022), 105825.
[39] J. Fu, Q. Yang, Y. Luo, S. Zhang, J. Tang, Y. Zhang, H. Zhang, H. Xu, F. Zhu, Label-free proteome quantification and evaluation, Briefings Bioinf. 23 (2022) bbac477.
[40] J. Tang, M. Mou, Y. Wang, Y. Luo, F. Zhu, MetaFS: performance assessment of biomarker discovery in metaproteomics, Briefings Bioinf. 22 (2021) bbaa105.
[41] Q. Yang, B. Li, S. Chen, J. Tang, Y. Li, Y. Li, S. Zhang, C. Shi, Y. Zhang, M. Mou, W. Xue, F. Zhu, MMEASE: online meta-analysis of metabolomic data by enhanced metabolite annotation, marker selection and enrichment analysis, J. Proteonomics 232 (2021), 104023.
[42] Q. Yang, B. Li, P. Wang, J. Xie, Y. Feng, Z. Liu, F. Zhu, LargeMetabo: an out-of-the-box tool for processing and analyzing large-scale metabolomic data, Briefings Bioinf. 23 (2022) bbac455.
[43] Y.H. Li, C.Y. Yu, X.X. Li, P. Zhang, J. Tang, Q. Yang, T. Fu, X. Zhang, X. Cui, G. Tu, Y. Zhang, S. Li, F. Yang, Q. Sun, C. Qin, X. Zeng, Z. Chen, Y.Z. Chen, F. Zhu, Therapeutic target database update 2018: enriched resource for facilitating bench-to-clinic research of targeted therapeutics, Nucleic Acids Res. 46 (2018) D1121–D1127.
[44] J. Yin, W. Sun, F. Li, J. Hong, X. Li, Y. Zhou, Y. Lu, M. Liu, X. Zhang, N. Chen, X. Jin, J. Xue, S. Zeng, L. Yu, F. Zhu, Varidt 1.0: variability of drug transporter database, Nucleic Acids Res. 48 (2020) D1042–D1050.
[45] J. Yin, F. Li, Y. Zhou, M. Mou, Y. Lu, K. Chen, J. Xue, Y. Luo, J. Fu, X. He, J. Gao, S. Zeng, L. Yu, F. Zhu, INTEDE: interactome of drug-metabolizing enzymes, Nucleic Acids Res. 49 (2021) D1233–D1243.
[46] T. Fu, F. Li, Y. Zhang, J. Yin, W. Qiu, X. Li, X. Liu, W. Xin, C. Wang, L. Yu, J. Gao, Q. Zheng, S. Zeng, F. Zhu, Varidt 2.0: structural variability of drug transporter, Nucleic Acids Res. 50 (2022) D1417–D1431.
[47] W. Xue, T. Fu, S. Deng, F. Yang, J. Yang, F. Zhu, Molecular mechanism for the allosteric inhibition of the human serotonin transporter by antidepressant escitalopram, ACS Chem. Neurosci. 13 (2022) 340–351.
[48] F. Li, J. Yin, M. Lu, M. Mou, Z. Li, Z. Zeng, Y. Tan, S. Wang, X. Chu, H. Dai, T. Hou, S. Zeng, Y. Chen, F. Zhu, DrugMAP: molecular atlas and pharma-information of all drugs, Nucleic Acids Res. 51 (2023) gkac813.
[49] J. Rung, A. Brazma, Reuse of public genome-wide gene expression data, Nat. Rev. Genet. 14 (2013) 89–99.
[50] D. Tsoucas, R. Dong, H. Chen, Q. Zhu, G. Guo, G.-C. Yuan, Accurate estimation of cell-type composition from gene expression data, Nat. Commun. 10 (2019) 2975.
[51] Q. Yang, B. Li, J. Tang, X. Cui, Y. Wang, X. Li, J. Hu, Y. Chen, W. Xue, Y. Lou, Y. Qiu, F. Zhu, Consistent gene signature of schizophrenia identified by a novel feature selection strategy from comprehensive sets of transcriptomic data, Briefings Bioinf. 21 (2020) 1058–1068.
[52] X. Nie, J. Wei, Y. Hao, J. Tao, Y. Li, M. Liu, B. Xu, B. Li, Consistent biomarkers and related pathogenesis underlying asthma revealed by systems biology approach, Int. J. Mol. Sci. 20 (2019) 4037.
[53] X. Li, X. Li, Y. Li, C. Yu, W. Xue, J. Hu, B. Li, P. Wang, F. Zhu, What makes species productive of anti-cancer drugs? Clues from drugs' species origin, druglikeness, target and pathway, Anti Cancer Agents Med. Chem. 19 (2019) 194–203.
[54] C. Khunsriraksakul, D. McGuire, R. Sauteraud, F. Chen, L. Yang, L. Wang, J. Hughey, S. Eckert, J. Dylan Weissenkampen, G. Shenoy, O. Marx, L. Carrel, B. Jiang, D.J. Liu, Integrating 3D genomic and epigenomic data to enhance target gene discovery and drug repurposing in transcriptome-wide association studies, Nat. Commun. 13 (2022) 3258.
[55] T.R. Hughes, M.J. Marton, A.R. Jones, C.J. Roberts, R. Stoughton, C.D. Armour, H.A. Bennett, E. Coffey, H. Dai, Y.D. He, M.J. Kidd, A.M. King, M.R. Meyer, D. Slade, P.Y. Lum, S.B. Stepaniants, D.D. Shoemaker, D. Gachotte, K. Chakraburtty, J. Simon, M. Bard, S.H. Friend, Functional discovery via a compendium of expression profiles, Cell 102 (2000) 109–126.
[56] A. Engelberg, Iconix Pharmaceuticals, Inc.–removing barriers to efficient drug discovery through chemogenomics, Pharmacogenomics 5 (2004) 741–744.
[57] B. Ganter, S. Tugendreich, C.I. Pearson, E. Ayanoglu, S. Baumhueter, K. A. Bostian, L. Brady, L.J. Browne, J.T. Calvin, G.J. Day, N. Breckenridge, S. Dunlea, B.P. Eynon, L.M. Furness, J. Ferng, M.R. Fielden, S.Y. Fujimoto, L. Gong, C. Hu, R. Idury, M.S. Judo, K.L. Kolaja, M.D. Lee, C. McSorley, J. M. Minor, R.V. Nair, G. Natsoulis, P. Nguyen, S.M. Nicholson, H. Pham, A. H. Roter, D. Sun, S. Tan, S. Thode, A.M. Tolley, A. Vladimirova, J. Yang, Z. Zhou, K. Jarnagin, Development of a large-scale chemogenomics database to improve drug candidate selection and to understand mechanisms of chemical toxicity and action, J. Biotechnol. 119 (2005) 219–244.
[58] J. Lamb, E.D. Crawford, D. Peck, J.W. Modell, I.C. Blat, M.J. Wrobel, J. Lerner, J. P. Brunet, A. Subramanian, K.N. Ross, M. Reich, H. Hieronymus, G. Wei, S. A. Armstrong, S.J. Haggarty, P.A. Clemons, R. Wei, S.A. Carr, E.S. Lander, T. R. Golub, The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease, Science 313 (2006) 1929–1935.
[59] G. Hu, P. Agarwal, Human disease-drug network based on genomic expression profiles, PLoS One 4 (2009), e6536.
[60] Y. Xiao, Y. Gong, Y. Lv, Y. Lan, J. Hu, F. Li, J. Xu, J. Bai, Y. Deng, L. Liu, G. Zhang, F. Yu, X. Li, Gene Perturbation Atlas (GPA): a single-gene perturbation repository for characterizing functional mechanisms of coding and non-coding genes, Sci. Rep. 5 (2015), 10889.

