M3NetFlow: a novel multi-scale multi-hop modular graph AI model for multi-omics data integration and signaling network inference

Heming Zhang[1], S. Peter Goedegebuure[2,3], Li Ding[3,4], Williams Hawkins[2,3], David DeNardo[3,4], Ryan C. Fields[2,3], Yixin Chen[6], Philip Payne[1], Fuhai Li[1,5#]

[1]Institute for Informatics (I2), [2]Department of Surgery, [3]Siteman Cancer Center, [4]Department of medicine, [5]Department of Pediatrics, Washington University School of Medicine, [6]Department of Computer Science and Engineering, Washington University in St. Louis, St. Louis, MO, USA.
[#]Correspondence: Fuhai.Li@wustl.edu

Abstract: The integration and interpretation of multi-omics data play a crucial role in systems biology for prioritizing key molecular targets and deciphering core signaling pathways of complex diseases, such as cancer, covid-19 and Alzheimer's disease. However, it remains a challenge that has not been adequately addressed. Graph neural networks (GNN) have been emerged as powerful artificial intelligence models for analyzing data with a graphical structure. Nevertheless, GNN models have not been sufficiently designed for integrative and interpretable analysis of multi-omics data. In this study, we propose a novel multi-scale multi-hop modular GNN model, M3NetFlow, to integrate and interpret multi-omics data to rank key targets and infer core signaling pathways. Specifically, we applied the M3NetFlow model to infer cell-line specific core signaling networks explaining drug combination response. The evaluation and comparison results on drug combination prediction showed that the M3NetFlow model achieved significantly higher prediction accuracy than existing GNN models. Furthermore, M3NetFlow can predict key targets and infer essential signaling networks regulating drug combination response. It is critical for investigating mechanisms of synergy, and guiding the development of personalized precision medicine for patients with drug resistance. This model can be applied to general multi-omics data-driven research.

1. Introduction

In diseases and treatments with complex natures, such as cancer and Alzheimer's disease (AD), it is widely accepted that intricate and fundamental molecular signaling networks play a significant role in disease progression and the emergence of drug resistance. To gain a comprehensive understanding of the core disease signaling pathways and drug responses, extensive multi-omics datasets encompassing genetic, epigenetic, transcriptomic, and proteomic information have been generate[1,2,3,4,5]. Despite the identification of numerous novel molecular targets and medications through the analysis of multi-omics data, our comprehension of the fundamental signaling pathways in complex diseases is still constrained. The integration and analysis of multi-omics data to uncover the crucial molecular targets and core signaling pathways continue to be an ongoing and arduous endeavor[6–8].

The current state of machine learning models and computational approaches has proven inadequate in terms of delivering sufficient interpretation regarding the mechanisms that drive synergy and accurately identifying the primary drug targets. Graph neural network (GNN) models have emerged as a potent framework for analyzing drug-gene interaction networks. However, traditional GNN models are limited by the issue of over-smoothing, which hampers their ability to incorporate a large number of layers. Consequently, these models can only consider a limited amount of multi-hop/N-hop information (N usually smaller than 4). The over-smoothing problem was firstly mentioned by the extending the propagation layers in GCN[9]. What's more, theoretical paper using Dirichlet energy showed the diminished discriminative power by increasing the propagation layers[10].

In this research, we introduce a novel solution called ***M3NetFlow (Multi-scale, Multi-Hop, Multi-omics NETwork Flow inference)*** to address the challenges mentioned earlier. Our approach tackles these challenges by first incorporating local K-hop information within each subgraph using an attention mechanism. Subsequently, we establish global-level bi-directional propagation to facilitate message passing between genes/proteins. In the interpretation phase, we leverage the attention mechanism to aggregate the weights of all connected paths for genes, and a reweighting process inspired by TF-IDF is employed to redistribute the importance scores of genes across different cell lines.

Afterwards, a visualization tool ***NetFlowVis*** was developed for the better analysis of mechanism of synergy and identify the important genes. To assess and demonstrate the effectiveness of our proposed model, ***M3NetFlow***, we applied it to extract cell-line specific signaling networks

that explain the synergy scores of drug combinations by integrating multi-omics data. Comparative results indicated the superior performance of the ***M3NetFlow*** model compared to state-of-the-art models. Moreover, the ***M3NetFlow*** model has the capability to infer personalized signaling targets and signaling networks, enabling the inference of key targets and core signaling networks, and guiding the advancement of personalized precision medicine.

2. Methodology

Table 1 provided the links from which we collected data on drug combination effects, multi-omics data, gene-gene interactions, and drug-gene interactions. These data were preprocessed and represented in a diagram depicted in **Figure 1**.

Table 1. 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[11]	https://wiki.nci.nih.gov/display/NCIDTPdata/NCI-ALMANAC
Multi-omics Data	Cell Model Passports[12] RNA-Seq	https://cog.sanger.ac.uk/cmp/download/rnaseq_20191101.zip
	Cell Model Passports[12] CNV	https://cog.sanger.ac.uk/cmp/download/cnv_20191101.zip
	CCLE[13] Methylation	https://data.broadinstitute.org/ccle/CCLE_RRBS_TSS1kb_20181022.txt.gz
	CCLE[13] Gene Amplification	https://data.broadinstitute.org/ccle_legacy_data/binary_calls_for_copy_number_and_mutation_data/CCLE_MUT_CNA_AMP_DEL_binary_Revealer.gct
	CCLE[13] Gene Deletion	https://data.broadinstitute.org/ccle_legacy_data/binary_calls_for_copy_number_and_mutation_data/CCLE_MUT_CNA_AMP_DEL_binary_Revealer.gct
Gene-Gene Interaction Data	KEGG[14] Gene Interactions Network	https://www.genome.jp/kegg/
Drug-Target Interaction Data	DrugBank[15]	https://go.drugbank.com/

2.1 Drug combination screening datasets

In this paper, we utilized the NCI ALMANAC and O'Neil[16] drug screening datasets to train the deep learning models. The NCI ALMANAC dataset consists of combo scores for various permutations of 104 FDA-approved drugs, representing their impact on tumor growth in NCI60

human tumor cell lines. For evaluating the synergy score of two drugs on a specific tumor cell line, we utilized the average combo-score of two drugs at different doses, employing a 4-element tuple: $<D_A, D_B, C_C, S_{ABC}>$. On the other hand, the O'Neil dataset, obtained from the DrugComb[17] platform, served as another drug combination screening dataset. We extracted the processed dataset from this platform, which provided average synergy Loewe scores. These scores were used to assess the synergy score of two drugs on a given tumor cell line, employing a 4-element tuple: $<D_A, D_B, C_C, S_{ABC}>$.

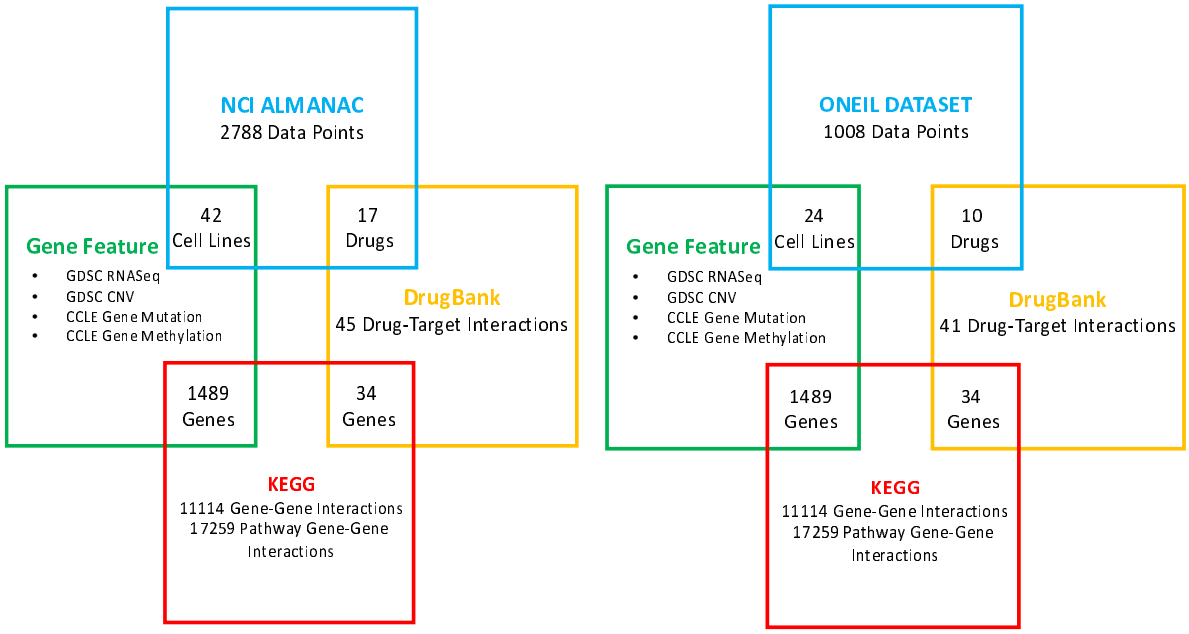

Figure 1. Preprocessed data of raw files collected from various databases

2.2 Multi-omics data of cancer cell lines

In this study, multi-omics data comprising RNA-seq, copy number variation, gene methylation, and gene mutation data for a total of 1,489 genes were incorporated into the model. These data were obtained from the Cell Model Passports[12] and CCLE[13] databases. Through the identification of overlapping cell lines from these databases, we identified 42 cell lines that were present in both the Cell Model Passports database and the NCI ALMANAC dataset (refer to **Table 2**). Additionally, we found 24 cell lines that were present in both the Cell Model Passports database and the O'Neil dataset (refer to **Table 3**).

Table 2. 42 cell lines in preprocessed NCI ALMANAC and Multi-omics datasets

A498	A549/ATCC	ACHN	BT-549	CAKI-1
DU-145	EKVX	HCT-116	HCT-15	HOP-62
HOP-92	HS 578T	IGROV1	K-562	KM12
LOX IMVI	MCF7	MDA-MB-231/ATCC	MDA-MB-468	NCI-H23
NCI-H460	NCI-H522	OVCAR-3	OVCAR-4	OVCAR-8
PC-3	RPMI-8226	SF-268	SF-295	SF-539
SK-MEL-2	SK-MEL-28	SK-MEL-5	SK-OV-3	SNB-75
SR	SW-620	T-47D	U251	UACC-257
UACC-62	UO-31			

Table 3: 24 human tumor cell lines included in Cell Model Passports database intersected with ONEIL

A2058	A2780	A375	CAOV3	HCT116
HT144	LOVO	MDAMB436	NCI-H460	NCIH1650
NCIH2122	NCIH23	OV90	OVCAR3	RKO
RPMI7951	SK-OV-3	SKMEL30	SKMES1	SW-620
SW837	T-47D	UACC62	VCAP	

2.3 KEGG database

About 59,241 gene-gene interactions for over 8,000 genes across different signaling pathways were collected from KEGG[14] database. And 48 signaling pathways in the KEGG dataset were selected as the ground truth of subgraphs in the gene-gene interactions, which were AGE-RAGE signaling pathway in diabetic complications, AMPK signaling pathway, Adipocytokine signaling pathway, Apelin signaling pathway, B cell receptor signaling pathway, C-type lectin receptor signaling pathway, Calcium signaling pathway, Chemokine signaling pathway, ErbB signaling pathway, Estrogen signaling pathway, Fc epsilon RI signaling pathway, FoxO signaling pathway, Glucagon signaling pathway, GnRH signaling pathway, HIF-1 signaling pathway, Hedgehog signaling pathway, Hippo signaling pathway, Hippo signaling pathway - multiple species, IL-17 signaling pathway, Insulin signaling pathway, JAK-STAT signaling pathway, MAPK signaling pathway, NF-kappa B signaling pathway, NOD-like receptor signaling pathway, Neurotrophin signaling pathway, Notch signaling pathway, Oxytocin signaling pathway, PI3K-Akt signaling pathway, PPAR signaling pathway, Phospholipase D signaling pathway,

Prolactin signaling pathway, RIG-I-like receptor signaling pathway, Rap1 signaling pathway, Ras signaling pathway, Relaxin signaling pathway, Signaling pathways regulating pluripotency of stem cells, Sphingolipid signaling pathway, T cell receptor signaling pathway, TGF-beta signaling pathway, TNF signaling pathway, Thyroid hormone signaling pathway, Toll-like receptor signaling pathway, VEGF signaling pathway, Wnt signaling pathway, cAMP signaling pathway, cGMP-PKG signaling pathway, mTOR signaling pathway, p53 signaling pathway. After preprocessing dataset, 17,259 gene-gene interactions for 1,489 genes were obtained.

