npj | systems biology and applications

Article

Published in partnership with the Systems Biology Institute

https://doi.org/10.1038/s41540-024-00421-w

Using DeepSignalingFlow to mine signaling flows interpreting mechanism of synergy of cocktails

Check for updates

Heming Zhang [1], Yixin Chen[2], Philip Payne[1] & Fuhai Li[1,3]

Complex signaling pathways are believed to be responsible for drug resistance. Drug combinations perturbing multiple signaling targets have the potential to reduce drug resistance. The large-scale multi-omic datasets and experimental drug combination synergistic score data are valuable resources to study mechanisms of synergy (MoS) to guide the development of precision drug combinations. However, signaling patterns of MoS are complex and remain unclear, and thus it is challenging to identify synergistic drug combinations in clinical. Herein, we proposed a novel integrative and interpretable graph AI model, *DeepSignalingFlow*, to uncover the MoS by integrating and mining multi-omic data. The major innovation is that we uncover MoS by modeling the signaling flow from multi-omic features of essential disease proteins to the drug targets, which has not been introduced by the existing models. The model performance was assessed utilizing four distinct drug combination synergy evaluation datasets, i.e., NCI ALMANAC, O'Neil, DrugComb, and DrugCombDB. The comparison results showed that the proposed model outperformed existing graph AI models in terms of synergy score prediction, and can interpret MoS using the core signaling flows. The code is publicly accessible via Github: https://github.com/FuhaiLiAiLab/DeepSignalingFlow

Complex signaling pathways play a crucial role in driving drug resistance in various types of cancer. These pathways involve a multitude of mechanisms such as the activation of signaling pathways promoting cell survival and proliferation, upregulation of drug efflux pumps, emergence of cancer stem cells, and genetic and epigenetic changes[1]. For example, the Wnt/β-catenin signaling pathway has been identified as a key player in resistance development in triple-negative breast cancer, interacting with receptor tyrosine kinase signaling pathways, epithelial-mesenchymal transition, and metastasis[2]. Additionally, signaling pathways like MAPK, PI3K/PKB, Wnt/β-catenin, and Notch have been associated with multidrug resistance in colorectal cancer, affecting cell survival, and drug metabolism[3]. Thus, targeting these intricate signaling pathways is essential for developing effective strategies to overcome drug resistance in cancer therapy. Previous research has demonstrated that the inhibition of a single component within a signaling network may not be sufficient to effectively treat cancer due to the robustness and resistance of network signaling[4]. Thus, drug combinations inhibiting more signaling targets, compared with single drug treatment, in cancer-related signaling networks have the potential to be effective in reducing drug resistance. For example, the RAS/ERK signaling pathway, which has been shown to inhibit the efficacy of certain anti-cancer drugs, was found to be synergistic with autophagy inhibitors in RAS-driven cancers. This is notable, given that autophagy signaling was activated after treatment with RAS inhibitors[5,6].

However, it is challenging to discover effective and synergistic drug combinations in clinical and laboratory settings because of the unclear and complex mechanism of synergy (MoS) for such drug combinations. Recent analyses of reported drug combinations indicate that there are complex, diverse, and heterogeneous MoS for different drug combinations in different diseases[7]. Despite this evolving body of research concerning the targeting of synergistic cancer therapeutics, there remain few effective drug combinations that are used broadly in the clinical environment. Therefore, novel and effective drug combinations are needed in order to expand the scope and impact of precision cancer therapy and ultimately improve cancer treatment outcomes.

To help facilitate the discovery of synergistic drug combinations for use in a variety of cancers, we sought to understand the associations between multi-omic features and drug response, as well as the mechanism of drug synergy, using computational methods. The National Cancer Institute

[1]Institute for Informatics, Data Science and Biostatistics (I2DB), Washington University School of Medicine, St. Louis, MO, USA. [2]Computer Science, Washington University School of Medicine, St. Louis, MO, USA. [3]Department of Pediatrics, Washington University School of Medicine, Washington University in St. Louis, St. Louis, MO, USA. e-mail: Fuhai.Li@wustl.edu

(NCI) has previously generated a comprehensive screening dataset of 5232 combinations from ~100 drugs in 60 human tumor cell lines (NCI-60), which is publicly available as the NCI-ALMANAC Drug Combination Data Set[8]. This dataset is valuable for developing computational models for drug combination prediction. Aside from that, O'Neil dataset[9] is another drug combination screening dataset containing 38 compounds and 39 cancer cell lines with 583 combinations and 22,737 experiments. In addition, DrugComb[10] is an innovative web-based data portal designed to standardize and harmonize drug combination screening data for cancer cell lines. It features a comprehensive collection of 437,923 drug combination experiments, encompassing 7,423,800 data points across 93 human cancer cell lines. DrugCombDB[11] is also a comprehensive dataset that includes 448,555 drug combinations derived from high-throughput screening (HTS) assays. It encompasses 2887 unique drugs tested across 124 human cancer cell lines. With over 6,000,000 quantitative dose responses, DrugCombDB computes multiple synergy scores to assess the overall synergistic or antagonistic effects of drug combinations, providing a valuable resource for advancing cancer treatment research.

There have been a variety of computational and deep-learning models described in the scientific literature concerning the application of machine learning techniques to such data. For example, random forest[12] (RF) and XGBoost[13] models have been used, integrating chemical and genomics features of source data, to predict drug combinations[14]. The regression model was used in a feedback system control (FSC) framework[15,16] for drug combination prediction. In addition, the use of multi-task learning with a transfer learning model has been proposed to predict drug combinations[17]. Then, deep-learning models have also been developed for drug combination prediction. For example, Matchmaker[18] was developed to analyze chemical and gene expression profiles using a deep-learning framework, which is similar to the DeepSynergy[19] and AuDNNsynergy[20], and DeepSignalingSynergy[21,22]. Multi-omic data were concatenated as integration in those deep-learning models[23] to improve the model robustness and accuracy[23]. Moreover, the Visible Neural NetworkVNN (VNN) models (e.g., DCell and DrugCell)[24–26] were proposed, using large hierarchical deep-learning architecture to model the hierarchical organization of biological processes, to predict drug response with important biomarkers. However, the signaling cascade level activity was not specifically investigated in the DCell and DrugCell models. These models concatenate a large number of genomics and chemical structure features using dense connection layers for drug combination prediction. However, the signaling pathway information was not incorporated. Thus, these models are less interpretable compared with models integrating omic data with signaling pathways.

Despite the promise of the existing body of work in this domain, there remain two limitations of the existing deep-learning models. First, the signaling interactions in gene regulatory signaling networks are often not incorporated in the existing deep-learning models, despite their criticality in terms of treatment efficacy. Second, complex mechanisms of synergy (MoS) are usually not investigated in these existing models. As such, these models are difficult to interpret and explain, which hinders the identification of MoS of synergistic drug combinations. Recently, Graph Neural Networks (GNNs) have gained prominence due to their capability to model relationships within graph-structured data[27–29] having been applied to numerous studies for the integration of multi-omics data. Compared with DNN models, GNN models can naturally integrate the signaling pathways or protein-protein interactions (networks/graphs—node and node interactions) with omic features (features of nodes). Then the target ranking and subnetwork or pathway inference can be achieved by using node and edge ranking operations in GNN. Therefore, GNN or graph AI models are ideal models for integrative and interpretable omic data analysis and for mining the MoS of synergistic drug combinations using the top-ranked signaling flow/cascades. For example, MOGONET[30] initially creates similarity graphs among samples by leveraging each omics data, then employs a graph convolutional network (GCN) to learn a label distribution from each omics data independently. Subsequently, a cross-omics discovery tensor is implemented to refine the prediction by learning the dependency among multi-omics data. MoGCN[31] adopts a similar approach by constructing a patient similarity network using multi-omics data and then using GCN to predict the cancer subtype of patients. GCN-SC[32] utilizes a GCN to combine single-cell multi-omics data derived from varying sequencing methodologies. MOGCL[33] takes this further by exploiting the potency of graph contrastive learning to pre-train the GCN on the multi-omics dataset, thereby achieving impressive results in downstream tasks with fine-tuning. DeepDDS[34] incorporated the chemical structure information of drugs into the GNN framework to achieve better performance on drug response prediction. Nevertheless, none of the aforementioned techniques contemplate incorporating structured signaling pathway data into the model to mine the detailed signaling cascades/flows associated with the MoS of drug cocktails.