[61] A. Subramanian, R. Narayan, S.M. Corsello, D.D. Peck, T.E. Natoli, X. Lu, J. Gould, J.F. Davis, A.A. Tubelli, J.K. Asiedu, D.L. Lahr, J.E. Hirschman, Z. Liu, M. Donahue, B. Julian, M. Khan, D. Wadden, I.C. Smith, D. Lam, A. Liberzon, C. Toder, M. Bagul, M. Orzechowski, O.M. Enache, F. Piccioni, S.A. Johnson, N. J. Lyons, A.H. Berger, A.F. Shamji, A.N. Brooks, A. Vrcic, C. Flynn, J. Rosains, D. Y. Takeda, R. Hu, D. Davison, J. Lamb, K. Ardlie, L. Hogstrom, P. Greenside, N. S. Gray, P.A. Clemons, S. Silver, X. Wu, W.N. Zhao, W. Read-Button, X. Wu, S. J. Haggarty, L.V. Ronco, J.S. Boehm, S.L. Schreiber, J.G. Doench, J.A. Bittker, D. E. Root, B. Wong, T.R. Golub, A next generation connectivity map: L1000 platform and the first 1,000,000 profiles, Cell 171 (2017) 1437–1452, e1417.

[62] T. Barrett, S.E. Wilhite, P. Ledoux, C. Evangelista, I.F. Kim, M. Tomashevsky, K. A. Marshall, K.H. Phillippy, P.M. Sherman, M. Holko, A. Yefanov, H. Lee, N. Zhang, C.L. Robertson, N. Serova, S. Davis, A. Soboleva, NCBI GEO: archive for functional genomics data sets–update, Nucleic Acids Res. 41 (2013) D991–D995.

[63] A. Athar, A. Füllgrabe, N. George, H. Iqbal, L. Huerta, A. Ali, C. Snow, N. A. Fonseca, R. Petryszak, I. Papatheodorou, U. Sarkans, A. Brazma, ArrayExpress update - from bulk to single-cell expression data, Nucleic Acids Res. 47 (2019) D711–D715.

