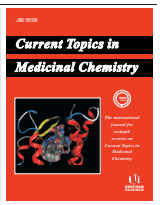
Current Topics in Medicinal Chemistry, 2022, *22*, 1897-1906

CURRENT FRONTIER



Analysis of Drug Repositioning and Prediction Techniques: A Concise Review

Shida He[1,2], Xin Liu[3*], Xiucai Ye[2*] and Sakurai Tetsuya[2]

[1]*Yangtze Delta Region Institute (Quzhou), University of Electronic Science and Technology of China, Quzhou, Zhejiang, China;* [2]*Department of Computer Science, University of Tsukuba, Tsukuba, Ibaraki, 305-8577, Japan;* [3]*Beidahuang Industry Group General Hospital, Harbin, China*

ARTICLE HISTORY

Received: October 19, 2021
Revised: January 06, 2022
Accepted: January 13, 2022

DOI:
10.2174/1568026622666220317164016

Abstract: High costs and risks are common issues in traditional drug research and development. Usually, it takes a long time to research and develop a drug, the effects of which are limited to relatively few targets. At present, studies are aiming to identify unknown new uses for existing drugs. Drug repositioning enables drugs to be quickly launched into clinical practice at a low cost because they have undergone clinical safety testing during the development process, which can greatly reduce costs and the risks of failed development. In addition to existing drugs with known indications, drugs that were shelved because of clinical trial failure can also be options for repositioning. In fact, many widely used drugs are identified *via* drug repositioning at present. This article reviews some popular research areas in the field of drug repositioning and briefly introduces the advantages and disadvantages of these methods, aiming to provide useful insights into future development in this field.

Keywords: Drug repositioning, Biomedicine, Drug discovery, Prediction techniques, Drug redistribution, Therapeutic conversion.

1. INTRODUCTION

Although great progress has been made in genomics and life science technology in recent decades, the development of new drugs involves high risks and costs and requires a prolonged time. In fact, 15 years and 800 million US dollars may be required to produce a new drug through research and development, regulatory approval, and market launch [1-3]. Despite the increasing investment in drug research and development, only a small number of new drugs are approved each year. In the US, for example, the Food and Drug Administration (FDA) approves approximately 20 drugs each year [4]. To overcome these problems in the development of traditional drugs, drug repositioning (also called drug repurposing or indication discovery) adopts related technologies to analyze existing drugs and discover new indications or new uses [5-7]. Because it greatly reduces the time and cost required for drug research and development and reduces the risk involved in the process, drug repositioning has attracted the attention of many pharmaceutical companies, researchers, clinicians, and even governments. Companies have resorted to projects known as “drug repositioning,” “drug redistribution,” or “therapeutic conversion” to identify new uses for existing or obsolete drugs, and some good examples are thalidomide, bupropion, and fluoxetine [8, 9]. Most approved drugs are known to be safe for humans (or animals) because they have passed relevant pharmacological verification tests. Therefore, compared with new drugs, repositioned drugs can enter clinical use more rapidly at a lower cost. Moreover, a large number of drugs that failed in clinical development can be repositioned, further explored, and possibly reused for new diseases. Of the 113 new drugs and biological agents approved or released in 2017, only seven were brand-new drugs (approved and released drugs with new mechanisms of action), and 36 were repositioned drugs [10]. Drug repositioning enables a drug to be available to patients in 3-12 years at an estimated total cost of 400,000 to 80 million US dollars [11, 12]. In view of the current research and development costs, it is impossible to start drug research and development from scratch for all current diseases, especially rare diseases. However, drug repositioning with the premise of discovering hidden links or establishing links between drugs and diseases is expected to become an important strategy for treating certain diseases. In addition, re-examining approved drugs to identify new indications can also help pharmaceutical companies extend the patent life of drugs by applying them to related diseases and also allow companies protect their intellectual property rights against infringement from competitors.

In addition to the traditional methods of researching drugs, massive amounts of biological data are generated

*Address correspondence to these authors at the Beidahuang Industry Group General Hospital, Harbin, China; E-mail: xinliu1074@163.com and Department of Computer Science, University of Tsukuba, , Tsukuba, Ibaraki 305-8577, Japan; E-mail: yexiucai@cs.tsukuba.ac.jp

1873-4294/22 $65.00+.00

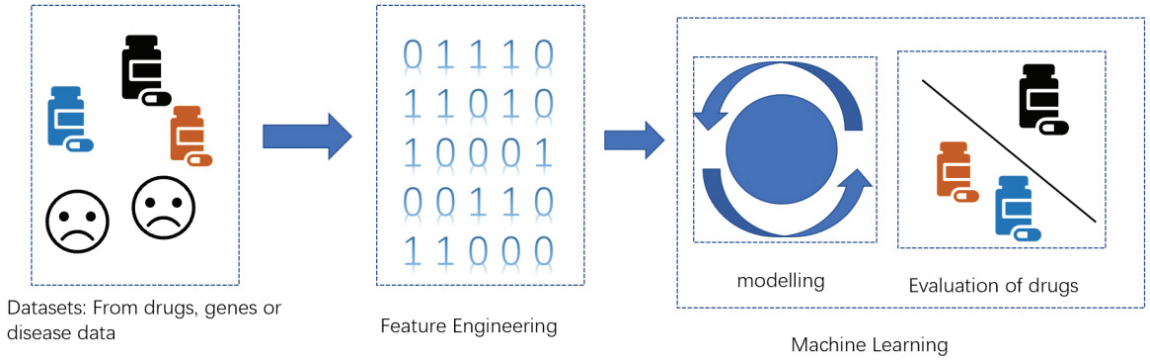


Fig. (1). Use of machine learning for drug repositioning. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

with scientific progress [13-16]. The emergence of these data sets and the development of computers have created new research fields. Many drug repositioning methods have emerged through the use of a hybrid of computer methods and statistical experimental screening. Computer-related methods such as data mining, machine learning, and web-based methods provide unprecedented opportunities. It is also possible to use the available diverse and heterogeneous data sources from the fields of genomics and biomedicine to predict all possible candidates for drug repositioning [17-23]. In this manner, predictive models can be built from protein targets [24-34], chemical structure or phenotypic information (*e.g*., side effects, gene expression), and other existing data. For example, Gottlieb *et al*. [35] used a logistic regression classifier by combining molecular structure, molecular activity, and disease semantic data to predict new indications for drugs. Yang and Agarwal [36] established a naïve Bayesian model to predict the indications of 145 diseases based on clinical side effects. Wang *et al*. [37] used the drug structure, drug-target protein, and side effects as features and trained a support vector machine (SVM) model to predict drugs for repositioning.

Wang *et al*. [38] proposed the TL_HGBI model to predict drug-disease associations and drug targets simultaneously. They established a heterogeneous network model composed of drug, disease, and target information and performed drug repositioning using an iterative algorithm that transmits information in a three-layer diagram. Martínez *et al*. [39] developed DrugNet, a network-based prioritization method that integrates information about diseases, drugs, and targets simultaneously, and it can query a list of candidate indications for drugs. In the following sections, we will review different methods for studying drug repositioning, including feature-matching technology, molecular-docking technology, genome-wide association technology, machine learning, deep learning, and network analysis strategies. Fig. (**1**) presents a flowchart of drug repositioning based on machine learning.

2. DRUG REPOSITIONING METHOD BASED ON CHARACTERISTICS AND STRUCTURE

2.1. Feature-matching Technologies

The characteristics of drugs mainly include three types of data: transcriptome (RNA), proteome, or metabolome data, chemical structure data, and data generated from adverse events.