In the specific context of cancer studies, signaling pathways playing important roles in terms of tumor development and drug response, are of critical importance when seeking to identify single or synergistic therapeutic targets. For example, in one set of studies, 10 signaling pathways were systematically analyzed across 9,125 tumor samples in 33 cancer types (64 subtypes)[35]. The results of this study indicated that 89% of tumor samples had at least one driver alteration in one of these pathways, and 57% and 30% had one and multiple potentially druggable targets, respectively. These results indicate the importance of identifying or repurposing effective drugs and drug combinations that target these signaling pathways. Therefore, in our study, we proposed a novel computational graph neural network (GNN) model, *DeepSignalingFlow*, to investigate the aforementioned issues as they relate to the use of deep-learning models for drug combination therapy prediction, by incorporating core cancer signaling pathways and multi-omic data into such analyses. What's more, the interpretability of the signaling flow was provided for analyzing the key signaling and targets for each cancer in our model.

Results

Architecture of *DeepSignalingFlow*

Figure 1 illustrates the schematic architecture of the proposed *DeepSignalingFlow* model. The model input parameters are: $X \in \mathbb{R}^{n\times d}$, $A \in \mathbb{R}^{n\times n}$, $D_{in} \in \mathbb{R}^{n\times n}$, $D_{out} \in \mathbb{R}^{n\times n}$, where X denotes the nodes features matrix with n nodes of d features. A is the adjacency matrix that links gene-gene interactions on the signaling pathways and gene-drug interactions, and the element of adjacency matrix A such as a_{ij} indicates an edge from i to j. D_{in} is an in-degree diagonal matrix for nodes in directed graph, and D_{out} is an out-degree diagonal matrix for nodes in directed graph. In details, n represents the total number of nodes in the graph, which consists of two distinct types of nodes: genes and drugs. Specifically, n_1 denotes the number of gene nodes, and n_2 denotes the number of drug nodes within the entire graph. Therefore, the total number of nodes n can be expressed as the sum of n_1 and n_2.

In the graph message passing[36–38] stages of our architecture, we added, WeB-GNN (Weighted Bi-direction Graph Neural Network) layer, the bi-directional nodes-flow with both 'upstream-to-downstream' (from upstream signaling to drug targets) and 'downstream-to-upstream' (from drug targets to downstream signaling), considering that multi-channel setting in the architecture can integrate the two-stream information and improve the model performances[39–43]. The trainable weight matrices are denoted as: $W_{up}^{(k)} \in \mathbb{R}^{n\times n}$, $W_{down}^{(k)} \in \mathbb{R}^{n\times n}$, $(k = 1, 2, \ldots,$ K, where K is the number of message propagation (WeB-GNN) layers in *DeepSignalingFlow*) respectively. Then the weight adjacency matrices will be denoted as:

$$A'^{(k)}_{up} = W^{(k)}_{up} \cdot A \tag{1}$$

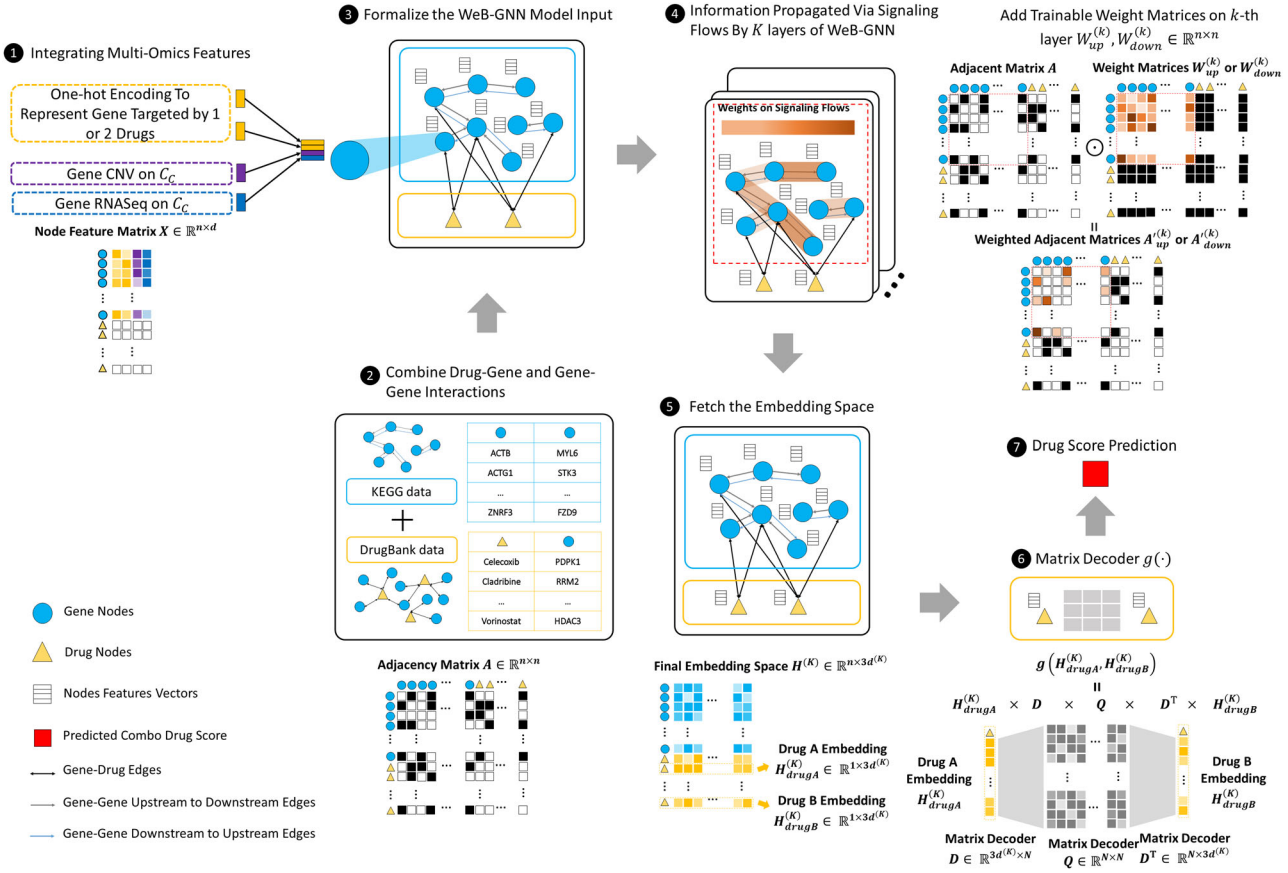


Fig. 1 | Model architecture of *DeepSignalingFlow*. ❶ Integrating the one-hot encoded and multi-omics features as a vector in each node. ❷ Combine the drug-gene and gene-gene interactions into the network as an adjacency matrix. ❸ Formalize the graph neural network input with feature matrix and edge index. ❹ Message passing layers of added trainable weighted matrices with bi-directional signaling flow (upstream and downstream). ❺ Fetching the embedded spaces from the pair of drug nodes as the initial vectors in decoder space. ❻ Decoding the pair of drug nodes with the trainable matrices D and Q.❼ Decoded value, $\hat{y}$, as the predicted drug combination synergistic score.

$$A'^{(k)}_{down} = W^{(k)}_{down} \cdot A^{\mathrm{T}}, \tag{2}$$

where for each $A'^{(k)}_{up} \in \mathbb{R}^{n \times n}, A'^{(k)}_{down} \in \mathbb{R}^{n \times n}, (k = 1, 2, \ldots, \mathrm{K})$, parameters for drug-gene relations values are the same with those in the original adjacency matrix. Complete model parameters are specified in Table 1. And the structural details of the model are discussed in Methods.

