OXFORD

Briefings in Bioinformatics, 2024, **25(6)**, bbae493
https://doi.org/10.1093/bib/bbae493
Problem Solving Protocol

BANDRP: a bilinear attention network for anti-cancer drug response prediction based on fingerprint and multi-omics

Cheng Cao‡, Haochen Zhao‡, Jianxin Wang*

Hunan Provincial Key Lab on Bioinformatics, School of Computer Science and Engineering, Central South University, Changsha 410083, China

*Corresponding author. School of Computer Science and Engineering, Central South University, Changsha 410083, Hunan, China. E-mail: jxwang@mail.csu.edu.cn
‡Cheng Cao and Haochen Zhao contributed equally.

Abstract

Predicting anti-cancer drug response can help with personalized cancer treatment and is an important topic in modern oncology research. Although some methods have been used for anti-cancer drug response prediction, how to effectively integrate various features related to cancer cell lines, drugs, and their known responses is still affected by the redundant information of input features and the complex interactions between features. In this study, we propose a bilinear attention model, named BANDRP, based on multiple omics data of cancer cell lines and multiple molecular fingerprints of drugs to predict potential anti-cancer drug responses. Compared with existing models, BANDRP uses gene expression data to calculate pathway enrichment scores to enrich the features of cancer cell lines and can automatically learn the interactive information of cancer cell lines and drugs through bilinear attention networks. Benchmarking and independent tests demonstrate that BANDRP surpasses baseline models and exhibits robust generalization performance. Ablation experiments affirm the optimality of the current model architecture and feature selection scheme for our prediction task. Furthermore, analytical experiments and case studies on unknown anti-cancer drug response predictions underscore BANDRP's potential as a potent and reliable framework for predicting anti-cancer drug response.

Keywords: anti-cancer drug response prediction; multi-source data; attention network

Introduction

The heterogeneity of the same cancer among different patients leads to variations in the effectiveness of the same drug across individuals. Consequently, tailoring therapeutic regimens for each patient based on individual genes, lifestyle, and environment poses a significant challenge in precision medicine [1]. Conventional methods of discovering anti-cancer drugs rely on *in vivo* animal experiments and in vitro drug screening, both expensive and labor-intensive [2]. In recent years, the progress in high-throughput screening technologies has allowed the archival of experimental results of anti-cancer drug response from several large-scale cancer genetic projects in public databases such as Cancer Cell Line Encyclopedia (CCLE) [3] and Genomics of Drug Sensitivity in Cancer (GDSC) [4]. These databases offer not only extensive anti-cancer drug response data but also provide genetic profiles of cancer cell lines. Furthermore, the PubChem database, serving as a repository of chemical information, contains validated structural details for 19 million unique compounds [5]. The advent of large-scale, high-throughput screening databases establishes a solid data foundation for developing computational prediction models. Using these models to forecast anti-cancer drug responses can assist in identifying potential therapeutic approaches in cancer treatment and advance research in predicting anti-cancer drug efficacy.

Over the decades, researchers have proposed a series of computational methods for drug response prediction based on machine learning techniques. For example, GUAN *et al.*[6] proposed a matrix factorization method to construct sparse drug similarity matrices and cell line similarity matrices to generate potential response matrices for drugs and cell lines. Liu *et al.*[7] framed the drug response prediction problem as a recommendation system issue, estimating the anti-cancer drug response of cell lines by integrating cell line similarity networks and drug similarity networks. Stanfield *et al.*[8] constructed heterogeneous networks of cancer cell lines and drugs, using link prediction methods to forecast drug responses. Wang *et al.*[9] established representations of cancer cell lines and drugs based on similarity, using support vector machines to infer associations between cancer cell lines and drugs. Wan *et al.*[10] applied a random forest-based approach to generate sensitivity predictions from each type of data and combined these predictions in a linear regression model to produce the final drug sensitivity prediction. Although machine learning-based computational methods for drug response prediction have achieved significant

Received: July 11, 2024. **Revised:** September 4, 2024. **Accepted:** September 23, 2024

results, the high dimensionality and noise in multi-omics data from cancer cell lines, along with the lack of consideration for the synergistic effects among different omics data, have limited the generalization and scalability of machine learning-based drug response prediction methods.

In recent years, deep learning-based methods have been applied in many fields due to their capacity to handle high-dimensional and large-scale complex data. Many researchers have introduced the technology to anti-cancer drug response prediction. Compared with other methods, deep learning-based methods can automatically learn the potential relationships between drugs and cancer cell lines, making it easier to process multi-source data. For instance, Liu *et al.* [11] proposed a method using drug molecular fingerprints and genomic mutation data of cancer cell lines as input, learning drug representations through two convolutional neural networks. Liu *et al.* [12] presented the DeepCDR method, which utilizes graph neural networks to acquire representations of drug molecular graphs and integrates multi-omics data to represent cancer cell lines. DeepCDR demonstrates the benefits of leveraging multi-omics data of cancer cell lines. Chawla *et al.* [13] proposed a predictive method to infer cancer drug responses using gene expression data of cancer cell lines and molecular descriptors of drugs. Their findings highlight the importance of pathway activity estimates. Jiang *et al.* [14] proposed the DeepTTA method. It adopts a Transformer-based architecture to improve the drug representation capability and uses a multi-layer perceptron (MLP) to learn representations of cancer cell lines. Then, it concatenates the representations of drugs and cancer cell lines and inputs them into a fully connected neural network to predict anti-cancer drug responses. Wang *et al.* [15] employ the similarity of drug-cell line pairs to build a sparse network and predicts anti-cancer drug response by graph convolutional networks and autoencoders.

Although the previous methods appear to be effective, they have certain limitations. First, most targeted therapies work through biological pathways, but many of these methods ignore the importance of biological pathway information. Second, some studies suggest that the contribution of existing drug features to model performance is greater than that of cancer cell line features [16]. This implies the need for a new perspective to learn cancer cell representations better to capture the differences among various cancer cell lines. To address these limitations, we propose a deep learning framework, named BANDRP for anti-cancer drug response prediction. BANDRP constructs a deep learning-based framework to integrate the multi-omics data and pathway enrichment scores of cancer cell lines and fingerprints of drugs. Specifically, we first collect the gene expression, genomic mutation, and DNA methylation data for cancer cell lines, as well as simplified molecular-input line-entry system (SMILES) strings for drugs. Then, we use Gene Set Variation Analysis (GSVA) [17] to obtain pathway enrichment scores for cancer cell lines and compile molecular fingerprints for the drugs. Afterward, we adopt some MLPs to capture representations of the input data, then employ a multi-head bilinear attention network to obtain fused representations, which are subsequently input into the MLP for anti-cancer drug response prediction. The experimental results demonstrate that BANDRP outperforms other state-of-the-art methods in anti-cancer drug response prediction. Independent test and *de novo* test indicate that BANDRP exhibits strong generalization capabilities. Furthermore, a set of ablation experiments demonstrates the optimality of our model's feature combinations. Finally, analytical experiments and case studies on unknown anti-cancer drug response prediction evaluate the reliability of our model. All results show that BANDRP effectively predicts drug responses.