2.1.1. Transcriptome Feature-matching Technology

This technology obtains the transcriptome characteristics of a specific drug by comparing the gene expression profiles of cells or tissues before and after drug treatment and then compares the generated differential gene expression characteristics with disease-related gene expression profiles. The extent of the negative correlation between the gene expression signal of the drug and the gene expression signal of the disease helps assess whether the drug has a potential impact on the disease. This calculation method has proven its applicability in metabolic disorders [40, 41], and it has succeeded in identifying new properties of drugs by Mirza *et al*. [42], who discovered that the drug Sitagliptin has antiepileptic effects. The disadvantage is that it relies too much on the principle of feature reversal. Gene expression patterns are unstable, and there can be situations where corresponding proteins are not translated. If the drug can reverse the expression pattern of the marker gene of a specific disease phenotype, it can restore the disease phenotype. Conversely, transcriptome feature-matching technology can help identify alternative targets of existing drugs and predict potential non-target effects for clinical application research. If two drugs share transcriptome characteristics, then it is possible that they share therapeutic applications regardless of the degree of similarity of their chemical structures. Thus, it is a one-sided method to a certain degree.

2.1.2. Chemical Structure Feature-matching Technology

Chemical similarity can indicate shared biological activity, and the potential characteristics of drugs can be identified by comparing the chemical characteristics of drugs. This process chooses a set of chemical characteristics for each drug and then builds a network based on the shared chemical characteristics. Keiser *et al*. [43] used statistical-based chemical methods to predict new targets for small-molecule drugs and pharmaceutical compounds (including Doralese, Kalgut, Prantal, *etc*.). Chen *et al*. [44] proposed a hybrid method combining the ontology information, chemical interaction information, and chemical structure information of drug compounds for the prediction of drug ATC

Table 1. Summary of drug repositioning method based on characteristics and structure.

Technology	Type	Characters	References (Not All)
Feature-matching	transcriptome	Compare gene expression profiles before and after drug treatment; Over-reliance on the principle of feature reversal, one-sided	Sirota *et al*. [40] Cheng *et al*. [41] Mirza *et al*. [42]
	chemical structure	Compare the chemical characteristics of the drugs and identify the potential properties of the drugs; Errors in chemical structure and limitations of physiological effects	Keiser *et al*. [43] Chen *et al*. [44]
	adverse reaction	Assuming that the drug may act on the same target or path, or the phenotype of the drug and the disease are similar; Difficulty in extracting information, lack of clear characteristics and causality assessment	Yang *et al*. [36] Campillosv *et al*. [45]
Molecular docking	--------	structure-based computational strategy; Data acquisition, and prediction effect is limited	Zulfiqar *et al*. [46] Dakshanamurthy *et al*. [47] Pagadala *et al*. [49]

classes. New drug-target associations can be identified by adopting a similarity coefficient to evaluate the two-dimensional structural similarity between each drug and its target. However, this method will also have disadvantages; chemical structural errors and physiological effects will limit the use of this method in drug repositioning.

2.1.3. Adverse Reaction Feature-matching Technology

Each drug has a relatively unique spectrum of adverse reactions. The adverse-reaction characteristics of matched drugs are based on the assumption that the two drugs causing the same adverse reactions may act on the same target or the same path. It is likely that the adverse-reaction phenotype of a certain drug is similar to the phenotype of a certain disease, indicating the existence of a shared pathway and physiology between the drug and disease. Campillos *et al*. [45] used the unified medical language system ontology to conduct medical research on approved drugs, extracted relevant adverse-reaction information from the drug package instruction, and evaluated the drugs based on frequency weighting and the similarity of adverse reactions, and the discovered drugs with new effects, including donepezil, zaleplon, acitretin, *etc*. This method confirmed previously known drug pairs with the same target protein and identified new shared targets for drug pairs. Yang *et al*. [36] adopted a different approach to match adverse drug reactions and diseases and predicted the indications of a drug by integrating the adverse-reaction information in the drug label with the drug-disease relationship. Although it is a logical method used for drug repositioning, the difficulty in extracting adverse-reaction information from drugs and the lack of clear adverse-reaction characteristics and causality assessment for many drugs limited this method in many ways.

2.2. Molecular Docking Technologies

Molecular docking is a structure-based computational strategy used to predict the complementarity of binding sites between drugs and targets [46]. If there is prior knowledge of the receptor target related to the disease, the specific target can be conventionally docked (one target to multiple ligands). Otherwise, a series of target receptors can be reversely docked (multiple targets to a ligand) to determine new interactions and achieve drug repositioning. Dakshanamurthy *et al*. [47] performed molecular fitting calculations for FDA-approved drugs using the crystal structures of human proteins and found that the anti-parasitic drug mebendazole could inhibit vascular endothelial growth factor receptor-2. There are some problems with the molecular docking technology for drug repositioning. First, the 3D structure of some protein targets of interest may not be available because drug targets are usually membrane proteins, such as G protein-coupled receptors [48]. Second, there is a lack of accurate macromolecular target databases. Third, the usefulness of the docking algorithm for predicting binding affinity remains doubtful. Despite constant optimization, there may be discrepancies between software packages, and there are some limitations regarding their predictive utility, including binding modes and effects of entropy [49].

The characteristics of this technology and feature-matching technology are shown in Table **1**.

3. GENOME-WIDE ASSOCIATION TECHNOLOGIES

Genome-wide association studies (GWAS) aim to identify genetic variants associated with common diseases, thereby providing a biological explanation for diseases [50]. Some of the targets can be shared between drug-treated diseases and disease phenotypes studied *via* GWAS, enabling drug repositioning. Many old drugs' new targets have been discovered through this technology. For example, succinylcholine [51] can treat coronary artery disease, and etodolac [52] can be used for breast cancer treatment. Grover *et al*. [51] used bioinformatics methods to integrate coronary artery disease gene targets with different databases, matching drug information to obtain new characteristics of drugs. Piñero *et al*. [53-56] integrated data from public databases, GWAS catalogues, animal models, and scientific literature, collected many mutations and genetic data related to human diseases and developed the DisGeNET database. DisGeNET can be used for different research purposes, including the study of the molecular basis of specific human diseases and

their complications, the analysis of the characteristics of pathogenic genes, the auxiliary construction of the hypothesis of drug treatment and adverse drug reaction, the verification of disease candidate genes and the evaluation of text mining methods performance. Rastegar *et al*. [57] used a literature-based knowledge discovery method to test the credibility of the two relationship chains of drug-gene and gene-disease and predict the relationship pair between the diseases that could potentially be treated by drugs. Yang *et al*. [52] analyzed various potential treatment drugs for high-risk diseases and predicted potential drug treatment targets. This research approach has other advantages. The study of genomic changes can advance personalized cancer treatment, and it is a useful method to evaluate potential drug targets. Pathway-based strategies can provide information about genes that are upstream or downstream of GWAS-related targets, and repositioning can be achieved using this information [58]. Iwata *et al*. [59] analyzed the chemically induced gene expression data of drugs in human cell lines, searched for pathways related to the growth of cancer cells, and screened potential anti-cancer drugs. The study by Greene *et al*. [60] combined the genetic variation information generated by GWAS with a tissue-specific functional interaction network, which can enrich drug targets to a greater extent. Although it has proven extremely successful to use GWAS information for drug repositioning, some problems still remain. Specifically, GWAS information with a strong imbalance of gene loci may make the identification of causal genes or gene variants more difficult, and GWAS cannot provide detailed pathophysiological information. It should also be considered that a greater number of new genes will be discovered as the understanding of the human genome continues to deepen.