[64] I. Papatheodorou, P. Moreno, J. Manning, A.M. Fuentes, N. George, S. Fexova, N. A. Fonseca, A. Füllgrabe, M. Green, N. Huang, L. Huerta, H. Iqbal, M. Jianu, S. Mohammed, L. Zhao, A.F. Jarnuczak, S. Jupp, J. Marioni, K. Meyer, R. Petryszak, C.A. Prada Medina, C. Talavera-López, S. Teichmann, J.A. Vizcaino, A. Brazma, Expression Atlas update: from tissues to single cells, Nucleic Acids Res. 48 (2020) D77–D83.

[65] Gtex-Consortium, Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans, Science 348 (2015) 648–660.

[66] M. Ghandi, F.W. Huang, J. Jané-Valbuena, G.V. Kryukov, C.C. Lo, E. R. McDonald 3rd, J. Barretina, E.T. Gelfand, C.M. Bielski, H. Li, K. Hu, A. Y. Andreev-Drakhlin, J. Kim, J.M. Hess, B.J. Haas, F. Aguet, B.A. Weir, M. V. Rothberg, B.R. Paolella, M.S. Lawrence, R. Akbani, Y. Lu, H.L. Tiv, P. C. Gokhale, A. de Weck, A.A. Mansour, C. Oh, J. Shih, K. Hadi, Y. Rosen, J. Bistline, K. Venkatesan, A. Reddy, D. Sonkin, M. Liu, J. Lehar, J.M. Korn, D. A. Porter, M.D. Jones, J. Golji, G. Caponigro, J.E. Taylor, C.M. Dunning, A. L. Creech, A.C. Warren, J.M. McFarland, M. Zamanighomi, A. Kauffmann, N. Stransky, M. Imielinski, Y.E. Maruvka, A.D. Cherniack, A. Tsherniak, F. Vazquez, J.D. Jaffe, A.A. Lane, D.M. Weinstock, C.M. Johannessen, M. P. Morrissey, F. Stegmeier, R. Schlegel, W.C. Hahn, G. Getz, G.B. Mills, J. S. Boehm, T.R. Golub, L.A. Garraway, W.R. Sellers, Next-generation characterization of the cancer cell line encyclopedia, Nature 569 (2019) 503–508.

[67] C. Hutter, J.C. Zenklusen, The cancer genome atlas: creating lasting value beyond its data, Cell 173 (2018) 283–285.

[68] J. Zhang, R. Bajari, D. Andric, F. Gerthoffert, A. Lepsa, H. Nahal-Bose, L.D. Stein, V. Ferretti, The international cancer genome consortium data portal, Nat. Biotechnol. 37 (2019) 367–369.

[69] R.M. Baldarelli, C.M. Smith, J.H. Finger, T.F. Hayamizu, I.J. McCright, J. Xu, D. R. Shaw, J.S. Beal, O. Blodgett, J. Campbell, L.E. Corbani, P.J. Frost, S. C. Giannatto, D.B. Miers, J.A. Kadin, J.E. Richardson, M. Ringwald, The mouse gene expression database (GXD): 2021 update, Nucleic Acids Res. 49 (2021) D924–D931.

[70] S.J. Park, B.H. Yoon, S.K. Kim, S.Y. Kim, GENT2: an updated gene expression database for normal and tumor tissues, BMC Med. Genom. 12 (2019) 101.

[71] X. Sheng, J. Wu, Q. Sun, X. Li, F. Xian, M. Sun, W. Fang, M. Chen, J. Yu, J. Xiao, MTD: a mammalian transcriptomic database to explore gene expression and regulation, Briefings Bioinf. 18 (2017) 28–36.

[72] A. Musa, L.S. Ghoraie, S.D. Zhang, G. Glazko, O. Yli-Harja, M. Dehmer, B. Haibe-Kains, F. Emmert-Streib, A review of connectivity map and computational approaches in pharmacogenomics, Briefings Bioinf. 19 (2018) 506–523.

[73] Y. Igarashi, N. Nakatsu, T. Yamashita, A. Ono, Y. Ohno, T. Urushidani, H. Yamada, Open Tg-Gates, A large-scale toxicogenomics database, Nucleic Acids Res. 43 (2015) D921–D927.

[74] R.L. Wang, A.D. Biales, N. Garcia-Reyero, E.J. Perkins, D.L. Villeneuve, G. T. Ankley, D.C. Bencic, Fish connectivity mapping: linking chemical stressors by their mechanisms of action-driven transcriptomic profiles, BMC Genom. 17 (2016) 84.

[75] Z. Wang, C.D. Monteiro, K.M. Jagodnik, N.F. Fernandez, G.W. Gundersen, A. D. Rouillard, S.L. Jenkins, A.S. Feldmann, K.S. Hu, M.G. McDermott, Q. Duan, N. R. Clark, M.R. Jones, Y. Kou, T. Goff, H. Woodland, F.M.R. Amaral, G.L. Szeto, O. Fuchs, S.M. Schussler-Fiorenza Rose, S. Sharma, U. Schwartz, X.B. Bausela, M. Szymkiewicz, V. Maroulis, A. Salykin, C.M. Barra, C.D. Kruth, N.J. Bongio, V. Mathur, R.D. Todoric, U.E. Rubin, A. Malatras, C.T. Fulp, J.A. Galindo, R. Motiejunaite, C. Juschke, P.C. Dishuck, K. Lahl, M. Jafari, S. Aibar, A. Zaravinos, L.H. Steenhuizen, L.R. Allison, P. Gamallo, F. de Andres Segura, T. Dae Devlin, V. Perez-Garcia, A. Ma'ayan, Extraction and analysis of signatures from the gene expression omnibus by the crowd, Nat. Commun. 7 (2016), 12846.