2.4 Drug-Target interactions from DrugBank database

Drug-target information was extracted from the DrugBank[15] database (version 5.1.5, released 2020-01-03). In total, 15,263 drug-target interactions were obtained for 5435 drugs/investigational agents and 2775 targets. Further, 17 drugs with known targets to the 1489 genes previously identified were selected for use in our model for NCI ALMANAC[11], specifically: Celecoxib, Cladribine, Dasatinib, Docetaxel, Everolimus, Fulvestrant, Gefitinib, Lenalidomide, Megestrol acetate, Mitotane, Nilotinib, Paclitaxel, Romidepsin, Sirolimus, Thalidomide, Tretinoin, Vorinostat. And 10 drugs with known targets to the 1489 genes previously identified were selected for use in our model, specifically: Dasatinib, Erlotinib, Lapatinib, Sorafenib, Sunitinib, Vorinostat, geldanamycin, metformin, paclitaxel, vinblastine.

2.5 Architecture design of M3NetFlow

Figure 2 shows the schematic architecture of the proposed ***M3NetFlow*** model. The model input parameters are: $\mathcal{X} = \{(X^{(1)}, T^{(1)}),\ (X^{(2)}, T^{(2)}), \dots, (X^{(m)}, T^{(m)}), \dots, (X^{(M)}, T^{(M)})\}\ (X^{(m)} \in \mathbb{R}^{n\times d}, T \in \mathbb{R}^{n\times 2}), A \in \mathbb{R}^{n\times n}, S = \{S_1,\ S_2, \dots, S_p, \dots, S_P\}(S_p \in \mathbb{R}^{n_p\times n_p}),\ D_{in} \in \mathbb{R}^{n\times n},\ D_{out} \in \mathbb{R}^{n\times n}$, where M represents number of data points in drug screening dataset. To predict the drug combination effect scores Y $(Y \in \mathbb{R}^{M\times 1})$, we would like to build up the machine learning model $f(\cdot)$ with $f(\mathcal{X}, A, S, D_{in}, D_{out}) = Y$, where $\mathcal{X}$ denotes all of the data points in the dataset and $(X^{(m)}, T^{(m)})$ is m-th data points in the dataset, where $X^{(m)}$ denotes the node features matrix with n nodes of d features and $T^{(m)}$ denotes the one-hot encoded of the number of drugs targeted on those n nodes. The matrix A is the adjacency matrix that demonstrates the node-node interactions , and the element in adjacency matrix A such as a_{ij} indicates an edge from i to j, and $\mathcal{S}$ is a set of subgraphs which partition the whole graph adjacent matrix A into multiple subgraphs with

$S_p \in \mathbb{R}^{n_p \times n_p}$ of nodes interactions between its internal n_p nodes and each subgraph has its own corresponding subgraph node feature matrix $X_p \in \mathbb{R}^{n_p \times d}$. 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 this certain model, we have 1,489 genes plus the 17 drugs, with $n = 1506$ nodes in NCI ALMANAC dataset and we have 1,489 genes plus the 10 drugs with $n = 1499$ nodes in O'Neil data set, where each node has $d = 8$ features.

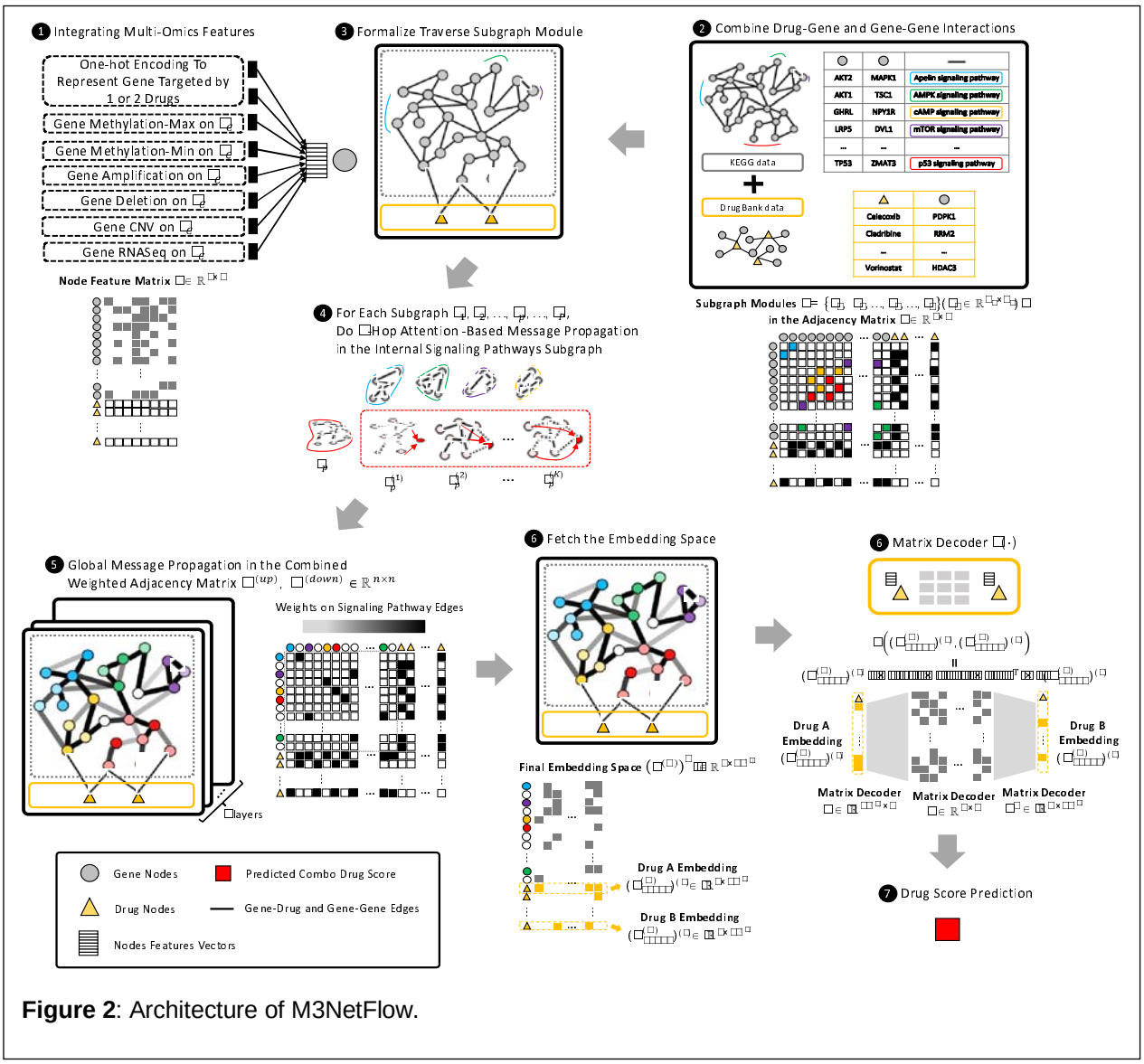

Figure 2: Architecture of M3NetFlow.

Multi-hop Subgraph Message Propagation In the graph message passing stages of our architecture (see **Figure 2** step3 and step4), we introduced the subgraph module message passing stage for each signaling pathway prior to the global message passing stage, which will ensure the message fully interacted in the internal subgraph. We opted for a K-hop attention-

based graph neural network because it allows for the consideration of longer distance information, which can be interpreted as signaling flow in a single message propagation layer. With those initial embedding features $X^{(m)} \in \mathbb{R}^{n\times d}, (m = 1, 2, \dots, M)$, the K-hop attention-based graph neural network was built to incorporate the long-distance information for each subgraph with

$$\left(\alpha_{ij}^{(m)}\right)_p^{(k)} = \frac{\exp\left(\mathrm{LeakyReLU}(a[W\left(X_p^{(m)}\right)_i || W\left(X_p^{(m)}\right)_j])\right)}{\sum_{q\in\mathcal{N}_i^{(k)}} \exp\left(\mathrm{LeakyReLU}(a[W\left(X_p^{(m)}\right)_i || W\left(X_p^{(m)}\right)_q])\right)} \tag{1}$$

$$(H_p^{(m)})_i = \frac{1}{KQ_i}\sum_{k}^{K}\sum_{q\in\mathcal{N}_i^{(k)}} [\left(\alpha_{iq}^{(m)}\right)_p^{(k)} W\left(X_p^{(m)}\right)_q] \tag{2}$$

, where $\left(\alpha_{ij}^{(m)}\right)_p^{(k)}$ is the attention score between node i and node j in the k-th kop of the subgraph S_p for the data point m and $\mathcal{N}_i^{(k)}$ is the neighbor nodes calculated by the adjacency matrix in k-th hop $A^{(k)}$ for the node i (See **Appendix A** for the algorithm of k-th hop adjacency matrix). The linear transformation vector $a \in \mathbb{R}^{2d'}$ was also defined. At the same time, the linear transformation for features of each node will be defined as $W \in \mathbb{R}^{d\times d'}$. And $\left(X_p^{(m)}\right)_t, \left(H_p^{(m)}\right)_t$ represent the feature and updated feature of node t $(t = 1, 2, \dots, n)$. Aside from that, Q_t represents the number of signaling pathway the node t $(t = 1, 2, \dots, n)$ belongs to. And above formula showed the calculation for the 1-head attention. The number of head h will be modified in the model, and the node embeddings for each subgraph will take the average of embeddings in every head attention.

Global Bi-directional Message Propagation Following the message propagation in the multiple internal subgraphs, the global weighted bi-directional message propagation will be performed, where nodes-flow contain both 'upstream-to-downstream' (from up-stream signaling to drug targets) and 'downstream-to-upstream' (from drug targets to down-stream signaling) (see **Figure 2** step 5). Before the global level message propagation, the node feature for data point m in each subgraph $H_p^{(m)}$ $(p = 1, 2, \dots, P)$ will be combined into a new unified node features matrix $\left(H^{(m)}\right)^{(0)}$ as the initial node features with

$$\left(H^{(m)}\right)^{(0)}{}_i = \frac{1}{\sum_{p=1}^{P} I[\mathcal{V}_i \in S_p]}\sum_{p=1}^{P}\left(H_p^{(m)}\right)_i \cdot I[\mathcal{V}_i \in S_p] \tag{3}$$

, where V_i represents the node/vertex i in the graph and $I[V_i \in S_p]$ is the indicator function, whose value will be one if $\mathcal{V}_i \in S_p$. In the end, the initial node features for global level propagation will be $\left(H^{(m)}\right)^{(0)} \in \mathbb{R}^{n \times d^{(0)}}, d^{(0)} = d' + d$. The weight matrices are denoted as: $W_{up}^{(l)} \in \mathbb{R}^{n \times n},\ W_{down}^{(l)} \in \mathbb{R}^{n \times n},\ (l = 1, 2, \dots, L)$ respectively for L global message propagation layers. And then the weight adjacency matrices will be denoted as:

$$A'^{(l)}_{up} = W_{up}^{(l)} \cdot A \ \#(4)$$

$$A'^{(l)}_{down} = W_{down}^{(l)} \cdot A^{\mathrm{T}} \ \#(5)$$

, where for each $A'^{(l)}_{up} \in \mathbb{R}^{n \times n},\ A'^{(l)}_{down} \in \mathbb{R}^{n \times n},\ (l = 1,2, \dots, L)$, parameters for drug-gene relations values are the same with those in original adjacency matrix. Afterwards, the mean aggregation of each node neighbors' features from upstream to downstream is achieved by:

$$\left(H^{(m)}_{up_neigh}\right)^{(l)} = D_{in}^{-1}(A'^{(l)}_{up})^{\mathrm{T}} \left(H^{(m)}\right)^{(l-1)} Z_{up}^{(l)} \ \#(6)$$

And the mean aggregation of each node neighbors' features from downstream to upstream is achieved by:

$$\left(H^{(m)}_{down_neigh}\right)^{(l)} = D_{out}^{-1}(A'^{(l)}_{down})^{\mathrm{T}} \left(H^{(m)}\right)^{(l-1)} Z_{down}^{(l)} \ \#(7)$$