4. DRUG REPOSITIONING METHOD BASED ON COMPUTER SCIENCE

4.1. Machine Learning

Recently, many machine learning methods have been applied in research on drug repositioning [61-65]. A common strategy in drug repositioning is the use of the known drug-disease associations, drug characteristics, and disease characteristics based on machine learning. Gottlieb *et al*. predicted new indications for drugs using a logistic regression classifier that combines molecular structure, molecular activity, and disease semantic data [35]. Yang and Agarwal established a naïve Bayesian model to predict the indications of 145 diseases based on clinical side effects [36]. Wang *et al*. used the drug structure, drug-target protein, and side effects as features and trained an SVM model to predict drug repositioning [32-34, 37, 66-68]. Oh *et al*. [69] predicted drug-disease associations by using classifiers such as random forest, multilayer perceptron, and C4.5 to select features from the integrated genetic network topology. Baricitinib is a drug for the treatment of rheumatoid arthritis. It has the potential for anti-cytokine properties and antiviral mechanisms against host proteins. In the latest research, Richardson *et al*. [70, 71] found that baricitinib may be used for COVID-19 through AI technology. Marconi *et al*. [72] also further proved the role of baricitinib through clinical trials. These computational modeling and machine learning methods have accelerated the development of drug repositioning technology, but there are certain problems and challenges associated with them. On the one hand, a one-sided focus on calculation results, ignoring the hidden deep relationship between compound structure and biological effects, without an overall and comprehensive grasp, can result in the loss of structure-activity relationships; on the other hand, the data noise caused by the heterogeneity of various biological data resources and the complicated etiology resulting from the diversity of disease pathogenesis, and phenotypes all pose huge challenges to drug repositioning. In addition, the continuously increasing cost of drug research, high standards of disease care, increased safety awareness, and high standards of regulatory agencies have also challenged the global pharmaceutical industry. It remains an unresolved and recognized problem to explore new curative effects using the potential properties of drugs and achieve personalized treatment for complex diseases. Fig. (**1**) presents a flowchart of the application of machine learning in drug repositioning. The data set may be obtained from the relevant information about drugs, genes, or patients. The data are first digitally processed by the feature extraction method in feature engineering, and then missing value processing, feature selection, and other methods are used to optimize the data features. These data are subsequently fed to the classifier. Finally, the drug is evaluated for its properties. The quantity and quality of training data are the limiting factors in the reliability of the prediction of any model. Specifically, the amount of data is too large, causing the model training time to be too long and possibly leading to overfitting. The small number of samples tends to make the model underfitting. Too worse data quality (for example, high noise and full of missing values) will worsen the generalization of the model. The characteristics of machine learning and the following two methods are shown in Table **2**.

4.2. Deep Learning

Deep learning is a branch of machine learning. Compared to traditional machine learning, deep learning contains more neural network structures with multiple hidden layers. This allows deep learning to process large data sets and identify complex rules from the learning process [73-75]. Deep learning-based methods have greatly improved the latest technologies in speech recognition [76], visual object recognition, and object detection, and they are currently being explored in the fields of biomedicine and genomics [77-97]. For example, Aliper *et al*. used deep learning methods with gene expression data to learn drug treatment categories and found that deep neural networks have greater predictive ability than traditional SVM methods after 10-fold cross-validation [98]. Zhao *et al*. [99] compared the deep neural network method with the one based on an SVM, and they also proved that the method based on deep learning has some advantages over traditional methods in drug discovery and development. The currently popular method deepDR [18] can successfully predict the effects of isoprenaline and aripiprazole on Alzheimer's disease.

Table 2. Summary of the application of drug relocation in the field of computer science.

Technology		Characters	References (Not All)	Methods (Name)
AI	Machine learning	Identify drugs with common efficacy and biological targets; reveal the functional relationship between compounds; accelerate the development of drug relocation technology	Gottlieb *et al*. [35] Yang *et al*. [36] Wang *et al*. [37] Oh *et al*. [69]	Based on logistic regression (PREDICT) Based on Naïve Bayesian Based on SVM Based on RandomForest, Perceptron, C4.5
		One-sided attention to calculate results; ignoring the deep layers of drug target relationship; susceptible to data quality and quantity		
	Deep learning	Generalization ability is higher than traditional machine learning methods; good at learning from large-scale data; able to handle complex unstructured data; affected by poor quality or unbalanced data;	Wang *et al*. [73] Ying *et al*. [76] Zeng *et al*. [18] Liu *et al*. [101]	based on Multi-head Self-attention Mechanism (DM3Loc) Based on HMM and LSTM based on Autoencoder (deepDR) based on GCN (Chemi-Net)
Network-based analysis		Based on knowledge or the use of various data resources; various expressions of different relationships; Integrate multi-source heterogeneous data, and reveal hidden or unidentified data based on the principle of internal correlation known drug-disease relationship	Yu *et al*. [103] Cheng *et al*. [104] Chen *et al*. [107]	Based on Issue-specific Pathways Based on Network Inference Based on RandomWalk

At present, a growing number of large-scale biochemical-related databases are used in drug research, and new attempts to develop drugs using deep neural networks have emerged [100]. The advantage of deep learning lies in its ability to learn the complex relationship between input features and output decisions from large-scale data. Its application in the fields of drug discovery and molecular informatics remains in its infancy, but it has displayed great potential. Compared with traditional machine learning methods, several commonly used deep neural network structures have been applied in drug-related studies, and substantial progress has been made. Convolutional neural networks (CNNs), which are special frame structures of deep neural networks, have succeeded in processing structured data (such as images). CNNs can automatically extract task-related functions from the original image through the convolution operator, thereby revealing the latest performance of image-related tasks. For such drugs and small molecules composed of atoms and chemical bonds, they have different types of structures, namely graphs, in which each atom is a node and each chemical bond is an edge. It is the same case when a direct attempt is made to adapt the convolution process for molecular graphs. However, unlike images, graphs have irregular shapes and sizes. There is no ordering among the spatial nodes, and the neighboring nodes are also related to their locations. Therefore, the regular grid structure cannot be directly applied in the traditional CNN. In fact, large amounts of structured data in the real world are usually formed in the form of graphs rather than images, making it significant to develop methods to handle irregular structures. Graph convolutional networks (GCNs) have been remarkable for extracting features from complex data structures. The current research hotspot is summing all feature vectors of all adjacent nodes in the graph for a direct expression of the convolution in the spatial domain or mapping it onto the spectrum theory to code the hidden expression and generate molecular graphs using the convolution process. There are many significant practical problems in the latest development of GCNs and their applications in drug discovery, as well as GCN processing graphs or network data forms, such as social networks, knowledge graphs, protein interaction networks, and molecular graphs. However, because of the uniqueness of these graphical data, deep learning can exert significant effects on predictive capabilities after receiving a large amount of training data, and it is extremely common that some tasks may not contain sufficient data to make meaningful predictions. By integrating one-time learning with GCNs, Altae-Tran *et al*. [101] proved that combining GCNs can greatly improve the ability to learn the appropriate distance metric for small molecules. This strategy aims to build a strong classifier for the test tasks by using the information in the training tasks and to use the similarity of the representations learned between molecules in different task groups; specifically, the label of the query molecule is the weighted sum of the support molecular label. In their proposed method, the molecular representation is obtained through graph convolutional layers. The embedding used to generate task-driven similarity measures is achieved in a residual network through iterative long-term short-term memory. The final prediction of the molecular marker can be drawn immediately after the similarity measure is obtained. Several commonly used network architectures [18, 73, 76, 101, 102] are shown in Fig. (**2**).

