Journal of Pharmaceutical Analysis 15 (2025) 101315

Contents lists available at ScienceDirect

Journal of Pharmaceutical Analysis

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

Original article

DTLCDR: A target-based multimodal fusion deep learning framework for cancer drug response prediction

Jie Yu [1], Cheng Shi [1], Yiran Zhou, Ningfeng Liu, Xiaolin Zong, Zhenming Liu [**], Liangren Zhang [*]

State Key Laboratory of Natural and Biomimetic Drugs, School of Pharmaceutical Sciences, Peking University, Beijing, 100191, China

ARTICLE INFO

Article history:
Received 2 December 2024
Received in revised form
27 March 2025
Accepted 16 April 2025
Available online 21 April 2025

Keywords:
Personalized medicine
Cancer drug response
Multimodal fusion
Deep learning
Drug-target interaction
Single-cell language model

ABSTRACT

Accurate prediction of drug responses in cancer cell lines (CCLs) and transferable prediction of clinical drug responses using CCLs are two major tasks in personalized medicine. Despite the rapid advancements in existing computational methods for preclinical and clinical cancer drug response (CDR) prediction, challenges remain regarding the generalization of new drugs that are unseen in the training set. Herein, we propose a multimodal fusion deep learning (DL) model called drug-target and single-cell language based CDR (DTLCDR) to predict preclinical and clinical CDRs. The model integrates chemical descriptors, molecular graph representations, predicted protein target profiles of drugs, and cell line expression profiles with general knowledge from single cells. Among these features, a well-trained drug-target interaction (DTI) prediction model is used to generate target profiles of drugs, and a pretrained single-cell language model is integrated to provide general genomic knowledge. Comparison experiments on the cell line drug sensitivity dataset demonstrated that DTLCDR exhibited improved generalizability and robustness in predicting unseen drugs compared with previous state-of-the-art baseline methods. Further ablation studies verified the effectiveness of each component of our model, highlighting the significant contribution of target information to generalizability. Subsequently, the ability of DTLCDR to predict novel molecules was validated through *in vitro* cell experiments, demonstrating its potential for real-world applications. Moreover, DTLCDR was transferred to the clinical datasets, demonstrating satisfactory performance in the clinical data, regardless of whether the drugs were included in the cell line dataset. Overall, our results suggest that the DTLCDR is a promising tool for personalized drug discovery.

© 2025 The Authors. Published by Elsevier B.V. on behalf of Xi'an Jiaotong University. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Personalized medicine aims to identify the most suitable therapy for each patient, and is becoming increasingly indispensable in cancer treatment [1]. Due to tumor heterogeneity among patients, individuals with the same type of cancer may show different therapeutic responses to the same treatment [2]. Accurate prediction of the response of each patient to a drug is a major challenge in personalized medicine [3]. As clinical experiments to determine drug responses are both time-consuming and costly, artificial intelligence (AI) has been increasingly applied in recent years to predict the potential responses of patients with cancer to drugs [4,5]. Ideally, computational models should be trained using clinical data and then used to predict patient responses. However, currently available clinical patient databases (e.g., The Cancer Genome Atlas (TCGA)) do not have enough drug response data for effective training [6]. Alternatively, the availability of large-scale cell line-based drug screening data (e.g., Genomics of Drug Sensitivity in Cancer (GDSC)) enables the development of effective preclinical computational models for predicting anticancer drug responses [7].

Based on preclinical and clinical drug response data combined with genetic profiles, several studies have applied machine learning (ML) and deep learning (DL) algorithms to develop cancer drug response (CDR) prediction models [8]. These models can be categorized into multi-drug and single-drug models depending on the number of drugs used for training [9]. Preclinical CDR prediction tasks generally utilize multi-drug models that integrate both drug and cell line features as inputs [8]. In contrast, almost all clinical CDR prediction tasks employ single-drug models that utilize only omics data to build individual models for each drug [9,10].

In preclinical CDR prediction research, many efforts have been made to construct integrative frameworks to enhance prediction

Peer review under responsibility of Xi'an Jiaotong University.
* Corresponding author.
** Corresponding author.
E-mail addresses: zmliu@bjmu.edu.cn (Z. Liu), liangren@bjmu.edu.cn (L. Zhang).
[1] Both authors contributed equally to this work.

https://doi.org/10.1016/j.jpha.2025.101315
2095-1779/© 2025 The Authors. Published by Elsevier B.V. on behalf of Xi'an Jiaotong University. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

accuracy, focusing on the effective extraction of drug and cell line features and their combinations for prediction [8]. Various methods have been used for drug feature extraction to obtain improved drug representations [11]. For example, DeepDRA [12] utilized deep neural networks (DNNs) to extract molecular descriptors and fingerprints, tCNNS [13] encoded simplified molecular input line entry system (SMILES) string sequences using convolutional neural networks (CNNs). DeepCDR [14] and GraphDRP [15] leveraged graph neural networks (GNNs) to learn molecular structural information. For cell line feature extraction, the complexity and high dimensionality of omics data make effective extraction a challenging task. DNN, CNN, and autoencoder (AE) have been widely used to extract features from single- or multi-omics data. For example, SWnet [16] uses a one-dimensional (1D) CNN to extract expression and mutation features, and DeepDSC [17] applies stacked AE to encode gene expression features. Moreover, from the perspective of drug sensitivity mechanisms, molecules interact with their targets and subsequently inhibit or activate targeted pathways, ultimately leading to phenotypic changes in cancer cells [18]. Therefore, in addition to drug and genetic profiles, features derived from pathway information, drug-target data, and protein-protein interactions (PPI) have also been considered in CDR prediction models [19,20]. For instance, HiDRA [20] introduced a hierarchical attention network model to represent biological pathways and their member genes related to drug responses. PathDNN [19] restructures the canonical DNN architecture by introducing a layer of pathway nodes connected to the input gene and drug target nodes. Notably, PathDNN only uses the confidence score from Search Tool for Interacting Chemicals (STITCH) as drug-target interactions (DTIs) to represent drug features, which is quite different from other CDR prediction models that rely on drug chemical structure features. However, PathDNN can only be used for molecules in the STITCH database. Another potential source for cell line features comes from single-cell data [21]. Zhao et al. [22] developed a large-scale cell representation learning model called CellLM based on single-cell RNA sequencing (scRNA-seq) data, which demonstrated enhanced semantic representation of cell line data. However, this study mainly focused on single-cell tasks and did not sufficiently address drug response studies in cell lines.

For clinical CDR prediction, many alternative methods have been proposed for training from preclinical to clinical, with the main difficulty being effectively bridging the gap between cell lines and clinical data. Specifically, these single-drug models are first trained to relate cell line omics to the response to a specific drug and then to estimate the response of patient tumors using clinical omics. For example, Huang et al. [23] systematically evaluated various ML algorithms using GDSC sample training to predict the CDR of TCGA patients, suggesting that tissue guided-least absolute shrinkage and selection operator (TG-LASSO), which employs LASSO regression, outperforms other ML methods. Another related study formulated a meta-learning approach called translation of cellular response prediction (TCRP), which was pretrained on cell-lines data and then utilized few-shot learning techniques to tune the general model for the Patient-Derived Tumor Xenograft (PDTX)-derived tumor cells (PDTC) dataset [24].

Despite significant advancements in CDR prediction models, many challenges remain in achieving accurate predictions. Many preclinical CDR models have not been validated for predicting drug responses to unseen drugs and cell lines, and most validated models demonstrate poor prediction performance, especially for unseen drugs [8]. This may be due to the limited number of drugs in the dataset and the small sample size of drugs providing limited structural information, leading to high uncertainty in the predictions. A major limitation of current clinical CDR models is that they can only validate drugs shared between clinical and cell line databases and cannot predict drugs outside the cell line datasets [9]. This is because these models were trained independently for each drug in the cell-line datasets and tested with each drug individually in the clinical datasets [25]. Therefore, to address these limitations, we need a model that can not only generalize well to unseen drugs in cell lines, but also demonstrate great accuracy and generalizability in clinical data regardless of whether the drugs are included in the cell line dataset.

It is well recognized that drugs interact with multiple targets, and drugs with similar targets often exhibit similar phenotypic effects [26]. Thus, we hypothesized that target information may be a powerful characteristic of drugs in predicting drug responses. However, only a few activities are known for each drug on the list of cancer-related targets. Therefore, it is necessary to obtain a complete target profile for each drug. In addition to target information, inspired by the large-scale pretrained language models single-cell bidirectional encoder representations from transformers (scBERT) [27] and CellLM [22], we propose a model called drug-target and single-cell language based CDR (DTLCDR) to predict preclinical and clinical CDR. Specifically, a well-trained DTI model was employed to generate the predicted drug target profiles, and a module based on pretrained scBERT was utilized to transfer knowledge from single cells to cell lines. In addition, DTLCDR integrates chemical descriptors, molecular graph representations, and cell line expression profiles. We conducted comprehensive evaluations of DTLCDR: i) The comparison experiments demonstrated that DTLCDR exhibits superior generalizability and robustness in predicting unseen drugs compared to other advanced methods. ii) The ablation experiments indicated that incorporating predicted drug target profiles and single-cell representations improved the performance of DTLCDR, and the evaluation of variants in other databases further validated the effectiveness of the drug target information. iii) *In vitro* cell assays for drug reproposing demonstrated that DTLCDR achieved an excellent hit rate for predicting and screening new potential anticancer active compounds. iv) Evaluation of transferred DTLCDR (DTLACDR) on the clinical datasets PDTC and TCGA demonstrated satisfactory performance. Unlike other methods, DTLACDR can be used to predict drugs that are not included in cell line datasets.