[76] H. Wu, J. Huang, Y. Zhong, Q. Huang, DrugSig: a resource for computational drug repositioning utilizing gene expression signatures, PLoS One 12 (2017), e0177743.

[77] C. Ye, D.J. Ho, M. Neri, C. Yang, T. Kulkarni, R. Randhawa, M. Henault, N. Mostacci, P. Farmer, S. Renner, R. Ihry, L. Mansur, C.G. Keller, G. McAllister, M. Hild, J. Jenkins, A. Kaykas, DRUG-seq for miniaturized high-throughput transcriptome profiling in drug discovery, Nat. Commun. 9 (2018) 4307.

[78] S.M. Corsello, J.A. Bittker, Z. Liu, J. Gould, P. McCarren, J.E. Hirschman, S. E. Johnston, A. Vrcic, B. Wong, M. Khan, J. Asiedu, R. Narayan, C.C. Mader, A. Subramanian, T.R. Golub, The Drug Repurposing Hub: a next-generation drug library and information resource, Nat. Med. 23 (2017) 405–408.

[79] M. Haghighi, J.C. Caicedo, B.A. Cimini, A.E. Carpenter, S. Singh, High-dimensional gene expression and morphology profiles of cells across 28,000 genetic and chemical perturbations, Nat. Methods 19 (2022) 1550–1557.

[80] S. Fang, L. Dong, L. Liu, J. Guo, L. Zhao, J. Zhang, D. Bu, X. Liu, P. Huo, W. Cao, Q. Dong, J. Wu, X. Zeng, Y. Wu, Y. Zhao, HERB: a high-throughput experiment- and reference-guided database of traditional Chinese medicine, Nucleic Acids Res. 49 (2021) D1197–D1206.

[81] P. Li, H. Zhang, W. Zhang, Y. Zhang, L. Zhan, N. Wang, C. Chen, B. Fu, J. Zhao, X. Zhou, S. Guo, J. Chen, TMNP: a transcriptome-based multi-scale network pharmacology platform for herbal medicine, Briefings Bioinf. 23 (2022).

[82] D. Gusenleitner, S.S. Auerbach, T. Melia, H.F. Gomez, D.H. Sherr, S. Monti, Genomic models of short-term exposure accurately predict long-term chemical carcinogenicity and identify putative mechanisms of action, PLoS One 9 (2014), e102579.

[83] D.A. Cusanovich, B. Pavlovic, J.K. Pritchard, Y. Gilad, The functional consequences of variation in transcription factor binding, PLoS Genet. 10 (2014), e1004226.

[84] W. Senkowski, M. Jarvius, J. Rubin, J. Lengqvist, M.G. Gustafsson, P. Nygren, K. Kultima, R. Larsson, M. Fryknas, Large-scale gene expression profiling platform for identification of context-dependent drug responses in multicellular tumor spheroids, Cell Chem. Biol. 23 (2016) 1428–1438.

[85] A. Suzuki, K. Onodera, K. Matsui, M. Seki, H. Esumi, T. Soga, S. Sugano, T. Kohno, Y. Suzuki, K. Tsuchihara, Characterization of cancer omics and drug perturbations in panels of lung cancer cells, Sci. Rep. 9 (2019), 19529.

[86] M. Zheng, S. Okawa, M. Bravo, F. Chen, M.-L. Martínez-Chantar, A. del Sol, ChemPert: mapping between chemical perturbation and transcriptional response for non-cancer cells, Nucleic Acids Res. 51 (2023) D877–D889.

[87] Q. Duan, C. Flynn, M. Niepel, M. Hafner, J.L. Muhlich, N.F. Fernandez, A. D. Rouillard, C.M. Tan, E.Y. Chen, T.R. Golub, P.K. Sorger, A. Subramanian, A. Ma'ayan, LINCS Canvas Browser, Interactive web app to query, browse and interrogate LINCS L1000 gene expression signatures, Nucleic Acids Res. 42 (2014) W449–w460.

[88] Q. Duan, S.P. Reid, N.R. Clark, Z. Wang, N.F. Fernandez, A.D. Rouillard, B. Readhead, S.R. Tritsch, R. Hodos, M. Hafner, M. Niepel, P.K. Sorger, J. T. Dudley, S. Bavari, R.G. Panchal, A. Ma'ayan, L1000CDS(2): LINCS L1000 characteristic direction signatures search engine, NPJ. Syst. Biol. Appl. 2 (2016), 16015.