Materials and methods

Materials and data preparation

GDSC [4] is a large public resource for information on drug sensitivity, which provides muti-omics data of cancer cell lines, molecular markers of drugs, and known anti-cancer drug responses. In this study, we collect half-maximal inhibitory concentration (IC50) values (log-transformed) between 969 cancer cell lines and 286 drugs from the GDSC v2 database. CCLE [3] performs large-scale deep sequencing of more than 1,000 cancer cell lines, providing muti-omics data including gene expression, genomic mutation, and DNA methylation data. We download gene expression, genomic mutation, and DNA methylation data related to cancer cell lines from CCLE. Following DeepCDR [12], to reduce dimensionality and redundant information, we only consider 733 human genes in COSMIC Cancer Gene Census [18] that are strongly associated with cancer development. Specifically, we first remove genes from the three omics data that are not in the COSMIC gene set. Due to the potential redundancy in having multiple mutation points on the same gene in genomic mutation data, we merge different mutation points on the same gene. After processing, the number of genes in the three omics data (i.e. gene expression, genomic mutation, and DNA methylation) is 714, 715, and 603, respectively. PubChem [5] database offers extensive information on drug molecules and is the world's largest chemical information database. We download the SMILES strings for all drugs from the PubChem database. Then, we exclude cancer cell lines with any missing omics data and drugs without PubChem ID or with incorrect PubChem ID. Due to the presence of multiple GDSC IDs for the same drug in the GDSC database resulting from database updates, we will retain only the entry with the most IC50 values. Finally, we compile a benchmark dataset containing 81,467 IC50 values across 169 drugs and 536 cancer cell lines. Among all the 90,584 (169×536) IC50 values, approximately 10% of the IC50 values are missing. Additionally, we download IC50 values from the CCLE database. By removing cancer cell line-drug pairs that duplicated with the benchmark dataset, we create an independent test set containing 8,909 IC50 values involving 551 cancer cell lines and 184 drugs.

Methods

Overview

In this study, we propose the BANDRP framework as shown in Fig. 1. First, based on the SMILES strings of drugs, we compute the Extended Connectivity Fingerprints (ECFP) [19], Explainable Substructure Partition Fingerprint (ESPF) [20], and PubChem Fingerprint (PubchemFP) for drugs. Second, we use the GSVA to calculate the pathway enrichment scores for each cell line based on gene expression. Third, we integrate the gene expression, genomic mutation, DNA methylation, and pathway enrichment scores into the cancer feature encoder, along with the molecular fingerprints of drugs into the drug feature encoder. MLPs tailored to specific features are designed within the feature encoder to acquire latent representations for drugs and cell lines. Fourth, we feed these latent representations into the bilinear attention module, which comprises the bilinear attention layer and bilinear pooling layer, to generate the fused representations of cancer cell lines and drugs. Finally, the fused representations are input into an MLP for prediction, resulting in the estimation of IC50 values.

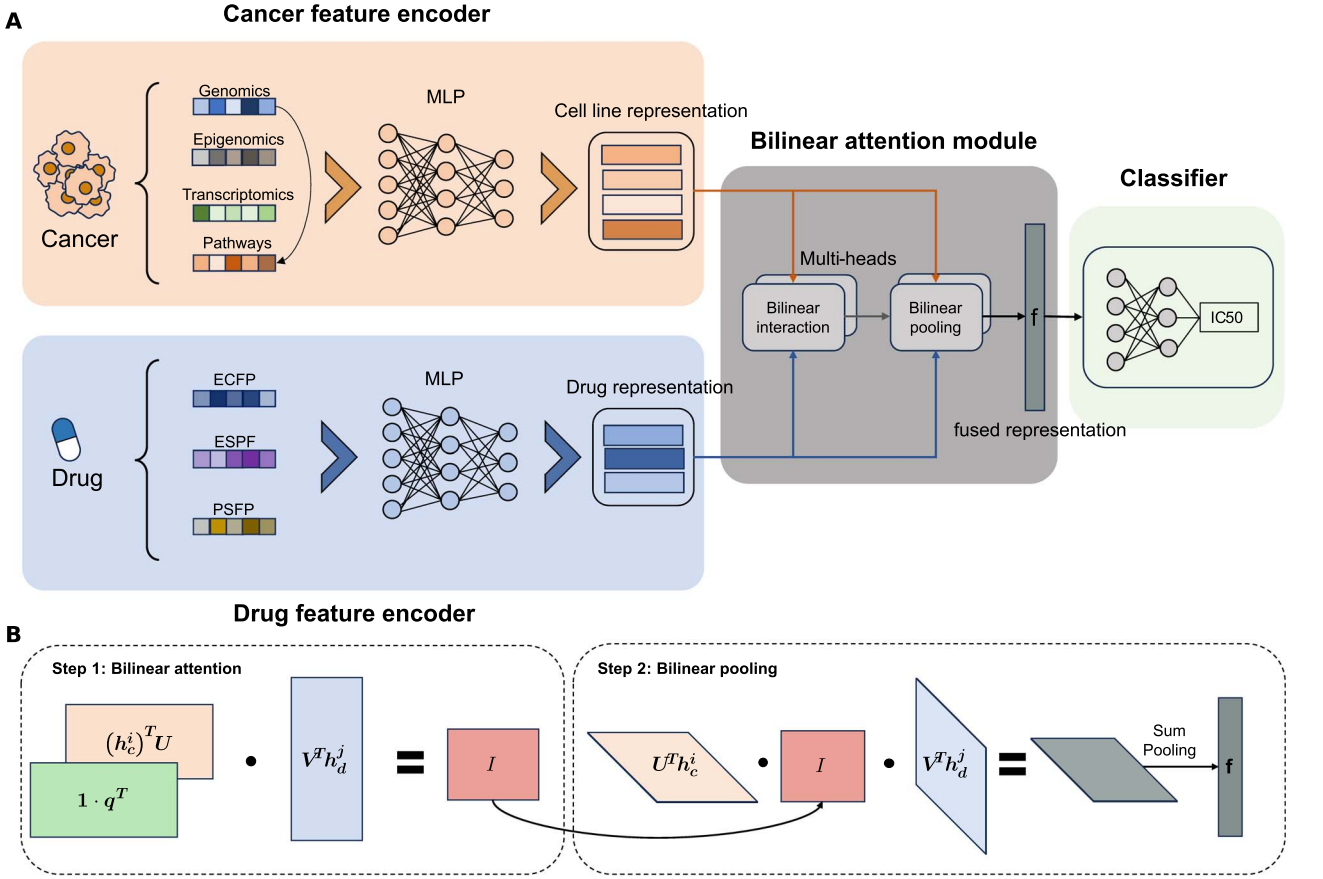

Figure 1. (A) Overview of BANDRP framework. The cancer feature encoder and drug feature encoder learn latent representations of cancer cell lines and drugs using different MLPs, respectively. The bilinear attention module utilizes the latent representation matrices of cancer cell lines and drugs. Through bilinear interaction and bilinear pooling, it produces a fused representation f, which is then fed into a classifier network composed of MLPs. (B) The bilinear attention module architecture. In the first step, bilinear interaction is used to obtain the bilinear attention mapping matrix I between the latent representations of cancer cell lines and drugs. In the second step, bilinear pooling is performed using I and shared transformation matrices U and V to obtain the fused representation f.