4.3. Network-based Analysis

Network-based methods have been widely applied to drug repositioning to identify new drug targets, interactions, and indications. Normally in these models, the nodes in the network represent drugs, diseases, or gene products, and the edges represent their interactions or interrelationships. These networks make computational inferences based on knowledge or various data resources. There are various ex-

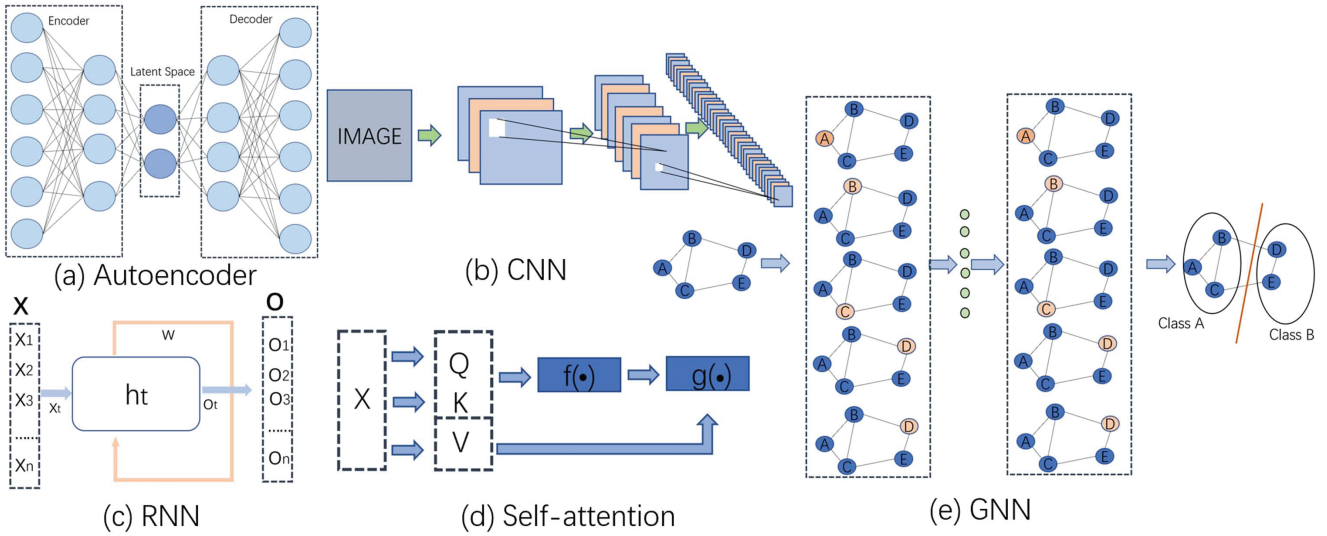


Fig. (2). The basic frameworks of the neural network. (**a**) Autoencoder, (**b**) Convolutional neural network, (**c**) Recurrent neural network, (**d**) Self-attention, (**e**) Graph convolutional network. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

pressions of different relationships, such as drug-drug, disease-disease, and gene-disease interactions. It is worth mentioning that text mining on the literature can help us generate a network. Commonly used tools include Python, Linguamatics' I2E application, SciBite's service, *etc*. There are two methods of network simulation, namely, drug-centric (identifying new indications for existing drugs) and disease-centric (identifying effective drugs as potential treatments for diseases). These methods serve as the basis of calculation by using available information about drugs, such as drug targets, chemical structures, adverse reactions, and pathways, and establishing a network computing model to predict unknown mechanisms, targets, or new biomarkers of diseases. The similarity or connection between drugs and diseases is calculated using data for signal transduction pathways, metabolic pathways, and protein interaction networks in the network-based methods [102]. For example, processed omics data from human patients or animal models can be used to reconstruct disease-specific pathways. In addition, those pathways can function as key targets for novel therapeutic discovery or drug repositioning. In this manner, better or even specific drug targets, as well as alternative drugs for diseases, can be identified. Cheng *et al*. calculated network similarity based on drugs and targets to predict the drug-target interaction in the dual network, finding that the network-based reasoning method is more effective in predicting the drug-target similarity, and they proved in experiments that Simvastatin and Ketoconazole (two old drugs) showed effective anti-proliferative activity on human MDA-MB-231 breast cancer cell line [103, 104]. The network-based analysis methods reveal hidden or unknown drug-disease relationships based on the principle of internal correlation by integrating multi-source heterogeneous data. Most network similarity-based methods are either drug-centric or disease-centric. By contrast, few methods use gene annotation and network clustering to establish drug-disease heterogeneous networks for drug repositioning. Campillos *et al*. [45] proposed a method to identify the correlation between drugs and targets based on the similarity of drug structure information and drug side effects. Luo *et al*. [105] established a similar network-based analysis method using heterogeneous data through the network diffusion process and inferred the predictive score of drug-target interaction by using diffusion distribution. Himmelstein *et al*. [106] integrated data from 29 public sources to identify discovered retargeted drug candidates and predicted hundreds of drug-disease repositioning possibilities. Chen *et al*. [107] proposed a random walk with a restart on the heterogeneous network method to predict possible drug-target interaction.

CONCLUSION

Traditional novel drug development is associated with low throughput, long cycles, high costs, and high risks. From a historical perspective, the reuse of drugs is mostly accidental. Once a non-targeting or new targeting effect is discovered for a drug, it can be used in clinical trials and development. In this paper, we reviewed different methods used in drug repositioning applications. Through these strategies, many potential drug candidates can be discovered, including some candidates in the late stage of clinical trials, and they may be used to treat both common and rare diseases. Repositioning is critical for the identification of drugs for the treatment of rare diseases, and sometimes it is the only strategy. However, some important technical, regulatory, and organizational challenges are hindering its expanded use.

CONSENT FOR PUBLICATION

Not applicable.

FUNDING

The work was supported by the Special Science Foundation of Quzhou (2021D004).

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

Declared none.

REFERENCES

[1] Drews, J. Drug discovery: A historical perspective. *Science,* **2000**, *287*(5460), 1960-1964. http://dx.doi.org/10.1126/science.287.5460.1960 PMID: 10720314

[2] Jakhar, R.; Dangi, M.; Khichi, A.; Chhillar, A.K. Relevance of molecular docking studies in drug designing. *Curr. Bioinform.,* **2020**, *15*(4), 270-278. http://dx.doi.org/10.2174/1574893615666191219094216

[3] Agarwal, S.; Agarwal, V.; Agarwal, M.; Singh, M. Exosomes: Structure, biogenesis, types and application in diagnosis and gene and drug delivery. *Curr. Gene Ther.,* **2020**, *20*(3), 195-206. http://dx.doi.org/10.2174/1566523220999200731011702 PMID: 32787759

[4] Yazdanian, M.; Briggs, K.; Jankovsky, C.; Hawi, A. The "high solubility" definition of the current FDA Guidance on Biopharmaceutical Classification System may be too strict for acidic drugs. *Pharm. Res.,* **2004**, *21*(2), 293-299. http://dx.doi.org/10.1023/B:PHAM.0000016242.48642.71 PMID: 15032311

[5] Yang, J.; Peng, S.; Zhang, B.; Houten, S.; Schadt, E.; Zhu, J.; Suh, Y.; Tu, Z. Human geroprotector discovery by targeting the converging subnetworks of aging and age-related diseases. *Geroscience,* **2020**, *42*(1), 353-372. http://dx.doi.org/10.1007/s11357-019-00106-x PMID: 31637571

[6] Liu, C.; Wei, D.; Xiang, J.; Ren, F.; Huang, L.; Lang, J.; Tian, G.; Li, Y.; Yang, J. An improved anticancer drug-response prediction based on an ensemble method integrating matrix completion and ridge regression. *Mol. Ther. Nucleic Acids,* **2020**, *21*, 676-686. http://dx.doi.org/10.1016/j.omtn.2020.07.003 PMID: 32759058

[7] Zhang, S.; Su, M.; Sun, Z.; Lu, H.; Zhang, Y. The signature of pharmaceutical sensitivity based on ctDNA mutation in eleven cancers. *Exp. Biol. Med. (Maywood),* **2020**, *245*(8), 720-732. http://dx.doi.org/10.1177/1535370220906518 PMID: 32050795

[8] Ashburn, T.T.; Thor, K.B.J.N.r.D.d. Drug repositioning: Identifying and developing new uses for existing drugs. **2004**, *3*(8), 673-683. http://dx.doi.org/10.1038/nrd1468

[9] Boguski, M.S.; Mandl, K.D.; Sukhatme, V.P.J.S. Repurposing with a difference. **2009**, *324*(5933), 1394-1395. http://dx.doi.org/10.1126/science.1169920

[10] Graul, A.I.; Cruces, E.; Stringer, M. The year's new drugs & biologics, 2013: Part I. *Drugs Today (Barc),* **2014**, *50*(1), 51-100. http://dx.doi.org/10.1358/dot.2014.50.1.2116673 PMID: 24524105