Experimental setup

To validate the proposed model, a 5-fold cross-validation was employed. Specifically, <D_A, D_B, C_C, S_{ABC}> (D_A, D_B are drug combinations, C_C is the cell line name, which will be combined with the multi-omics data as the details and S_{ABC} is the synergistic drug effect on the cell line C_C), a 4-element tuple, represents an input data point. In total, there are 2788, 1008, 27,781, and 21,399 model input data points for NCI ALMANAC[8], O'Neil[9], DrugComb[10], and DrugCombDB[11] datasets. In each dataset, 4 folds of the data points will be used as the training dataset, and the rest 1 fold will be used as the test dataset. Considering all possible 2 drug permutations, each generated graph comprised 2016 gene nodes (n_1=2016), and the number of drug nodes (n_2) varied across different datasets with each node in these graphs having 4 initial features. For those 2016 genes, their 4 features indicate the RNA sequence number, copy number variation, and whether they have a connection spanning D_A and D_B respectively. For each 2 drug pair, their 4 features are initiated with zeros. Additionally, to indicate connections between nodes, adjacency matrices are also created, being formed from the KEGG[44] database which contains gene pairs with sources and destinations, resulting in directed graphs. For drug-gene edges, the connections are bi-directional, meaning that in adjacency matrices, these elements are symmetric. Concurrently, in-degree and out-degree matrices are formed according to the original input adjacency matrix $A \in \mathbb{R}^{n \times n}$.

Hyperparameters setting

The number of training epochs was empirically set to 500 using a step-leaning rate scheduler, which allowed for optimal results. Three GNN layers ($K = 3$) were used with an initial input feature dimension of $d^{(0)} = 4$. The feature dimensions varied at the different layers and were denoted by: $(d^{(1)}, d^{(2)}, \ldots, d^{(k)}, \ldots, d^{(K)}), k = 1, 2, 3$, concatenating biased node features and transformed node features of the previous layer for both upstream and downstream flows to generate combined dimensions of the GNN layers, resulting $3 \times d^{(k)}$ for k-th layer in output dimensions. The output dimensions of the previous layer served as the input dimensions for the current layer, as follows:

(1) First layer (input dims, output dims): $(d^{(0)}, 3d^{(1)})$
(2) Second layer (input dims, output dims): $(3d^{(1)}, 3d^{(2)})$
(3) Third layer (input dims, output dims): $(3d^{(2)}, 3d^{(3)})$

The final embedded drug node dims were $3d^{(3)}$. The decoder utilized trainable transformation matrices with dimensions $D \in \mathbb{R}^{3d^{(K)} \times N}$ and $Q \in \mathbb{R}^{N \times N}$ with N set to 150 to adapt model performance. Additionally, the LeakyReLU function was employed with a parameter value of 0.1.

Synergistic score prediction evaluation

To evaluate model performance in terms of synergy score prediction for drug combinations, a 5-fold cross-validation was conducted. As shown in

Table 1 | Model parameters and notations

Input Matrices	When $k = 2, 3, \ldots, K$	When $k = 1$
Nodes Feature Matrix	$H^{(k-1)} \in \mathbb{R}^{n \times 3d^{(k-1)}}$	$H^{(0)} \in \mathbb{R}^{n \times d^{(0)}}$
New Nodes Adjacency Matrices	$A'^{(k)}_{up} \in \mathbb{R}^{n \times n}, A'^{(k)}_{down} \in \mathbb{R}^{n \times n}$	Same
Nodes In-degree Diagonal Matrix	$D_{in} \in \mathbb{R}^{n \times n}$	Same
Nodes Out-degree Diagonal Matrix	$D_{out} \in \mathbb{R}^{n \times n}$	Same
Trainable Feature Transformation Matrix	$M^{(k)}_{up} \in \mathbb{R}^{3d^{(k-1)} \times d^{(k)}}$ $M^{(k)}_{down} \in \mathbb{R}^{3d^{(k-1)} \times d^{(k)}}$	$M^{(1)}_{up} \in \mathbb{R}^{d^{(0)} \times d^{(1)}} M^{(1)}_{down} \in \mathbb{R}^{d^{(0)} \times d^{(1)}}$
Trainable Biased Feature Transformation Matrix	$B^{(k)} \in \mathbb{R}^{3d^{(k-1)} \times d^{(k)}}$	$B^{(1)} \in \mathbb{R}^{d^{(0)} \times d^{(1)}}$

Table 2 | Model comparison using average Pearson correlation of 5-fold cross-validation using both NCI ALMANAC, O'Neil, DrugComb, and DrugCombDB datasets

Dataset	NCI ALMANAC	O'Neil	DrugComb	DrugCombDB
GCN	44.88% ± 4.87%	50.07% ± 8.16%	46.40% ± 1.58%	44.87% ± 3.65%
GAT	20.37% ± 7.08%	31.54% ± 15.34%	31.61% ± 6.18%	43.41% ± 2.26%
GraphFormer	36.39% ± 4.51%	28.66% ± 14.12%	52.49% ± 2.04%	44.83% ± 4.18%
MixHop	48.13% ± 14.36%	6.93% ± 5.08%	46.77% ± 2.60%	6.62% ± 2.53%
GIN	52.94% ± 5.28%	23.34% ± 9.37%	31.95% ± 13.93%	42.59% ± 5.52%
BGNN	48.32% ± 5.79%	61.36% ± 3.09%	55.35% ± 0.91%	55.39% ± 2.36%
DeepSignalingFlow	63.75% ± 1.58%	65.81% ± 2.83%	63.68% ± 0.92%	61.51% ± 0.64%

Supplementary Tables 1-4, the average predictions (using the Pearson correlation coefficient) were 63.7523%, 65.8141%, 63.6843%, 61.5092% Pearson correlation for NCI ALMANAC, O'Neil, DrugComb, and DrugCombDB on their test datasets respectively. These prediction results are comparable with existing machine learning models[12,13], indicating the feasibility of drug combination synergy score prediction using a graph neural network with a small set of core signaling pathways genes. Moreover, the proposed model *DeepSignalingFlow* was compared with other graph neural networks (see Table 2 and Fig. 2a), which included the GNN model with mean aggregation in transductive model[28] (no sampling for neighborhood function) and graph attention network[29] (GAT), GraphFormer[45], MixHop[46], and GIN[47]. Each of these GNN models is implemented with the decoder matrices to predict the drug effects. Furthermore, to validate the impacts of trainable edge weight on model performance, the unweighted bi-directional signaling flow (BGNN) layer was compared with the weighted bi-directional signaling flow (WeBGNN) layer in *DeepSignalingFlow*.