2. Materials and methods

2.1. Model architecture

As shown in Fig. 1A, our DL framework consists of three components: i) a DTI prediction model that is constructed to generate the complete drug-target profile as a drug feature input for CDR tasks, ii) a CDR model for predicting drug response in cell lines, and iii) transfer of the CDR model to clinical prediction. The source code is available at https://github.com/yujie0317/DTLCDR.git.

First, we developed a novel DTI prediction model called graph convolution attention drug-target interaction (GCADTI), which consists of graph convolution networks (GCNs) and CNN for feature extraction, along with an attention module for feature fusion. Specifically, GCADTI leverages GCN and CNN to encode drugs and proteins as 2D molecular graphs and 1D protein sequences, respectively. A mutual attention network was then applied to generate drug- and protein-related context matrices for each pair. We concatenated these matrices and fed them to a fully connected block to predict DTI using binary cross-entropy as a loss function.

Second, the DTLCDR implemented in the cell lines consist of two side inputs. On the drug side, we used the well-trained GCADTI from the previous step to generate potential activity against cancer-

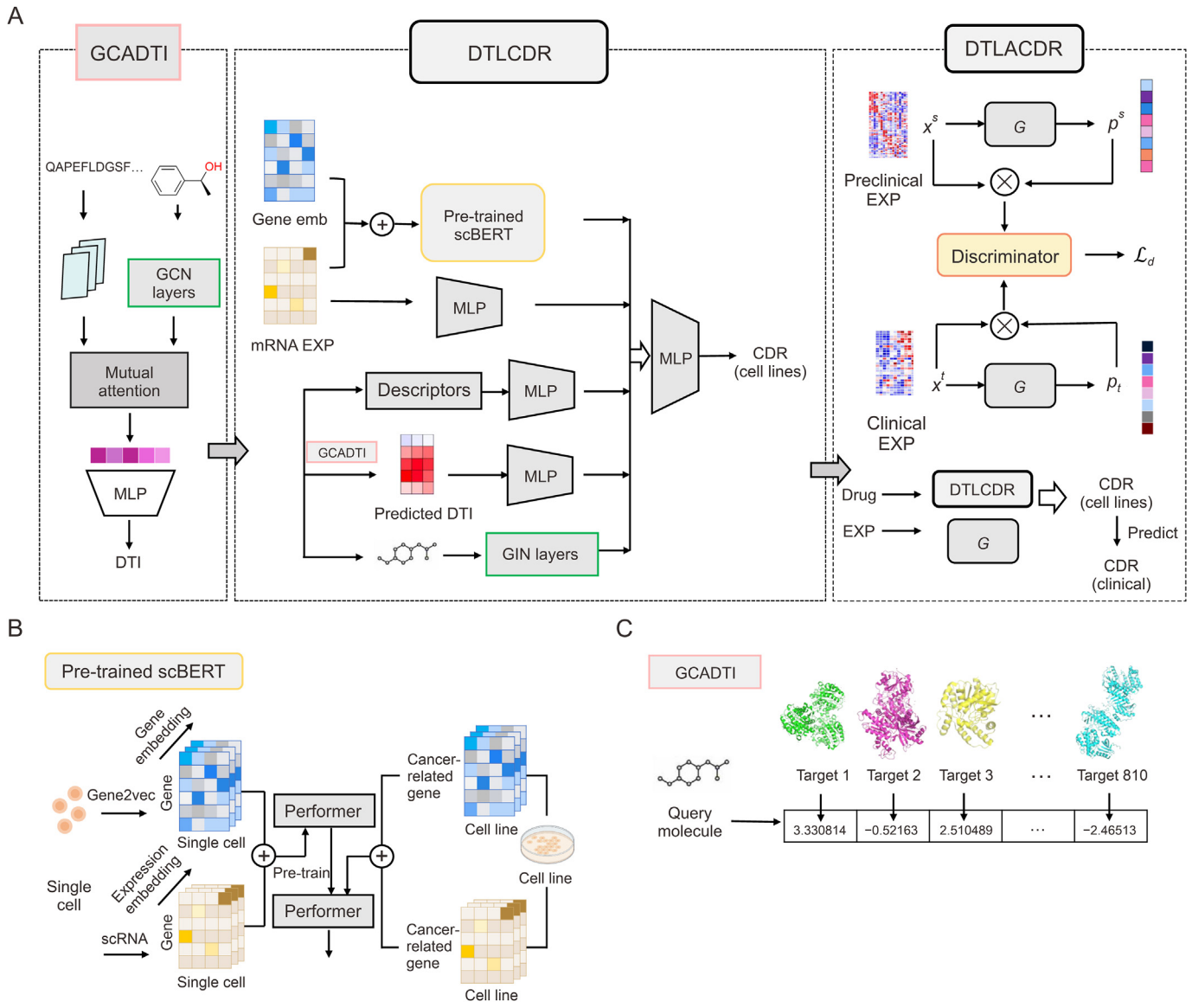


Fig. 1. Overview of the deep learning (DL) framework. (A) The framework consists of three components: the graph convolution attention drug-target interaction (GCADTI) architecture for generating the target profiles of drugs, the drug-target and single-cell language based cancer drug response (CDR) (DTLCDR) architecture for predicting drug response in cell lines, and the transferred DTLCDR (DTLACDR) architecture for clinical CDR prediction. (B) The pretrained single-cell bidirectional encoder representations from transformers (scBERT) architecture in DTLCDR. (C) The target profiles of drugs predicted by GCADTI. GCN: graph convolution network; MLP: multilayer perceptron; DTI: drug-target interaction; Gene emb: gene embedding; mRNA: messenger RNA; EXP: expression; scRNA: single-cell RNA. Fig. 1B was drawn by using BioRender.com and Fig. 1C was drawn by using PyMOL software (pymol.org).

related protein targets for each drug. This series of potential activity values forms the input features with the feature dimension equal to the number of proteins. Additionally, graph isomorphic networks (GINs) [28] were used to extract structural information from molecular graphs, and a multilayer perceptron (MLP) was used to extract molecular descriptors. On the cell line side, a pretrained model called scBERT was fine-tuned to provide cell line gene expression features with generic knowledge of single cells. Additionally, MLP was used to extract gene expression features. Subsequently, all the extracted features were concatenated and fed to a fully connected block to predict the CDR using the mean square error (MSE) as the loss function.

Finally, through integrating the conditional adversarial neural network (CDAN) [29] to align cell line genomics with clinical genomics information, we transformed DTLCDR to DTLACDR. Additionally, the loss function was modified to transform the regression model into a classification model, which was trained on preclinical CDR data to predict the clinical CDR (Fig. 1A).

2.1.1. Pretrained scBERT

As illustrated in Fig. 1B, scBERT is a large-scale deep language model that is pretrained on millions of unlabeled scRNA-seq data points and then transferred to downstream tasks through fine-tuning [27]. The scBERT embedding includes gene and expression embedding. Gene embeddings are derived from Gene2Vec [30], which provide unique embeddings for each gene to represent its identity. These embeddings can be considered as relative embeddings that capture the semantic similarities between any two genes. For expression embeddings, a term-frequency analysis method was applied, in which continuous expression variables were discretized into bins and then converted into 200D vectors. Subsequently, expression and gene embeddings were added as

inputs and fed into the Performer blocks. Through self-supervised pretraining on unlabeled data, the model can learn general deep representations of gene expression patterns and gene-gene interactions, which might enhance cell representation for downstream tasks. In this study, we download the scBERT model that pretrained on PanglaoDB (https://github.com/TencentAILabHealthcare/scBERT) [27] and then fine-tuned the model on cell line datasets. Unlike the single-cell gene expression used during the pre-training stage, the expression embeddings used in the fine-tuning stage were derived from approximately 600 cancer-related gene expression in cell lines.

2.1.2. GCADTI for predicting the target profiles of drugs

As shown in Fig. 1C, the GCADTI model generated potential activity values for the cancer-related protein targets of each drug. Each drug is represented as an 810D vector, with each value corresponding to its predicted activity against a specific cancer-related target. Subsequently, the drug-target activity profile predicted by GCADTI was fed into a three-layer MLP for feature extraction of target information (Fig. 1A).

2.1.3. CDAN module

ML models generally perform well on data from a similar distribution but perform worse on data from a different distribution [29]. Because the fundamental challenge in clinical CDR prediction is aligning the distribution between cell-line data and clinical data, CDAN was used to enhance generalization from the preclinical (source domain) to the clinical (target domain). As illustrated in Fig. 1A, the output of the expression-MLP (EXP-MLP) module in the DTLCDR was used as the input for the CDAN module. The CDAN module comprises feature extractors G, a classifier C, and a domain discriminator D. The goal of the CDAN is to construct a classifier that performs well in both the source and target domains, whereas the classifier in the CDAN is equivalent to that in DTLCDR.

In Fig. 1A, the source domain data (x^s, y^s) and target domain data x^t are defined as follows: x^s and x^t represent the features of preclinical and clinical gene expression, respectively, extracted from the EXP-MLP module in DTLCDR. The y^s denotes the source domain labels, whereas the target domain label y^t remains unseen during the training stage.

First, the extractor G extracts features from the source and target domain data: $f^s = G(x^s)$ and $f^t = G(x^t)$. Classifier C then uses feature f^s for classification, generates predictions $p^s = C(f^s)$, and uses f^t to generate predictions $p^t = C(f^t)$. By taking the outer product of features f and predictions p, the conditional information h is generated: $h = f \times p$. The conditional information h encapsulates the joint information of features and predictions, making it more effective at capturing distribution differences between domains than using features alone. The domain discriminator is trained through adversarial learning to align the distributions of the source and target domain data based on the conditional information h. The loss function of the domain discriminator D is $\mathscr{L}_d$ and the final optimization objective is achieved by minimizing the weighted sum of the classification and domain discriminator losses.