[89] P. Smirnov, Z. Safikhani, N. El-Hachem, D. Wang, A. She, C. Olsen, M. Freeman, H. Selby, D.M. Gendoo, P. Grossmann, A.H. Beck, H.J. Aerts, M. Lupien, A. Goldenberg, B. Haibe-Kains, PharmacoGx: an R package for analysis of large pharmacogenomic datasets, Bioinformatics 32 (2016) 1244–1246.

[90] L. Huang, F. Li, J. Sheng, X. Xia, J. Ma, M. Zhan, S.T.C. Wong, DrugComboRanker: drug combination discovery based on target network analysis, Bioinformatics 30 (2014) i228–i236.

[91] Z. Wang, N.R. Clark, A. Ma'ayan, Drug-induced adverse events prediction with the LINCS L1000 data, Bioinformatics 32 (2016) 2338–2345.

[92] D. Carrella, F. Napolitano, R. Rispoli, M. Miglietta, A. Carissimo, L. Cutillo, F. Sirci, F. Gregoretti, D. Di Bernardo, Mantra 2.0: an online collaborative resource for drug mode of action and repurposing by network analysis, Bioinformatics 30 (2014) 1787–1788.

[93] C. Pacini, F. Iorio, E. Gonçalves, M. Iskar, T. Klabunde, P. Bork, J. Saez-Rodriguez, DvD, An R/Cytoscape pipeline for drug repurposing using public repositories of gene expression data, Bioinformatics 29 (2013) 132–134.

[94] F. Napolitano, F. Sirci, D. Carrella, D. di Bernardo, Drug-set enrichment analysis: a novel tool to investigate drug mode of action, Bioinformatics 32 (2016) 235–241.

[95] Z. Wang, A. Lachmann, A.B. Keenan, A. Ma'ayan, L1000FWD, Fireworks visualization of drug-induced transcriptomic signatures, Bioinformatics 34 (2018) 2150–2152.

[96] A.S. Brown, S.W. Kong, I.S. Kohane, C.J. Patel, ksRepo: a generalized platform for computational drug repositioning, BMC Bioinf. 17 (2016) 78.

[97] F. Napolitano, D. Carrella, B. Mandriani, S. Pisonero-Vaquero, F. Sirci, D. L. Medina, N. Brunetti-Pierri, D. di Bernardo, gene2drug, A computational tool for pathway-based rational drug repositioning, Bioinformatics 34 (2018) 1498–1505.

[98] M. Pilarczyk, M. Kouril, B. Shamsaei, J. Vasiliauskas, W. Niu, N. Mahi, L. Zhang, N. Clark, Y. Ren, S. White, Connecting Omics Signatures of Diseases, Drugs, and Mechanisms of Actions with iLINCS, BioRxiv, 2020, 826271.

[99] B.K. Lee, K.H. Tiong, J.K. Chang, C.S. Liew, Z.A. Abdul Rahman, A.C. Tan, T. F. Khang, S.C. Cheong, DeSigN: connecting gene expression with therapeutics for drug repurposing and development, BMC Genom. 18 (2017) 934.

[100] Z. Jia, Y. Liu, N. Guan, X. Bo, Z. Luo, M.R. Barnes, Cogena, a novel tool for co-expressed gene-set enrichment analysis, applied to drug repositioning and drug mode of action discovery, BMC Genom. 17 (2016) 414.

[101] J. Chen, E.E. Bardes, B.J. Aronow, A.G. Jegga, ToppGene Suite for gene list enrichment analysis and candidate gene prioritization, Nucleic Acids Res. 37 (2009) W305–w311.

[102] F. Li, J. Yin, M. Lu, Q. Yang, Z. Zeng, B. Zhang, Z. Li, Y. Qiu, H. Dai, Y. Chen, F. Zhu, ConSIG: consistent discovery of molecular signature from OMIC data, Briefings Bioinf. 23 (2022) bbac253.

[103] F. Li, Y. Zhou, Y. Zhang, J. Yin, Y. Qiu, J. Gao, F. Zhu, POSREG: proteomic signature discovered by simultaneously optimizing its reproducibility and generalizability, Briefings Bioinf. 23 (2022) bbac040.

[104] C.H. Huang, M.Y. Wu, P.M. Chang, C.Y. Huang, K.L. Ng, In silico identification of potential targets and drugs for non-small cell lung cancer, IET Syst. Biol. 8 (2014) 56–66.