[11] Hurle, M.R.; Yang, L.; Xie, Q.; Rajpal, D.K.; Sanseau, P.; Agarwal, P. Computational drug repositioning: From data to therapeutics. *Clin. Pharmacol. Ther.,* **2013**, *93*(4), 335-341. http://dx.doi.org/10.1038/clpt.2013.1 PMID: 23443757

[12] Liu, K.; Chen, W. iMRM: A platform for simultaneously identifying multiple kinds of RNA modifications. *Bioinformatics,* **2020**, *36*(11), 3336-3342. http://dx.doi.org/10.1093/bioinformatics/btaa155 PMID: 32134472

[13] Su, R.; Wu, H.; Xu, B.; Liu, X.; Wei, L. Developing a multi-dose computational model for drug-induced hepatotoxicity prediction based on toxicogenomics data. *IEEE/ACM Trans. Comput. Biol. Bioinformatics,* **2019**, *16*(4), 1231-1239. http://dx.doi.org/10.1109/TCBB.2018.2858756 PMID: 30040651

[14] Cheng, L.; Qi, C.; Zhuang, H.; Fu, T.; Zhang, X. gutMDisorder: A comprehensive database for dysbiosis of the gut microbiota in disorders and interventions. *Nucleic Acids Res.,* **2020**, *48*(D1), D554-D560. http://dx.doi.org/10.1093/nar/gkz843 PMID: 31584099

[15] Li, T.; Huang, T.; Guo, C.; Wang, A.; Shi, X.; Mo, X.; Lu, Q.; Sun, J.; Hui, T.; Tian, G.; Wang, L.; Yang, J. Genomic variation, origin tracing, and vaccine development of SARS-CoV-2: A systematic review. *Innovation (N Y),* **2021**, *2*(2), 100116. http://dx.doi.org/10.1016/j.xinn.2021.100116 PMID: 33997827

[16] Li, M. Identification of genes of four malignant tumors and a novel prediction model development based on PPI data and support vector machines. **2020**, *27*(9), 715-725. http://dx.doi.org/10.1038/s41417-019-0143-5

[17] Liu, X.; Yang, J.; Zhang, Y.; Fang, Y.; Wang, F.; Wang, J.; Zheng, X.; Yang, J. A systematic study on drug-response associated genes using baseline gene expressions of the Cancer Cell Line Encyclopedia. *Sci. Rep.,* **2016**, *6*(1), 22811. http://dx.doi.org/10.1038/srep22811 PMID: 26960563

[18] Zeng, X.; Zhu, S.; Liu, X.; Zhou, Y.; Nussinov, R.; Cheng, F. deepDR: A network-based deep learning approach to *in silico* drug repositioning. *Bioinformatics,* **2019**, *35*(24), 5191-5198. http://dx.doi.org/10.1093/bioinformatics/btz418 PMID: 31116390

[19] An, Q.; Yu, L. A heterogeneous network embedding framework for predicting similarity-based drug-target interactions. *Brief. Bioinform.,* **2021**, *22*(6), bbab275. http://dx.doi.org/10.1093/bib/bbab275 PMID: 34373895

[20] Zeng, X.; Zhu, S.; Lu, W.; Liu, Z.; Huang, J.; Zhou, Y.; Fang, J.; Huang, Y.; Guo, H.; Li, L.; Trapp, B.D.; Nussinov, R.; Eng, C.; Loscalzo, J.; Cheng, F. Target identification among known drugs by deep learning from heterogeneous networks. *Chem. Sci. (Camb.),* **2020**, *11*(7), 1775-1797. http://dx.doi.org/10.1039/C9SC04336E PMID: 34123272

[21] Zeng, X.; Liao, Y.; Liu, Y.; Zou, Q. Prediction and validation of disease genes using HeteSim scores. *IEEE/ACM Trans. Comput. Biol. Bioinformatics,* **2017**, *14*(3), 687-695. http://dx.doi.org/10.1109/TCBB.2016.2520947 PMID: 26890920

[22] Zeng, X. Predicting disease-associated circular RNAs using deep forests combined with positive-unlabeled learning methods. *Brief. Bioinform.,* **2020**, *21*(4), 1425-1436. http://dx.doi.org/10.1093/bib/bbz080 PMID: 31612203

[23] Yu, L. Prediction of drug response in multilayer networks based on fusion of multiomics data. *Methods,* **2021**, *192*, 85-92.

[24] Wang, J.; Shi, Y.; Wang, X.; Chang, H. A drug target interaction prediction based on LINE-RF learning. *Curr. Bioinform.,* **2020**, *15*(7), 750-757. http://dx.doi.org/10.2174/1574893615666191227092453

[25] Wang, J.; Wang, H.; Wang, X.; Chang, H. Predicting drug-target interactions *via* FM-DNN learning. *Curr. Bioinform.,* **2020**, *15*(1), 68-76. http://dx.doi.org/10.2174/1574893614666190227160538

[26] Yu, L.S.; Shi, Y.; Zou, Q.; Wang, S.; Zheng, L.; Gao, L.Y.; Zou, Q.; Wang, S.; Zheng, L.; Gao, L. Exploring drug treatment patterns based on the action of drug and multilayer network model. *Int. J. Mol. Sci.,* **2020**, *21*(14), 5014. http://dx.doi.org/10.3390/ijms21145014

[27] Mo, F.; Luo, Y.; Fan, D.; Zeng, H.; Zhao, Y.; Luo, M.; Liu, X.; Ma, X. Integrated analysis of mRNA-seq and miRNA-seq to Identify c-MYC, YAP1 and miR-3960 as major players in the anticancer effects of caffeic acid phenethyl ester in human small cell lung cancer Cell Line. *Curr. Gene Ther.,* **2020**, *20*(1), 15-24. http://dx.doi.org/10.2174/1566523220666200523165159 PMID: 32445454

[28] Liu, M.L.; Su, W.; Wang, J.S.; Yang, Y.H.; Yang, H.; Lin, H. Predicting preference of transcription factors for methylated DNA using sequence information. *Mol. Ther. Nucleic Acids,* **2020**, *22*, 1043-1050. http://dx.doi.org/10.1016/j.omtn.2020.07.035 PMID: 33294291

[29] Cai, L. ITP-Pred: An interpretable method for predicting, therapeutic peptides with fused features low-dimension representation. *Brief. Bioinform.,* **2021**, *22*(4), bbaa367. PMID: 33313672

[30] Fu, X.; Cai, L.; Zeng, X.; Zou, Q. StackCPPred: A stacking and pairwise energy content-based prediction of cell-penetrating peptides and their uptake efficiency. *Bioinformatics,* **2020**, *36*(10), 3028-3034. http://dx.doi.org/10.1093/bioinformatics/btaa131 PMID: 32105326

[31] Yu, L.; Xia, M.; An, Q. A network embedding framework based on integrating multiplex network for drug combination prediction. *Brief. Bioinform.,* **2021**, *23*(1), bbab364. PMID: 34505623

[32] Hu, Y.; Zhang, H.; Liu, B.; Gao, S.; Wang, T.; Han, Z.; Ji, X.; Liu, G. rs34331204 regulates TSPAN13 expression and contributes to Alzheimer's disease with sex differences. *Brain,* **2020**, *143*(11), e95. http://dx.doi.org/10.1093/brain/awaa302 PMID: 33175954

[33] Hu, Y.; Sun, J.Y.; Zhang, Y.; Zhang, H.; Gao, S.; Wang, T.; Han, Z.; Wang, L.; Sun, B.L.; Liu, G. rs1990622 variant associates with Alzheimer's disease and regulates TMEM106B expression in human brain tissues. *BMC Med.,* **2021**, *19*(1), 11. http://dx.doi.org/10.1186/s12916-020-01883-5 PMID: 33461566