2.2. Data set and preparation

2.2.1. Drug targets

To obtain the target information used for drug features in CDR prediction, 1572 cancer-related targets were collected from a Therapeutic Target Database (TTD). Bioactivity data for each target were collected from three publicly available databases: CHEMBL, BindingDB, and TTD. Subsequently, the preprocessing followed these protocols: i) data with unit of half maximal inhibitory concentration (IC_{50}) was selected and converted to pIC_{50} ($pIC_{50} = -\log(IC_{50})$) using the negative logarithm of IC_{50}; ii) duplicated molecules were merged based on their canonical SMILES, with the arithmetic mean of pIC_{50} retained if the fluctuation was less than 0.3, otherwise dropped; iii) the pIC_{50} values were converted into two classes: active ($pIC_{50} \geq 6$) and inactive ($pIC_{50} < 6$); and iv) proteins with fewer than 20 bioactive data points were excluded. The final merged dataset contained 994,953 drug-protein pairs composed of 978 proteins and 646,452 compounds. The number of positive samples was 594,238 and the number of negative samples was 400,715.

2.2.2. Drug response on cell lines

The two most commonly used drug sensitivity datasets, GDSC and Cancer Cell Line Encyclopedia (CCLE), were used to derive the data. Drug-cell line pairs were derived from Genomics of Drug Sensitivity in Cancer version 2 (GDSC2) (https://www.cancerrxgene.org/), whereas gene expression profiles were obtained from CCLE (https://portals.broadinstitute.org/ccle). Subsequently, preprocessing was performed in two steps: i) remove drugs without PubChem ID and retrieve SMILES from the PubChem database, and ii) cell lines were excluded from the GDSC whose Catalogue of Somatic Mutations in Cancer (COSMIC) IDs were not included in the CCLE. A more detailed preprocessing protocol can be found in a previously reported study [31]. The final dataset contained 117,665 drug-cell line pairs with natural log-transformed IC_{50} values, comprising 800 cell lines and 175 drugs. Owing to the high dimensionality of the expression profiles, we performed variance analysis for feature selection, retaining the top 3000 genes ranked by variance. Genes with lower variance were excluded because their expression differences across cell lines were minimal, and were considered insignificant for further analysis. The preprocessing methods for the GDSC1 and Cancer Therapeutics Response Portal Version 2 (CTRPv2) databases were similar to those for the GDSC2 [32].

For clinical CDR prediction, the training set of GDSC responses was categorized into binary labels based on the area under the curve (AUC), with the median AUC value of each drug as the threshold.

2.2.3. TCGA patients

The TCGA dataset provides various patient-related information such as treatment methods, new tumor events, days after treatment, and survival time [6]. Referring to the data preprocessing approach of the baseline model context-aware deconfounding autoencoder (CODE)-AE in the comparison experiments [10], we classified TCGA samples into responders and non-responders based on the number of new tumor event days after treatment, using the median number of days to a new tumor event as the threshold. Only patients who received single-drug therapy throughout the treatment duration were included and drugs with more than 20 labeled samples were retained. The TCGA dataset also provides the gene expression profiles of patients. To align with the training dataset, GDSC, only the overlapping genes between the TCGA and GDSC gene sets were retained, and the top 3000 genes ranked by variance were selected.

2.2.4. PDTC dataset

PDTC are a dataset of cancer cells derived from patient-derived xenografts (PDX) in breast cancer [33]. The intra-tumoral genomic clonal architecture present in the originating breast cancer was mostly preserved upon serial passaging in xenografts and short-term cultured PDTCs. Drug responses in PDTCs were assessed using a high-throughput platform. PDTC responses were categorized into binary labels based on the AUC values with the median AUC value of each drug as the threshold. Similar to TCGA, only overlapping genes between the PDTC and GDSC gene sets were retained, and the top 3000 genes ranked by variance were selected (https://doi.org/10.6084/m9.figshare.2069274.v16).

3. Results and discussion

One strategy of the DTLCDR model is to extract drug features more effectively by integrating target information, thereby enhancing its generalization capability. The mechanisms of action of drugs are usually complex and multi-targeted, but the drug-target profiles obtained from bioactivity experiments are incomplete. Using sparse matrices as drug-target profiles introduces considerable erroneous information, thus we plan to train a high-accuracy DTI prediction model to generate potential target activity profiles to provide drug features for CDR tasks. Another strategy of the DTLCDR is to attempt to enhance cell representation by integrating a pretrained language model that has learned general scRNA-seq knowledge. We conducted a series of experiments to assess these strategies. First, we evaluated the accuracy and generalization performance of the proposed DTI model (GCADTI) through several comparative experiments. Second, we assessed the accuracy and generalization of DTLCDR in comparison experiments. We then performed ablation studies to validate the contribution of the target information and pretrained cell-language representations to model improvements. Additionally, evaluation of variants of GDSC1 and CTRPv2 further validated the effectiveness of the target information. Third, we used DTLCDR for drug repurposing and validated it using cell viability experiments. Finally, we evaluated the performance of transferring DTLCDR to PDTC and TCGA data.

3.1. Evaluation of GCADTI in comparison experiments

To ensure the accuracy of the target activity of drugs generated using the GCADTI model, we conducted systematic performance comparison experiments using GCADTI with three state-of-the-art baseline methods (GraphDTA [34], DeepConv-DTI [35], and DrugBAN [36]). The experimental settings used were the following four dataset splitting strategies (Fig. S1). i) Warm start: compound-protein pairs are randomly split into training and test sets in a 9:1 ratio, with the test set including some compounds or proteins present in the training data. ii) Cold start for drugs: compounds are randomly split into training and test sets in a 9:1 ratio, ensuring that compounds in the test set are unseen in the training set, while some proteins are present in both sets. iii) Cold start for proteins: proteins are randomly split into training and test sets in a 9:1 ratio, ensuring that proteins in the test set are unseen in the training set. iv) Cold start for both drugs and proteins: proteins were randomly split into training and test sets in an 8:2 ratio, and compounds already present in the training set were removed from the test set, ensuring no overlap of compounds or proteins between the two sets. In all settings, a part of the training set was then divided into a final ratio of 8:1:1 for the training, validation, and test sets, and the validation set was used to decide when to stop the training. For each setting, we conducted three independent runs using three random seeds for dataset partitioning, followed by model training and reporting of the average results along with the standard deviation. In this study, DTI prediction was formulated as a classification problem, with the area under the receiver operating curve (AUROC) and area under the precision-recall curve (AUPRC) as the evaluation metrics.

The overall performances of all methods under the four dataset split settings are shown in Fig. 2A and Table S1. We can see that even though all models experience a performance drop from the warm start scenario to the cold star scenario, especially for unseen protein scenarios, our proposed GCADTI still outperforms all other

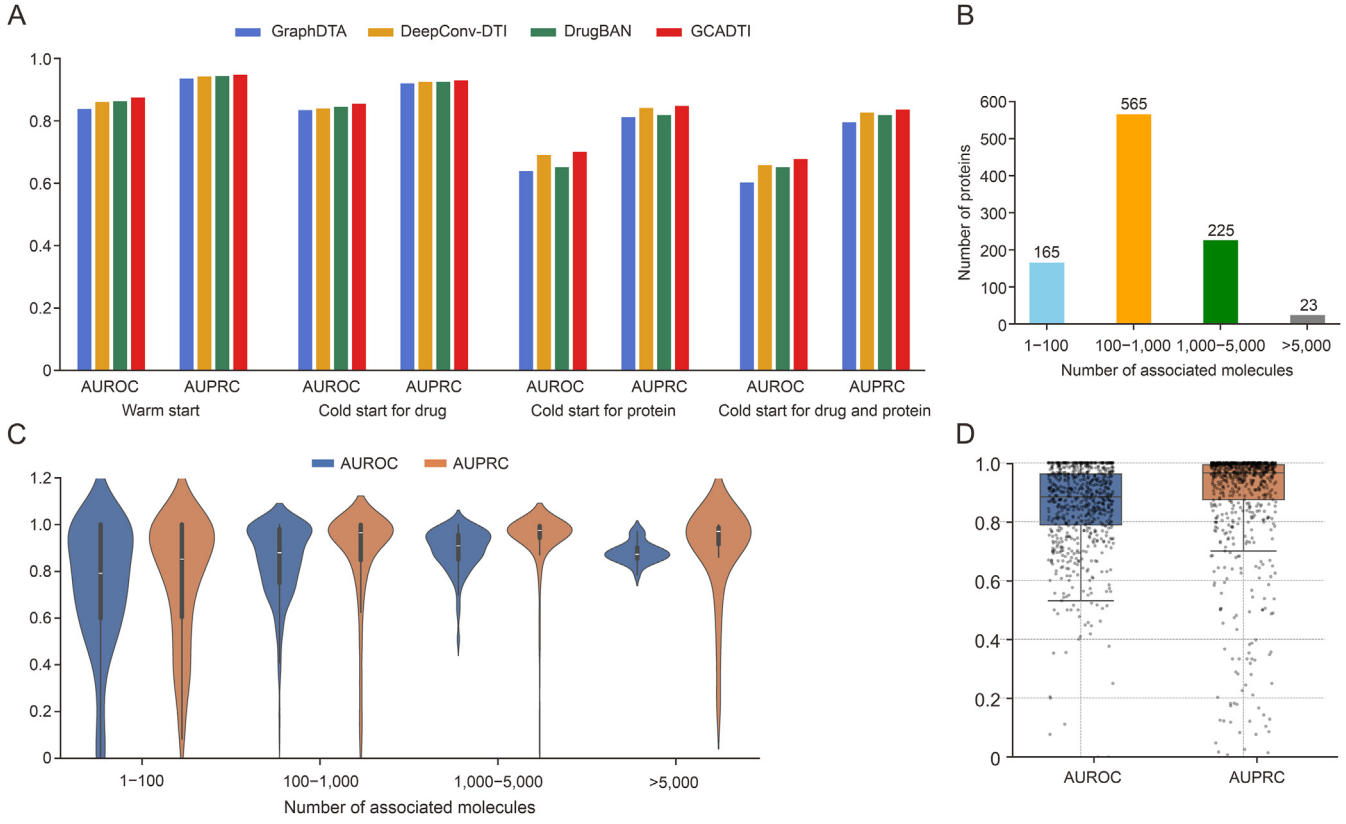


Fig. 2. Performance evaluation of graph convolution attention drug-target interaction (GCADTI). (A) Performance comparison between GCADTI and baseline methods under four dataset split settings, each value of area under the receiver operating curve (AUROC) and area under the precision-recall curve (AUPRC) was the average result of three independent tests. (B) The counts of bioactivity data points associated with each protein target in the training set. (C) The performance of GCADTI across targets with different sample sizes. (D) Distribution of AUROC and AUPRC across different proteins in the drug cold start scenario. Each point represents a protein. DTI: drug-target interaction.

state-of-the-art baseline models. GCADTI achieved AUROC of 0.8744, 0.8544, 0.7001, and 0.6773 in the warm start, drug cold start, protein cold start, and both cold start scenarios, respectively.