And biased transformation of nodes themselves and their features is achieved by:

$$\left(H^{(m)}_{self}\right)^{(l)} = \left(H^{(m)}\right)^{(l-1)} B^{(l)} \ \#(8)$$

, where the $Z_{up}^{(l)} \in \mathbb{R}^{d^{(l-1)} \times d^{(l)}}, Z_{down}^{(k)} \in \mathbb{R}^{d^{(l-1)} \times d^{(l)}}$ are the linear transformation matrix for each layer l, $(l = 1, 2, \dots, L)$. For $l = 1, 2, \dots, L$, the model used normalization function $N: \mathbb{R}^{n \times 3d^{(l)}} \rightarrow \mathbb{R}^{n \times 3d^{(l)}}$ and LeakyReLU function with parameter α to map concatenated nodes features $\left(H^{(m)}_{concat}\right)^{(l)} = \left(\left(H^{(m)}_{up_neigh}\right)^{(l)} \mid \left(H^{(m)}_{down_neigh}\right)^{(l)} \mid \left(H^{(m)}_{self}\right)^{(l)}\right) \in \mathbb{R}^{n \times 3d^{(l)}}$ to $H^{(l)} \in \mathbb{R}^{n \times 3d^{(l)}}$ with the equation:

$$(H^{(m)})^{(l)} = \mathrm{LeakyReLU}(N(\left(H^{(m)}_{concat}\right)^{(l)}))\#(9)$$

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

$$v'_{ij} = \frac{v_{ij}}{\sqrt{\sum_{j=1}^{q} v_{ij}^2}} \#(10)$$

where v'_{ij} is the element of new matrix $V' \in \mathbb{R}^{p \times q}$. Therefore, the output of l-th layer for global message propagation is $(H^{(m)})^{(l)}$. After obtaining the embedded nodes features with $(H^{(m)})^{(L)} \in \mathbb{R}^{n \times 3d^{(L)}}$, the embedded nodes features of drug_A are obtained as: $(H_{drugA}^{(m)})^{(L)} \in \mathbb{R}^{1 \times 3d^{(L)}}$, and the drug_B embedded nodes features are obtained as: $(H_{drugB}^{(m)})^{(L)} \in \mathbb{R}^{1 \times 3d^{(L)}}$. Based on the idea of decagon decoder[18], the prediction of combo score will be calculated in following equation:

$$g\left((H_{drugA}^{(m)})^{(L)}, (H_{drugB}^{(m)})^{(L)}\right) = (H_{drugA}^{(m)})^{(L)} D U D^{\mathrm{T}} ((H_{drugB}^{(m)})^{(L)})^{\mathrm{T}} \#(11)$$

where $D \in \mathbb{R}^{3d^{(L)} \times E}$ and $U \in \mathbb{R}^{E \times E}$ are trainable decoder matrices.

2.6 Drug combination response interpretation using the subgraph attention

As aforementioned, we added the K-hop subgraph message propagation with attention in each subgraph, and the attentions or weights on each edge can potentially indicate and interpret the signaling flow on the signaling network to affect the drug combination response. Specifically, putting the attentions/weights on each edge will use the trained parameters $a \in \mathbb{R}^{2d'}$ and $W \in \mathbb{R}^{d \times d'}$. And for cell lines set, the set $\mathcal{C} = \{C_1, C_2, \dots, C_r, \dots, C_R\}$ contains $R = 39$ cell lines, and we can choose the input node feature matrix within a specific cell line to form the cell-line specific attention matrix in k-hop by

$$(\alpha_{ij}^{(C_r)})^{(k)} = \sum_{p=1}^{P} \left(\alpha_{ij}^{(m)}\right)_p^{(k)}, \left(\forall X^{(m)} \in C_r\right) \quad (12)$$

, where $(\alpha_{ij}^{(C_r)})^{(k)}$ will be the element in $\left(\mathcal{A}^{(C_r)}\right)^{(k)}$, the k-hop attention matrix for cell line C_r. And above formula showed the calculation for 1-head attention matrix. For multi-head attention mechanism, the averaged value will be used for the attention matrix.

3. Results

3.1 Experimental Setup

In order to validate the proposed model, a 5-fold cross-validation was used. Specifically, <drug_A, drug_B, cell_line name, combo score>, a 4-element vector, represents a model input data point. In total, there are 2788 model input data points for NCI ALMANAC dataset and 1008 model input data points for O'Neil dataset. In each fold model, 4 folds of the data points will be used as the training dataset, and the rest 1-fold will be used as the testing dataset. For those 1489 genes, their 8 features indicate the RNA sequence number, copy number variation, gene amplification, gene deletion, gene methylation maximum and minimum value and whether they have connection spanning drug_A and drug_B respectively. For each 2 drug pair, their 8 features were initiated with zeros. Furthermore, to indicate connections between nodes, adjacency matrices were also created. Those matrices were formed from the KEGG database, which contains gene pairs with sources and destinations. In this way, the graphs generated were directed graphs**.** For drug-gene edges, the connections are bidirectional, which means that in adjacency matrices, those elements are symmetric.

3.2 Hyperparameters

Subsequently, a model was developed by using pytorch and torch geomteric. Using the Adam optimizer, learning rate started with 0.002 and reduced equally within each batch for certain epoch stage. For epochs range [1, 10], [11, 20], [21, 30], [31, 60], the learning rate reduced from [0.002, 0.0008], [0.0008, 0.0006], [0.0006, 0.0005], [0.0005, 0.0001] separately in each range. And the epochs after 60 will keep the learning rate at 0.0001. Adam optimizer was chosen for optimization with eps=1e-7 and weight_decay=1e-20. We empirically set ***K-hop Subgraph Message Propagation*** part with $K = 3$ and ***Global Bi-directional Message Propagation*** part with $L = 3$, and the initial input feature dimension was: $d = 8$, $d' = 1$ with setting number of head $h = 3$. Afterwards, the feature dimensions will vary at the different layers and will be denoted by: $(d^{(1)}, d^{(2)}, \ldots, d^{(l)}, \ldots, d^{(L)})$. At the global message propagation part, the layer will concatenate biased node features and transformed nodes features of previous layer for both upstream and down-stream, generating concatenated dimensions for the output dims being $3 \times d^{(l)}$ in l-th layer. Output dims of the previous layer served as the input dims for 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 nodes dims were $3d^{(3)}, (L = 3)$. Hence, for dimensions of node embedding in each layer, we have $d^{(0)} = 8,\ d^{(1)} = 8, d^{(2)} = 8, d^{(3)} = 24$ and decoder trainable transformation matrix dims: $D \in \mathbb{R}^{3d^{(L)} \times E}$ and $U \in \mathbb{R}^{E \times E}$ were used as trainable decoder matrices, with changeable parameters of E to adapt model performance and the model used $E = 150$. As for the LeakyReLU function, the parameter $\alpha = 0.1$.

3.3 M3NetFlow significantly improves drug combination prediction.

To evaluate the model performance in terms of synergy score prediction for drug combinations, we conducted 5-fold cross validation. As shown in **Table 4**, the average prediction (using the Pearson correlation coefficient), was about 61% Pearson correlation using the test data in NCI ALMANAC dataset and was about 64% Pearson correlation using the test data in O'Neil dataset. These prediction results are comparable with existing deep learning models[19,20]. **Figure S1-S5** shows the scatterplots and boxplots of the predicted and experimental synergy scores on the 5 training and test datasets. The results indicated the feasibility of drug combination synergy score prediction using a graph neural network.

Table 4. 5-fold Cross Validation of proposed M3NetFlow model in NCI ALMANAC dataset

Dataset	Training MSE Loss	Training Pearson Correlation	Test MSE Loss	Test Pearson Correlation
1st dataset	24.910785	78.8935%	42.420434	60.0966%
2nd dataset	34.32198	69.0788%	42.366917	60.1075%
3rd dataset	32.42659	71.7287%	38.854818	61.7941%
4th dataset	34.696549	69.2506%	40.94755	60.0938%
5th dataset	31.875111	70.3966%	47.63544	61.5159%
Average	**31.646203**	**71.86964%**	**42.4450318**	**60.72158%**

Table 5. 5-fold Cross Validation of proposed M3NetFlow model in O'Neil dataset

Dataset	Training MSE Loss	Training Pearson Correlation	Test MSE Loss	Test Pearson Correlation
1st dataset	168.836656	68.2072%	134.640407	62.9468%
2nd dataset	144.522245	73.3936%	144.548363	63.8356%
3rd dataset	69.703043	85.8798%	283.912207	61.6114%
4th dataset	87.725931	83.6393%	180.135809	69.0636%
5th dataset	151.858177	71.0429%	183.728567	64.3278%
Average	**124.52921**	**76.4326%**	**185.393071**	**64.3571%**

Moreover, we also compared our proposed model M3NetFlow with other deep learning models (see **Table 6** and **Figure 3**), which included the GCN[21] and Graph Attention network[22] (GAT). Based on this GCN and GAT model, both models were implemented with the decoder matrices to predict the drug effects, which improved the performance of both models, named as GCN Decoder and GAT Decoder.

Table 6. Comparisons between GCN Decoder, GAT Decoder and M3NetFlow of average MSE loss and Pearson correlation of 5-fold cross-validation in NCI ALMANAC dataset.

Model	Training MSE Loss	Training Pearson Correlation	Test MSE Loss	Test Pearson Correlation
GCN Decoder	43.0061078	**58.9086%**	48.5379624	**51.93114%**
GAT Decoder	43.1162382	**58.65168%**	50.9773424	**49.1606%**
M3NetFlow	31.646203	**71.86964%**	42.4450318	**60.72158%**

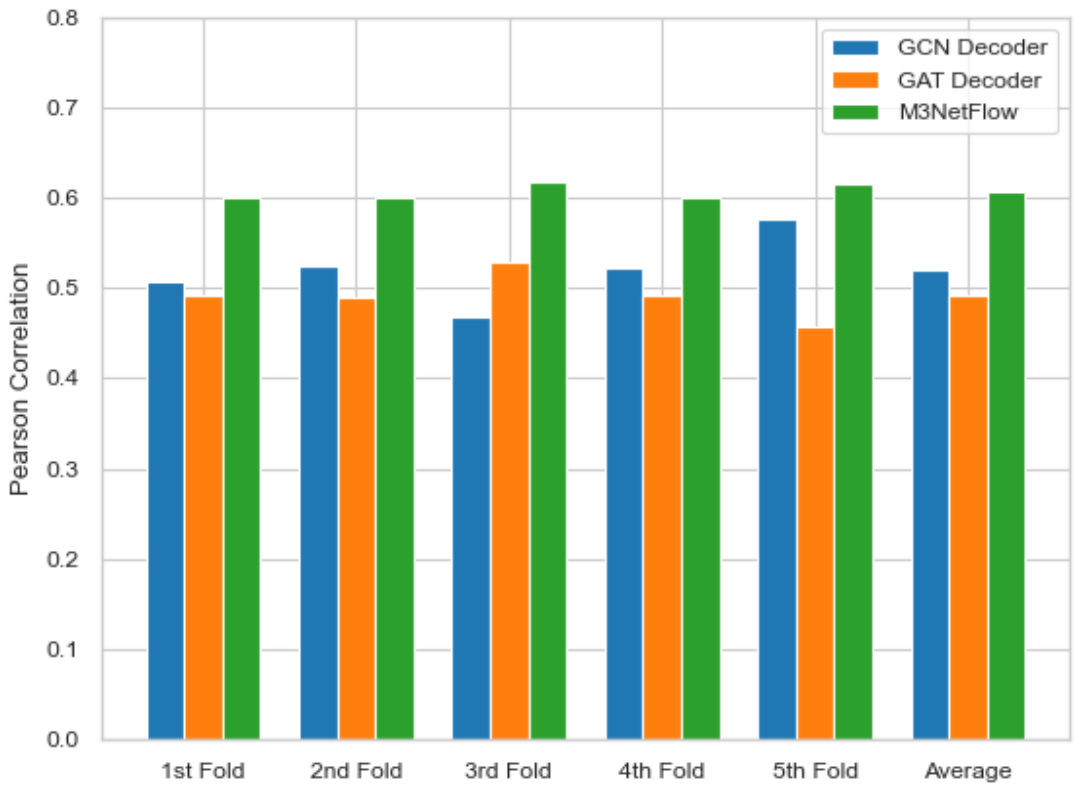

Figure 3. Pearson correlation comparisons for GCN Decoder, GAT Decoder and M3NetFlow models in each split of 5-fold cross-validation for NCI ALMANAC dataset

3.4 M3NetFlow interprets mechanism of synergy of drug combinations using signaling flow.

To interpret the mechanism of synergy, the top or most important genes in each cell line were selected as the potential drug targets from NCI ALMANAC[11] dataset. Based on the attention matrix of each cell line from 1-hop neighbor nodes in each fold of cross validation, the average attention matrix in 1-hop was calculated. Therefore, the weighted degree of each node of each cell line will be calculated based on the attention with

, where is the element of 1-hop averaged 5-fold attention matrix from cell line in the -th row and -th column and is the node degree (importance score) for gene in the

cell line C_r. Combining all the node degree vector from each cell line, the matrix $\mathcal{D} \in \mathbb{R}^{n \times R}$ will be generated. However, some of genes may show the importance with relatively higher scores in every cell line, which weaken the analysis of the cell line specific drug targets. In this way, the idea of reweighting the gene importance score was created based on the TF-IDF to reweight the node degree in each cell line with

$$\mathcal{W}_g = Log\left(\frac{R+1}{S_g + 0.01}\right) \tag{14}$$

$$\mathcal{D}'_g = \mathcal{W}_g \cdot \mathcal{D}_g \tag{15}$$

, where S_g is the number of cell lines has higher node degree than the threshold $\mathcal{S}$ in all 39 cell lines. Here, 95 percentile of node degree values in whole matrix $\mathcal{D}$ was used for $\mathcal{S}$. Finally, the reweighted node degree (importance score) matrix $\mathcal{D}'$ will be generated.