Signaling flow interpreting the mechanism of synergy

To investigate the potential MoS, a core signaling subnetwork was generated based on the trained model. Specifically, the integrated signaling flow networks were derived from the large signaling network using the averaged trained directional weight matrices from the 5 splits of test datasets. Afterward, the cell line-specific gene degrees were calculated using the following formula:

$$d_i = \left(\sum_{j=1}^{n} a^{(ij)}_{bind_final} \right)$$

$$d_i^{(C_k)} = d_i \cdot x_i^{(C_k)},$$

where $a^{(ij)}_{bind_final}$ is the element in the i-th row and j-th column of the matrix A'_{bind_final}. Hence, d_i is the weighted degree for gene i and $x_i^{(C_k)}$ is the i-th row of node feature matrix X with the feature of RNA-seq for cell line C_k. This allows for the calculation of the cell line-specific gene degree, $d_i^{(C_k)}$, for gene i on the cell line C_k. Specifically, the cell line-specific core signaling subnetwork were obtained by selecting nodes that meet the criteria of setting a weighted degree threshold of $d_i^{(C_k)}$. In this scenario, edge weight lower than 0.1 in A'_{bind_final} were filtered out, resulting in the core signaling subnetwork for each individual cell lines.

In these cell line-specific core signaling pathways, the summation of weights on the shortest paths (with a distance smaller than 5) between 2 genes targeted by drug combinations in the NCI ALMANAC dataset was analyzed in detail. Supplementary Tables 5 and 6 compare the summation of weights between the top 10 and bottom 10 drug combinations for each cell line (excluding for SK-MEL-2 due to an insufficient number of data points) using F-test p values, mean values, and median values. Notably, 21 out of 41 cell lines exhibited a condition where the top 10 drug combinations had both a higher mean and median value than the bottom 10 combinations, with a p-value smaller than 0.1, indicating significant differences (see Supplementary Figs. 1 and 2 for details). Specifically, the core signaling pathways for cell lines T-47D and PC-3 generated the cell line-specific network with interactions of 88 genes (see Figs. 3 and 4). To mark the critical genes in the core signaling subnetwork, weighted node degree $d_i^{(C_k)}$ lower than 1 were marked red in the above figures.

Analysis of the top-ranked drug combinations revealed that these combinations inhibit multiple essential targets within the core signaling subnetwork under shorter paths. The potential drug targets were identified within a range of 3 or 4 hop neighboring nodes in the selected core signaling subnetwork. This finding demonstrates that these core signaling networks can accurately measure the importance of pathways. Consequently, drug combinations that can perturb the core signaling subnetwork are more likely to be effective.

Discussion

Drug resistance is a common challenge in cancer therapy due to the complex signaling pathways of cancers. Drug combinations inhibiting more and multiple signaling targets, compared with single drug treatment, in cancer-related signaling networks have the potential to be effective in reducing drug resistance. However, it remains an open problem and challenging because of the mysterious, heterogeneous molecular mechanisms of synergy of drug combinations. Deep-learning models have been reported to predict drug

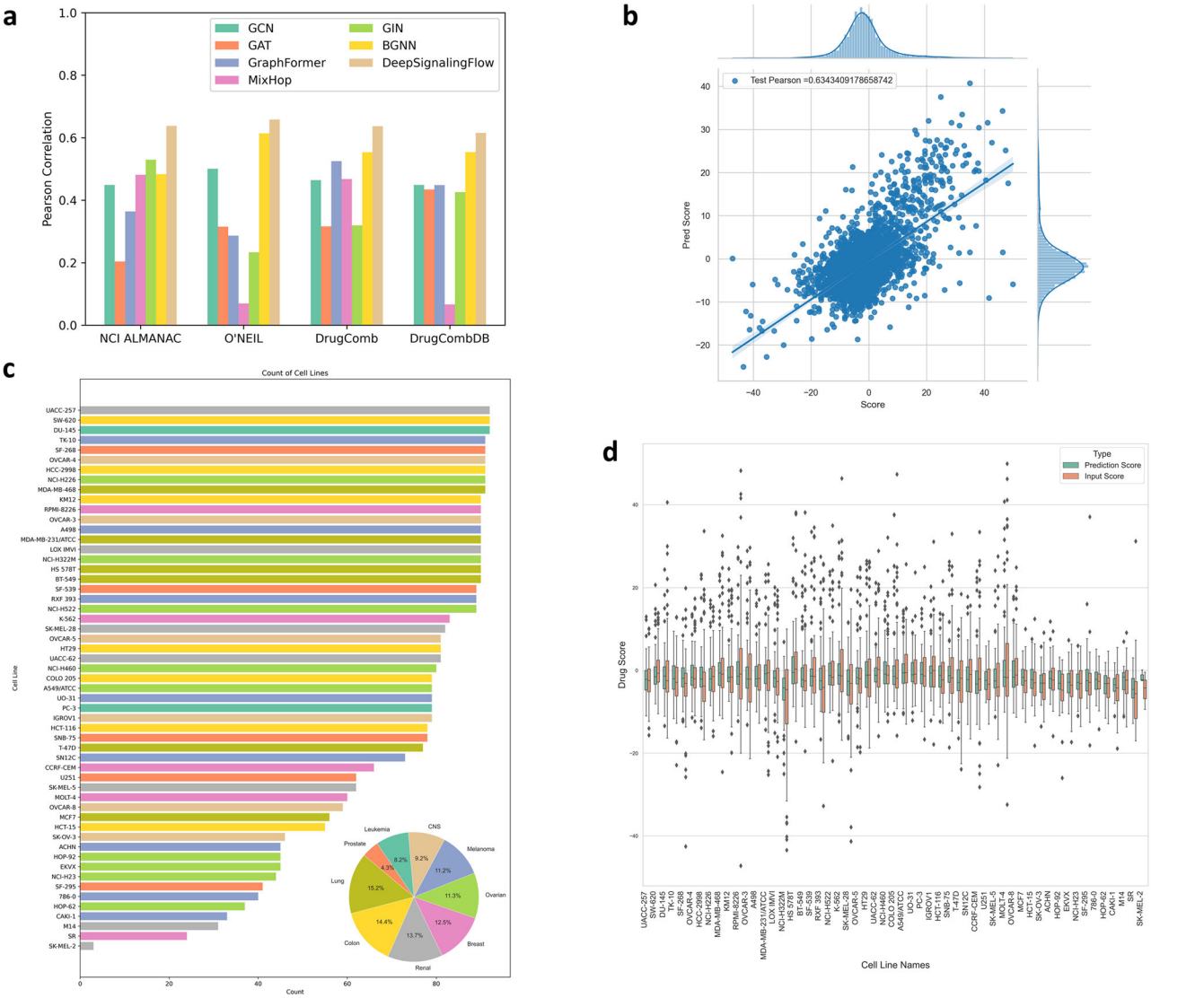


Fig. 2 | Model performance and overview of input datasets NCI ALMANAC, O'Neil, DrugComb, and DrugCombDB. a Pearson correlation comparison for GCN, GAT, GraphFormer, MixHop, GIN Decoder, and *DeepSignalingFlow* models for NCI ALMANAC, O'Neil, DrugComb, and DrugCombDB datasets. **b** Scatter plot of the model with data points in the whole NCI ALMANAC dataset. **c** Distributions of all cell lines in the whole NCI ALMANAC dataset. **d** Box plots across all cell lines in the NCI ALMANAC dataset.

combinations by simply concatenating a large number of multi-omic data and chemical structure information using the deep neural network (DNN) model. One limitation of the existing DNN models is the missing of complex and biologically meaningful gene regulatory relationships. Thus, it is hard to interpret these models to support real clinical use, which requires clinical experts to understand the molecular mechanisms used in the AI prediction.