Because not all drugs in the datasets of the CDR prediction task can be found in the DTI datasets, prediction of target profiles for the following DTLCDR model tends to correspond to the drug cold start scenario. Fig. 2B shows the counts of the bioactivity data points associated with each protein target in the training set. Among the 978 cancer-related targets in the preprocessed dataset, 165 targets had fewer than 100 training samples, most targets had between 100 and 5000 training samples, and only 23 targets had more than 5000 training samples. It is widely recognized that predicting DTIs for targets associated with less training data is more challenging [37]. We evaluated GCADTI's DTI prediction capability of GCADTI for 978 proteins in a drug cold start scenario. As shown in Fig. 2C, GCADTI achieved high prediction accuracy across targets with varying sample sizes. For targets with fewer than 100 training samples, the median AUROC in the test set was slightly lower than 0.8. For targets within the other three sample size ranges, the median AUROC exceeded 0.85. Although GCADTI's performance of the GCADTI on targets with small sample sizes (<100) was slightly lower than that on targets with larger sample sizes, it still demonstrated satisfactory prediction accuracy and generalization capability.

Additionally, the scatter plot in Fig. 2D illustrates that GCADTI achieve AUROC values between 0.8 and 1.0 for most proteins, with a small portion between 0.6 and 0.8, and only a very few below 0.6. The median AUROC and AUPRC values were 0.8846 and 0.9658, respectively. Therefore, we suggest that the GCADTI can provide a reliable predicted target activity profile for subsequent CDR prediction tasks. To further refine the predictions, protein targets with AUROC values below 0.5 were excluded. Ultimately, 810 protein targets with superior prediction performances were selected for DTLCDR, with each drug represented by an 810D feature.

3.2. Evaluation of DTLCDR in comparison experiments

For performance comparison experiments, we compared DTLCDR with three state-of-the-art baseline methods on the GDSC2 dataset: DeepTTA [38], GraphDRP [15], and twin graph neural networks for drug response prediction (TGDRP) [39]. DeepTTA is a transformer-based model that creates drug embeddings from the SMILES substructure and uses an MLP module to extract features from the gene expression data [38]. GraphDRP uses GNNs for drug feature extraction, and 1D CNN for genomic feature extraction [15]. TGDRP abstracts cell lines as graphs based on Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) PPI networks and employs GNNs for both cell lines and drug representation learning [39].

Similar to the three scenarios for DTI prediction mentioned above, the experimental settings used the following dataset split strategies. i) Warm start: randomly split drug-cell line pairs. ii) Cold start for cell lines: prediction of unseen cell lines. iii) Cold start for drugs: prediction of unseen drugs. For each setting, we conducted three independent runs using three random seeds for dataset partitioning, followed by model training and reporting of the average results along with the standard deviation. In this study, CDR prediction was formulated as a regression problem using three evaluation metrics: Pearson's correlation coefficient (PCC), Spearman's correlation coefficient (SPM), and R^2. All competitive method experiments were conducted on the GDSC2 dataset with a 5-fold cross-validation procedure, while a 10-fold cross-validation was implemented in the drug cold-start scenario.

The overall performances of all methods under the three dataset split settings are shown in Fig. 3A and Table 1. Compared to the warm start scenario, the performances of all models were slightly lower in the cold start scenario. Our proposed DTLCDR model achieved the highest PCC value (0.9492) in the warm start experiments, but slightly underperformed TGDRP in the SPM and R^2 metrics, while outperforming the other two baseline models, DeepTTA and GraphDRP. In cell-line cold start experiments, DTLCDR showed a PCC value of 0.8667, ranking second in comparison, whereas it ranked third in the SPM and R^2 metrics, underperforming DeepTTA and TGDRP. Moreover, the performance gaps between the different models were insignificant in both scenarios. Most models can reach a high PCC value > 0.8 during the cold start for cell lines, indicating that predicting the effect of a certain drug on a new cell line is relatively easy when its effect on other cell lines is known.

As shown in Fig. 3A, all models experienced a significant performance drop from the warm start to the cold start for drug. This was predictable because the training set contained only 175 compounds, and as a result, the models only learned limited chemical information, which was insufficient for accurately predicting molecules with novel structures. Despite the significant challenge of predicting unseen drugs, our model outperformed all baseline models on the three metrics. As shown in Table 1, compared to the best baseline method, DeepTTA, DTLCDR improved PCC by 0.04, SPM by 0.02, and R^2 by 0.17 in the drug cold start scenario. Additionally, the standard deviations of all three metrics for DTLCDR are lower than those of the other baseline models, indicating that DTLCDR demonstrates stable performance in predicting different novel structures. In summary, the comparison experiments indicated that DTLCDR has improved generalizability and robustness in predicting drug responses to new molecular structures. This may be because DTLCDR incorporates drug target information, whereas the other models focus solely on the extraction of more efficient drug structural information.

3.3. Ablation studies

We conducted an ablation experiment to verify the effectiveness of the individual components of the proposed DTLCDR model. Specifically, we removed one or two components from the DTLCDR framework, observed the subsequent performance, and analyzed their contributions to the overall model performance. We implemented five variants of the DTLCDR: DTLCDR (Drug-GIN), DTLCDR (Drug-descriptors (Drug-desc)), DTLCDR (GCADTI), DTLCDR (pretrained scBERT language model (P-LM)), and DTLCDR (Exp-MLP). For example, DTLCDR (Drug-GIN) uses only the GIN module for drug feature extraction, removing both the drug descriptors and target information from GCADTI. Similarly, DTLCDR (Drug-desc): Drug-desc + P-LM + Exp-MLP; DTLCDR (GCADTI): GCADTI + P-LM + Exp-MLP; DTLCDR (P-LM): Drug-GIN + Drug-desc + GCADTI + P-LM; and DTLCDR (Exp-MLP): Drug-GIN + Drug-desc + GCADTI + Exp-MLP. To better compare the contributions of drug target information and drug structural information to the model, all ablation experiments were conducted in the drug cold start scenario.

As shown in Table 2 and Fig. 3B, each variant exhibited a decrease in all three metrics compared to DTLCDR, indicating that each module made a positive contribution to the performance of DTLCDR. Comparing the extraction between the drug and cell features, it was observed that the contribution of drug features was generally greater than that of the cell features. When cell features were removed, the PCC value decreased by 0.0168 and 0.0362 from DTLCDR to the variants, whereas the PCC value decreased by 0.0848, 0.0729, and 0.0584 when drug features were removed. This suggests that drug features contribute more to the overall model performance than cell line features.

For drug features, the DTLCDR (GCADTI) variant exhibited higher values than DTLCDR (Drug-GIN) and DTLCDR (Drug-desc) for the