Cancer feature encoder

Let $S_{cell} = \{c_1, c_2, \ldots, c_n\}$ represent the set of all n cancer cell lines. The cancer feature encoder calculates the latent representations of multi-omics data. Taking cancer cell line $c_i \in S_{cell}$ as an example, we obtain three feature vectors based on 733 human genes and their associated gene expression, genomic mutation, and DNA methylation data, denoted as a^i_{EXP}, a^i_{MUT}, and a^i_{METH}, respectively. In a^i_{EXP}, each bit signifies a log-normalized Transcripts Per Million (TPM) value of gene expression. For a^i_{MUT}, each bit indicates whether the gene has a mutation (1 for mutated, 0 otherwise). Similarly, each bit in a^i_{METH} represents the methylation value of the corresponding gene.

The pathway enrichment score incorporates predefined biological pathway information and reveals underlying biological mechanisms by considering the collective behavior of genes. Therefore, we introduced the pathway enrichment score as a feature of cancer cell lines. We first downloaded the c2 gene set from the Molecular Signatures Database (MSigDB) [21], which contains canonical pathway gene sets related to humans. Next, we intersect the 714 genes from the gene expression data of cancer cell lines with each biological pathway gene set, and biological pathway gene sets with fewer than 5 genes in the intersection are removed. Finally, 1,283 biological pathway gene sets are selected as background gene sets, denoted S_{PATH}. S_{PATH} contains 1,283 elements, each representing a predefined biological pathway and the genes it includes. Then, we use the GSVA R package [17] to calculate the pathway enrichment scores. GSVA is a non-parametric, unsupervised method used to calculate sample-wise gene set enrichment scores. The input parameters for the GSVA method include the gene expression vectors of the samples to be analyzed, the background gene sets representing biological pathways, and an optional kernel function. The output is the enrichment score for each sample in the biological pathways of the background gene sets. Taking the cancer cell line c_i as an example, we use a^i_{EXP} and S_{PATH} as inputs and set the kernel function to Gaussian. Next, GSVA calculates the probability density of a^i_{EXP} using kernel density estimation. Then, GSVA transforms these probability densities into normalized ranks to minimize errors. Subsequently, S_{PATH} and normalized ranks are used to calculate the Kolmogorov–Smirnov (KS) like random walk statistic for each biological pathway in S_{PATH} relative to c_i. Finally, the normalized KS statistic is used as the pathway enrichment score, denoted as a^i_{PATH}. We utilize a^i_{PATH} as the pathway enrichment feature vector for c_i.

Next, we use different MLPs based on a^i_{EXP}, a^i_{MUT}, a^i_{METH} and a^i_{PATH} to calculate the latent representations for c_i. In the cancer feature encoder, each MLP consists of two hidden layers and each layer in the MLP is expressed as follows:

$$t^i_{mlp} = \sigma\left(w_{mlp} a^i_{mlp} + b_{mlp}\right) \quad (1)$$

where a^i_{mlp} represents the input data, t^i_{mlp} is the latent representation, w_{mlp} and b_{mlp} are learnable parameters, and σ is the Rectified Linear Unit (ReLU) activation function. Additionally, we apply Batch Normalization (BN) to standardize activation values within each batch, ensuring a zero mean and unit variance, expressed as follows:

$$t^i_{BN} = \mathrm{BN}\left(t^i_{mlp}\right) = \frac{t^i_{mlp} - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \tag{2}$$

where μ_B and σ_B^2 denote the mean and variance of all data within a batch, and ϵ is a constant to prevent division by zero. Finally, we obtain 4 latent representations of the cell line c_i with the same dimension: t^i_{EXP}, t^i_{MUT}, t^i_{METH}, and t^i_{PATH}. These representations are then concatenated vertically into a matrix $h^i_c \in \mathbb{R}^{4\times l}$:

$$h^i_c = t^i_{EXP} \oplus t^i_{MUT} \oplus t^i_{METH} \oplus t^i_{PATH} \tag{3}$$

where h^i_c is the latent representation matrix for the cancer cell line c_i, and l is a hyperparameter of the dimension of the latent representations.

Drug feature encoder

Let $S_{drug} = \{d_1, d_2, \ldots, d_m\}$ represent the set of all m drugs, and then the ECFP, ESPF, and PubchemFP are calculated for each drug using the RDKit Python package [22], respectively. ECFP is a classical method for calculating molecular fingerprints that represent molecules by constructing bit strings based on atoms and their surrounding molecular structures, incorporating both the topological information and local chemical information of the molecules. ESPF decomposes drugs into a set of moderately sized discrete drug substructures and then encodes the drugs using a byte-pair encoding algorithm. According to the principle of similarity [23], molecules inducing similar effects in cancer cell lines often have similar substructures or pharmacophore signatures. Thus, the substructure information provided by ESPF enhances the ability of the model to identify and predict the biological activity of new compounds by leveraging the structural similarities and common pharmacophores among known effective molecules. PubchemFP, provided by the PubChem database, is a bit-string encoding based on chemical molecular fragments. Each bit position in the encoding corresponds to a specific molecular or chemical feature. We compute the three fingerprints above for each drug in default settings, with dimensions of 2,048, 2,586, and 881, respectively. Similar to the representations for cancer cell lines, let the three molecular fingerprints for drug d_j be denoted as a^j_{ECFP}, a^j_{ESPF}, and $a^j_{PubchemFP}$. Then, they are input into the drug feature encoder to obtain latent representations of the same dimension: t^j_{ECFP}, t^j_{ESPF}, and $t^j_{PubchemFP}$. The MLP in the drug feature encoder consists of two hidden layers. Finally, the latent representations of d_j are concatenated vertically into a matrix $h^j_d \in \mathbb{R}^{3\times l}$:

$$h^j_d = t^j_{ECFP} \oplus t^j_{ESPF} \oplus t^j_{PubchemFP} \tag{4}$$

where h^j_d is the latent representation matrix for drug d_j.

Bilinear attention module

We construct the bilinear attention module using a multi-head bilinear attention network. In the bilinear attention module, for a given cancer cell line and a drug, different latent representations of the cancer cell line and the drug are combined in pairs. Firstly, the attention weights of each combined latent representation for the cancer cell line and the drug are generated according to Equations (5) and (6). Subsequently, these drug-cancer cell line representation pairs, along with their corresponding attention weights, are fed into the bilinear pooling layer (Equations 7 and 8) to produce the final latent representation f for the drug-cancer cell line pair.

Using a single-head bilinear attention network as an example, the first part is a bilinear attention layer used to calculate the interaction intensity between the latent representations of cancer cell lines and drugs. For cancer cell line c_i and drug d_j, we construct a bilinear attention map using their latent representations h^i_c and h^j_d to obtain a pairwise interaction matrix $I \in \mathbb{R}^{4\times 3}$, which can be expressed as follows:

$$I = \left(\left(1 \cdot q^T\right) \circ \sigma\left(\left(h^i_c\right)^T U\right)\right) \cdot \sigma\left(V^T h^j_d\right) \tag{5}$$

where $U \in \mathbb{R}^{3\times l}$ and $V \in \mathbb{R}^{4\times l}$ are learnable weight matrices, $q \in \mathbb{R}^l$ is a learnable weight vector, 1 is the unit vector, σ represents the activation function ReLU and $\circ$ denotes the Hadamard product. More specifically, let $h^i_c(o)$ be one of the four latent representations of cancer cell line c_i, and $h^j_d(p)$ be one of the three latent representations of drug d_j. The interaction between $h^i_c(o)$ and $h^j_d(p)$ can be formally represented by the element $i_{o,p}$ in I, as expressed below:

$$i_{o,p} = q^T\left(\sigma\left(U^T h^i_c(o)\right) \circ \sigma\left(V^T h^j_d(p)\right)\right) \tag{6}$$

where $h^i_c(o)$ and $h^j_d(p)$ are mapped into a common feature space with weight matrices U and V, and then learn the element-wise interaction between $h^i_c(o)$ and $h^j_d(p)$ using Hadamard product. Finally, we get pairwise interactions between different latent representations.