[34] Hu, Y.; Qiu, S.; Cheng, L. Integration of multiple-Omics data to analyze the population-specific differences for coronary artery disease. *Comput. Math. Methods Med.,* **2021**, *2021*, 7036592. http://dx.doi.org/10.1155/2021/7036592 PMID: 34447459

[35] Gottlieb, A.; Stein, G.Y.; Ruppin, E.; Sharan, R. PREDICT: A method for inferring novel drug indications with application to personalized medicine. *Mol. Syst. Biol.,* **2011**, *7*(1), 496. http://dx.doi.org/10.1038/msb.2011.26 PMID: 21654673

[36] Yang, L.; Agarwal, P. Systematic drug repositioning based on clinical side-effects. *PLoS One,* **2011**, *6*(12), e28025. http://dx.doi.org/10.1371/journal.pone.0028025 PMID: 22205936

[37] Wang, Y.; Chen, S.; Deng, N.; Wang, Y. Drug repositioning by kernel-based integration of molecular structure, molecular activity, and phenotype data. *PLoS One,* **2013**, *8*(11), e78518. http://dx.doi.org/10.1371/journal.pone.0078518 PMID: 24244318

[38] Wang, W.; Yang, S.; Zhang, X.; Li, J. Drug repositioning by integrating target information through a heterogeneous network model. *Bioinformatics,* **2014**, *30*(20), 2923-2930. http://dx.doi.org/10.1093/bioinformatics/btu403 PMID: 24974205

[39] Martinez, V. DrugNet: Network-based drug-disease prioritization by integrating heterogeneous data. **2015**, *63*(1), 41-49. http://dx.doi.org/10.1016/j.artmed.2014.11.003

[40] Sirota, M.; Dudley, J.T.; Kim, J.; Chiang, A.P.; Morgan, A.A.; Sweet-Cordero, A.; Sage, J.; Butte, A.J. Discovery and preclinical validation of drug indications using compendia of public gene expression data. *Sci. Transl. Med.,* **2011**, *3*(96), 96ra77. http://dx.doi.org/10.1126/scitranslmed.3001318 PMID: 21849665

[41] Cheng, L.; Yang, H.; Zhao, H.; Pei, X.; Shi, H.; Sun, J.; Zhang, Y.; Wang, Z.; Zhou, M. MetSigDis: A manually curated resource for the metabolic signatures of diseases. *Brief. Bioinform.,* **2019**, *20*(1), 203-209. http://dx.doi.org/10.1093/bib/bbx103 PMID: 28968812

[42] Mirza, N.; Sills, G.J.; Pirmohamed, M.; Marson, A.G. Identifying new antiepileptic drugs through genomics-based drug repurposing. *Hum. Mol. Genet.,* **2017**, *26*(3), 527-537. http://dx.doi.org/10.1093/hmg/ddw410 PMID: 28053048

[43] Keiser, M.J.; Setola, V.; Irwin, J.J.; Laggner, C.; Abbas, A.I.; Hufeisen, S.J.; Jensen, N.H.; Kuijer, M.B.; Matos, R.C.; Tran, T.B.; Whaley, R.; Glennon, R.A.; Hert, J.; Thomas, K.L.; Edwards, D.D.; Shoichet, B.K.; Roth, B.L. Predicting new molecular targets for known drugs. *Nature,* **2009**, *462*(7270), 175-181. http://dx.doi.org/10.1038/nature08506 PMID: 19881490

[44] Chen, L. A hybrid method for prediction and repositioning of drug anatomical therapeutic chemical classes. **2014**, *10*(4), 868-877. http://dx.doi.org/10.1039/c3mb70490d

[45] Campillos, M.; Kuhn, M.; Gavin, A.C.; Jensen, L.J.; Bork, P. Drug target identification using side-effect similarity. *Science,* **2008**, *321*(5886), 263-266. http://dx.doi.org/10.1126/science.1158140 PMID: 18621671

[46] Zulfiqar, H.; Masoud, M.S.; Yang, H.; Han, S-G.; Wu, C-Y.; Lin, H. Screening of prospective plant compounds as H1R and CL1R inhibitors and its antiallergic efficacy through molecular docking approach. *Comput. Math. Methods Med.,* **2021**, *2021*, 6683407. http://dx.doi.org/10.1155/2021/6683407

[47] Dakshanamurthy, S.; Issa, N.T.; Assefnia, S.; Seshasayee, A.; Peters, O.J.; Madhavan, S.; Uren, A.; Brown, M.L.; Byers, S.W. Predicting new indications for approved drugs using a proteochemometric method. *J. Med. Chem.,* **2012**, *55*(15), 6832-6848. http://dx.doi.org/10.1021/jm300576q PMID: 22780961

[48] Cooke, R.M.; Brown, A.J.; Marshall, F.H.; Mason, J.S. Structures of G protein-coupled receptors reveal new opportunities for drug discovery. *Drug Discov. Today,* **2015**, *20*(11), 1355-1364. http://dx.doi.org/10.1016/j.drudis.2015.08.003 PMID: 26303408

[49] Pagadala, N.S.; Syed, K.; Tuszynski, J. Software for molecular docking: A review. *Biophys. Rev.,* **2017**, *9*(2), 91-102. http://dx.doi.org/10.1007/s12551-016-0247-1 PMID: 28510083

[50] Zhang, T.; Hu, Y.; Wu, X.; Ma, R.; Jiang, Q.; Wang, Y. Identifying liver cancer-related enhancer SNPs by integrating GWAS and histone Modification ChIP-seq Data. *BioMed Res. Int.,* **2016**, *2016*, 2395341. http://dx.doi.org/10.1155/2016/2395341 PMID: 27429976

[51] Grover, M.P.; Ballouz, S.; Mohanasundaram, K.A.; George, R.A.; Goscinski, A.; Crowley, T.M.; Sherman, C.D.; Wouters, M.A. Novel therapeutics for coronary artery disease from genome-wide association study data. *BMC Med. Genomics,* **2015**, *8*(S2)(Suppl. 2), S1. http://dx.doi.org/10.1186/1755-8794-8-S2-S1 PMID: 26044129

[52] Yang, J.; Li, Z.; Fan, X.; Cheng, Y. Drug-disease association and drug-repositioning predictions in complex diseases using causal inference-probabilistic matrix factorization. *J. Chem. Inf. Model.,* **2014**, *54*(9), 2562-2569. http://dx.doi.org/10.1021/ci500340n PMID: 25116798

[53] Piñero, J.; Bravo, À.; Queralt-Rosinach, N.; Gutiérrez-Sacristán, A.; Deu-Pons, J.; Centeno, E.; García-García, J.; Sanz, F.; Furlong, L. I. DisGeNET: a comprehensive platform integrating information on human disease-associated genes and variants. *Nucleic Acids Res*., **2017**, *45*(D1): D833-D839.

[54] Piñero, J.; Queralt-Rosinach, N.; Bravo, À.; Deu-Pons, J.; Bauer-Mehren, A.; Baron, M.; Sanz, F.; Furlong, L. I. DisGeNET: A discovery platform for the dynamical exploration of human diseases and their genes. *Database,* **2015**, *2015*, bav028. http://dx.doi.org/10.1093/database/bav028

[55] Piñero, J.; Bravo, À.; Queralt-Rosinach, N.; Gutiérrez-Sacristán, A.; Deu-Pons, J.; Centeno, E.; García-García, J.; Sanz, F.; Furlong, L. I. The DisGeNET knowledge platform for disease genomics: 2019 update. *Nucleic Acids Res*., **2020**; *48* (D1): D845-D855.

[56] Piñero, J.; Sauch, J.; Sanz, F.; Furlong L.I. The DisGeNET cytoscape app: Exploring and visualiz-ing disease genomics data. *Comput Struct Biotechnol J.,* **2021**, *19*, 2960-2967.