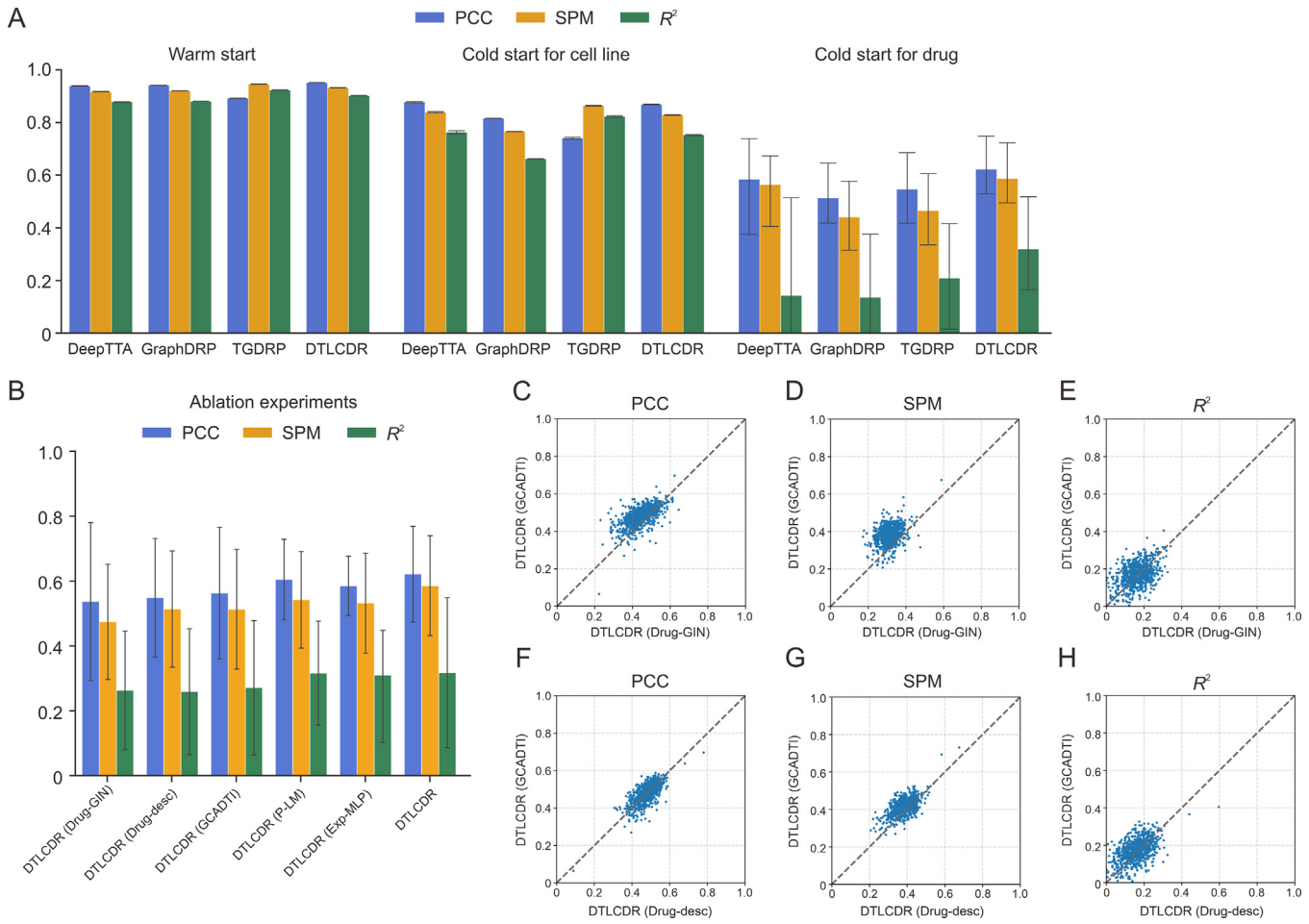


Fig. 3. Performance evaluation of drug-target and single-cell language based cancer drug response (CDR) (DTLCDR). (A) Performance comparison between DTLCDR and baseline methods under three dataset split settings, each value of Pearson's correlation coefficient (PCC), Spearman's correlation coefficient (SPM), and R^2 was the average result of the cross-validation tests. (B) Performance comparison between DTLCDR and five variants in ablation experiments. (C–E) Scatter plots of the PCC (C), SPM (D), and R^2 (E) results of the variants DTLCDR (graph convolution attention drug-target interaction (GCADTI)) and DTLCDR (Drug-graph isomorphic networks (GIN)) across each cell line. Each point represents a cell line. (F–H) Scatter plots of the PCC (F), SPM (G), and R^2 (H) results of the variants DTLCDR (GCADTI) and DTLCDR (Drug-descriptors (Drug-desc)) across each cell line. TGDRP: twin graph neural networks for drug response prediction; DTLCDR (Drug-desc): DTLCDR (Drug-descriptors); DTLCDR (P-LM): DTLCDR (pretrained single-cell bidirectional encoder representations from transformers (scBERT) language model); DTLCDR (Exp-MLP): DTLCDR (expression-MLP (EXP-MLP)).

Tabel 1
Performance comparison of drug-target and single-cell language based cancer drug response (CDR) (DTLCDR) and other baseline models.

Model	Warm start			Cold start for cell line			Cold start for drug		
	PCC	SPM	R^2	PCC	SPM	R^2	PCC	SPM	R^2
DeepTTA	0.9368 ± 0.0004	0.9149 ± 0.0009	0.8760 ± 0.0013	**0.8754 ± 0.0027**	0.8368 ± 0.0031	0.7614 ± 0.0067	0.5822 ± 0.2415	0.5622 ± 0.1782	0.1418 ± 0.6222
GraphDRP	0.9393 ± 0.0007	0.9182 ± 0.0007	0.8778 ± 0.0007	0.8138 ± 0.0007	0.7639 ± 0.0007	0.6604 ± 0.0007	0.5111 ± 0.1519	0.4393 ± 0.1681	0.1347 ± 0.2613
TGDRP	0.8898 ± 0.0011	**0.9437 ± 0.0003**	**0.9209 ± 0.0005**	0.7388 ± 0.0046	**0.8616 ± 0.0021**	**0.8207 ± 0.0028**	0.5447 ± 0.1776	0.4629 ± 0.1764	0.2079 ± 0.2664
DTLCDR	**0.9492 ± 0.0001**	0.9299 ± 0.0004	0.8996 ± 0.0008	0.8667 ± 0.0015	0.8268 ± 0.0017	0.7510 ± 0.0025	**0.6210 ± 0.1470**	**0.5855 ± 0.1537**	**0.3179 ± 0.2304**

The bold data indicates the maximum value in each column of the table. PCC: Pearson's correlation coefficient; SPM: Spearman's correlation coefficient; TGDRP: twin graph neural networks for drug response prediction.

three metrics, demonstrating that drug target information is more effective than drug chemical structure information when predicting drug responses. This can be explained by the ability of the target information to reflect the biological mechanisms of drugs, providing a perspective that extends beyond the molecular structure. The importance of target information becomes particularly apparent when drug data are limited, as it serves as a valuable complement to chemical features, capturing functional or mechanistic similarities between drugs, and thereby enhancing the accuracy of drug response predictions for novel drugs. With respect to drug structural and physicochemical features, although their contribution is less significant than that of the target information, removing either of them still leads to a decrease in the overall performance of the DTLCDR. Therefore, drug structural information and physicochemical features remain important drug representations in CDR prediction tasks and have been widely applied in previous studies. For example, the baseline models DeepTTA, GraphDRP, and TGDRP used in previous comparison experiments employed advanced neural network architectures such as GNNs and transformers for drug structural feature learning.

For cell line features, the DTLCDR (P-LM) variant showed a higher value than DTLCDR (Exp-MLP) on the three metrics,

Table 2
Ablation study on Genomics of Drug Sensitivity in Cancer version 2 (GDSC2).

Method	PCC	SPM	R^2
DTLCDR (Drug-GIN)	0.5362 ± 0.2433	0.4741 ± 0.1773	0.2630 ± 0.1825
DTLCDR (Drug-desc)	0.5481 ± 0.1825	0.5134 ± 0.1790	0.2596 ± 0.1934
DTLCDR (GCADTI)	0.5626 ± 0.2023	0.5129 ± 0.1843	0.2718 ± 0.2061
DTLCDR (P-LM)	0.6042 ± 0.1238	0.5421 ± 0.1486	0.3163 ± 0.1597
DTLCDR (Exp-MLP)	0.5848 ± 0.0911	0.5317 ± 0.1541	0.3099 ± 0.2072
DTLCDR	0.6210 ± 0.1470	0.5855 ± 0.1537	0.3179 ± 0.2304

PCC: Pearson's correlation coefficient; SPM: Spearman's correlation coefficient; GIN: graph isomorphic network; DTLCDR (Drug-desc): drug-target and single-cell language based cancer drug response (CDR) (DTLCDR) (Drug-descriptors); GCADTI: graph convolution attention drug-target interaction; DTLCDR (P-LM): DTLCDR (pre-trained single-cell bidirectional encoder representations from transformers (scBERT) language model); DTLCDR (Exp-MLP): DTLCDR (expression-MLP (EXP-MLP)).

indicating that integration of the pretrained language model enhances prediction accuracy. This is likely because the pretrained scBERT language model acquired general knowledge from large-scale single-cell expression data, such as general co-expression patterns and gene-gene interactions. Fine-tuned scBERT in DTLCDR can capture the underlying gene expression patterns in cell lines, thereby improving cell-line representation.

To further illustrate the results, we present scatter plots showing the results of the variants DTLCDR (GCADTI), DTLCDR (Drug-GIN), and DTLCDR (Drug-desc) across each cell line (Figs. 3C–H). In Figs. 3C–E, the performance of DTLCDR (GCADTI) is compared with that of DTLCDR (Drug-GIN). The results showed that most cell lines were distributed above the diagonal line, particularly for the PCC and SPM metrics, indicating that DTLCDR (GCADTI) significantly outperformed DTLCDR (Drug-GIN) for most cell lines. In Figs. 3F–H, the performance of DTLCDR (GCADTI) is compared with that of DTLCDR (Drug-desc). The results demonstrated that DTLCDR (GCADTI) performed better than DTLCDR (Drug-desc) with respect to most cell lines, particularly for the SPM and R^2 metrics. This further demonstrates that predicting the response to new drugs using drug target information is more accurate than using drug structure and descriptors for most cell lines.

3.4. Evaluation of DTLCDR variants on GDSC1 and CTRPv2 databases

Although the GDSC2 is recognized as a high-quality drug sensitivity database, its limited drug coverage presents constraints for comprehensive model evaluation. To further validate the effectiveness of drug target information in predicting drug response, we evaluated the performance of two variants, DTLCDR (GCADTI) and DTLCDR (Drug-GIN), in a drug cold start test using the GDSC1 and CTRPv2 databases. DTLCDR (Drug-GIN) was selected instead of DTLCDR (Drug-desc) because of the broader application of structural drug information in drug response prediction.

After data preprocessing, the GDSC1 dataset contained 170,498 drug-cell line pairs, covering 343 drugs and 985 cell lines, whereas the CTRPv2 dataset contained 241,670 drug-cell line pairs, covering 544 drugs and 887 cell lines. Compared with GDSC2, these two databases included a relatively larger number of drugs. Both datasets were split into training, validation, and test sets using the drug cold start method, and 10-fold cross-validation was performed to calculate the average values and standard deviations of the evaluation metrics.