The second part is the bilinear pooling layer, which processes the interaction matrix I to generate a fused representation f'. The kth element of f' can be expressed as follows:

$$\begin{aligned} f'_k &= \sigma\left(\left(h^i_c\right)^T U\right)_k^T \cdot I \cdot \sigma\left(\left(h^j_d\right)^T V\right)_k \\ &= \sum_{o=1}^{4}\sum_{p=1}^{3} i_{o,p}\left(\left(h^i_c(o)\right)^T\left(U_k V_k^T\right)h^j_d(p)\right) \end{aligned} \tag{7}$$

where U_k and V_k are the kth columns of learnable weight matrices U and V, respectively. U and V are shared with the previous bilinear attention layer to decrease the number of parameters and alleviate overfitting. Then, the fused representation is fed into the pooling layer to obtain a compact feature map:

$$f = \mathrm{SumPool}\left(f', s\right) \tag{8}$$

where $\mathrm{SumPool}(\cdot)$ is a 1D non-overlapping pooling operation with stride s of 3. BANDRP adopts a multi-head form to learn the fused representations between drugs and cancer cell lines, and the final fused representation vector is a sum of individual heads.

Classifier

We employ an MLP classifier to further extract features from the output f and perform the final prediction. The input layer of the MLP classifier receives f. The intermediate layers of the MLP classifier consist of three sets of fully connected layers with ReLU activations and BN, with 512 and 128 neurons, respectively. Ultimately, a fully connected layer with one neuron serves as the output layer for IC50 value prediction.

Model training

To minimize the difference between the predicted and known drug response values, we use the mean squared error loss function for model training. The mean squared error loss function is defined as follows:

$$Loss = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2 \tag{9}$$

where n represents the number of samples, Y_i represents the true value, and ($\hat{Y}_i$) represents the predicted value. Additionally, our model includes some important hyper-parameters, namely the learning rate, the batch size, the number of attention heads, the dropout rate, and l. These hyper-parameters are determined through grid search. In the grid search, the learning rate is in $\{1e-2, 1e-3, 1e-4, 1e-5\}$, the batch size is in $\{32, 64, 128, 256\}$, the number of attention heads is in $\{1, 2, 3, 4\}$, the dropout rate is in $\{0.1, 0.2, 0.3, 0.4, 0.5\}$ and l is in $\{64, 128, 256, 512\}$. Specifically, we first determine the learning rate and batch size. After the learning rate and batch size are fixed, we further select the number of attention heads and dropout rate to enhance the model's performance and robustness. Finally, the optimized hyper-parameters for learning rate, batch size, number of attention heads, dropout rate, and l are $1e-3$, 128, 3, 0.5, and 128, respectively. Detailed experimental results are shown in Supplementary Section 1. In addition, we utilize the Adam optimization algorithm [24] to update the model parameters. During the training process, we save the model with the best performance in the validation set.

Results

Performance evaluation metrics

To evaluate the performance of BANDRP, we employ metrics including Root Mean Square Error (*RMSE*), Mean Absolute Error (*MAE*), Coefficient of Determination (R^2), Pearson Correlation Coefficient (*PCC*), and Spearman Correlation Coefficient (*SCC*). *RMSE* can quantify the deviation between the observed values and the true values. *MAE* measures the average absolute difference between predicted values and true values. R^2 characterizes the accuracy of the fit of the model. *PCC* reflects the correlation between the predicted values and the true values. *SCC* is a non-parametric metric used to describe the *PCC* correlation between variables after ranking. The formulas for these metrics are as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2} \tag{10}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|Y_i - \hat{Y}_i\right| \tag{11}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}{\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2} \tag{12}$$

$$PCC = \frac{\sum_{i=1}^{n}\left(\hat{Y}_i - \bar{\hat{Y}}\right)\left(Y_i - \bar{Y}\right)}{\sqrt{\sum_{i=1}^{n}\left(\hat{Y}_i - \bar{\hat{Y}}\right)^2}\sqrt{\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2}} \tag{13}$$

$$SCC = 1 - \frac{6\sum_{i=1}^{n}\delta_i^2}{n\left(n^2-1\right)} \tag{14}$$

where n is the size of data,Y_i represents the predicted IC50 value, $\hat{Y}_i$ represents the real IC50 value. δ_i denotes the difference in ranks between Y_i and $\hat{Y}_i$, $\bar{Y}$ represents the mean of all Y, and $\bar{\hat{Y}}$ represents the mean of $\hat{Y}$.

Performance comparison of BANDRP and existing methods

We first design a comparative experiment to assess the performance of BANDRP in anti-cancer drug response prediction. We compare BANDRP against five state-of-the-art deep learning methods, including GADRP [15], DeepTTA [14], precily [13], DeepCDR [12], and tCNNs [11]. All methods are based on the benchmark dataset, and each model uses default or best-performing parameters. All baseline models preprocess the relevant data according to their input features and preprocessing methods. The benchmark dataset is divided into training, validation, and test sets in an 8:1:1 ratio. We evaluate the performance of different models based on the test set.

The results of the comparative experiments are shown in Table 1. BANDRP outperforms other baseline models on all metrics. Compared to the baseline models, BANDRP exhibits the lowest *RMSE* (0.9305) and *MAE* (0.6769), as well as the highest R^2 (0.8802), *PCC* (0.9382), and *SCC* (0.9168). The reason for the superior performance of BANDRP may be attributed to the fact that BANDRP not only integrates multi-omics data types and molecular information of drugs but also incorporates biological pathway information. We further evaluate the performance of BANDRP using the datasets provided by DeepCDR, DeepTTA, and GADRP, respectively, to demonstrate the robustness of our model (Supplementary Section 2).

De novo test and independent test

To evaluate the predictive performance of BANDRP on new cancer cell-drug responses, we conduct the *de novo* test with two experimental settings for a comprehensive comparison. The specific experimental settings are as follows:

1. ES_1: S_{cell} is divided into training set, validation set, and test set in an 8:1:1 ratio. The IC50 values related to the cancer cell lines in the validation set and test set will not appear in the training set.
2. ES_2: S_{drug} is divided into training set, validation set, and test set in an 8:1:1 ratio. The IC50 values related to the drugs in the validation set and test set will not appear in the training set.

Tables 2 and 3 present the results under ES_1 and ES_2, respectively. In both ES_1 and ES_2, BANDRP exhibits the best performance. In ES_1, BANDRP outperforms DeepCDR, which also utilizes three omics data. Compared to the second-ranked DeepTTA, BANDRP achieves a lower *RMSE* and a higher *PCC*. These results underscore the beneficial impact of incorporating multiple omics data and pathway information on the model's excellent generalization ability. In ES_2, there is a slight decrease in performance for all models in ES_2, which can be attributed to the diversity of drug molecular structures and the presence of similar genetic information in cancer cell lines. These findings suggest that BANDRP demonstrates robust adaptability and generalization to new cancer cell-drug responses.