In this study, we aim to explore the feasibility of building an interpretable AI model using the proposed *DeepSignalingFlow* model. The model is built on a core cancer signaling network consisting of 2016 genes, with gene expression and copy number data derived from 48 core cancer signaling pathways. The novel upstream signaling flow (from upstream signaling to drug targets), and the downstream signaling flow (from drug targets to downstream signaling), were designed using trainable weights of network edges. The selected signaling interaction edges with large weights indicate (1) the important biomarker genes; (2) the signaling cascades among them, which can potentially interpret the model prediction results; and (3) the potential synergistic drug targets in the cell line-specific core signaling pathways. In conclusion, the first-time proposed upstream and downstream signaling flow design is novel and can potentially interpret the mechanism of drug combination response and offer the potential targets of future drug targets. Understanding the upstream and downstream signaling pathways/mechanisms of drugs is crucial, as drugs can perturb disease-related dysfunctional signaling by either 'blocking' the signaling cascades/flow from upstream dysfunctional signaling or 'inhibiting' the downstream dysfunctional signaling. By mapping the signaling flow dynamics, this model helps elucidate how signals propagate through the network to achieve a synergistic response. Gaining such an understanding of the mechanistic basis of drug combinations is instrumental in designing drug combinations that can overcome resistance, ultimately leading to more effective therapies. The model can be also applied to general multi-omic data analysis studies.

This is still an exploratory study to predict and investigate drug combination response using deep graph neural network models. There are some limitations to be further investigated. For example, additional large-scale signaling pathways and protein-protein interactions can cover more proteins and drug targets and will include more drugs, which will be further investigated. In addition, our current analysis is a global analysis or pan-cancer analysis by considering all drugs and all cell lines together. However, the drug combination and cell line-specific signaling flows might be different. Thus, it is interesting to investigate the drug- and cell-specific signaling flows that are informative for the drug combination response

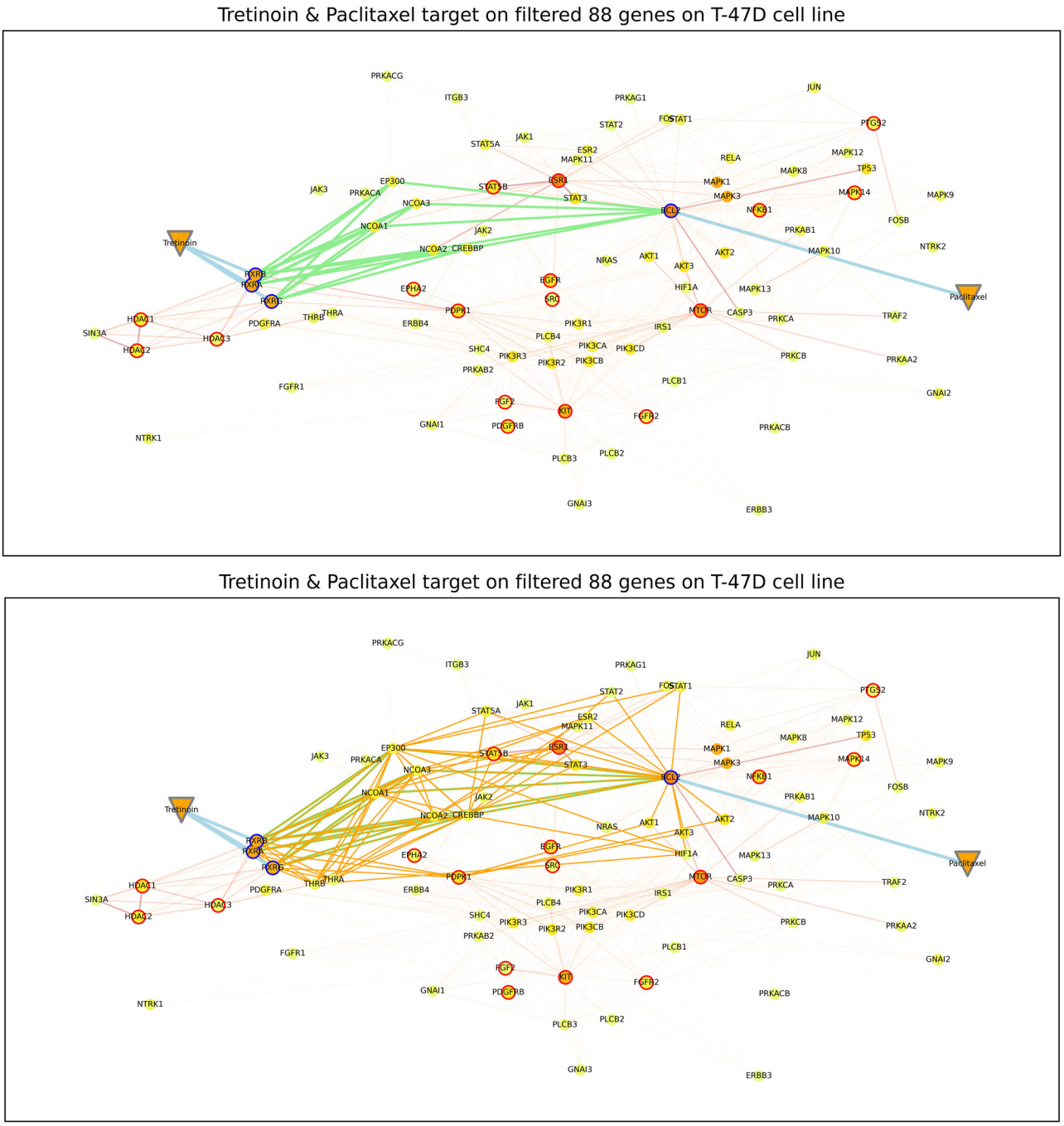


Fig. 3 | T-47D cell line-specific effective combinations perturbing the core signaling subnetwork targeted by the drug Tretinoin and Romedepsin in the core signaling pathways. All nodes considered as important targets in *DeepSignalingFlow* are circled with red and nodes targeted by Celecoxib and Sirolimus will be circled with blue. The upper panel shows the critical paths between 2 drugs within a distance of 4 (green edges) and the lower panel shows the paths within a distance of 5 (orange edges). In particular, the depth of color for nodes and edges indicates node degrees $d_i^{(T47D)}$ and edge weights in matrix A'_{bind_final}.

prediction. Therefore, it would be ideal to interpret the MoS of synergistic drug combinations using the detailed signaling flows/interactions that are specific to cell lines and specific to drug combinations. We will investigate these challenges in future work. Moreover, considering the presence of various edge types in signaling pathway databases (e.g., KEGG), it is possible to incorporate edge information into graph neural network models to enable a more detailed analysis of the edges, considering that edge information has been incorporated into various graph AI framework[33,48–52]. In addition, the model is trained using a given set of synergistic scores of drug combinations, thus, currently, it is limited to predicting the synergy effects of new unseen drug pairs. When more cell lines and drug pairs (drug-targets) are available, the model can facilitate the discovery of new drug pairs by investigating the patterns of the drug-target and disease gene signaling in the learned MoS. In future work, we will further investigate these questions.