Table S2 summarizes the PCC, SPM, and R^2 values of DTLCDR (GCADTI) and DTLCDR (Drug-GIN) for the GDSC1 and CTRPv2 datasets. Fig. 4 compares the performances of the two variants across all cell lines in the GDSC1 and CTRPv2 datasets for the three metrics. In the GDSC1 dataset, a larger number of scatter points

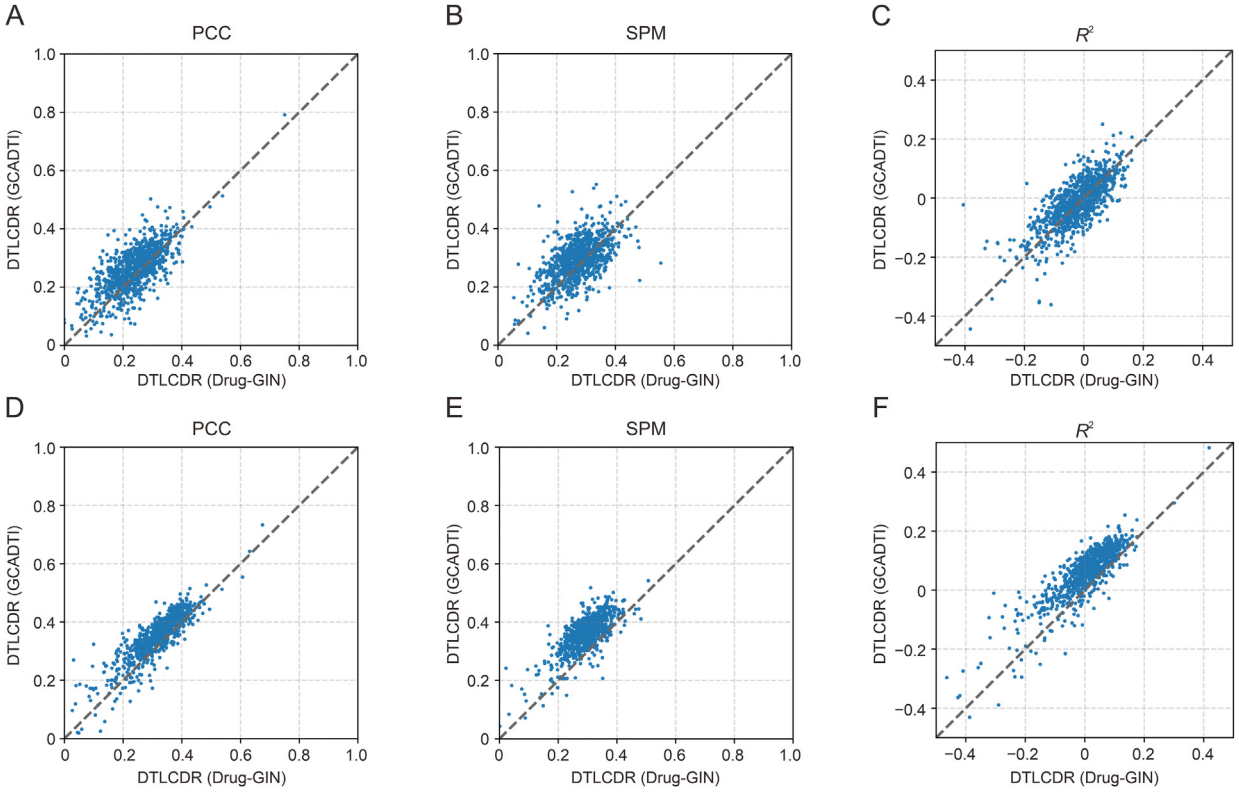


Fig. 4. Scatter plots of the Pearson's correlation coefficient (PCC), Spearman's correlation coefficient (SPM), and R^2 results of the variants drug-target and single-cell language based cancer drug response (CDR) (DTLCDR) (graph convolution attention drug-target interaction (GCADTI)) and DTLCDR (Drug-graph isomorphic networks (GIN)) across each cell line in the two datasets. (A–C) Genomics of drug sensitivity in cancer version 1 (GDSC1) dataset: Pearson's correlation coefficient (PCC) (A), Spearman's correlation coefficient (SPM) (B), and R^2 values (C). (D–F) Cancer Therapeutics Response Portal Version 2 (CTRPv2) dataset: PCC (D), SPM (E), and R^2 values (F). Each point represents a cell line.

were located above the diagonal line (Figs. 4A–C). In the CTRPv2 dataset, nearly all scatter points were above the diagonal line (Figs. 4D–F). Thus, DTLCDR (GCADTI) outperformed DTLCDR (Drug-GIN) in most cell lines in both the GDSC1 and CTRPv2 datasets, further highlighting the effectiveness of drug target information.

3.5. Drug repositioning using DTLCDR

To evaluate the performance of the DTLCDR model in real-world scenarios, we conducted virtual screening-based drug repurposing using the MedChemExpress (MCE) bioactive compound library. This library contained 19,729 molecules with known targets and bioactivity information commonly used for drug repurposing.

First, the DTLCDR model was used to predict drug responses to 19,729 molecules across a panel of cell lines, including the breast cancer cell line (CCL) MCF-7, non-small cell lung CCL A-549, and lung squamous carcinoma cell lines NCI-H520 and HCC95. The top 100 ranked compounds were selected based on their prediction results. To ensure that the candidate compounds had no previously reported anticancer activity, the relevant literature and patent information for these 100 compounds were searched using the SciFinder platform, and those with documented cancer-related activity were excluded. Subsequently, 42 compounds were purchased.

The 42 compounds (Y1–Y42) were tested on the four CCLs using the Cell Counting Kit-8 (CCK-8) assay. Each compound was incubated at a concentration of 10 μM for 24 h, and their inhibitory effects on cell proliferation were measured (Fig. S2). Based on these results, the top five compounds (Y17, Y23, Y24, Y27, Y33, and Y41) were selected for further experiments to obtain fitted IC_{50} curves. The structures of these five compounds are shown in Fig. S3. These compounds are initially known for their non-cancer-related therapeutic applications: Y17 is a compound used for osteoarthritis treatment; Y23 is an antimalarial agent; Y33 has previously been studied for thrombosis treatment; and Y24 and Y41 are used to suppress inflammatory responses.

The four CCLs were then treated with a concentration gradient of the five compounds for 48 h and detailed information is provided in the Supplementary data. The dose-response curves were fitted to calculate the IC_{50} values (Fig. 5). Four of the five compounds exhibited micromolar-level inhibitory activity in the MCF-7 cell line (Fig. 5A), and all five compounds achieved micromolar-level inhibitory activity in three lung CCLs (Figs. 5B–D). The drug repurposing screening experiment identified five potential anticancer compounds, demonstrating that DTLCDR achieved an impressive hit rate in predicting and identifying active compounds with anticancer potential, highlighting its promising application in future cancer drug discovery and repurposing.

3.6. Evaluation of transferring DTLCDR to PDTC data

Because adversarial networks are widely used to address systematic biases between preclinical and clinical genomic information [40], we transformed DTLCDR into DTLACDR using the domain adaptation approach CDAN. The DTLACDR was trained on the GDSC dataset and tested on the *ex vivo* PDTC dataset. In the PDTC dataset, drugs were divided into "shared drug sets" ($PDTC_{in\ GDSC}$) and "non-shared drug sets" ($PDTC_{out\ GDSC}$) based on whether the drugs in PDTC were included in the GDSC dataset. The PDTC dataset contained 1570 samples covering 99 drugs. Among them, the $PDTC_{in\ GDSC}$ contained 906 samples covering 50 drugs, whereas the $PDTC_{out\ GDSC}$ contained 664 samples covering 49 drugs (Table S3).

To evaluate the performance of DTLACDR on the PDTC, we compared it with several baseline models. The outstanding model CODE-AE was used as a baseline [10], which learns shared and private information between cell lines and patient expression data and then uses the shared encoder from the pretrained CODE-AE to generate new representations for CDR

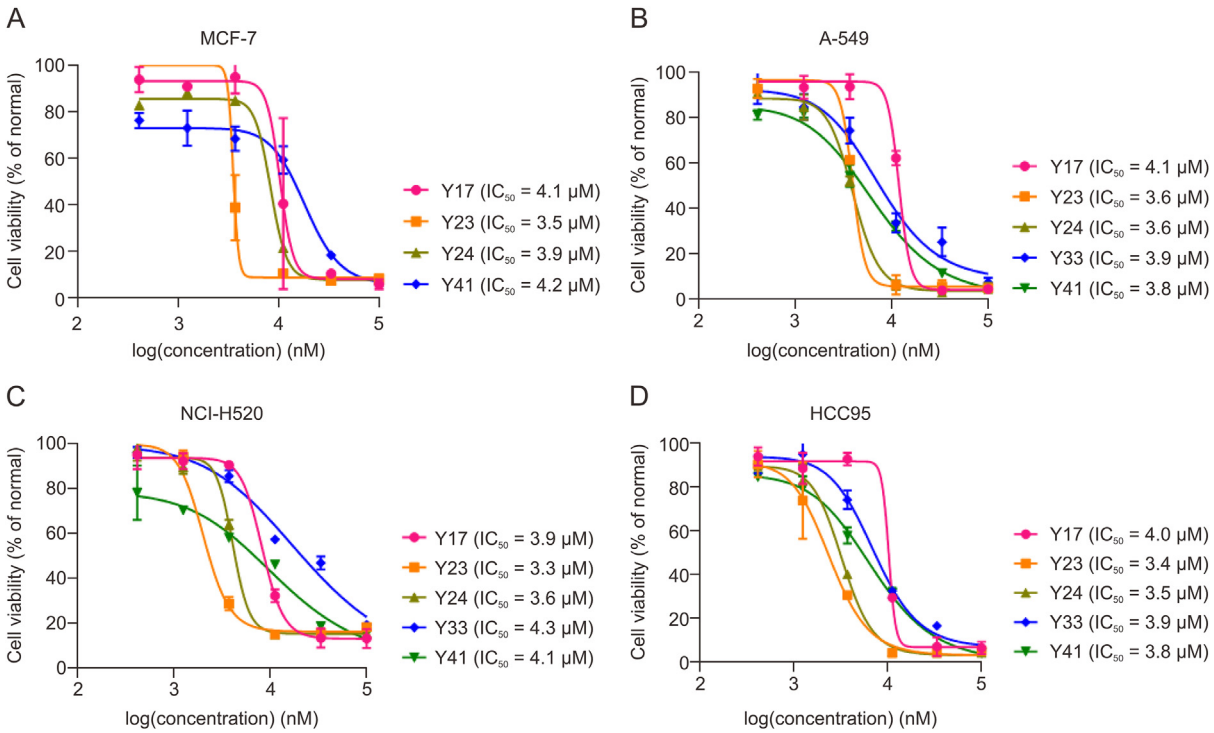


Fig. 5. Cell proliferation curve in (A) MCF-7, (B) A-549, (C) NCI-H520, and (D) HCC95 cell lines treated with screened top five ranked compounds. IC_{50}: half maximal inhibitory concentration.

prediction in PDTC and TCGA. In addition, we used other baseline models from the CODE-AE study for comparison, which demonstrated that CODE-AE outperformed the other baseline models [10].