Based on the reweighted matrix $\mathcal{D}'$, the node degree (gene importance score) will be defined in each cell line. To validation our conclusions on the gene importance generated from the proposed model and reweighting process borrowed from TF-IDF, following statistical analysis was made. For each cell line, the highest 5, lowest 5 drug scores and their corresponding drugs and drug targeted genes were collected. Comparing the targeted genes by drugs from highest 5 drug scores and lowest drug scores, the differences are statistically significant with p values in more than half cell lines. Since SK-MEL-2 cell line only have 3 drug screen data points and all of them are smaller than 0, this cell line was excluded from the statistical analysis. In total, there are 27 out of 41 (~65%) cell lines have p values smaller than 0.1 (filled with deep orange color in **Table 6**) and 32 out of 41 (~78%) cell lines have p value smaller than 0.3 (filled with light orange color in **Table 7**). And 37 out of 41 (~90%) cell lines have higher gene importance scores on both mean and median value (filled with light green in **Table 7**).

Table 7. Comparisons of mean and median importance scores of the targeted genes by drugs from highest 5 drug scores and lowest 5 drug scores

Cancer Name	Cell Line Name	T-test P-value	Top 5 Mean	Low 5 Mean	Top 5 Median	Low 5 Median
Breast	BT-549	0.1891838	9.89209091	6.87665385	9.7295	4.7625
Breast	HS 578T	0.00037475	10.9233846	2.91441935	10.9285	1.015
Breast	MCF7	0.00633313	8.41165	3.779625	9.7145	2.987
Breast	MDA-MB-231/ATCC	0.01164589	9.25486364	4.65006452	11.1535	2.446
Breast	MDA-MB-468	0.00856875	10.1535455	3.92812903	10.719	1.456

Breast	T-47D	0.0412549	9.1465	4.10940909	8.999	1.2975
CNS	SF-268	0.26470404	9.00357143	6.45484615	9.976	2.993
CNS	SF-295	0.06452436	4.5337931	2.24575	3.38	0.638
CNS	SF-539	0.00260949	11.5421739	3.59255556	13.699	1.1785
CNS	SNB-75	0.00426485	10.9184783	4.57329412	11.708	1.7435
CNS	U251	0.96860118	5.05774286	5.115625	1.782	3.32
Colon	HCT-116	0.01060546	11.0535909	4.86373684	12.072	0.968
Colon	HCT-15	0.68943374	6.64127778	5.99260465	6.0465	4.337
Colon	KM12	0.01962015	10.0042381	4.00055556	6.732	1.78
Colon	SW-620	0.09481587	11.952	7.60654839	11.832	6.941
Leukemia	K-562	0.05677624	8.10402941	4.78274074	5.2215	3.699
Leukemia	RPMI-8226	0.0905662	8.20359091	4.76770455	4.3245	3.1255
Leukemia	SR	0.80550572	4.640625	5.06011111	2.233	6.353
Lung	A549/ATCC	0.0003007	9.69596	2.657	10.709	1.063
Lung	EKVX	0.06179127	8.11203846	4.634625	10.0715	0.793
Lung	HOP-62	0.41290086	7.49063636	5.95386957	8.4205	2.403
Lung	HOP-92	0.60890571	7.08108333	6.04933333	3.541	1.152
Lung	NCI-H23	0.50492165	6.54445833	5.36216	5.2885	3.57
Lung	NCI-H460	0.00866585	9.82470833	4.04656522	10.2615	0.995
Lung	NCI-H522	0.00381137	10.3553913	4.10762069	10.431	1.112
Melanoma	LOX IMVI	0.49161966	8.93631579	7.24984211	6.739	6.739
Melanoma	SK-MEL-28	0.01487613	8.823	3.91362857	10.068	3.142
Melanoma	SK-MEL-5	0.00826385	8.74145833	3.65655556	8.3605	1.907
Melanoma	UACC-257	0.06250272	9.99073913	5.55445	13.737	1.475
Melanoma	UACC-62	0.00107127	10.2899583	3.59857895	11.868	1.73
Ovarian	IGROV1	0.03065246	6.5681	2.92847059	5.255	1.162
Ovarian	OVCAR-3	0.30277752	10.0375	7.48319355	8.322	4.721
Ovarian	OVCAR-4	0.02253953	10.5576111	4.38482051	10.734	2.716
Ovarian	OVCAR-8	0.0200668	10.575913	5.27962963	11.749	4.012
Ovarian	SK-OV-3	0.50431896	7.19541667	5.89610526	5.4435	5.753
Prostate	DU-145	0.04073121	11.0988182	6.51108571	13.492	4.613
Prostate	PC-3	0.01682485	7.69476667	3.68580769	7.8885	1.091
Renal	A498	0.2626584	8.24572222	5.56285185	8.1115	1.179
Renal	ACHN	0.20042872	8.06225	5.263	5.836	1.85
Renal	CAKI-1	0.18968805	7.71534615	4.7425625	4.777	1.214
Renal	UO-31	0.86750804	3.98037037	4.18062069	1.985	1.985

Figure S6-S8 provided clearer evidence of a trend where genes targeted by drugs with higher scores exhibited higher node degree (gene importance scores). In each cell line, the left boxplots compared the node degree distribution of genes targeted by the top 5 drugs with the highest scores against those targeted by the bottom 5 drugs with the lowest scores. The right boxplots in each cell line represented the overall comparisons between the top 5 and bottom 5 drug scores. Consequently, we were able to assign prior targets to genes with the highest node degree (gene importance scores) using our calculations. This allowed us to rank the genes in each cell line according to their node degree (importance scores), which can be found in ' **TableS1.csv**'. We generated a list of the top 20 genes in each cell line, available in '**TableS2.csv**'. Subsequently, we identified overlapping genes among the cell lines of the same cancer type based on the top 20 genes in each line, as depicted in **Figure 4**.

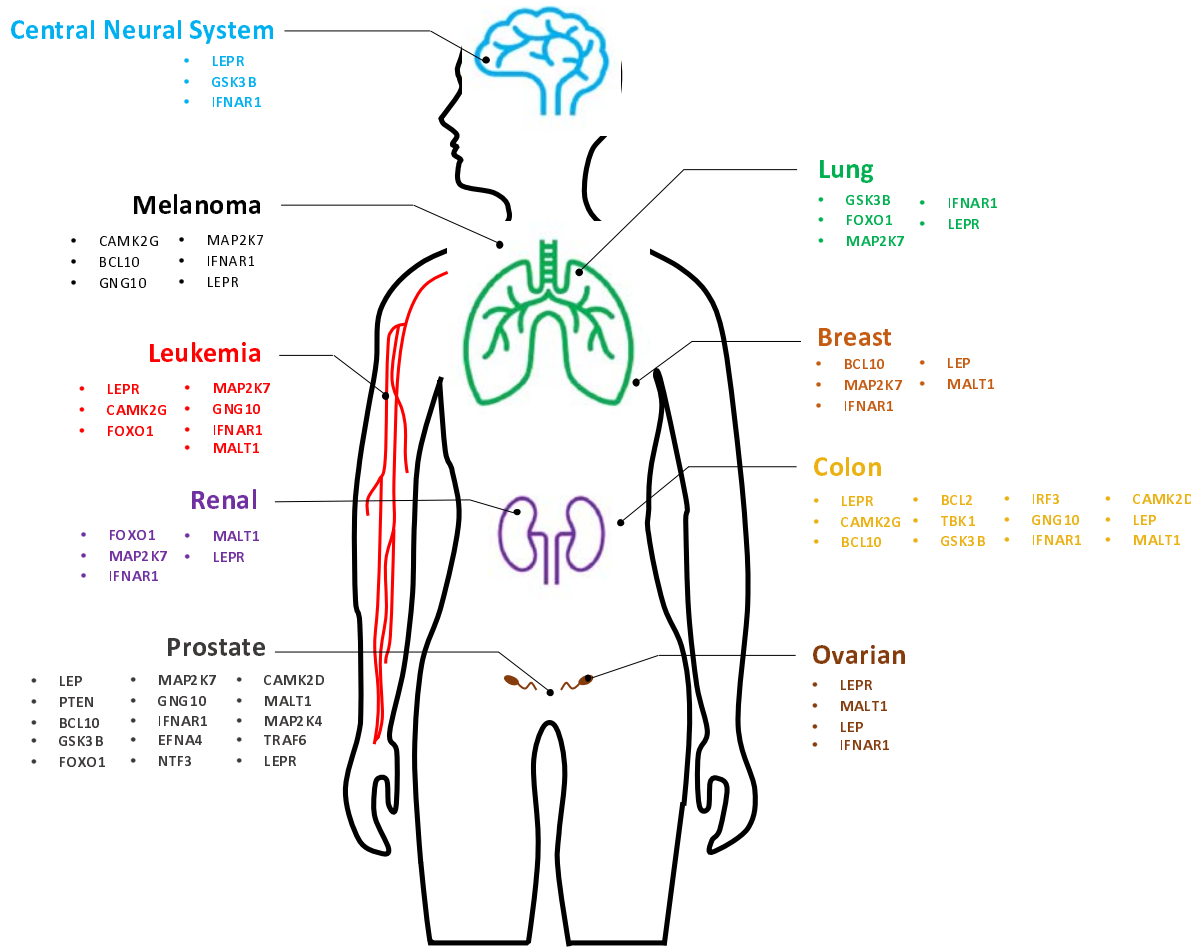

Figure 4. Top ranked key target genes across cell lines in the same type of cancer.

In order to visualize the results and gain a better understanding of the underlying mechanism of drug effects, we generated a core network of signaling interactions by applying a threshold. To enhance the interactions, we utilized the RShiny package and developed a visualization tool called ***NetFlowVis***. (Video link for how to use this tool is attached in the **Appendix B**) This tool allows users to control the edge threshold and set a minimum number of nodes in each network component. Users have the option to select the desired cell line for visualization and mark specific signaling pathways for detailed analysis. As an example, **Figure 5** demonstrates the visualization of the cell line DU-145, while **Figure 6** displays the comparisons of gene degree distributions for the top 5 and bottom 5 drug scores in the same cell line. Furthermore, by

referring to **Table 8**, we observed that the majority of gene targets were included in the filtered signaling network interactions depicted in **Figure 5**, where the nodes representing these genes were marked in red. It is worth noting that similar analyses were performed for all cell lines, and the results can be found in **'Table-S1.csv'**.