Methods

Datasets

Table 3 lists the datasets used in this study. Specifically, 4 drug combination synergistic score datasets were used to train and evaluate the model performance. The gene expression and copy number variation data of cancer cell lines were obtained from Cell Model Passports[53]. Signaling pathways were collected from KEGG[44] signaling pathways. Drug target information

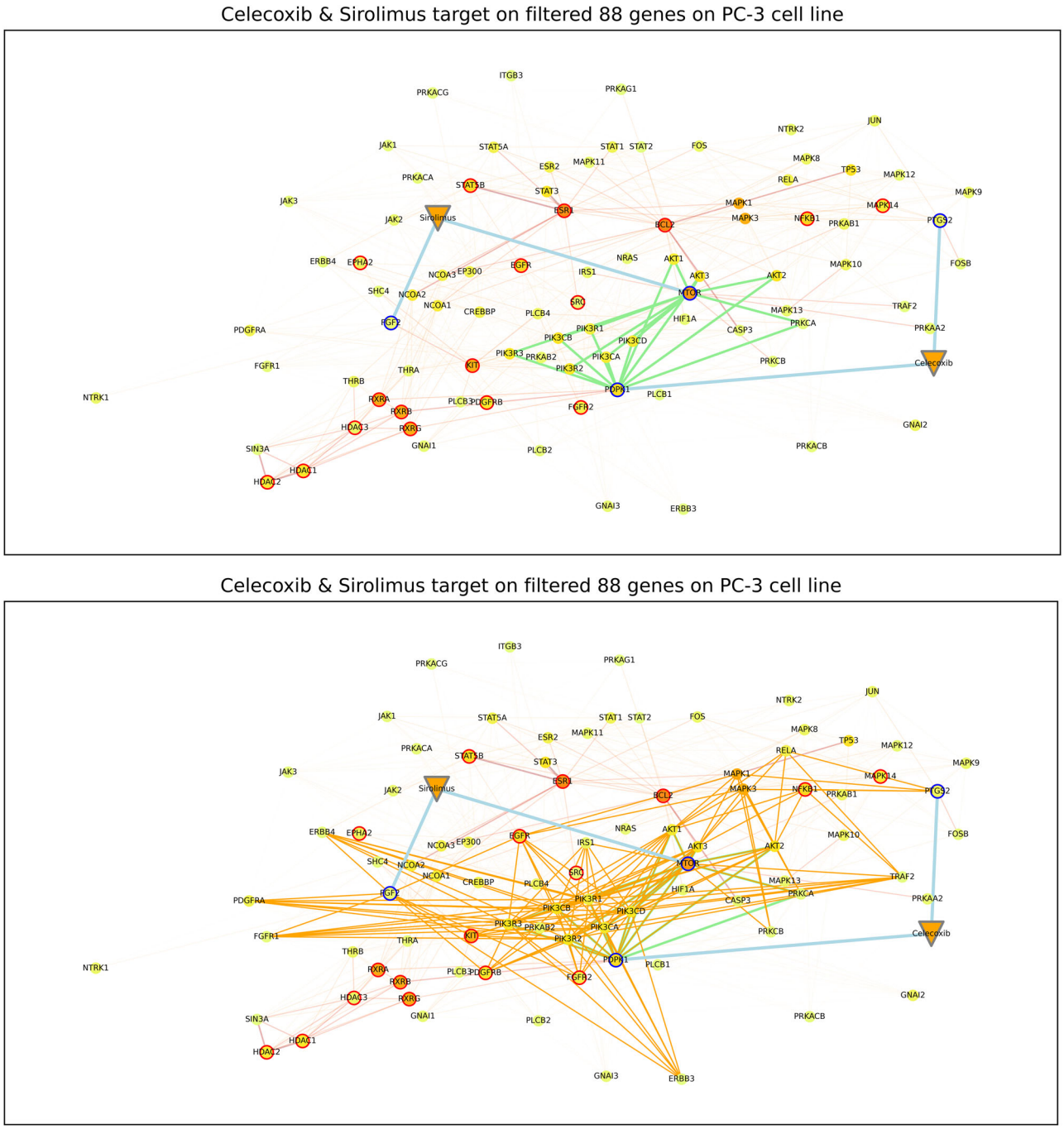


Fig. 4 | PC-3 cell line-specific effective combinations perturbing the core signaling subnetwork targeted by the drug Celecoxib and Sirolimus in the core signaling pathways. All nodes considered as important targets in *DeepSignalingFlow* are circled with red and nodes targeted by Tretionoin and Paclitaxel will be circled with blue. The upper panel shows the critical paths between 2 drugs within a distance of 4 (green edges) and the lower panel shows the paths within a distance of 5 (orange edges). In particular, the depth of color for nodes and edges indicates node degrees $d_i^{(PC-3)}$ and edge weights in matrix A'_{bind_final}.

was extracted from DrugBank[54]. The detailed information on the data and preprocessing results can be found in supplementary note.

Weighted bi-directional signaling flow construction

For the structural details of *DeepSignalingFlow*, the process begins with the mean aggregation of each node's neighboring features from upstream to downstream, accomplished as follows:

$$H_{up_neigh}^{(k)} = D_{in}^{-1}(A'^{(k)}_{up})^{\mathrm{T}} H^{(k-1)} M_{up}^{(k)} \quad (3)$$

The mean aggregation of each node's neighboring features from downstream to upstream is achieved by:

$$H_{down_neigh}^{(k)} = D_{out}^{-1}\left(A'^{(k)}_{down}\right)^{\mathrm{T}} H^{(k-1)} M_{down}^{(k)} \quad (4)$$

The biased transformation of nodes and their features is achieved by:

$$H_{self}^{(k)} = H^{(k-1)} B^{(k)} \quad (5)$$

Table 3 | Four sources and file types of databases leveraged in the deep-learning model

Database type	Database name	Website link
Drug Combination Effect Data	NCI ALMANAC[8]	https://wiki.nci.nih.gov/display/NCIDTPdata/NCI-ALMANAC
	O'Neil[9]	https://drugcomb.fimm.fi/
	DrugComb[10]	https://drugcomb.fimm.fi/
	DrugCombDB[11]	http://drugcombdb.denglab.org/main
Multi-omics Data	Cell Model Passports[53] RNA-Seq	https://cog.sanger.ac.uk/cmp/download/rnaseq_20191101.zip
	Cell Model Passports[53] CNV	https://cog.sanger.ac.uk/cmp/download/cnv_20191101.zip
Gene-Gene Interaction Data	KEGG[45] Gene Interactions Network	https://www.genome.jp/kegg/
Drug-Target Interaction Data	DrugBank[54]	https://go.drugbank.com/

For $= 1, 2, \ldots,$ K, the model uses normalization function $N: \mathbb{R}^{n\times 3d^{(k)}} \to \mathbb{R}^{n\times 3d^{(k)}}$ and LeakyReLU function with parameter α to map concatenated nodes features $H^{(k)}_{concat} = \left(H^{(k)}_{up_neigh} | H^{(k)}_{down_neigh} | H^{(k)}_{self}\right) \in \mathbb{R}^{n\times 3d^{(k)}}$ to $H^{(k)} \in \mathbb{R}^{n\times 3d^{(k)}}$ with the equation:

$$H^{(k)} = \text{LeakyReLU}(\text{N}(\text{H}^{(\text{k})}_{\text{concat}})), \tag{6}$$

where the normalization function N enable L2 normalization for demo matrix $V \in \mathbb{R}^{p\times q}$ along the row axis with the equation:

$$v'_{ij} = \frac{v_{ij}}{\sqrt{\sum_{j=1}^{q} v_{ij}^2}}, \tag{7}$$

where $v\prime_{ij}$ is the element of new matrix $V' \in \mathbb{R}^{p\times q}$. Therefore, the output of k-th layer is $H^{(k)}$. After obtaining the embedded nodes features with $H^{(K)} \in \mathbb{R}^{n\times 3d^{(K)}}$, the embedded nodes features of drug A are obtained as: $H^{(K)}_{drugA} \in \mathbb{R}^{1\times 3d^{(K)}}$, and the drug B embedded nodes features are obtained as: $H^{(K)}_{drugB} \in \mathbb{R}^{1\times 3d^{(K)}}$. Based on the idea of decagon decoder[53], the prediction of combo score will be calculated using the following equation:

$$g\left(H^{(K)}_{drugA}, H^{(K)}_{drugB}\right) = H^{(K)}_{drugA} DQD^{\mathrm{T}} \left(H^{(K)}_{drugB}\right)^{\mathrm{T}}, \tag{8}$$

where $D \in \mathbb{R}^{3d^{(K)}\times N}$ and $Q \in \mathbb{R}^{N\times N}$ are trainable decoder matrices.