Because the baseline models mentioned above are single-drug models that cannot predict drugs not included in the GDSC training set, comparison experiments were conducted on the $PDTC_{in\ GDSC}$ dataset. Figs. 6A and B displays the performance of DTLACDR and other baseline models for each drug in $PDTC_{in\ GDSC}$ by AUROC and AUPRC metrics, with each point representing the prediction result of a model for a specific drug. In terms of the AUROC, DTLACDR ranked fourth among the 18 models. The mean AUROC of all shared drugs was 0.610, which was lower than those of CODE-AE (0.6987), VAE (0.6263), and ADAE (0.6208). As shown in Fig. 6A, most points representing the other models were located to the left of the red points, indicating that DTLACDR's AUROC for most drugs was higher than that of most baseline models. A similar pattern was observed in Fig. 6B. Overall, DTLACDR outperformed majority of the baseline models for the shared drugs PDTC and GDSC.

Although our model did not achieve the best performance for $PDTC_{in\ GDSC}$, DTLACDR can be used to predict drugs in $PDTC_{out\ GDSC}$, whereas the other models cannot. Figs. 6C and D presents the performance of DTLACDR on drugs in the $PDTC_{out\ GDSC}$ dataset. The mean AUROC of the 49 drugs was 0.549 and the mean AUPRC was 0.656. Compared to the results for $PDTC_{in\ GDSC}$ (AUROC = 0.6101 and AUPRC = 0.6533), DTLACDR showed a slight decrease in the AUROC for $PDTC_{out\ GDSC}$, but its AUPRC performance remained comparable. Among the 49 drugs, 35 achieved a mean AUROC > 0.5, and 42 achieved a mean AUPRC > 0.5. These results indicate that, for most drugs, DTLACDR demonstrates a relatively effective predictive performance, which highlights its potential for predicting non-shared drugs between PDTC and GDSC.

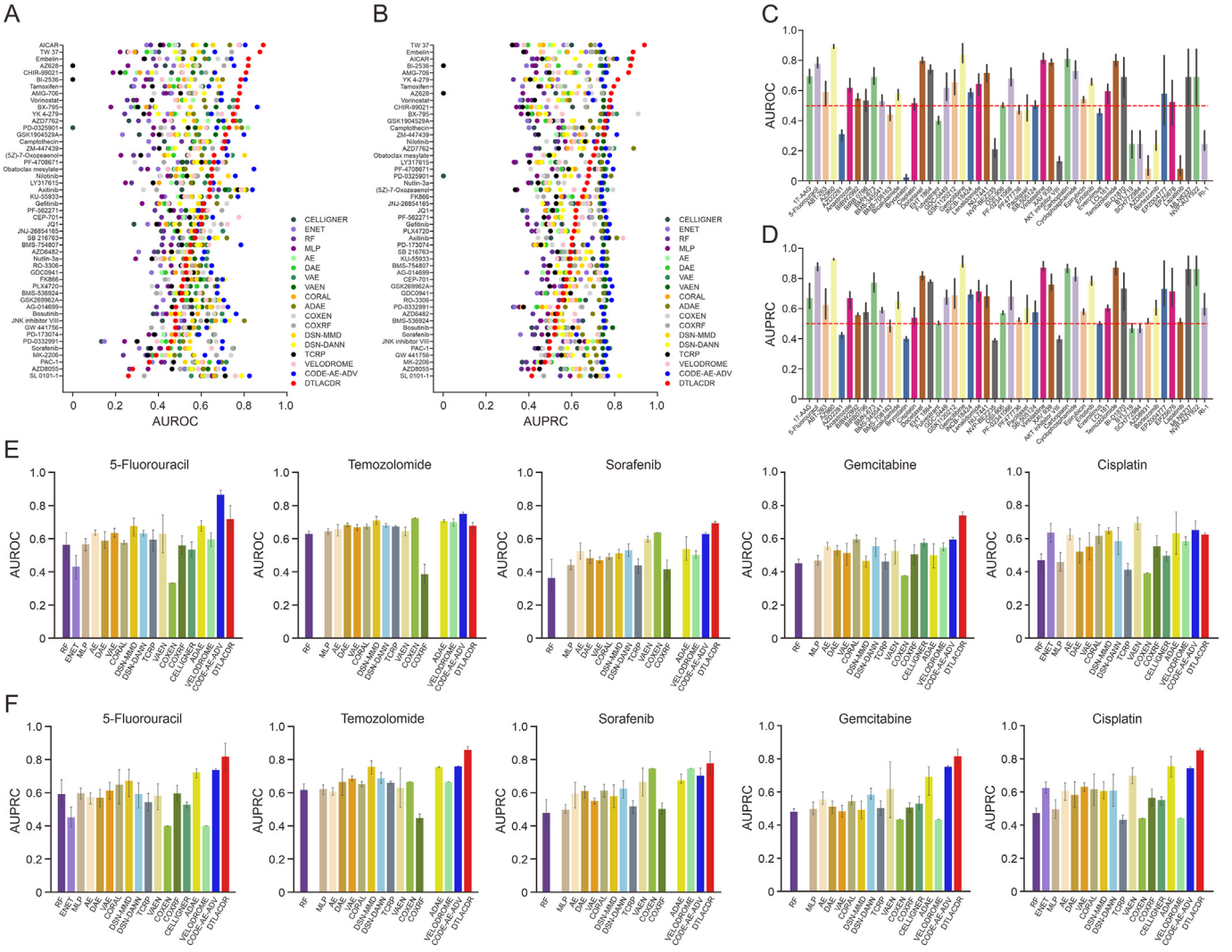


Fig. 6. Performance evaluation of transferred DTLCDR (DTLACDR). (A, B) Performance comparison of DTLACDR with other baseline models for each drug shared between patient-derived tumor xenograft (PDTX)-derived tumor cells (PDTC) and Genomics of Drug Sensitivity in Cancer (GDSC) datasets by area under the receiver operating curve (AUROC) (A) and area under the precision-recall curve (AUPRC) (B) metrics. (C, D) AUROC (C) and AUPRC (D) of DTLACDR on non-shared drugs that are present in PDTC but not present in GDSC. (E, F) Performance comparison of DTLACDR with other baseline models for five drugs in The Cancer Genome Atlas (TCGA) dataset by AUROC (E) and AUPRC (F) metrics. CELLIGNER: cell line alignment; ENET: elastic net; RF: random forest; MLP: multilayer perceptron; AE: autoencoder; DAE: deconfounding AE; VAE: variational AE; VAEN: variational AE followed by elastic net; CORAL: correlation alignment for deep domain adaptation; ADAE: adversarial deconfounding AE; COXEN: co-expression extrapolation; COXRF: co-expression extrapolation and random forest; DSN-MMD: domain separation network-minimization of their maximum mean discrepancy; DANN: domain-adversarial training of neural networks; TCRP: translation of cellular response prediction; VELODROME: semi-supervised method of out-of-distribution generalization; CODE: context-aware deconfounding AE; ADV: adversarial.

3.7. Evaluation of transferring DTLCDR to TCGA data

To further validate the effectiveness of DTLACDR in predicting patient drug response and its potential in personalized medicine, we evaluated its performance on TCGA dataset. TCGA provides extensive molecular data and clinical information from patients across dozens of cancer types, and has been widely employed in previous clinical CDR prediction studies [10,14,41]. After data pre-processing, only five drugs were associated with more than 20 samples (Table S3).

Figs. 6E and F shows the performance of the DTLACDR and baseline models for the five drugs. We observed that the AUROC of DTLACDR are over 0.6 for all five drugs, and the AUPRC are over 0.8 for all five drugs. Compared with the other baseline models, DTLACDR's predictions for the five drugs mostly achieved the best or second-best performance, with particularly superior performance for gemcitabine and sorafenib. Notably, DTLACDR significantly outperformed all other baseline models in terms of the AUPRC for all five drugs, demonstrating the potential of the model to predict positive samples.

4. Conclusions

Existing preclinical CDR studies have often overlooked the evaluation of unseen drugs, and focusing solely on warm start scenarios may lead to inflated estimates of the model performance. The limited number of drugs in preclinical datasets makes it difficult to predict new drugs. Previously reported clinical CDR prediction models within a single-drug pattern require that the drugs to be predicted be included in the training cell line datasets. However, many of these drugs lack sufficient response data from cell lines, making it difficult to effectively train models.