[105] J.T. Dudley, M. Sirota, M. Shenoy, R.K. Pai, S. Roedder, A.P. Chiang, A.A. Morgan, M.M. Sarwal, P.J. Pasricha, A.J. Butte, Computational repositioning of the anticonvulsant topiramate for inflammatory bowel disease, Sci. Transl. Med. 3 (2011), 96ra76.
[106] D. Toro-Domínguez, M.E. Alarcón-Riquelme, P. Carmona-Sáez, Drug repurposing from transcriptome data: methods and applications, in: K. Roy (Ed.), In Silico Drug Design, Academic Press, 2019, pp. 303–327.
[107] F. Iorio, T. Rittman, H. Ge, M. Menden, J. Saez-Rodriguez, Transcriptional data: a new gateway to drug repositioning? Drug Discov. Today 18 (2013) 350–357.
[108] Y.F. Huang, H.Y. Yeh, V.W. Soo, Inferring drug-disease associations from integration of chemical, genomic and phenotype data using network propagation, BMC Med. Genom. 6 (Suppl 3) (2013) S4.
[109] V. van Noort, S. Schölch, M. Iskar, G. Zeller, K. Ostertag, C. Schweitzer, K. Werner, J. Weitz, M. Koch, P. Bork, Novel drug candidates for the treatment of metastatic colorectal cancer through global inverse gene-expression profiling, Cancer Res. 74 (2014) 5690–5699.
[110] J.S. Pang, Z.K. Li, P. Lin, X.D. Wang, G. Chen, H.B. Yan, S.H. Li, The underlying molecular mechanism and potential drugs for treatment in papillary renal cell carcinoma: a study based on TCGA and Cmap datasets, Oncol. Rep. 41 (2019) 2089–2102.
[111] R. Iyengar, Complex diseases require complex therapies, EMBO Rep. 14 (2013) 1039–1042.
[112] J.B. Fitzgerald, B. Schoeberl, U.B. Nielsen, P.K. Sorger, Systems biology and combination therapy in the quest for clinical efficacy, Nat. Chem. Biol. 2 (2006) 458–466.
[113] A.I. Casas, A.A. Hassan, S.J. Larsen, V. Gomez-Rangel, M. Elbatreek, P.W. M. Kleikers, E. Guney, J. Egea, M.G. López, J. Baumbach, H. Schmidt, From single drug targets to synergistic network pharmacology in ischemic stroke, Proc. Natl. Acad. Sci. U.S.A. 116 (2019) 7129–7136.
[114] X. Sun, Y. Zhang, Y. Zhou, X. Lian, L. Yan, T. Pan, T. Jin, H. Xie, Z. Liang, W. Qiu, J. Wang, Z. Li, F. Zhu, X. Sui, NPCDR: natural product-based drug combination and its disease-specific molecular regulation, Nucleic Acids Res. 50 (2022) D1324–D1333.
[115] J.D. Jordan, E.M. Landau, R. Iyengar, Signaling networks: the origins of cellular multitasking, Cell 103 (2000) 193–200.
[116] L. Wu, Y. Wen, D. Leng, Q. Zhang, C. Dai, Z. Wang, Z. Liu, B. Yan, Y. Zhang, J. Wang, S. He, X. Bo, Machine learning methods, databases and tools for drug combination prediction, Briefings Bioinf. 23 (2022) bbab355.
[117] J.G.C. van Hasselt, R. Iyengar, Systems pharmacology: defining the interactions of drug combinations, Annu. Rev. Pharmacol. Toxicol. 59 (2019) 21–40.
[118] E.K. Gupta, M.K. Ito, Lovastatin and extended-release niacin combination product: the first drug combination for the management of hyperlipidemia, Heart Dis. 4 (2002) 124–137.
[119] M. Tyers, G.D. Wright, Drug combinations: a strategy to extend the life of antibiotics in the 21st century, Nat. Rev. Microbiol. 17 (2019) 141–155.
[120] Y. Acharya, S. Bhattacharyya, G. Dhanda, J. Haldar, Emerging roles of glycopeptide antibiotics: moving beyond gram-positive bacteria, ACS Infect. Dis. 8 (2022) 1–28.
[121] B. Al-Lazikani, U. Banerji, P. Workman, Combinatorial drug therapy for cancer in the post-genomic era, Nat. Biotechnol. 30 (2012) 679–692.
[122] C.T. Huang, C.H. Hsieh, Y.H. Chung, Y.J. Oyang, H.C. Huang, H.F. Juan, Perturbational gene-expression signatures for combinatorial drug discovery, iScience 15 (2019) 291–306.
[123] R.M. Webster, Combination therapies in oncology, Nat. Rev. Drug Discov. 15 (2016) 81–82.
[124] J.H. Lee, D.G. Kim, T.J. Bae, K. Rho, J.T. Kim, J.J. Lee, Y. Jang, B.C. Kim, K. M. Park, S. Kim, CDA: combinatorial drug discovery using transcriptional response modules, PLoS One 7 (2012), e42573.
[125] A. Anighoro, J. Bajorath, G. Rastelli, Polypharmacology: challenges and opportunities in drug discovery, J. Med. Chem. 57 (2014) 7874–7887.