[57] Rastegar-Mojarad, M. A new method for prioritizing drug repositioning candidates extracted by literature-based discovery. *IEEE International Conference on Bioinformatics and Biomedicine - Medical Informatics and Decision Making,* Washington, DC, **2015**. http://dx.doi.org/10.1109/BIBM.2015.7359766

[58] Greene, C.S.; Voight, Greene, C.S.; Voight, B.F. Pathway and network-based strategies to translate genetic discoveries into effective therapies, *Human Molec Genet*, **2016**, *25* (R2), R98-R98. http://dx.doi.org/10.1093/hmg/ddw160

[59] Iwata, M.; Hirose, L.; Kohara, H.; Liao, J.; Sawada, R.; Akiyoshi, S.; Tani, K.; Yamanishi, Y. Pathway-based drug repositioning for cancers: Computational prediction and experimental validation. *J. Med. Chem.,* **2018**, *61*(21), 9583-9595. http://dx.doi.org/10.1021/acs.jmedchem.8b01044 PMID: 30371064

[60] Greene, C.S.; Krishnan, A.; Wong, A.K.; Ricciotti, E.; Zelaya, R.A.; Himmelstein, D.S.; Zhang, R.; Hartmann, B.M.; Zaslavsky, E.; Sealfon, S.C.; Chasman, D.I.; FitzGerald, G.A.; Dolinski, K.; Grosser, T.; Troyanskaya, O.G. Understanding multicellular function and disease with human tissue-specific networks. *Nat. Genet.,* **2015**, *47*(6), 569-576. http://dx.doi.org/10.1038/ng.3259 PMID: 25915600

[61] Ding Y.T., Jijun; Guo, ; Fei, Identification of drug-target interactions *via* dual laplacian regularized least squares with multiple kernel fusion. *Knowl. Base. Syst.,* **2020**, 204.

[62] Ding, Y.; Tang, J.; Guo, F. Identification of drug-target interactions *via* fuzzy bipartite local model. *Neural Comput. Appl.,* **2020**, *23*(14), 10303-10319. http://dx.doi.org/10.1007/s00521-019-04569-z

[63] Ding, Y.; Tang, J.; Guo, F. Identification of drug-target interactions *via* multiple information integration. *Inf. Sci.,* **2017**, *418*, 546-560.

http://dx.doi.org/10.1016/j.ins.2017.08.045

[64] Cheng, L.; Hu, Y.; Sun, J.; Zhou, M.; Jiang, Q. DincRNA: A comprehensive web-based bioinformatics toolkit for exploring disease associations and ncRNA function. *Bioinformatics,* **2018**, *34*(11), 1953-1956.
http://dx.doi.org/10.1093/bioinformatics/bty002 PMID: 29365045

[65] Zeng, X.; Song, X.; Ma, T.; Pan, X.; Zhou, Y.; Hou, Y.; Zhang, Z.; Li, K.; Karypis, G.; Cheng, F. Repurpose open data to discover therapeutics for COVID-19 using deep learning. *J. Proteome Res.,* **2020**, *19*(11), 4624-4636.
http://dx.doi.org/10.1021/acs.jproteome.0c00316
PMID: 32654489

[66] Wu, Y.; Lu, X.; Shen, B.; Zeng, Y. The therapeutic potential and role of miRNA, lncRNA, and circRNA in osteoarthritis. *Curr. Gene Ther.,* **2019**, *19*(4), 255-263.
http://dx.doi.org/10.2174/1566523219666190716092203
PMID: 31333128

[67] Cai, L.; Lu, C.; Xu, J.; Meng, Y.; Wang, P.; Fu, X.; Zeng, X.; Su, Y. Drug repositioning based on the heterogeneous information fusion graph convolutional network. *Brief. Bioinform.,* **2021**, *22*(6), bbab319.
http://dx.doi.org/10.1093/bib/bbab319 PMID: 34378011

[68] Song, B.; Li, F.; Liu, Y.; Zeng, X. Deep learning methods for biomedical named entity recognition: A survey and qualitative comparison. *Brief. Bioinform.,* **2021**, *22*(6), bbab282.
http://dx.doi.org/10.1093/bib/bbab282 PMID: 34308472

[69] Oh, M.; Ahn, J.; Yoon, Y. A network-based classification model for deriving novel drug-disease associations and assessing their molecular actions. *PLoS One,* **2014**, *9*(10), e111668.
http://dx.doi.org/10.1371/journal.pone.0111668 PMID: 25356910

[70] Richardson, P. Baricitinib as potential treatment for 2019-nCoV acute respiratory disease. **2020**, *395*(10223), 30.

[71] Stebbing, J. Mechanism of baricitinib supports artificial intelligence-predicted testing in COVID-19 patients. **2020**, *12*(8), 12697.
http://dx.doi.org/10.15252/emmm.202012697

[72] Marconi, V.C. Efficacy and safety of baricitinib for the treatment of hospitalised adults with COVID-19 (COV-BARRIER): A randomised, double-blind, parallel-group, placebo-controlled phase 3 trial. **2021**, *9*(12), 1407-1418.

[73] Wang, D.; Zhang, Z.; Jiang, Y.; Mao, Z.; Wang, D.; Lin, H.; Xu, D. DM3Loc: Multi-label mRNA subcellular localization prediction and analysis based on multi-head self-attention mechanism. *Nucleic Acids Res.,* **2021**, *49*(8), e46.
http://dx.doi.org/10.1093/nar/gkab016 PMID: 33503258

[74] Dao, F.Y.; Lv, H.; Zhang, D.; Zhang, Z.M.; Liu, L.; Lin, H. DeepYY1: A deep learning approach to identify YY1-mediated chromatin loops. *Brief. Bioinform.,* **2021**, *22*(4), bbaa356.
http://dx.doi.org/10.1093/bib/bbaa356 PMID: 33279983

[75] Liu, Y. A deep learning approach for filtering structural variants in short read sequencing data. *Brief. Bioinform.,* **2021**, *22*(4), bbaa370.
PMID: 33378767

[76] Ying, W.; Zhang, L.; Deng, H. Sichuan dialect speech recognition with deep LSTM network. *Front. Comput. Sci.,* **2020**, *14*(2), 378-387.
http://dx.doi.org/10.1007/s11704-018-8030-z

[77] Zhang, Y.; Yan, J.; Chen, S.; Gong, M.; Gao, D.; Zhu, M.; Gan, W. Review of the applications of deep learning in bioinformatics. *Curr. Bioinform.,* **2020**, *15*(8), 898-911.
http://dx.doi.org/10.2174/1574893615999200711165743

[78] Cui, F.; Zhang, Z.; Zou, Q. Sequence representation approaches for sequence-based protein prediction tasks that use deep learning. *Brief. Funct. Genomics,* **2021**, *20*(1), 61-73.
http://dx.doi.org/10.1093/bfgp/elaa030 PMID: 33527980

[79] Zhang, T.; Wei, X.; Li, Z.; Shi, F.; Xia, Z.; Lian, M.; Chen, L.; Zhang, H. Natural scene nutrition information acquisition and analysis based on deep learning. *Curr. Bioinform.,* **2020**, *15*(7), 662-670.
http://dx.doi.org/10.2174/1574893614666190723121610

[80] Long, H.; Sun, Z.; Li, M.; Fu, H.Y.; Lin, M.C. Predicting protein phosphorylation sites based on deep learning. *Curr. Bioinform.,* **2020**, *15*(4), 300-308.
http://dx.doi.org/10.2174/1574893614666190902154332

[81] Lv, Z.; Ao, C.; Zou, Q. Protein function prediction: From traditional classifier to deep learning. *Proteomics,* **2019**, *19*(14), e1900119.
http://dx.doi.org/10.1002/pmic.201900119 PMID: 31187588

[82] Townes, F.W.; Irizarry, R.A. Quantile normalization of single-cell RNA-seq read counts without unique molecular identifiers. *Genome Biol.,* **2020**, *21*(1), 160.
http://dx.doi.org/10.1186/s13059-020-02078-0 PMID: 32620142