Furthermore, we conduct independent tests using a separate test set from CCLE. As hyper-parameter tuning and model training occur independently of the benchmark dataset, this independent test provides a more robust assessment of the model's ability

Table 1. Performance comparison of BANDRP and other methods on the benchmark dataset.

Method	RMSE	MAE	R^2	PCC	SCC
tCNNs	1.1574	0.8509	0.8166	0.9042	0.8776
DeepCDR	1.0189	0.7538	0.8579	0.9270	0.9003
Precily	1.1136	0.8345	0.8302	0.9125	0.8792
DeepTTA	0.9634	0.7182	0.8272	0.9105	0.8770
GADRP	0.9953	0.7263	0.8402	0.9298	0.9097
BANDRP	**0.9305**	**0.6769**	**0.8802**	**0.9382**	**0.9168**

Table 2. Comparison results of BANDRP and other methods on the benchmark dataset under the setting ES_1.

Method	RMSE	MAE	R^2	PCC	SCC
tCNNs	1.6429	1.2348	0.6293	0.7952	0.7408
DeepCDR	1.4068	1.0552	0.7296	0.8604	0.8200
Precily	1.4399	1.0750	0.7154	0.8484	0.8030
DeepTTA	1.3512	1.0114	0.7506	0.8664	0.8288
GADRP	1.3528	1.0166	0.6589	0.8670	0.8289
BANDRP	**1.2998**	**0.9692**	**0.7692**	**0.8783**	**0.8445**

Table 3. Comparison results of BANDRP and other methods on the benchmark dataset under the setting ES_2.

Method	RMSE	MAE	R^2	PCC	SCC
tCNNs	2.0739	1.6611	0.2422	0.5064	0.4298
DeepCDR	1.8569	1.3699	0.5171	0.7399	0.6470
Precily	1.8641	1.1292	0.4929	0.7151	0.6647
DeepTTA	1.7055	1.2975	0.5543	0.7507	0.6646
GADRP	1.7965	1.2970	0.4680	0.7501	0.6789
BANDRP	**1.6348**	**1.2739**	**0.5926**	**0.7738**	**0.6969**

to generalize to unknown data. The experimental results are in Table 4. In independent tests, BANDRP achieves an RMSE that is 0.0514 lower than the second-ranked DeepTTA and an SCC that is 0.0156 higher than the second-ranked GADRP. In summary, BANDRP consistently outperforms state-of-the-art deep learning methods, showcasing its superior adaptability and generalization across various experimental settings and independent test scenarios.

Ablation study

To investigate the contributions of different modules and features within the BANDRP framework, we perform ablation experiments on the input data of the model and the fusion methods of features. In addition, we analyze the correlation between the features of cancer cell lines and the features of drugs, as well as the impact of other potential drug features on the model to validate the optimality of the current feature selection and framework.

The impact of different input features

To further investigate the contributions of different input data in BANDRP, we design the following BANDRP variants that retain only one feature from drugs or cell lines or remove only one feature from drugs or cell lines:

- **BANDRP without gene expression (w/o EXP)** removes the gene expression feature.
- **BANDRP without genomic mutation (w/o MUT)** removes the genomic mutation feature.
- **BANDRP without DNA methylation (w/o METH)** removes the DNA methylation feature.
- **BANDRP without pathway enrichment (w/o PATH)** removes the pathway enrichment feature.
- **BANDRP without ECFP (w/o ECFP)** removes the ECFP.
- **BANDRP without ESPF (w/o ESPF)** removes the ESPF.
- **BANDRP without PubchemFP (w/o PubchemFP)** removes the PubchemFP.
- **BANDRP keep gene expression(K_EXP)** keep the gene expression feature and all drug features.
- **BANDRP keep genomic mutation (K_MUT)** keep the genomic mutation feature and all drug features.
- **BANDRP keep DNA methylation (K_METH)** keep the DNA methylation feature and all drug features.
- **BANDRP keep pathway enrichment (K_PATH)** keep the pathway enrichment feature and all drug features.
- **BANDRP keep ECFP (K_ECFP)** keep the ECFP and all features of the cancer cell line.
- **BANDRP keep ESPF (K_ESPF)** keep the ESPF and all features of the cancer cell line.
- **BANDRP keep PubchemFP (K_PubchemFP)** keep the PubchemFP and all features of the cancer cell line.

All variant models of BANDRP are constructed to benchmark the dataset with consistent hyper-parameters and training strategies. Table 5 shows the results of variant models with one feature of a cancer cell line or drug removed, while Table 6 provides the results of variant models retaining one feature of cancer cell lines and all features of drugs, or retaining one feature of drugs and all features of cancer cell lines. The *RMSE* of these variant models ranges from 0.9313 to 0.9573, which is higher than that of BANDRP (0.9305), indicating the usefulness of all individual features. For

Table 4. Comparison results of BANDRP and other methods on the independent dataset.

Method	RMSE	MAE	R^2	PCC	SCC
tCNNs	2.0174	1.5780	0.2395	0.3706	0.3354
DeepCDR	1.6519	1.2948	0.4399	0.5026	0.4739
Precily	1.9919	1.5618	0.2881	0.4091	0.3698
DeepTTA	1.5436	1.2420	0.5118	0.6469	0.4739
GADRP	1.6465	1.2970	0.4514	0.4717	0.4217
BANDRP	**1.3981**	**1.0844**	**0.5167**	**0.7561**	**0.5565**

Table 5. Ablation experiment results of variant models in which one feature of cancer cell line or drug is removed.

Method	RMSE	MAE	R^2	PCC	SCC
w/o EXP	0.9573	0.6925	0.8732	0.9351	0.9164
w/o MUT	0.9494	0.6904	0.8753	0.9356	0.9142
w/o METH	0.9458	0.6845	0.8762	0.9362	0.9156
w/o PATH	0.9359	0.6764	0.8788	0.9375	0.9168
w/o ECFP	0.9423	0.6810	0.8771	0.9368	0.9149
w/o ESPF	0.9313	0.6764	0.8800	0.9384	0.9187
w/o PubchemFP	0.9503	0.6911	0.8750	0.9356	0.9145
BANDRP	**0.9305**	**0.6769**	**0.8802**	**0.9382**	**0.9168**

Table 6. Ablation experiment results of variant models that retain one feature of cancer cell lines and all features of drugs or retain one feature of drugs and all features of cancer cell lines.

Method	RMSE	MAE	R^2	PCC	SCC
K_EXP	0.9565	0.7009	0.8734	0.9347	0.9134
K_MUT	0.9544	0.7023	0.8739	0.9350	0.9123
K_METH	0.9607	0.7013	0.8723	0.9340	0.9127
K_PATH	0.9707	0.7225	0.8696	0.9335	0.9108
K_ECFP	0.9376	0.6798	0.8783	0.9372	0.9176
K_ESPF	0.9405	0.6850	0.8776	0.9372	0.9164
K_PubchemFP	0.9448	0.6912	0.8765	0.9365	0.9146
BANDRP	**0.9305**	**0.6769**	**0.8802**	**0.9382**	**0.9168**

the three features of drugs, the variant (w/o PubchemFP) performs the worst, while the variant (K_ECFP) exhibits the best performance. However, the disparity in performance among the variant models is not substantial. This may be attributed to the presence of some overlapping chemical and structural information in the molecular fingerprints. For the four features of cancer cell lines, the result of the variants with only one feature removed shows that (w/o EXP) is the most degraded performance, indicating that the gene expression feature provides more information than other features. Among the variants retaining only one cancer cell line feature, the lowest RMSE is achieved by (K_MUT), while the highest is by (K_PATH). The experimental results suggest that each feature contributes uniquely to the overall performance of the model, highlighting the importance of integrating multiple features for optimal predictive accuracy. To further investigate the contribution of different cancer cell line features, we extract the attention map from the bilinear attention module in the trained BANDRP model and analyze it using the attention weights within the attention map. Detailed information can be found in Supplementary Section 3.