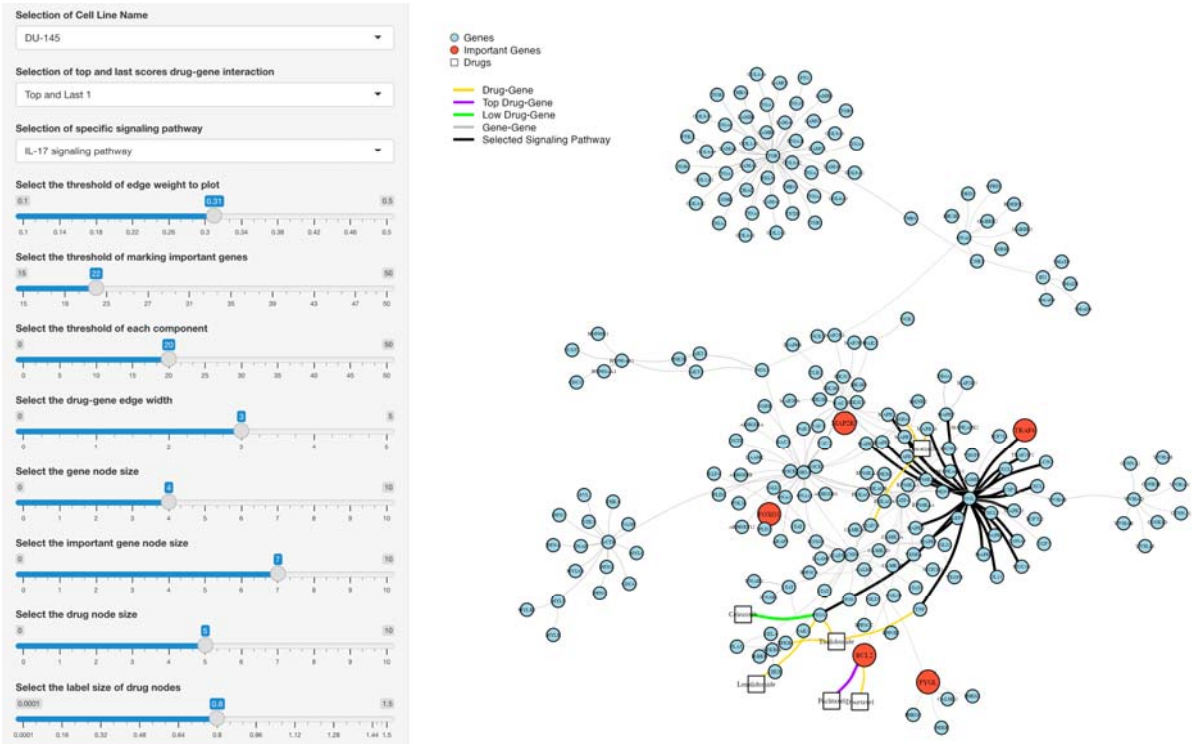

Figure 5. Visualization tool ***NetFlowVis*** for core signaling network interactions of cell line DU-145

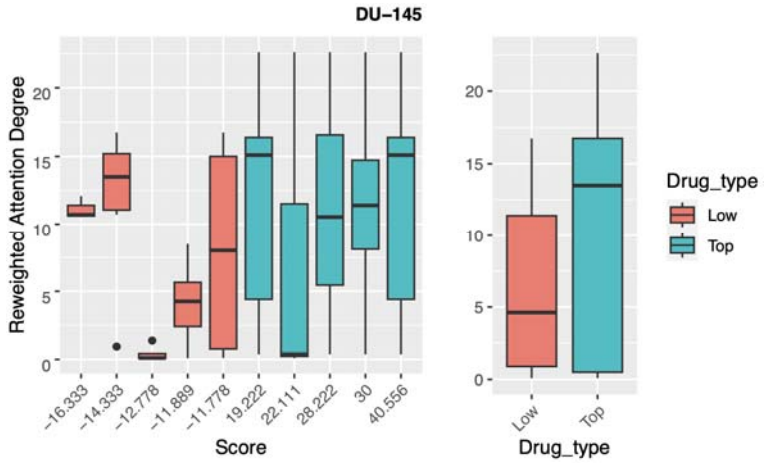

Figure 6. Cell line DU-145 gene degree distribution targeted by highest 5 and lowest 5 drugs scores

Table 8. Top 20 prior gene targets and their corresponding gene degree (node importance scores) for cell line DU-145 (genes highlighted in orange are shown in filtered subgraph in **Figure 5**)

Cell Line Name	Top1 Gene	Top2 Gene	Top3 Gene	Top4 Gene	Top5 Gene
DU-145	**PYGL**	**MAP2K7**	**TRAF6**	**FOXO1**	**LEP**
	25.1939102	23.8104848	23.3673616	22.8896179	22.6935263
	BCL2	**KRT18**	**LEPR**	**ZAP70**	**BCL10**
	22.627587	22.3407961	22.1099223	22.0225113	21.8549317
	MALT1	**GSK3B**	**IFNAR1**	**NTF3**	**EFNA4**
	21.7879979	21.593142	21.5865511	21.4453806	21.1022174
	CAMK2D	**GNG10**	**MAP2K4**	**PTEN**	**TSC2**
	21.0612515	20.7186959	20.5855408	20.5002708	20.4714351

By utilizing the ***NetFlowVis*** visualization tool, users were able to take control of the size of the core signaling network. In the case of the DU-145 cell line, an edge threshold of 0.31 was selected, which corresponded to approximately the 98.5th percentile for edge weight among all edges. Additionally, a threshold of 23 was chosen for marking important nodes/genes in red, representing ~99.9th percentile for node importance scores across all nodes. To gain a deeper understanding of the underlying mechanism, the interactions between drugs and genes were

analyzed, with the highest and lowest drug-gene interactions being represented by the colors purple and green, respectively.

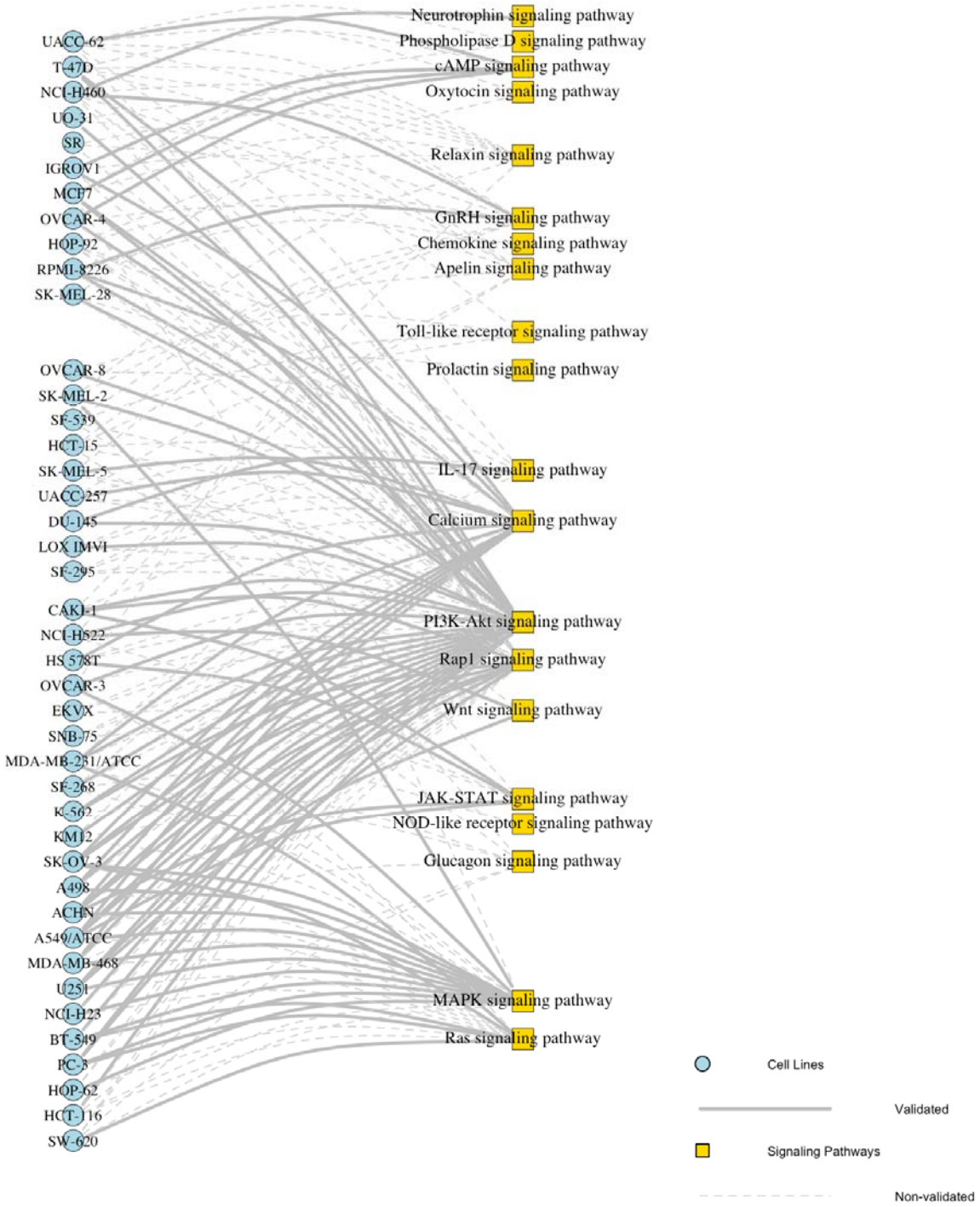

Figure 7. Important signaling pathways for each cell line (including validated and non-validated relations)

The analysis revealed that the drug ***Paclitaxel*** targeted the important gene ***BCL2***, while the drug ***Celecoxib*** targeted less important genes. This finding demonstrated that targeting the important genes selected by the ***M3NetFlow*** model led to higher drug scores, as depicted in **Figure 6**. Moreover, by selecting a specific signaling pathway within the core signaling network, it was observed that the ***IL-17 signaling pathway*** was downstream of the critical gene ***TRAF6***. Interestingly, experimental validation has shown the involvement of the ***IL-17 signaling pathway*** in promoting migration and invasion of ***DU-145*** cell lines through the upregulation of ***MTA1*** expression in prostate cancer[23]. For more comprehensive information, all the relationships between cell lines and signaling pathways are provided in **Table 9**, and the visualized plots can be referenced in **Figure 7**.

In order to further analyze the mechanism of synergy, we utilized our visualization tool to identify all paths shorter than a distance of 4, which were marked in red. This analysis specifically focused on selected drug pairs to investigate their synergy. **Figure 8** illustrates the analysis of the synergy mechanism between the drug pair ***Paclitaxel*** and ***Tretinon***. In this analysis, all paths shorter than a distance of 4 for this drug pair were included. Notably, we discovered that the gene ***BCL2*** played a crucial role in connecting these shortest paths. Additionally, it was found that the downstream ***Sphingolipid*** signaling pathway should be considered as one of the most significant signaling pathways associated with this particular drug pair.