Drug combination response interpretation using the bi-directional signaling flow

As aforementioned, the directional signaling interaction weights are added, to indicate the potential signaling flow from upstream signaling to drug targets and from drug targets to downstream signaling. The trainable directional weights can potentially indicate and interpret the signaling flow within the signaling network to affect the drug combination response. Specifically, the signaling flow derived from the directional weighted matrices is defined as follows:

Each model had K layers, and each layer had upstream weight adjacency parameters and downstream weight adjacency parameters, which are $A'^{(1)}_{up}, A'^{(2)}_{up}, \ldots, A'^{(k)}_{up}, \ldots, A'^{(K)}_{up}$ and $A'^{(1)}_{down}, A'^{(2)}_{down}, \ldots, A'^{(k)}_{down}, \ldots, A'^{(K)}_{down}$. Through aggregating K layers upstream and downstream to weight adjacency matrices, the model arrives at an aggregated upstream, downstream, and bind-stream for m th M-fold cross valuation model with m th model's $A'^{(1)}_{up}, A'^{(2)}_{up}, \ldots, A'^{(k)}_{up}, \ldots, A'^{(K)}_{up}$ and $A'^{(1)}_{down}, A'^{(2)}_{down}, \ldots, A'^{(k)}_{down}, \ldots, A'^{(K)}_{down}$ using the formula:

$$A'^{(m)}_{upagg} = \frac{1}{K}\sum_{k=1}^{K} A'^{(k)}_{up} \tag{9}$$

$$A'^{(m)}_{downagg} = \frac{1}{K}\sum_{k=1}^{K} A'^{(k)}_{down} \tag{10}$$

$$A'^{(m)}_{bindagg} = \frac{1}{K}\left[\frac{1}{2}\sum_{k=1}^{K}(A'^{(k)}_{up} + A'^{(k)}_{down})\right] \tag{11}$$

Using the upstream, downstream, and bi-directional weight adjacency matrix parameters, the weight adjacency matrices of those M models in M-fold cross-validation are aggregated by the following formula:

$$A'_{up_final} = \frac{1}{M}\sum_{m=1}^{M} A'^{(m)}_{upagg} \tag{12}$$

$$A'_{down_final} = \frac{1}{M}\sum_{m=1}^{M} A'^{(m)}_{downagg} \tag{13}$$

$$A'_{bind_final} = \frac{1}{M}\sum_{m=1}^{M} A'^{(m)}_{bindagg} \tag{14}$$

Leveraging the final upstream, downstream, and bind-stream weight adjacency matrices parameters for 2016 genes of $A'_{up_final}, A'_{down_final}, A'_{bind_final}$, the signaling flow in the network is analyzed.

Data availability

Publicly accessible data, as detailed in the Datasets section, were employed for this study.

Code availability

The implementation of the model, DeepSignalingFlow, is available at the GitHub repository: https://github.com/FuhaiLiAiLab/DeepSignalingFlow

Received: 13 February 2024; Accepted: 5 August 2024;
Published online: 21 August 2024

References

1. Tufail, M., Cui, J. & Wu, C. Breast cancer: molecular mechanisms of underlying resistance and therapeutic approaches. *Am. J. Cancer Res*. **12**, 2920–2949 (2022).
2. Merikhian, P., Eisavand, M. R., & Farahmand, L. Triple-negative breast cancer: understanding Wnt signaling in drug resistance. *Cancer Cell Int*. **21**, https://doi.org/10.1186/s12935-021-02107-3 (2021)
3. Zahedipour, F., Kesharwani, P. & Sahebkar, A. In *Aptamers Engineered Nanocarriers for Cancer Therapy*. 51–83 (Elsevier, 2023).
4. Shin, S. Y. & Nguyen, L. K. SynDISCO: a mechanistic modeling-based framework for predictive prioritization of synergistic drug