[83] Tang, W.; Bertaux, F.; Thomas, P.; Stefanelli, C.; Saint, M.; Marguerat, S.; Shahrezaei, V. bayNorm: Bayesian gene expression recovery, imputation and normalization for single-cell RNA-sequencing data. *Bioinformatics,* **2020**, *36*(4), 1174-1181.
PMID: 31584606

[84] Jin, Q.; Meng, Z.; Pham, T.D.; Chen, Q.; Wei, L.; Su, R. DUNet: A deformable network for retinal vessel segmentation. *Knowl. Base. Syst.,* **2019**, *178*, 149-162.
http://dx.doi.org/10.1016/j.knosys.2019.04.025

[85] Manavalan, B.; Basith, S.; Shin, T.H.; Wei, L.; Lee, G. Meta-4mCpred: A sequence-based meta-predictor for accurate DNA 4mC site prediction using effective feature representation. *Mol. Ther. Nucleic Acids,* **2019**, *16*, 733-744.
http://dx.doi.org/10.1016/j.omtn.2019.04.019 PMID: 31146255

[86] Su, R.; Hu, J.; Zou, Q.; Manavalan, B.; Wei, L. Empirical comparison and analysis of web-based cell-penetrating peptide prediction tools. *Brief. Bioinform.,* **2020**, *21*(2), 408-420.
http://dx.doi.org/10.1093/bib/bby124 PMID: 30649170

[87] Su, R.; Liu, X.; Wei, L.; Zou, Q. Deep-Resp-Forest: A deep forest model to predict anti-cancer drug response. *Methods,* **2019**, *166*, 91-102.
http://dx.doi.org/10.1016/j.ymeth.2019.02.009 PMID: 30772464

[88] Wei, L.; Liao, M.; Gao, Y.; Ji, R.; He, Z.; Zou, Q. Improved and promising identification of human microRNAs by incorporating a high-quality negative set. *IEEE/ACM Trans. Comput. Biol. Bioinformatics,* **2014**, *11*(1), 192-201.
http://dx.doi.org/10.1109/TCBB.2013.146 PMID: 26355518

[89] Wei, L.; Wan, S.; Guo, J.; Wong, K.K. A novel hierarchical selective ensemble classifier with bioinformatics application. *Artif. Intell. Med.,* **2017**, *83*, 82-90.
http://dx.doi.org/10.1016/j.artmed.2017.02.005 PMID: 28245947

[90] Wei, L.; Xing, P.; Zeng, J.; Chen, J.; Su, R.; Guo, F. Improved prediction of protein-protein interactions using novel negative samples, features, and an ensemble classifier. *Artif. Intell. Med.,* **2017**, *83*, 67-74.
http://dx.doi.org/10.1016/j.artmed.2017.03.001 PMID: 28320624

[91] Jin, Q.; Cui, H.; Sun, C.; Meng, Z.; Su, R. Free-form tumor synthesis in computed tomography images *via* richer generative adversarial network. *Knowl. Base. Syst.,* **2021**, *218*, 106753.
http://dx.doi.org/10.1016/j.knosys.2021.106753

[92] Liu, J. Classification and gene selection of triple-negative breast cancer subtype embedding gene connectivity matrix in deep neural network. *Brief. Bioinform.,* **2021**, 1477-4054.
http://dx.doi.org/10.1093/bib/bbaa395

[93] Su, R.; Liu, X.; Jin, Q.; Liu, X.; Wei, L. Identification of glioblastoma molecular subtype and prognosis based on deep MRI features. *Knowl. Base. Syst.,* **2021**, *232*, 107490.
http://dx.doi.org/10.1016/j.knosys.2021.107490

[94] Cheng, L. gutMGene: A comprehensive database for target genes of gut microbes and microbial metabolites. *Nucleic Acids Res.,* **2022**, *50*(D1), D795-D800.
PMID: 34500458

[95] Zhao, T.; Hu, Y.; Peng, J.; Cheng, L. DeepLGP: A novel deep learning method for prioritizing lncRNA target genes. *Bioinformatics,* **2020**, *36*(16), 4466-4472.
http://dx.doi.org/10.1093/bioinformatics/btaa428
PMID: 32467970

[96] Zhang, L.; Xiao, X.; Xu, Z.C. iPromoter-5mC: A novel fusion decision predictor for the identification of 5-methylcytosine sites in genome-wide DNA promoters. *Front. Cell Dev. Biol.,* **2020**, *8*, 614.
http://dx.doi.org/10.3389/fcell.2020.00614 PMID: 32850787

[97] Xu, Z. DLpTCR: An ensemble deep learning framework for predicting immunogenic peptide recognized by T cell receptor. *Brief Bioinform,* **2021**, *22*(6), bbab335.
http://dx.doi.org/10.1093/bib/bbab335

[98] Aliper, A. Deep learning applications for predicting pharmacological properties of drugs and drug repurposing using transcriptomic data. *Mol. Pharmaceutics.,* **2016**, *13*(7), 2524-2530. http://dx.doi.org/10.1021/acs.molpharmaceut.6b00248

[99] Zhao, K.; So, H.-C. Drug repositioning for schizophrenia and depression/anxiety disorders: A machine learning approach leveraging expression data. *IEEE J. Biomed. Health Inform.,* **2018**, *99*, 1-1.

[100] Shang, Y.; Gao, L.; Zou, Q.; Yu, L. Prediction of drug-target interactions based on multi-layer network representation learning. *Neurocomputing,* **2021**, *434*, 80-89. http://dx.doi.org/10.1016/j.neucom.2020.12.068

[101] Liu, K.; Sun, X.; Jia, L.; Ma, J.; Xing, H.; Wu, J.; Gao, H.; Sun, Y.; Boulnois, F.; Fan, J. Chemi-Net: A molecular graph convolutional network for accurate drug property prediction. *Int. J. Mol. Sci.,* **2019**, *20*(14), E3389. http://dx.doi.org/10.3390/ijms20143389 PMID: 31295892

[102] Min, S.; B. Lee; and S. Yoon. Deep learning in bioinformatics. Briefings in Bioinformatics, 2017. 18(5): p. 851-869.

[103] Yu, L.; Wang, M.; Yang, Y.; Xu, F.; Zhang, X.; Xie, F.; Gao, L.; Li, X. Predicting therapeutic drugs for hepatocellular carcinoma based on tissue-specific pathways. *PLOS Comput. Biol.,* **2021**, *17*(2), e1008696. http://dx.doi.org/10.1371/journal.pcbi.1008696 PMID: 33561121

[104] Cheng, F.; Liu, C.; Jiang, J.; Lu, W.; Li, W.; Liu, G.; Zhou, W.; Huang, J.; Tang, Y. Prediction of drug-target interactions and drug repositioning *via* network-based inference. *PLOS Comput. Biol.,* **2012**, *8*(5), e1002503. http://dx.doi.org/10.1371/journal.pcbi.1002503 PMID: 22589709

[105] Luo, Y.; Zhao, X.; Zhou, J.; Yang, J.; Zhang, Y.; Kuang, W.; Peng, J.; Chen, L.; Zeng, J. A network integration approach for drug-target interaction prediction and computational drug repositioning from heterogeneous information. *Nat. Commun.,* **2017**, *8*(1), 573. http://dx.doi.org/10.1038/s41467-017-00680-8 PMID: 28924171

[106] Himmelstein, D.S.; Lizee, A.; Hessler, C.; Brueggeman, L.; Chen, S.L.; Hadley, D.; Green, A.; Khankhanian, P.; Baranzini, S.E. Systematic integration of biomedical knowledge prioritizes drugs for repurposing. *eLife,* **2017**, *6*, 6. http://dx.doi.org/10.7554/eLife.26726 PMID: 28936969

[107] Chen, X.; Liu, M-X.; Yan, G-Y. Drug-target interaction prediction by random walk on the heterogeneous network. *Mol. Biosyst.,* **2012**, *8*(7), 1970-1978. http://dx.doi.org/10.1039/c2mb00002d PMID: 22538619