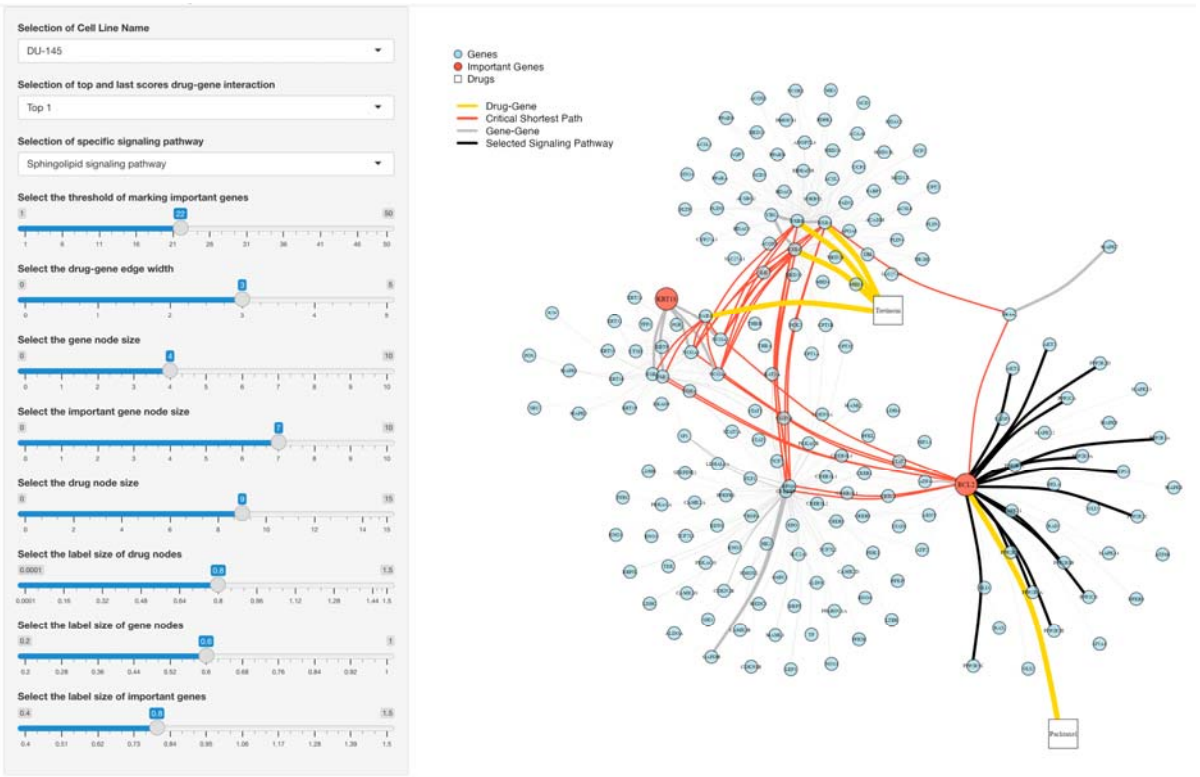

Figure 8. Visualization tool ***NetFlowVis*** for analysis of mechanism of synergy with interactions of cell line DU-145

4. Discussion and conclusion

Along with the advancement of next-generation sequencing (NGS) technology, multi-omics data, like gene expression, copy number variation, methylation, genetic mutation, microRNA, are being generated to characterize the dysfunctional molecules and complex signaling pathways in cancer patients. The valuable multi-omics data provides the basis for personalized medicine or precision medicine prediction. Drug combinations are being investigated to reduce the drug resistance influenced by the complex signaling networks in cancer. 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 combinations by simply concatenating a large number of multi-omics 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

biological meaningful gene regulatory relationships. Thus, it is hard to interpret these models to support the real clinical use, which requires the 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 ***M3NetFlow*** model. The model is built on a core cancer signaling network consisting of 1489 genes, with gene expression and copy number data derived from 48 core cancer signaling pathways. To incorporate longer range of node information and form the cell-line-specific network, the (multi-)***K-hop Subgraph Message Propagation*** was created. Aside from that, ***Global Bi-directional Message Propagation*** was created to build the novel up-stream signaling-flow (from up-stream signaling to drug targets), and the down-stream signaling-flow (from drug targets to down-stream signaling), were designed using trainable weights of network edges. In the end, the most important genes for each cell line were calculated based on the attention weight from their connected edges in both upstream and downstream directions and reweighting process was also used to reorder the important genes. From the statistical perspective, our model can sort the important genes aligning with drug scores.

Aside from that, visualization tool ***NetFlowVis*** was developed for the better analysis of synergy of drug combinations. By selecting the specific cell lines or drug pairs in certain cell line, the critical paths with edge threshold and mechanism of drug combinations could be easily analyzed. In conclusion, the K-hop subgraph propagation in the internal signaling pathways and the up-stream and down-stream signaling flow design for global propagation are novel and can potentially interpret the drug effects.

This is still an exploratory study to predict and investigate drug combination response using deep neural graph network models. There are some limitations to be further investigated. For example, more signaling pathways and protein-protein interactions should be considered to include more drugs and the mechanism of drug synergy should be investigated in detail.

Author contributions

FL conceived this study. FL, HZ designed the model, prepared the manuscript. HZ implemented the mode and analyzed the data and results. FL, HZ, PG, LD, WK, DD, RF, YC, PP discussed the method and results.

Supplementary Figures

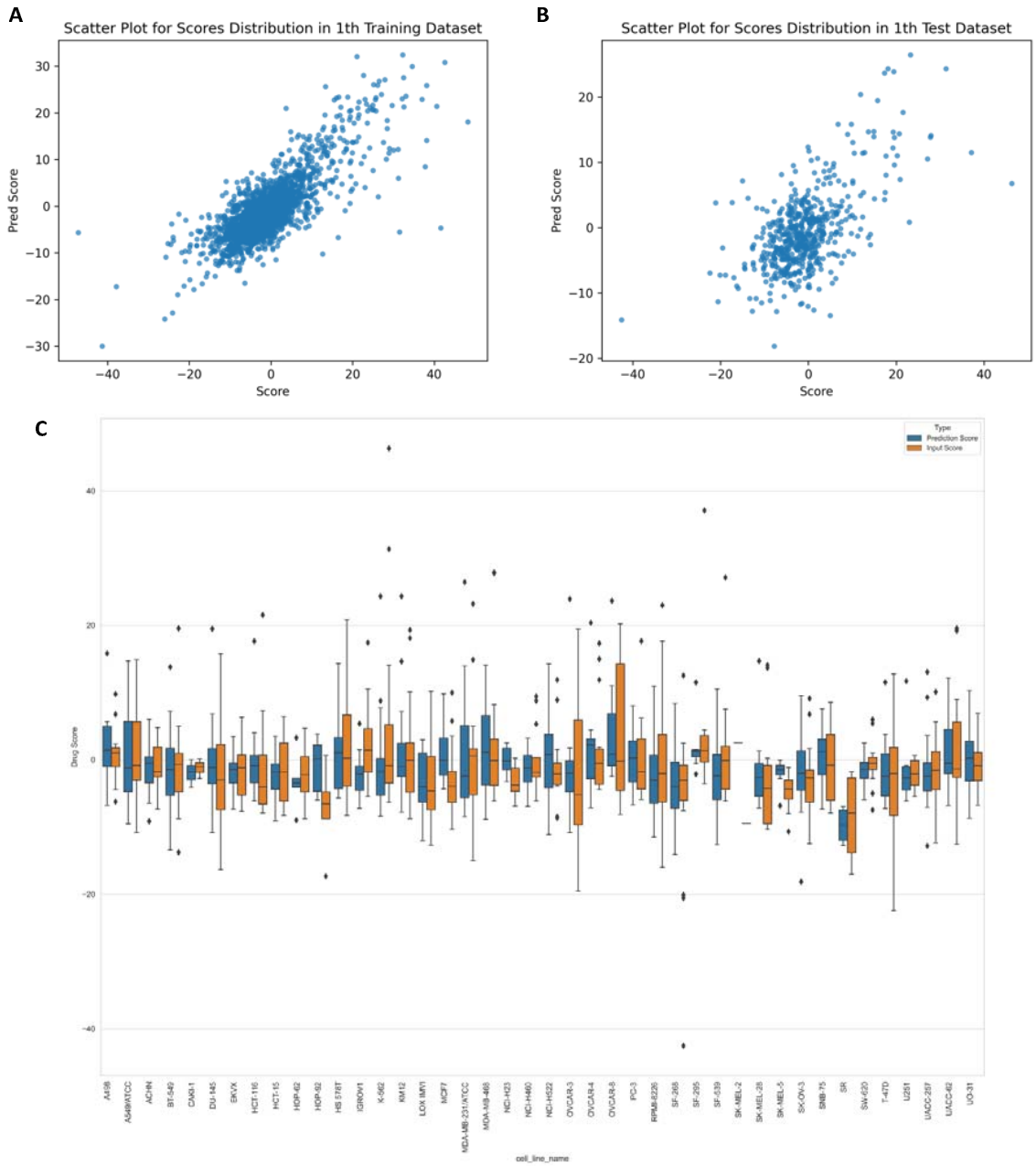

Figure S1 Model Performances on the 1st training and test dataset **A.** Scatter plot for score distribution in 1st training and test dataset **B.** Scatter plot for score distribution in 1st test dataset **C.** Box plots for distributions comparisons between predicted scores and original scores

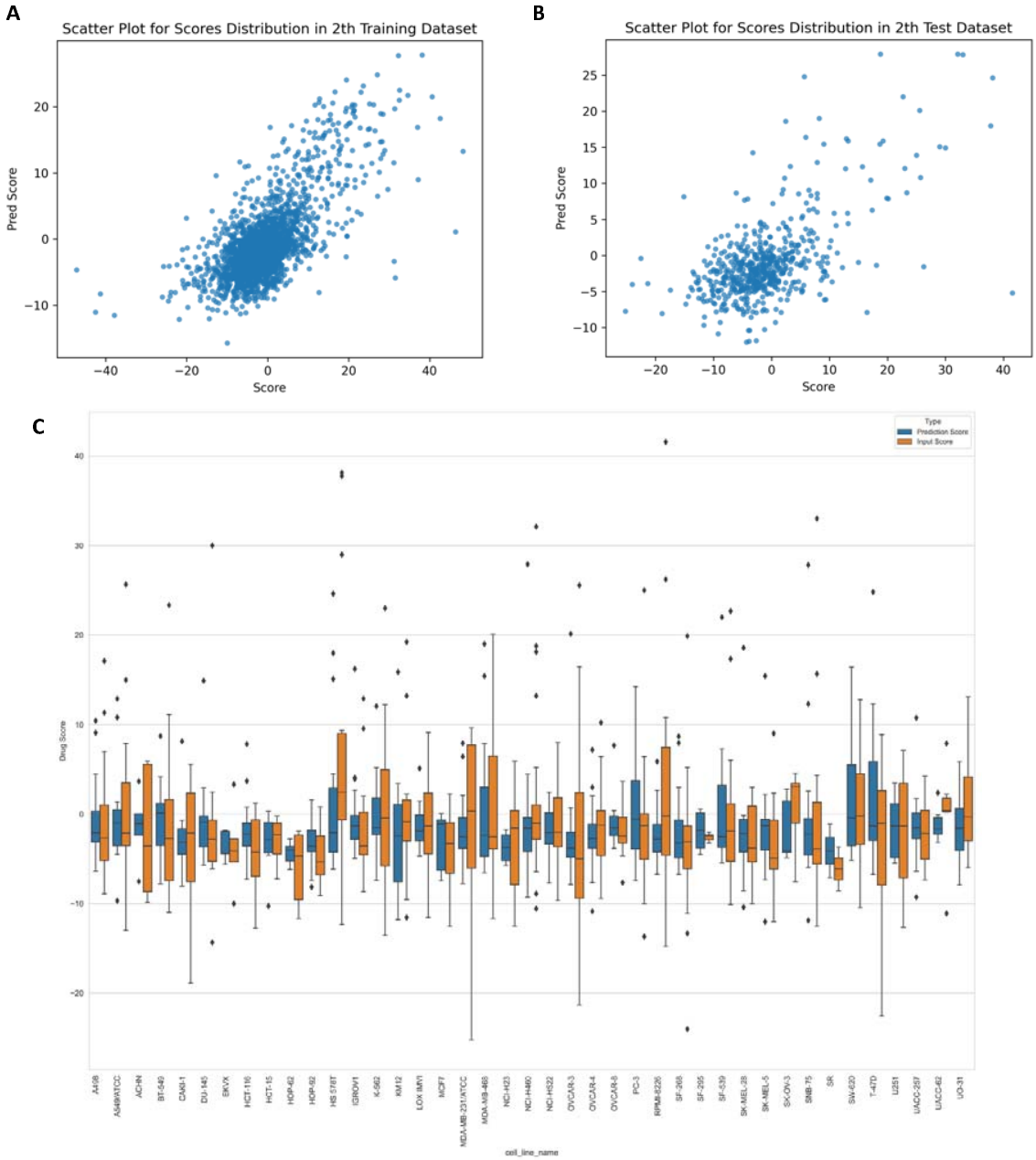

Figure S2 Model Performances on the 2nd training and test dataset **A.** Scatter plot for score distribution in 2nd training and test dataset **B.** Scatter plot for score distribution in 2nd test dataset **C.** Box plots for distributions comparisons between predicted scores and original scores