In addition, we analyze the impact of molecular graphs and physicochemical properties of drugs as drug features on the model. We design the following BANDRP variant models with different drug feature inputs while keeping other structures unchanged.

- **U_DMG**: Using molecular graphs and convolutional neural networks instead of molecular fingerprints and MLPs in BANDRP.
- **U_PC**: Using physicochemical properties instead of fingerprints in BANDRP.
- **U_DMG_PC**: Using molecular graphs and convolutional neural networks, as well as physicochemical properties and an MLP, instead of molecular fingerprints and MLPs in BANDRP.
- **U_DMG_FP**: Adding molecular graphs and convolutional neural networks to BANDRP.
- **U_PC_FP**: Adding physicochemical properties and an MLP to BANDRP.
- **U_DMG_PC_FP**: Adding molecular graphs and convolutional neural networks, as well as physicochemical properties and an MLP, to BANDRP.

Specifically, following Liu *et al.* [25], we generate molecular graphs based on SMILES sequences of drugs and obtain drug representations through a convolutional neural network graph. We use the RDKit Python package to calculate the physicochemical properties of drugs, resulting in a physicochemical property vector

Table 7. Ablation experiment results of different drug input.

Method	RMSE	MAE	R^2	PCC	SCC
U_DMG	1.0603	0.7700	0.8559	0.9277	0.9032
U_PC	1.0254	0.7512	0.8607	0.9286	0.9053
U_DMG_PC	1.2394	0.8544	0.8316	0.9228	0.8994
U_DMG_FP	1.1225	0.7996	0.8475	0.9245	0.9000
U_PC_FP	0.9834	0.7259	0.8686	0.9323	0.9092
U_DMG_PC_FP	1.0164	0.7412	0.8619	0.9292	0.9071
BANDRP	**0.9305**	**0.6769**	**0.8802**	**0.9382**	**0.9168**

Table 8. Ablation experiment results of different feature fusion methods.

Method	RMSE	MAE	R^2	PCC	SCC
BANDRP_st	1.0279	0.7348	0.8607	0.9279	0.9050
BANDRP_mlp	1.0279	0.7348	0.8607	0.9279	0.9050
BANDRP	**0.9305**	**0.6769**	**0.8802**	**0.9382**	**0.9168**

of dimension 209 for each drug, and construct an MLP module to obtain the drug representation. For variant models using multiple features, such as U_DMG_PC, similar to BANDRP, we concatenate the drug representations obtained from different methods into a matrix, which serves as the input for the Bilinear attention module. The experimental results, as shown in Table 7, indicate that the best performance is achieved when only molecular fingerprints are used within the current task and model framework.

In general, BANDRP leverages diverse information from cancer cell lines and drugs for predictions of anti-cancer drug response. The removal of features may reduce its predictive capabilities, underscoring the effectiveness of the feature selection scheme for the prediction task.

The impact of feature fusion methods

To validate the impact of the bilinear attention network on feature fusion in the model, we construct two variant models. BANDRP_st, which uses a self-attention network for feature fusion, and BANDRP_mlp, which uses an MLP network for feature fusion. Under the same experimental conditions, the results are shown in Table 8. The experimental results indicate that the feature fusion method using a bilinear attention network is optimal.

Prediction of unknown drug responses

In this section, we employ the known IC50 values from the benchmark dataset to train the model and predict the missing 9,117 IC50 values. The result illustrates the distribution of predicted values for missing drug responses, organized by drugs, with a focus on the top 10 sensitive drugs and the last 10 resistant drugs (see Fig. 2). Notably, Bortezomib ranks as the top 'sensitive' drug. Functioning as a proteasome inhibitor, Bortezomib induces apoptosis in cancer cells by inhibiting angiogenesis and reducing cancer cell adhesion to the matrix [26]. Numerous studies have substantiated its efficacy in multiple myeloma and various hematologic and solid tumors [27, 28]. Specifically, the predicted IC50 value for Bortezomib in the esophageal squamous cell carcinoma cell line TE-15 is −5.12, underscoring its therapeutic effect against this cancer cell line, as corroborated by Lioni et al.'s study [29]. Daporinad (also known as FK886) secures the second rank. As an anti-tumor and anti-angiogenic agent, it depletes the energy reserves of metabolically active tumor cells by inhibiting the synthesis of nicotinamide adenine dinucleotide (NAD+), inducing apoptosis [30]. In the prediction results, Daporinad shows lower average IC50 values across multiple types of cancer cell lines. Temozolomide ranks first among 'resistant' drugs, primarily designed for the treatment of specific types of brain tumors [31]. Therefore, it exhibits a higher average IC50 value in the results. Collectively, this evidence underscores the accurate predictive capabilities of BANDRP.

Case study

Acute myeloid leukemia (AML) is distinguished by the abnormal proliferation of blood cells in the bone marrow, exhibiting resistance to treatment, high relapse rates, and challenges in achieving a cure [32]. Despite the utilization of some effective drugs for AML, there is a pressing need for innovative therapeutic approaches to enhance patient survival rates. In this study, we conduct a *De novo* test, focusing on AML cell lines using a benchmark dataset. More specifically, nine AML cell lines are selected, and the drug responses associated with these cell lines are designated as the test set. The remaining drug responses are utilized for model training. The outcome involves the prediction of IC50 values for 169 drugs across 9 AML cell lines (see Fig. 3). Upon ranking the predicted drug values in AML cell lines by their mean values, the top 3 drugs exhibiting sensitivity are Vinblastine, Docetaxel, and Bortezomib. Notably, existing biological experiments [33–35] affirm the efficacy of these drugs in AML treatment. These results suggest that BANDRP can aid in discovering novel therapeutic effects of anti-cancer drugs and provide potential treatment options for cancer therapy.