combinations directed at cell signaling networks. *Methods Mol. Biol*. **2634**, 357–381 (2023).
5. Bryant, K. L., Stalnecker, C. A., Zeitouni, D., Klomp, J. E., & Peng, S. et al. Combination of ERK and autophagy inhibition as a treatment approach for pancreatic cancer. *Nat Med*. https://doi.org/10.1038/s41591-019-0368-8 (2019).
6. Kinsey, C. G., Camolotto, S. A., Boespflug, A. M., Guillen, K. P., & Foth, M. et al. Protective autophagy elicited by RAF→MEK→ERK inhibition suggests a treatment strategy for RAS-driven cancers. *Nat. Med*. https://doi.org/10.1038/s41591-019-0367-9 (2019).
7. Jia, J. et al. Mechanisms of drug combinations: interaction and network perspectives. *Nat. Rev. Drug Discov.* **8**, 111–128 (2009).
8. Holbeck, S. L., Camalier, R., Crowell, J. A., Govindharajulu, J. P. & Hollingshead, M. et al. The National Cancer Institute ALMANAC: a comprehensive screening resource for the detection of anticancer drug pairs with enhanced therapeutic activity. *Cancer Res* **77**, 3564–3576 (2017).
9. O'Neil, J., Benita, Y., Feldman, I., Chenard, M. & Roberts, B. et al. An unbiased oncology compound screen to identify novel combination strategies. *Mol. Cancer Ther* **15**, 1155–1162 (2016).
10. Zagidullin, B., Aldahdooh, J., Zheng, S., Wang, W. & Wang, Y. et al. DrugComb: an integrative cancer drug combination data portal. *Nucleic Acids Res* **47**, W43–W51 (2019).
11. Liu, H. et al. DrugCombDB: a comprehensive database of drug combinations toward the discovery of combinatorial therapy. *Nucleic Acids Res.* **48**, D871–D881 (2020).
12. Louppe, G. Understanding Random Forests: from theory to practice. http://arxiv.org/abs/1407.7502 (2014).
13. Chen, T. & Guestrin, C. XGBoost: a scalable tree boosting system. https://doi.org/10.1145/2939672.2939785 (2016).
14. Sidorov, P., Naulaerts, S., Ariey-Bonnet, J., Pasquier, E. & Ballester, P. J. Predicting synergism of cancer drug combinations using NCI-ALMANAC data. *Front. Chem.* **7**, 1–13 (2019).
15. Weiss, A., Ding, X., van Beijnum, J. R., Wong, I. & Wong, T. J. et al. Rapid optimization of drug combinations for the optimal angiostatic treatment of cancer. *Angiogenesis* **18**, 233–244 (2015).
16. Nowak-Sliwinska, P., Weiss, A., Ding, X., Dyson, P. J. & van den Bergh, H. et al. Optimization of drug combinations using feedback system control. *Nat. Protoc.* **11**, 302–315 (2016).
17. Gilvary, C., Dry, J. R., & Elemento, O. Multi-task learning predicts drug combination synergy in cells and in the clinic. *bioRxiv*. 576017. https://doi.org/10.1101/576017 (2019).
18. Kuru, H. I, Tastan, O. & Cicek, A. E. MatchMaker: a deep learning framework for drug synergy prediction. https://doi.org/10.1101/2020.05.24.113241 (2020).
19. Preuer, K. et al. DeepSynergy: predicting anti-cancer drug synergy with deep learning. *Bioinformatics* **34**, 1538–1546 (2018).
20. Zhang, T., Zhang, L., Payne, P.R.O. & Li, F. Synergistic drug combination prediction by integrating multiomics data in deep learning models. In: *Translational Bioinformatics for Therapeutic Development. Methods in Molecular Biology*, (ed. Markowitz, J.) vol 2194. https://doi.org/10.1007/978-1-0716-0849-4_12 (Humana, New York, NY, 2021).
21. Zhang, H., Feng, J., Zeng, A., Payne, P, & Li, F. Predicting tumor cell response to synergistic drug combinations using a novel simplified deep learning model. *AMIA 2020 Annual Symposium*. https://doi.org/10.1101/2020.04.10.036491 (2020).
22. Feng, J., Zhang, H., & Li F. Investigate the relevance of major signaling pathways in cancer survival using a biologically meaningful deep learning model. *bioRxiv* https://doi.org/10.1101/2020.04.13.039487 (2020).
23. Bersanelli, M., Mosca, E., Remondini, D., Giampieri, E., & Sala, C. et al. Methods for the integration of multi-omics data: mathematical aspects. *BMC Bioinformatics* **17**. https://doi.org/10.1186/s12859-015-0857-9 (2016).
24. Ma, J., Yu, M. K., Fong, S., Ono, K. & Sage, E. et al. Using deep learning to model the hierarchical structure and function of a cell. *Nat. Methods* **15**, 290–298 (2018).
25. Kuenzi, B. M., Park, J., Fong, S. H., Sanchez, K. S., & Lee, J. et al. Predicting drug response and synergy using a deep learning model of human cancer cells. *Cancer Cell* **38**, 672–684.e6 (2020).
26. Rosen-Zvi, M., Griffiths, T., Steyvers, M. & Smyth P. *The Author-Topic Model for Authors and Documents*. Accessed September 4. https://arxiv.org/abs/1207.4169 (2020).
27. Kipf, T. N. & Welling, M. Semi-supervised classification with graph convolutional networks. *5th International Conference on Learning Representations, ICLR 2017—Conference Track Proceedings*. Published online:1–14. (2017).
28. Hamilton, W. L., Ying, R. & Leskovec, J. Inductive representation learning on large graphs. *Adv. Neural. Inf. Process. Syst*. 1025–1035. 2017 *31st Conference on Neural Information Processing Systems (NIPS 2017).* Long Beach, CA, USA (NIPS, 2017).
29. Veličković, P. et al. Graph attention networks. *6th International Conference on Learning Representations, ICLR 2018—Conference Track Proceedings*. 1–12. https://arxiv.org/abs/1710.10903 (2018).
30. Wang, T., Shao, W., Huang, Z., Tang, H., & Zhang, J. et al. MOGONET integrates multi-omics data using graph convolutional networks allowing patient classification and biomarker identification. *Nat. Commun.* **12**, 3445 (2021).
31. Li, X., Ma, J., Leng, L., Han, M., & Li, M. et al. MoGCN: a multi-omics integration method based on graph convolutional network for cancer subtype analysis. *Front. Genet.* **13**, 806842 (2022).
32. Gao, H. et al. A universal framework for single-cell multi-omics data integration with graph convolutional networks. *Brief. Bioinform.* **24**, bbad081 (2023).
33. Zhou, Y., Huo, H., Hou, Z., Bu, L., & Mao, J. et al. Co-embedding of edges and nodes with deep graph convolutional neural networks. *Sci. Rep.* **13**. https://doi.org/10.1038/s41598-023-44224-1 (2023).
34. Wang, J., Liu, X., Shen, S., Deng, L. & Liu. H. DeepDDS: Deep graph neural network with attention mechanism to predict synergistic drug combinations. *Brief Bioinform.* **23**. https://doi.org/10.1093/bib/bbab390 (2022).
35. Sanchez-Vega, F., Mina, M., Armenia, J., Chatila, W. K., & Luna, A. et al. Oncogenic signaling pathways in the cancer genome atlas. *Cell* **173**, 321–337.e10, https://doi.org/10.1016/j.cell.2018.03.035 (2018).
36. Gilmer, J., Schoenholz, S. S., Riley, P. F., Vinyals, O. & Dahl, G. E. Neural message passing for quantum chemistry. http://arxiv.org/abs/1704.01212 (2017).
37. Withnall, M., Lindelöf, E., Engkvist, O. & Chen H. Building attention and edge message passing neural networks for bioactivity and physical-chemical property prediction. *J. Cheminform*. **12**. https://doi.org/10.1186/s13321-019-0407-y (2020).
38. Garg, V. K., Jegelka, S. & Jaakkola T. Generalization and representational limits of graph neural networks. https://arxiv.org/abs/1710.10903 (2020).
39. Akkus, C., Chu, L., Djakovic, V., Jauch-Walser, S., & Koch, P. et al. Multimodal deep learning. http://arxiv.org/abs/2301.04856 (2023).
40. Baltrušaitis, T., Ahuja, C. & Morency, L. P. Multimodal machine learning: a survey and taxonomy. http://arxiv.org/abs/1705.09406 (2017).
41. Simonyan, K. & Zisserman, A. Two-stream convolutional networks for action recognition in videos. http://arxiv.org/abs/1406.2199, (2014).
42. Kaiser, L., Gomez, A. N., Shazeer, N., Vaswani, A., & Parmar, N. et al. One model to learn them all. http://arxiv.org/abs/1706.05137 (2017).
43. Vielzeuf, V., Lechervy, A., Pateux, S. & Jurie, F. CentralNet: a multilayer approach for multimodal fusion. http://arxiv.org/abs/1808.07275 (2018).
44. Kanehisa, M. G. S. & Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).

45. Shi, Y. et al. Masked label prediction: unified message passing model for semi-supervised classification. http://arxiv.org/abs/2009.03509 (2020)
46. Abu-El-Haija, S., Perozzi, B., Kapoor, A., Alipourfard, N., & Lerman, K. et al. Mixhop: Higher-order graph convolutional architectures via sparsified neighborhood mixing. In: *International Conference on Machine Learning,* 21–29 (PMLR, 2019).
47. Thekumparampil, K. K., Wang, C., Oh, S. & Li, L. J. Attention-based graph neural network for semi-supervised learning. http://arxiv.org/abs/1803.03735 (2018).
48. Corso, G., Cavalleri, L., Beaini, D., Liò, P. & Veličković P. Principal neighbourhood aggregation for graph nets. http://arxiv.org/abs/2004.05718 (2020)
49. Rong, Y., Huang, W., Xu, T. & Huang, J. DropEdge: towards deep graph convolutional networks on node classification. http://arxiv.org/abs/1907.10903 (2019).
50. Schlichtkrull, M. et al. Modeling relational data with graph convolutional networks. http://arxiv.org/abs/1703.06103 (2017).
51. Gong, L. & Cheng Q. Exploiting edge features for graph neural networks. https://arxiv.org/abs/1809.02709 (2019).
52. Goyal, P., Hosseinmardi, H., Ferrara, E., & Galstyan, A. Embedding networks with edge attributes. In: *HT 2018—Proceedings of the 29th ACM Conference on Hypertext and Social Media*. 38–42 https://doi.org/10.1145/3209542.3209571 (Association for Computing Machinery, Inc, 2018).
53. van der Meer, D., Barthorpe, S., Yang, W., Lightfoot, H., & Hall, C. et al. Cell Model Passports—a hub for clinical, genetic and functional datasets of preclinical cancer models. *Nucleic Acids Res.* **47**, D923–D929 (2019).
54. Wishart, D. S., Feunang, Y. D., Guo, A. C., Lo, E. J., & Marcu, A. et al. DrugBank 5.0: A major update to the DrugBank database for 2018. *Nucleic Acids Res.* **46**, D1074–D1082 (2018).

Acknowledgements

We thank Emily Chen for helping us with paper proofreading and editing. This study was partially supported by Children's Discovery Institute (CDI) M-II-2019-802 (to Li), and 1R01LM013902-01A1 (to Li).

Author contributions

H.Z. performed the computational modeling, data analysis, and drafted and revised the manuscript. H.Z., Y.C., P.P., and F.L. designed the computational method. F.L. conceived the study and drafted and revised the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41540-024-00421-w.

Correspondence and requests for materials should be addressed to Fuhai Li.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2024