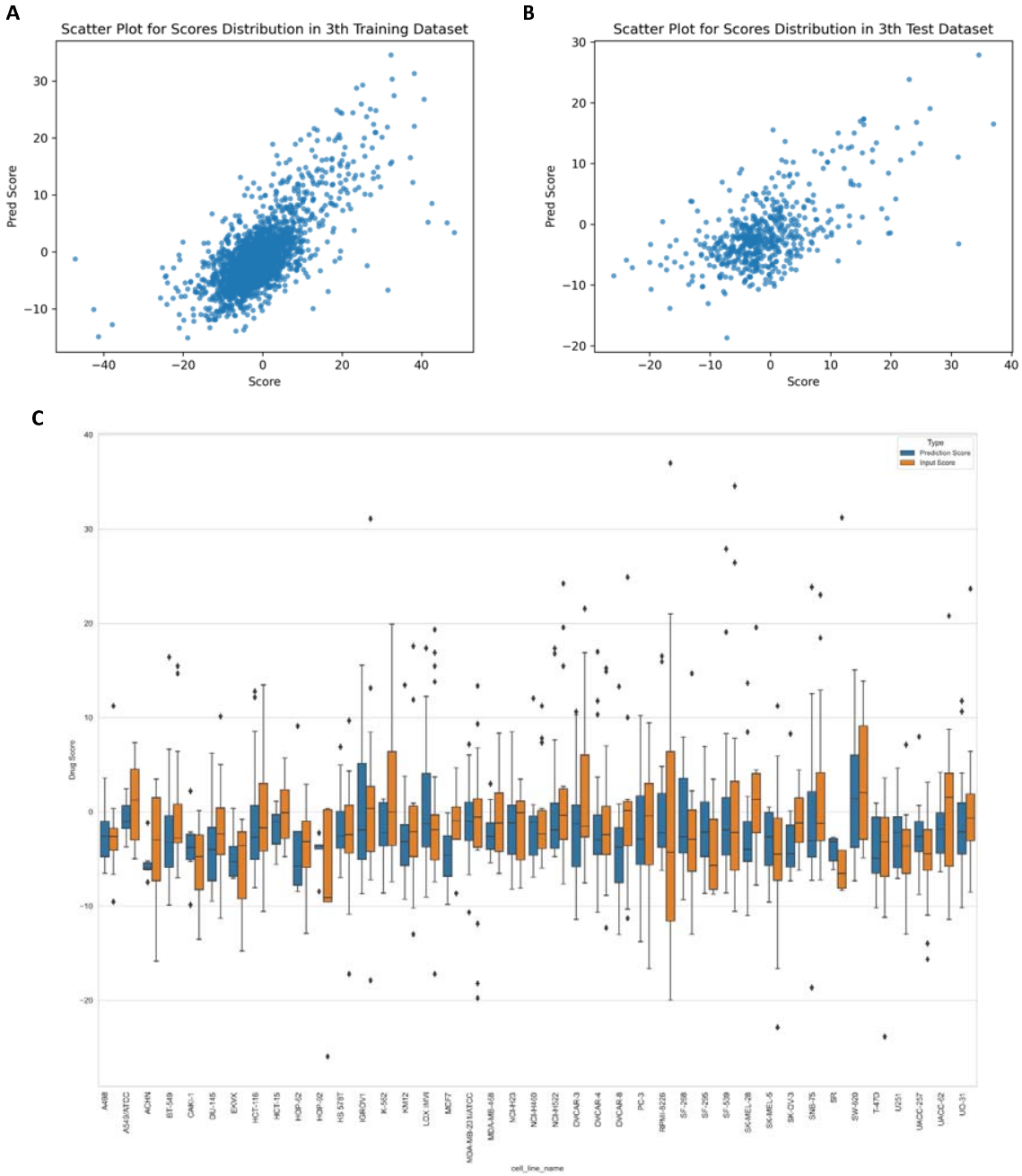

Figure S3 Model Performances on the 3rd training and test dataset **A.** Scatter plot for score distribution in 3rd training and test dataset **B.** Scatter plot for score distribution in 3rd test dataset **C.** Box plots for distributions comparisons between predicted scores and original scores

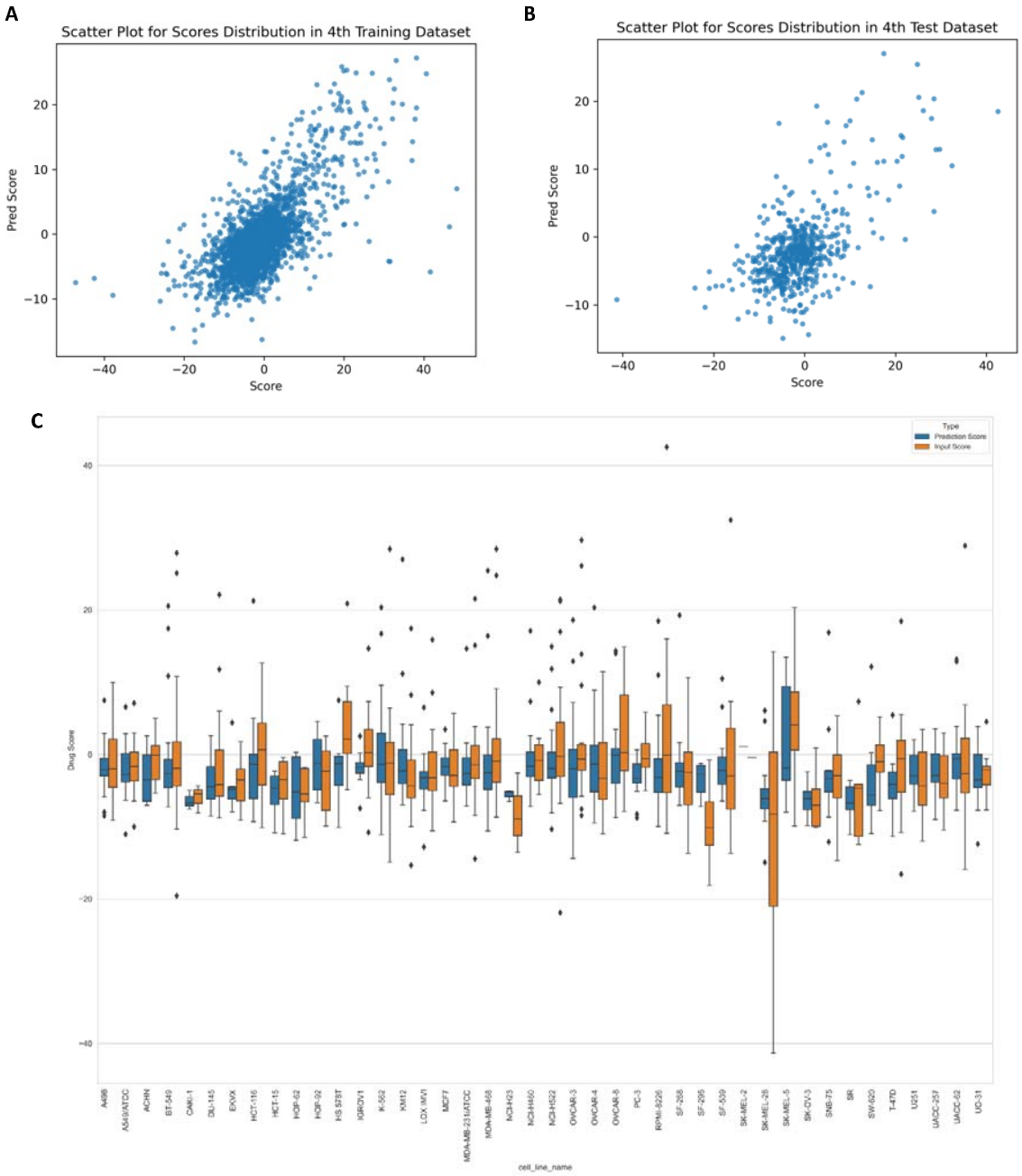

Figure S4 Model Performances on the 4th training and test dataset **A.** Scatter plot for score distribution in 4th training and test dataset **B.** Scatter plot for score distribution in 4th test dataset **C.** Box plots for distributions comparisons between predicted scores and original scores

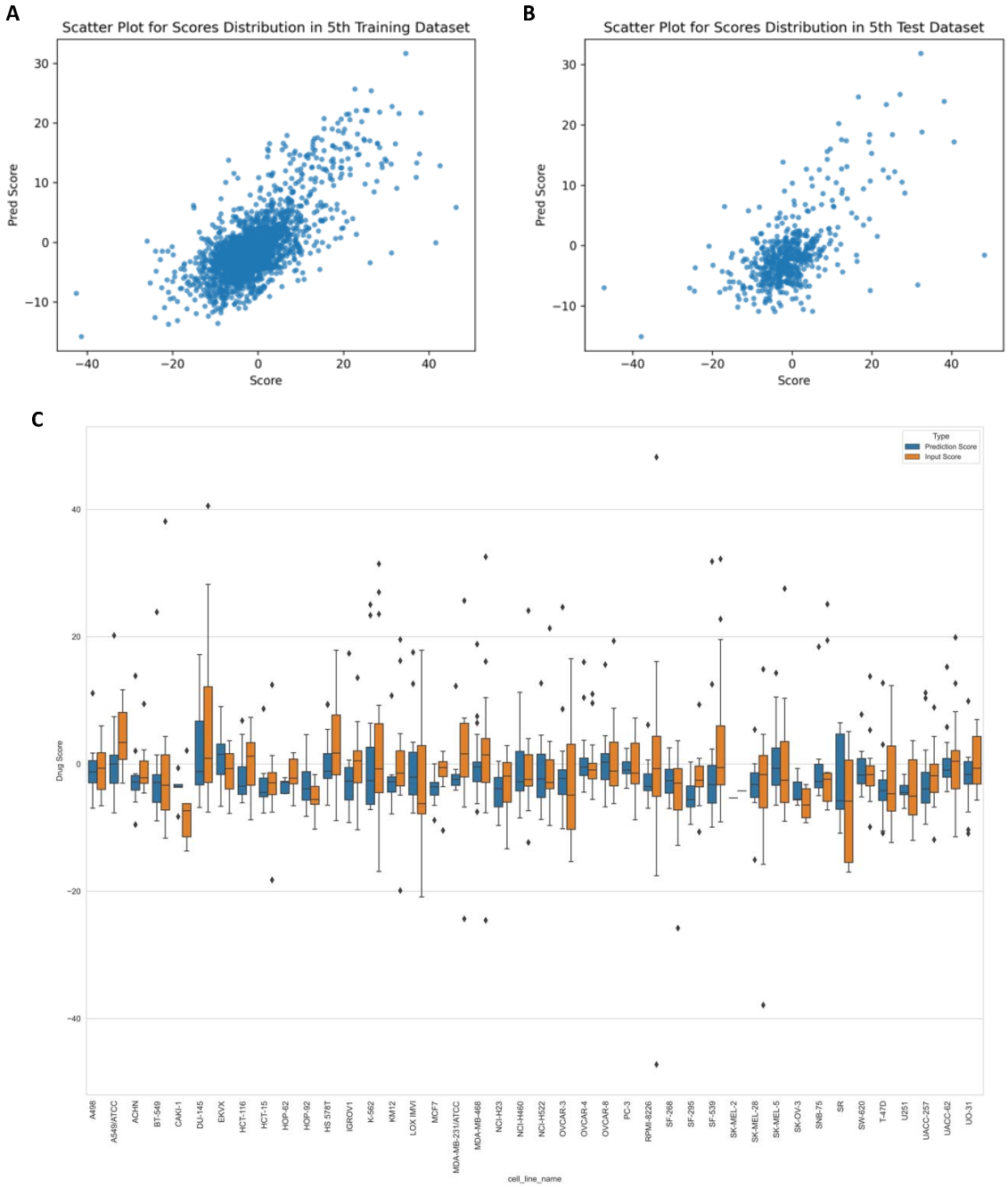

Figure S5 Model Performances on the 5th training and test dataset **A.** Scatter plot for score distribution in 5th training and test dataset **B.** Scatter plot for score distribution in 5th test dataset **C.** Box plots for distributions comparisons between predicted scores and original scores