Discussion and conclusion

In this study, we present BANDRP, a novel approach for predicting anti-cancer drug response by integrating multi-modal data from cancer cell lines and drugs. BANDRP captures the features of cancer cell lines across four views: transcriptomics, genomics, epigenetics, and biological pathways. Similarly, it learns the features of drugs using diverse molecular fingerprints. Notably, BANDRP incorporates a bilinear attention module to effectively fuse features from cancer cell lines and drugs. Experimental results on benchmark datasets and an independent test set consistently

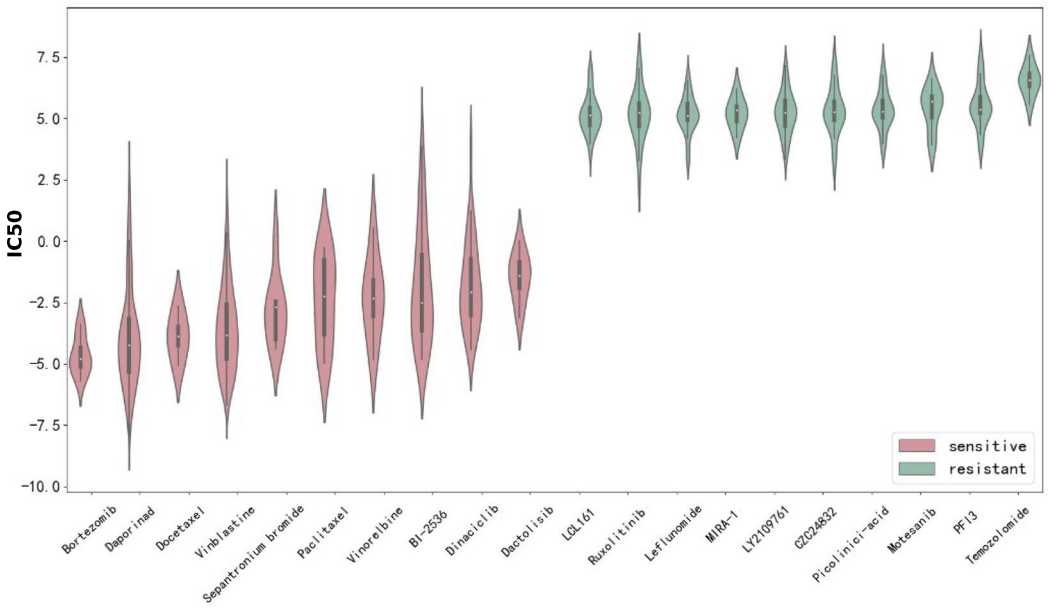

Figure 2. Predicting unknown IC50 values and ranking the resulting top 10 'sensitive' and last 10 'resistant' drugs, plotted by the mean for each drug.

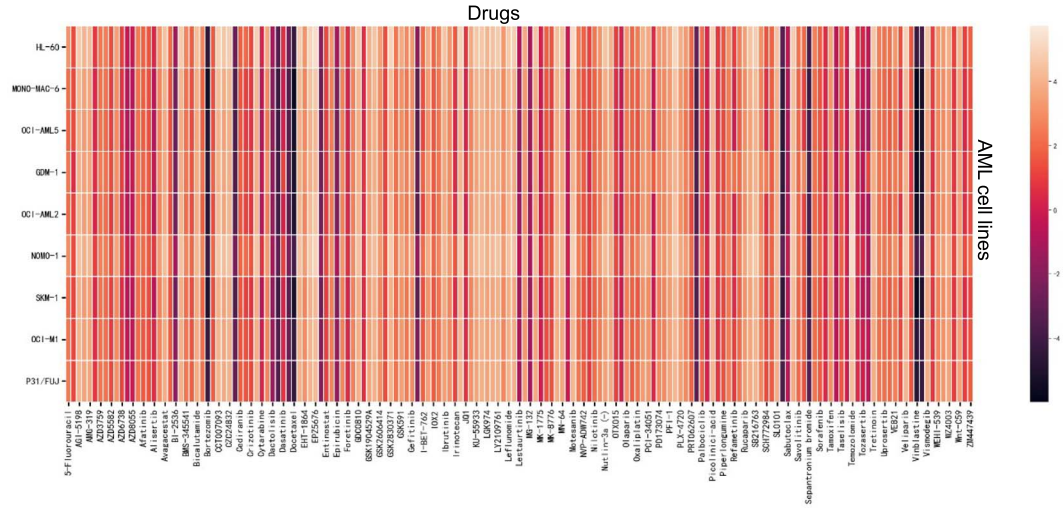

Figure 3. Heatmap of drug responses across 9 AML cell lines and 169 drugs. The horizontal axis represents drugs, the vertical axis represents AML cell lines, and the darker the color, the lower the predicted IC50 value.

demonstrate the superiority of BANDRP over competing methods. Feature analysis and ablation experiments provide insights into the viability of our feature selection scheme for the prediction task. Additionally, analytical experiments and case studies on unknown anti-cancer drug response predictions underscore the predictive capabilities of BANDRP.

Despite the promising performance of BANDRP in anti-cancer drug response prediction, there are areas for potential improvement. Firstly, the model does not currently account for biological entities such as targets and diseases. Exploring the inclusion of biological entities and managing the challenges of data sparsity in this context is an avenue for future exploration. Secondly, while the use of molecular fingerprint-based drug representation proves effective, further exploration is needed to enhance the model's predictive performance when dealing with unknown drugs. Third, BANDRP focuses on predicting drug responses in single-drug scenarios. Considering clinical drug use scenarios, the prediction of synergistic drug combinations deserves further exploration, and some computational methods to predict synergistic drug combinations have already been proposed [36, 37]. In future research, we plan to develop our model on larger datasets and explore the integration of pre-trained models to enhance drug representations. Additionally, we will aim to focus on various combinations of drugs and develop prediction models for drug combination responses to better reflect real clinical treatment strategies, optimize combination therapies, and improve therapeutic outcomes.

Key Points

- This study proposes a bilinear attention network, named BANDRP, for anti-cancer drug response prediction based on multi-omics data of cancer cell lines and molecular fingerprints of drugs.

- This study utilizes pathway enrichment scores to enhance the features of cancer cell lines. It also automatically learns the interactive information between cancer cell lines and drugs through bilinear attention networks, improving prediction accuracy.
- Experiments demonstrate that BANDRP surpasses existing models in predictive performance. Feature analysis and ablation experiments confirm the optimality of the model's architecture and feature selection scheme.

Supplementary data

Supplementary data are available at *Briefings in Bioinformatics* online.

Conflict of interest: None declared.

Funding

This work was supported in part by the National Key Research and Development Program of China (No. 2021YFF1201200), the National Natural Science Foundation of China (No. U22A2041, No. 62272309, and No. 62472202), the National Funded Postdoctoral Program of China (No. G2C20233162), the Natural Science Foundation of Hunan Province (No. 2022JJ30750), the Science and Technology Major Project of Changsha (No. kh2402004). This work was also carried out in part using computing resources at the High Performance Computing Center of Central South University.

Data availability

The codes of BANDRP are available at https://github.com/heckletbot/BANDRP.

Availability of data and code

All datasets utilized in this study are publicly available, and the source code is available online at https://github.com/heckletbot/BANDRP.