To address the challenge of drug response prediction in personalized medicine, we propose a multimodal fusion framework called DTLCDR. This model employs two key strategies: first, it enhances drug feature extraction by incorporating target information derived from a well-trained GCADTI model; second, it improves cell representation by fine-tuning a pretrained language model that captures general knowledge from single-cell data. However, these strategies have rarely been considered in previous CDR prediction studies. DTLCDR demonstrated outstanding performance through a comprehensive evaluation, particularly in predicting unseen drugs. Ablation studies and evaluations across multiple databases have highlighted the effectiveness of integrating predicted target profiles and pretrained cell-language representations. *In vitro* cell viability assays further validated their potential for real-world applications. Moreover, DTLCDR was modified to DTLACDR and applied to the clinical datasets PDTC and TCGA, where it achieved impressive performance. Unlike traditional single-drug models, the multidrug pattern of DTLACDR enables the prediction of responses to drugs that are not covered in the training cell line datasets. In conclusion, our results suggest that DTLCDR is a promising tool for personalized drug discovery.

CRediT authorship contribution statement

Jie Yu: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Cheng Shi:** Visualization, Validation, Methodology, Investigation, Writing – original draft. **Yiran Zhou:** Writing – review & editing, Supervision, Data curation. **Ningfeng Liu:** Software, Methodology. **Xiaolin Zong:** Investigation, Project administration. **Zhenming Liu:** Supervision, Resources, Project administration, Funding acquisition, Conceptualization. **Liangren Zhang:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This research was supported by the National Key Research and Development Program of China (Grant No.: 2023YFC2605002), the National Key R&D Program of China (Grant No.: 2022YFF1203003), Beijing AI Health Cultivation Project, China (Grant No.: Z221100003522022), the National Natural Science Foundation of China (Grant No.: 82273772), and the Beijing Natural Science Foundation, China (Grant No.: 7212152).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jpha.2025.101315.

References

[1] W. Sadee, D. Wang, K. Hartmann, et al., Pharmacogenomics: Driving personalized medicine, Pharmacol. Rev. 75 (2023) 789–814.
[2] I. Dagogo-Jack, A.T. Shaw, Tumour heterogeneity and resistance to cancer therapies, Nat. Rev. Clin. Oncol. 15 (2018) 81–94.
[3] F. Azuaje, Computational models for predicting drug responses in cancer research, Brief. Bioinform. 18 (2017) 820–829.
[4] B. Bhinder, C. Gilvary, N.S. Madhukar, et al., Artificial intelligence in cancer research and precision medicine, Cancer Discov. 11 (2021) 900–915.
[5] A. Partin, T.S. Brettin, Y. Zhu, et al., Deep learning methods for drug response prediction in cancer: Predominant and emerging trends, Front. Med. (Lausanne) 10 (2023), 1086097.
[6] N. Cancer Genome Atlas Research, J.N. Weinstein, E.A. Collisson, et al., The Cancer Genome Atlas Pan-Cancer analysis project, Nat. Genet. 45 (2013) 1113–1120.
[7] W. Yang, J. Soares, P. Greninger, et al., Genomics of drug sensitivity in cancer (GDSC): A resource for therapeutic biomarker discovery in cancer cells, Nucleic Acids Res. 41 (2013) D955–D961.
[8] D. Baptista, P.G. Ferreira, M. Rocha, Deep learning for drug response prediction in cancer, Brief. Bioinform. 22 (2021) 360–379.
[9] J. Kim, S.H. Park, H. Lee, PANCDR: Precise medicine prediction using an adversarial network for cancer drug response, Brief. Bioinform. 25 (2024), bbae088.
[10] D. He, Q. Liu, Y. Wu, et al., A context-aware deconfounding autoencoder for robust prediction of personalized clinical drug response from cell-line compound screening, Nat. Mach. Intell. 4 (2022) 879–892.
[11] X. An, X. Chen, D. Yi, et al., Representation of molecules for drug response prediction, Brief. Bioinform. 23 (2022), bbab393.
[12] T. Mohammadzadeh-Vardin, A. Ghareyazi, A. Gharizadeh, et al., DeepDRA: Drug repurposing using multi-omics data integration with autoencoders, PLoS One 19 (2024), e0307649.
[13] P. Liu, H. Li, S. Li, et al., Improving prediction of phenotypic drug response on cancer cell lines using deep convolutional network, BMC Bioinformatics 20 (2019), 408.
[14] Q. Liu, Z. Hu, R. Jiang, et al., DeepCDR: A hybrid graph convolutional network for predicting cancer drug response, Bioinformatics 36 (2020) i911–i918.
[15] T. Nguyen, G.T.T. Nguyen, T. Nguyen, et al., Graph convolutional networks for drug response prediction, IEEE/ACM Trans. Comput. Biol. Bioinform. 19 (2022) 146–154.
[16] Z. Zuo, P. Wang, X. Chen, et al., SWnet: A deep learning model for drug response prediction from cancer genomic signatures and compound chemical structures, BMC Bioinformatics 22 (2021), 434.
[17] M. Li, Y. Wang, R. Zheng, et al., DeepDSC: A deep learning method to predict drug sensitivity of cancer cell lines, IEEE/ACM Trans. Comput. Biol. Bioinform. 18 (2021) 575–582.
[18] Y.C. Tang, A. Gottlieb, Explainable drug sensitivity prediction through cancer pathway enrichment, Sci. Rep. 11 (2021), 3128.
[19] L. Deng, Y. Cai, W. Zhang, et al., Pathway-guided deep neural network toward interpretable and predictive modeling of drug sensitivity, J. Chem. Inf. Model. 60 (2020) 4497–4505.
[20] I. Jin, H. Nam, HiDRA: Hierarchical network for drug response prediction with attention, J. Chem. Inf. Model. 61 (2021) 3858–3867.
[21] C.Y. Hsieh, J.H. Wen, S.M. Lin, et al., scDrug: From single-cell RNA-seq to drug response prediction, Comput. Struct. Biotechnol. J. 21 (2023) 150–157.
[22] S. Zhao, J. Zhang, Z. Nie, Large-scale cell representation learning via Divide-and-Conquer contrastive learning, arXiv. 2023. https://doi.org/10.48550/arXiv.2306.04371.
[23] E.W. Huang, A. Bhope, J. Lim, et al., Tissue-guided LASSO for prediction of clinical drug response using preclinical samples, PLoS Comput. Biol. 16 (2020), e1007607.

[24] E. So, F. Yu, B. Wang, et al., Reusability report: Evaluating reproducibility and reusability of a fine-tuned model to predict drug response in cancer patient samples, Nat. Mach. Intell. 5 (2023) 792–798.
[25] D.E. Hostallero, L. Wei, L. Wang, et al., Preclinical-to-clinical anti-cancer drug response prediction and biomarker identification using TINDL, Genomics Proteomics Bioinformatics 21 (2023) 535–550.
[26] F. Vincent, A. Nueda, J. Lee, et al., Phenotypic drug discovery: Recent successes, lessons learned and new directions, Nat. Rev. Drug Discov. 21 (2022) 899–914.
[27] F. Yang, W. Wang, F. Wang, et al., scBERT as a large-scale pretrained deep language model for cell type annotation of single-cell RNA-seq data, Nat. Mach. Intell. 4 (2022) 852–866.
[28] K. Xu, W. Hu, J. Leskovec, et al., How powerful are graph neural networks? arXiv. 2018. https://doi.org/10.48550/arXiv.1810.00826.
[29] M. Long, Z. Cao, J. Wang, et al., Conditional adversarial domain adaptation, arXiv. 2017. https://doi.org/10.48550/arXiv.1705.10667.
[30] J. Du, P. Jia, Y. Dai, et al., Gene2vec: Distributed representation of genes based on co-expression, BMC Genomics 20 (2019), 82.
[31] X. Liu, W. Zhang, A subcomponent-guided deep learning method for interpretable cancer drug response prediction, PLoS Comput. Biol. 19 (2023), e1011382.
[32] M.G. Rees, B. Seashore-Ludlow, J.H. Cheah, et al., Correlating chemical sensitivity and basal gene expression reveals mechanism of action, Nat. Chem. Biol. 12 (2016) 109–116.
[33] A. Bruna, O.M. Rueda, W. Greenwood, et al., A biobank of breast cancer explants with preserved intra-tumor heterogeneity to screen anticancer compounds, Cell 167 (2016) 260–274.e22.
[34] T. Nguyen, H. Le, T.P. Quinn, et al., GraphDTA: Predicting drug-target binding affinity with graph neural networks, Bioinformatics 37 (2021) 1140–1147.
[35] I. Lee, J. Keum, H. Nam, DeepConv-DTI: Prediction of drug-target interactions via deep learning with convolution on protein sequences, PLoS Comput. Biol. 15 (2019), e1007129.
[36] P. Bai, F. Miljković, B. John, et al., Interpretable bilinear attention network with domain adaptation improves drug–target prediction, Nat. Mach. Intell. 5 (2023) 126–136.
[37] H. Wu, J. Liu, R. Zhang, et al., A review of deep learning methods for ligand based drug virtual screening, Fundam. Res. 4 (2024) 715–737.
[38] L. Jiang, C. Jiang, X. Yu, et al., DeepTTA: A transformer-based model for predicting cancer drug response, Brief. Bioinform. 23 (2022), bbac100.
[39] Y. Zhu, Z. Ouyang, W. Chen, et al., TGSA: Protein-protein association-based twin graph neural networks for drug response prediction with similarity augmentation, Bioinformatics 38 (2022) 461–468.
[40] R. Peres da Silva, C. Suphavilai, N. Nagarajan, TUGDA: Task uncertainty guided domain adaptation for robust generalization of cancer drug response prediction from *in vitro* to *in vivo* settings, Bioinformatics 37 (2021) i76–i83.
[41] B. Shen, F. Feng, K. Li, et al., A systematic assessment of deep learning methods for drug response prediction: From *in vitro* to clinical applications, Brief. Bioinform. 24 (2023), bbac605.