[126] A. Talevi, Multi-target pharmacology: possibilities and limitations of the "skeleton key approach" from a medicinal chemist perspective, Front. Pharmacol. 6 (2015) 205.
[127] A. Flemming, Finding the perfect combination, Nat. Rev. Drug Discov. 14 (2015), 13-13.
[128] X. Sun, Y. Zhang, H. Li, Y. Zhou, S. Shi, Z. Chen, X. He, H. Zhang, F. Li, J. Yin, M. Mou, Y. Wang, Y. Qiu, F. Zhu, DRESIS: the first comprehensive landscape of drug resistance information, Nucleic Acids Res. 51 (2023) gkac812.
[129] K.L. Yeo, Y.L. Chen, H.Y. Xu, H. Dong, Q.Y. Wang, F. Yokokawa, P.Y. Shi, Synergistic suppression of dengue virus replication using a combination of nucleoside analogs and nucleoside synthesis inhibitors, Antimicrob. Agents Chemother. 59 (2015) 2086–2093.
[130] S.C.M. Kumar, Drug resistance in malaria, in: G. Arora, A. Sajid, V.C. Kalia (Eds.), Drug Resistance in Bacteria, Fungi, Malaria, and Cancer, Springer International Publishing, Cham, 2017, pp. 429–447.
[131] M. Cokol, H.N. Chua, M. Tasan, B. Mutlu, Z.B. Weinstein, Y. Suzuki, M.E. Nergiz, M. Costanzo, A. Baryshnikova, G. Giaever, C. Nislow, C.L. Myers, B.J. Andrews, C. Boone, F.P. Roth, Systematic exploration of synergistic drug pairs, Mol. Syst. Biol. 7 (2011) 544.
[132] P.J. Yeh, M.J. Hegreness, A.P. Aiden, R. Kishony, Drug interactions and the evolution of antibiotic resistance, Nat. Rev. Microbiol. 7 (2009) 460–466.
[133] A. Alani, R. Seymour, Systemic medication and the inflammatory cascade, Periodontol 2000 (64) (2014) 198–210.
[134] B. Singh, S. Maharjan, D.C. Pan, Z. Zhao, Y. Gao, Y.S. Zhang, S. Mitragotri, Imiquimod-gemcitabine nanoparticles harness immune cells to suppress breast cancer, Biomaterials 280 (2022), 121302.
[135] H. Zhao, K.K.W. To, H. Lam, X. Zhou, J.F.-W. Chan, Z. Peng, A.C.Y. Lee, J. Cai, W.-M. Chan, J.D. Ip, C.C.-S. Chan, M.L. Yeung, A.J. Zhang, A.W.H. Chu, S. Jiang, K.-Y. Yuen, Cross-linking peptide and repurposed drugs inhibit both entry pathways of SARS-CoV-2, Nat. Commun. 12 (2021) 1517.
[136] H. Seo, D. Tkachuk, C. Ho, A. Mammoliti, A. Rezaie, S.A. Madani Tonekaboni, B. Haibe-Kains, SYNERGxDB: an integrative pharmacogenomic portal to identify synergistic drug combinations for precision oncology, Nucleic Acids Res. 48 (2020) W494–W501.
[137] H. Liu, W. Zhang, B. Zou, J. Wang, Y. Deng, L. Deng, DrugCombDB: a comprehensive database of drug combinations toward the discovery of combinatorial therapy, Nucleic Acids Res. 48 (2020) D871–D881.
[138] A. Luna, F. Elloumi, S. Varma, Y. Wang, V.N. Rajapakse, M.I. Aladjem, J. Robert, C. Sander, Y. Pommier, W.C. Reinhold, CellMiner Cross-Database (CellMinerCDB) version 1.2: exploration of patient-derived cancer cell line pharmacogenomics, Nucleic Acids Res. 49 (2021) D1083–D1093.
[139] Y. Liu, Q. Wei, G. Yu, W. Gai, Y. Li, X. Chen, DCDB 2.0: a Major Update of the Drug Combination Database, 2014, Database, Oxford, 2014, p. bau124.
[140] A. Ianevski, A.K. Giri, T. Aittokallio, SynergyFinder 2.0: visual analytics of multi-drug combination synergies, Nucleic Acids Res. 48 (2020) W488–W493.
[141] B. Güvenç Paltun, S. Kaski, H. Mamitsuka, Machine learning approaches for drug combination therapies, Briefings Bioinf. 22 (2021) bbab293.
[142] M. Jeon, S. Kim, S. Park, H. Lee, J. Kang, In silico drug combination discovery for personalized cancer therapy, BMC Syst. Biol. 12 (2018) 16.
[143] O.I. Pulkkinen, P. Gautam, V. Mustonen, T. Aittokallio, Multiobjective optimization identifies cancer-selective combination therapies, PLoS Comput. Biol. 16 (2020), e1008538.
[144] X. Zeng, Z. Jia, Z. He, W. Chen, X. Lu, H. Duan, H. Li, Measure clinical drug-drug similarity using Electronic Medical Records, Int. J. Med. Inf. 124 (2019) 97–103.
[145] A.S. Brown, C.J. Patel, MeSHDD: literature-based drug-drug similarity for drug repositioning, J. Am. Med. Inf. Assoc. 24 (2017) 614–618.