References

1. Hodson R. Precision medicine. *Nature* 2016;**537**:S49–9. https://doi.org/10.1038/537S49a.
2. Li K, Du Y, Li L. *et al.* Bioinformatics approaches for anticancer drug discovery. *Curr Drug Targets* 2020;**21**:3–17. https://doi.org/10.2174/1389450120666190923162203.
3. Barretina J, Caponigro G, Stransky N. *et al.* The cancer cell line encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* 2012;**483**:603–7. https://doi.org/10.1038/nature11003.
4. Iorio F, Knijnenburg TA, Vis DJ. *et al.* A landscape of pharmacogenomic interactions in cancer. *Cell* 2016;**166**:740–54. https://doi.org/10.1016/j.cell.2016.06.017.
5. Kim S, Chen J, Cheng T. *et al.* Pubchem 2019 update: improved access to chemical data. *Nucleic Acids Res* 2019;**47**:D1102–9. https://doi.org/10.1093/nar/gky1033.
6. Guan N, Zhao Y, Wang C. *et al.* Anticancer drug response prediction in cell lines using weighted graph regularized matrix factorization. *Mol Ther-Nucleic Acids* 2019;**17**:164–74. https://doi.org/10.1016/j.omtn.2019.05.017.
7. Liu H, Zhao Y, Zhang L. *et al.* Anti-cancer drug response prediction using neighbor-based collaborative filtering with global effect removal. *Mol Ther-Nucleic Acids* 2018;**13**:303–11. https://doi.org/10.1016/j.omtn.2018.09.011.
8. Stanfield Z, Coşkun M, Koyutürk M. Drug response prediction as a link prediction problem. *Sci Rep* 2017;**7**:40321. https://doi.org/10.1038/srep40321.
9. Wang Y, Fang J, Chen S. Inferences of drug responses in cancer cells from cancer genomic features and compound chemical and therapeutic properties. *Sci Rep* 2016;**6**:32679. https://doi.org/10.1038/srep32679.
10. Wan Q, Pal R. An ensemble based top performing approach for nci-dream drug sensitivity prediction challenge. *PloS One* 2014;**9**:e101183. https://doi.org/10.1371/journal.pone.0101183.
11. Liu P, Li H, Li S. *et al.* Improving prediction of phenotypic drug response on cancer cell lines using deep convolutional network. *BMC Bioinform* 2019;**20**:1–14. https://doi.org/10.1186/s12859-019-2910-6.
12. Liu Q, Hu Z, Jiang R. *et al.* DeepCDR: a hybrid graph convolutional network for predicting cancer drug response. *Bioinformatics* 2020;**36**:i911–8. https://doi.org/10.1093/bioinformatics/btaa822.
13. Chawla S, Rockstroh A, Lehman M. *et al.* Gene expression based inference of cancer drug sensitivity. *Nat Commun* 2022;**13**:5680. https://doi.org/10.1038/s41467-022-33291-z.
14. Jiang L, Jiang C, Yu X. *et al.* DeepTTA: a transformer-based model for predicting cancer drug response. *Brief Bioinform* 2022;**23**:bbac100. https://doi.org/10.1093/bib/bbac100.
15. Wang H, Dai C, Wen Y. *et al.* Gadrp: Graph convolutional networks and autoencoders for cancer drug response prediction. *Brief Bioinform* 2023;**24**:bbac501. https://doi.org/10.1093/bib/bbac501.
16. Shen B, Feng F, Li K. *et al.* A systematic assessment of deep learning methods for drug response prediction: from in vitro to clinical applications. *Brief Bioinform* 2023;**24**:bbac605. https://doi.org/10.1093/bib/bbac605.
17. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinform* 2013;**14**:1–15. https://doi.org/10.1186/1471-2105-14-7.
18. Sondka Z, Bamford S, Cole CG. *et al.* The cosmic cancer gene census: describing genetic dysfunction across all human cancers. *Nat Rev Cancer* 2018;**18**:696–705. https://doi.org/10.1038/s41568-018-0060-1.
19. Rogers D, Hahn M. Extended-connectivity fingerprints. *J Chem Inf Model* 2010;**50**:742–54. https://doi.org/10.1021/ci100050t.
20. Huang K, Xiao C, Glass L. *et al.* Explainable Substructure Partition Fingerprint for protein, drug, and more. In: *NeurIPS Learning Meaningful Representation of Life Workshop*, MIT Press, Vancouver, Canada, 2019.
21. Liberzon A, Subramanian A, Pinchback R. *et al.* Molecular Signatures Database (MSigDB) 3.0. *Bioinformatics* 2011;**27**:1739–40. https://doi.org/10.1093/bioinformatics/btr260.
22. Bento AP, Hersey A, Félix E. *et al.* An open source chemical structure curation pipeline using RDKit. *Journal of Cheminformatics*, 2020;**12**:1–16. https://doi.org/10.1186/s13321-020-00456-1.
23. Willett P. The calculation of molecular structural similarity: principles and practice. *Mol Inform* 2014;**33**:403–13. https://doi.org/10.1002/minf.201400024.
24. Reyad M, Sarhan A M, Arafa M. A modified Adam algorithm for deep neural network optimization. *Neural Computing and Applications*, 2023;**35**:17095–17112. https://doi.org/10.1007/s00521-023-08568-z.
25. Liu X, Song C, Huang F. *et al.* GraphCDR: a graph neural network method with contrastive learning for cancer drug response

prediction. *Brief Bioinform* 2022;**23**:bbab457. https://doi.org/10.1093/bib/bbab457.
26. Mujtaba T, Dou QP. Advances in the understanding of mechanisms and therapeutic use of bortezomib. *Discov Med* 2011;**12**:471–80.
27. Roccaro AM, Vacca A, Ribatti D. Bortezomib in the treatment of cancer. *Recent Pat Anticancer Drug Discov* 2006;**1**:397–403. https://doi.org/10.2174/157489206778776925.
28. Curran MP, McKeage K. Bortezomib: a review of its use in patients with multiple myeloma. *Drugs* 2009;**69**:859–88. https://doi.org/10.2165/00003495-200969070-00006.
29. Lioni M, Noma K, Snyder A. *et al.* Bortezomib induces apoptosis in esophageal squamous cell carcinoma cells through activation of the p38 mitogen-activated protein kinase pathway. *Mol Cancer Ther* 2008;**7**:2866–75. https://doi.org/10.1158/1535-7163.MCT-08-0391.
30. Hasmann M, Schemainda I. FK866, a highly specific noncompetitive inhibitor of nicotinamide phosphoribosyltransferase, represents a novel mechanism for induction of tumor cell apoptosis. *Cancer Res* 2003;**63**:7436–42.
31. Zhang J, Stevens MFG, Bradshaw TD. Temozolomide: mechanisms of action, repair and resistance. *Curr Mol Pharmacol* 2012;**5**:102–14. https://doi.org/10.2174/1874467211205010102.
32. Khwaja A, Bjorkholm M, Gale RE. *et al.* Acute myeloid leukaemia. *Nat Rev Dis Primers* 2016;**2**:1–22. https://doi.org/10.1038/nrdp.2016.10.
33. De La Garza F, Pedraza PRC, De Leon AG. *et al.* The leukoreductive power of vinblastine in acute myeloid leukemia with hyperleukocytosis: a pilot study. *Blood* 2020;**136**:35–6. https://doi.org/10.1182/blood-2020-143404.
34. Szatrowski T, Hensley M, Ely S. Phase I trial of docetaxel (taxotere) for patients with relapsed or refractory acute myeloid leukemia (AML). In: *Acute Leukemias VIII: Prognostic Factors and Treatment Strategies*. Springer, Berlin, Heidelberg; 2001, p. 477–81, https://doi.org/10.1007/978-3-642-18156-6_77.
35. Conticello C, Adamo L, Vicari L. *et al.* Antitumor activity of bortezomib alone and in combination with trail in human acute myeloid leukemia. *Acta Haematol* 2008;**120**:19–30. https://doi.org/10.1159/000151511.
36. Li TH, Wang CC, Zhang L. *et al.* SNRMPACDC: computational model focused on siamese network and random matrix projection for anticancer synergistic drug combination prediction. *Brief Bioinform* 2023;**24**:bbac503. https://doi.org/10.1093/bib/bbac503.
37. Chen X, Ren B, Chen M. *et al.* NLLSS: predicting synergistic drug combinations based on semi-supervised learning. *PLoS Comput Biol* 2016;**12**:e1004975. https://doi.org/10.1371/journal.pcbi.1004975.