[146] F. Iorio, R. Bosotti, E. Scacheri, V. Belcastro, P. Mithbaokar, R. Ferriero, L. Murino, R. Tagliaferri, N. Brunetti-Pierri, A. Isacchi, D. di Bernardo, Discovery of drug mode of action and drug repositioning from transcriptional responses, Proc. Natl. Acad. Sci. U.S.A. 107 (2010) 14621–14626.
[147] K. Cha, M.S. Kim, K. Oh, H. Shin, G.S. Yi, Drug similarity search based on combined signatures in gene expression profiles, Healthc. Inform. Res. 20 (2014) 52–60.
[148] G. Lim, C.J. Lim, J.H. Lee, B.H. Lee, J.Y. Ryu, K.S. Oh, Identification of new target proteins of a Urotensin-II receptor antagonist using transcriptome-based drug repositioning approach, Sci. Rep. 11 (2021), 17138.
[149] A. Özdemİr, M. Ark, A novel ROCK inhibitor: off-target effects of metformin, Turk. J. Biol. 45 (2021) 35–45.
[150] N.A. Pabon, Y. Xia, S.K. Estabrooks, Z. Ye, A.K. Herbrand, E. Süß, R.M. Biondi, V. A. Assimon, J.E. Gestwicki, J.L. Brodsky, C.J. Camacho, Z. Bar-Joseph, Predicting protein targets for drug-like compounds using transcriptomics, PLoS Comput. Biol. 14 (2018), e1006651.
[151] J.A. Wagner, Strategic approach to fit-for-purpose biomarkers in drug development, Annu. Rev. Pharmacol. Toxicol. 48 (2008) 631–651.
[152] A.P. Chiang, A.J. Butte, Systematic evaluation of drug-disease relationships to identify leads for novel drug uses, Clin. Pharmacol. Ther. 86 (2009) 507–510.
[153] L. Yu, S. Yao, L. Gao, Y. Zha, Conserved disease modules extracted from multilayer heterogeneous disease and gene networks for understanding disease mechanisms and predicting disease treatments, Front. Genet. 9 (2018) 745.
[154] J. Menche, A. Sharma, M. Kitsak, S.D. Ghiassian, M. Vidal, J. Loscalzo, A. L. Barabasi, Uncovering disease-disease relationships through the incomplete interactome, Science 347 (2015), 1257601.
[155] E. Tambuyzer, B. Vandendriessche, C.P. Austin, P.J. Brooks, K. Larsson, K.I. Miller Needleman, J. Valentine, K. Davies, S.C. Groft, R. Preti, T.I. Oprea, M. Prunotto, Therapies for rare diseases: therapeutic modalities, progress and challenges ahead, Nat. Rev. Drug Discov. 19 (2020) 93–111.
[156] S. Suthram, J.T. Dudley, A.P. Chiang, R. Chen, T.J. Hastie, A.J. Butte, Network-based elucidation of human disease similarities reveals common functional modules enriched for pluripotent drug targets, PLoS Comput. Biol. 6 (2010), e1000662.
[157] N.R. Clark, K.S. Hu, A.S. Feldmann, Y. Kou, E.Y. Chen, Q. Duan, A. Ma'ayan, The characteristic direction: a geometrical approach to identify differentially expressed genes, BMC Bioinf. 15 (2014) 79.
[158] O.S. Kwon, W. Kim, H.J. Cha, H. Lee, In silico drug repositioning: from large-scale transcriptome data to therapeutics, Arch Pharm. Res. (Seoul) 42 (2019) 879–889.
[159] J. Hong, Y. Luo, Y. Zhang, J. Ying, W. Xue, T. Xie, L. Tao, F. Zhu, Protein functional annotation of simultaneously improved stability, accuracy and false discovery rate achieved by a sequence-based deep learning, Briefings Bioinf. 21 (2020) 1437–1447.
[160] W. Xia, L. Zheng, J. Fang, F. Li, Y. Zhou, Z. Zeng, B. Zhang, Z. Li, H. Li, F. Zhu, PFmulDL: a novel strategy enabling multi-class and multi-label protein function annotation by integrating diverse deep learning methods, Comput. Biol. Med. 145 (2022), 105465.
[161] J. Hong, Y. Luo, M. Mou, J. Fu, Y. Zhang, W. Xue, T. Xie, L. Tao, Y. Lou, F. Zhu, Convolutional neural network-based annotation of bacterial type IV secretion

system effectors with enhanced accuracy and reduced false discovery, Briefings Bioinf. 21 (2020) 1825–1836.
[162] H. Zhang, Y. Wang, Z. Pan, X. Sun, M. Mou, B. Zhang, Z. Li, H. Li, F. Zhu, ncRNAInter: a novel strategy based on graph neural network to discover interactions between lncRNA and miRNA, Briefings Bioinf. 23 (2022) bbac411.
[163] J.M. Levin, T.I. Oprea, S. Davidovich, T. Clozel, J.P. Overington, Q. Vanhaelen, C. R. Cantor, E. Bischof, A. Zhavoronkov, Artificial intelligence, drug repurposing and peer review, Nat. Biotechnol. 38 (2020) 1127–1131.