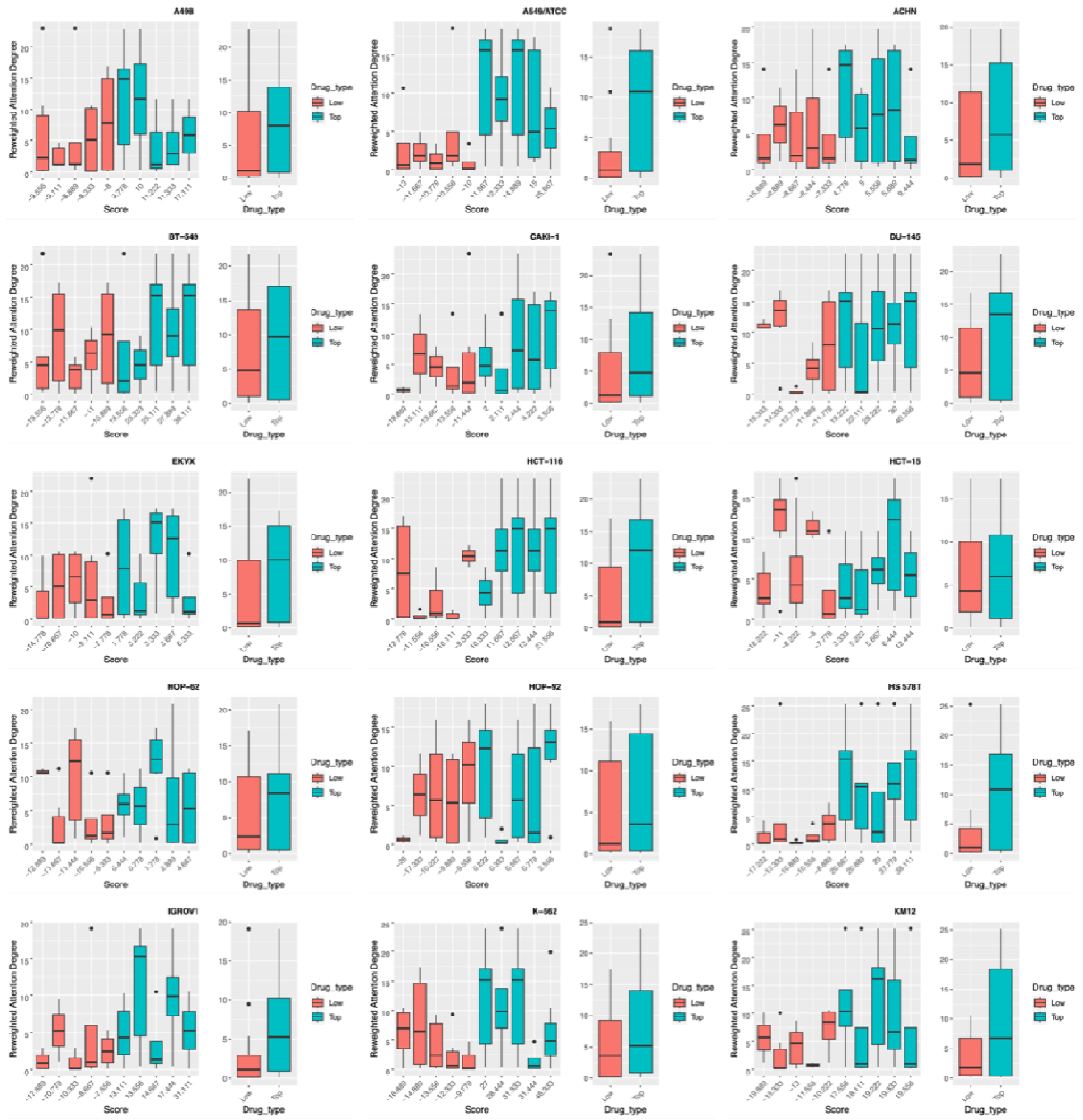

Figure S6 Cell line A498, A549/ATCC, ACHN, BT-549, CAKI-1, DU-145, EKVX, HCT-116, HCT-15, HOP-62, HOP-92, HS 578T, IGROV1, K-562, KM12 gene degree distribution targeted by highest 5 and lowest 5 drugs scores

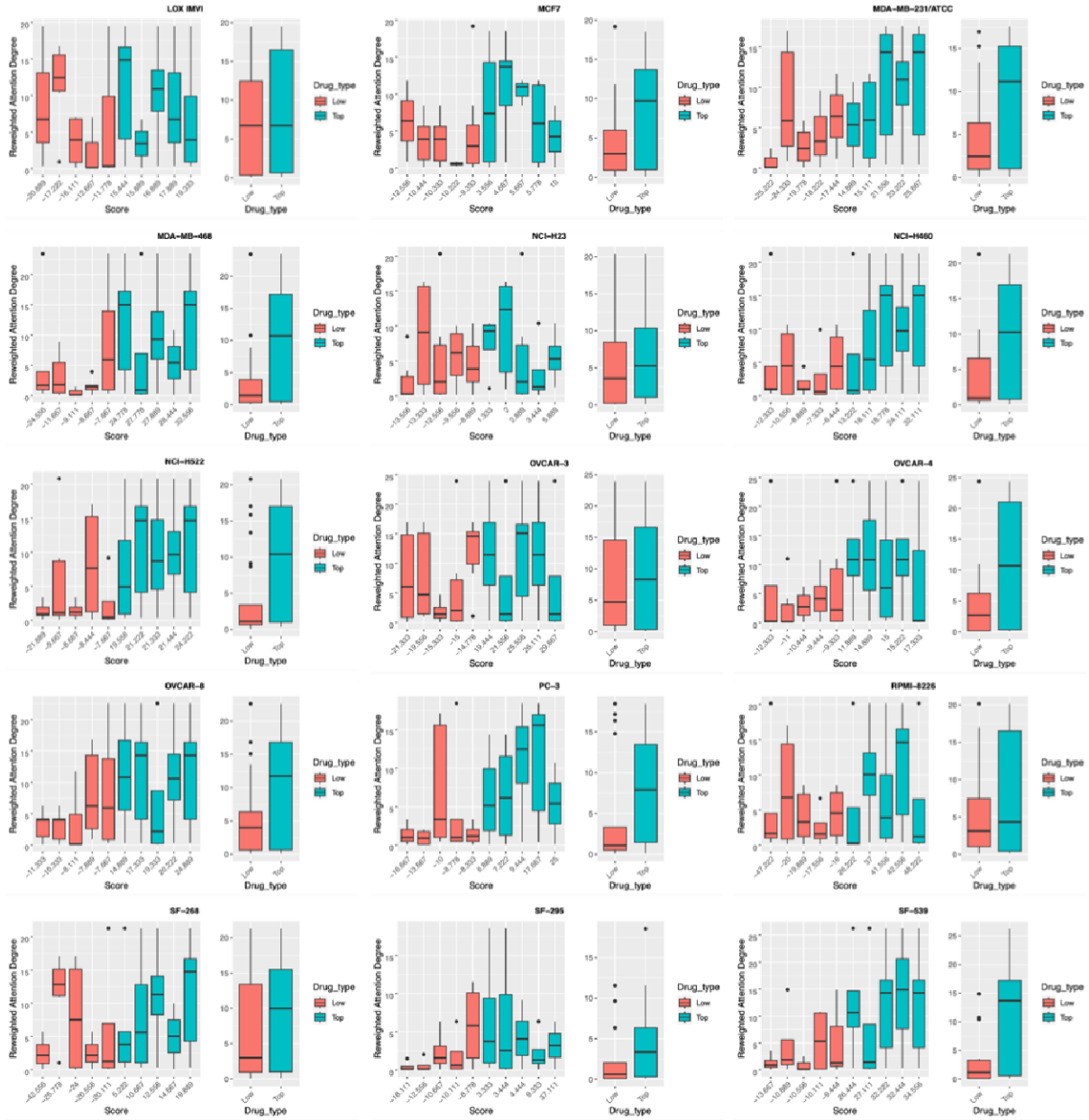

Figure S7 Cell line LOX IMVI, MCF7, MDA-MB-231/ATCC, MDA-MB-468, NCI-H23, HCI-H460, NCI-H522, OVCAR-3, OVCAR-4, OVCAR-8, PC-3, RPMI-8226, SF-268, SF-295, SF-539 gene degree distribution targeted by highest 5 and lowest 5 drugs scores

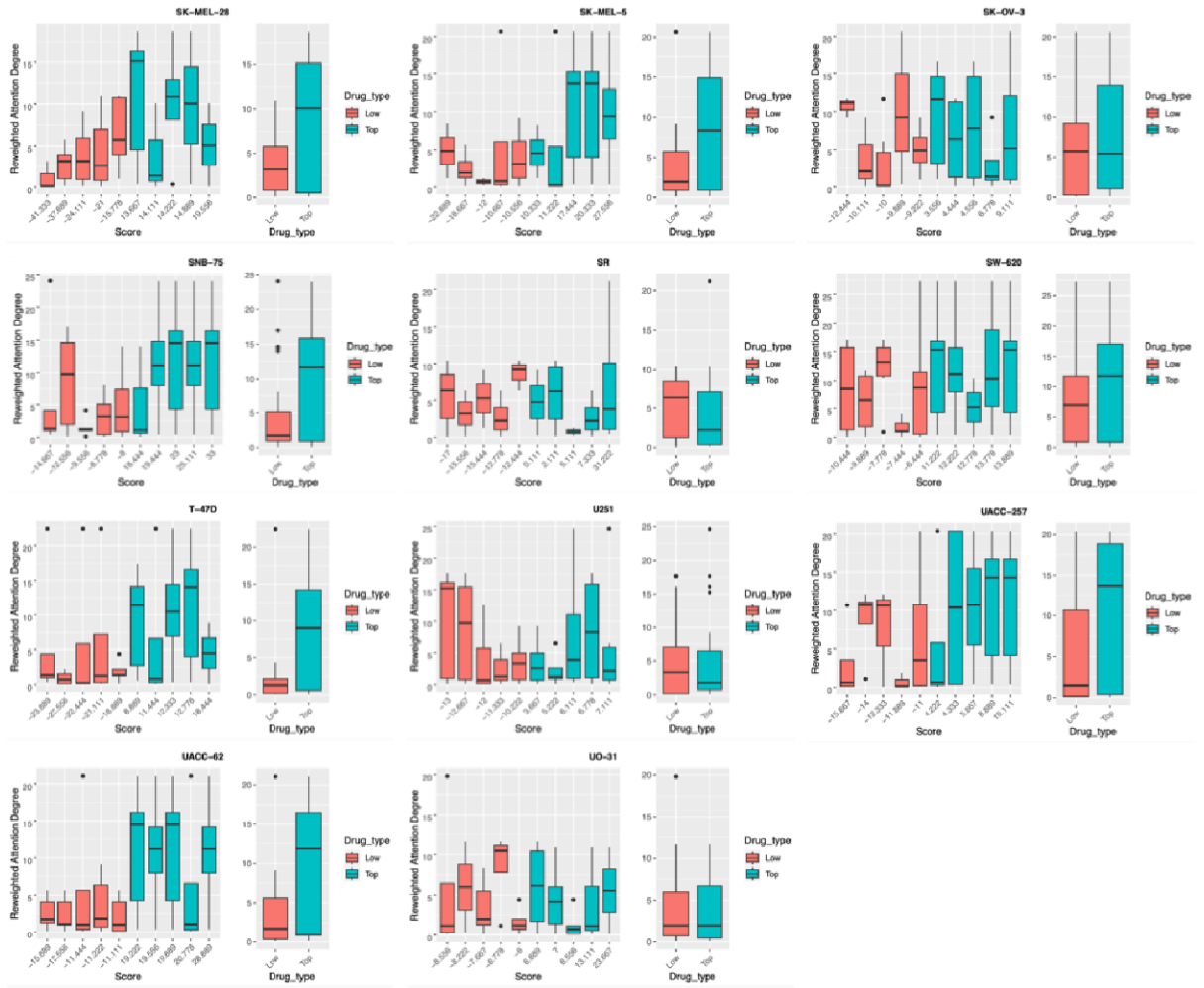

Figure S8 Cell line SK-MEL-28, SK-MEL-5, SK-OV-3, SNB-75, SR, SW-620, T-47D, U251, UACC-257, UACC-62, UO-31 gene degree distribution targeted by highest 5 and lowest 5 drugs scores

Table 9. Important signaling pathways corresponding to each cancer cell line

Cancers	Cell Lines	Signaling Pathways	Edges Weight Ratio	Validation Paper Citations
Breast	BT-549	MAPK signaling pathway	6.82%	24 25
Breast	BT-549	PI3K-Akt signaling pathway	19.96%	26 27 28 29
Breast	BT-549	Rap1 signaling pathway	8.85%	NONE
Breast	BT-549	Ras signaling pathway	6.65%	30
Breast	HS 578T	Calcium signaling pathway	5.56%	31 32 33
Breast	HS 578T	JAK-STAT signaling pathway	6.37%	34
Breast	HS 578T	PI3K-Akt signaling pathway	10.69%	NONE
Breast	HS 578T	Rap1 signaling pathway	5.92%	NONE
Breast	MCF7	cAMP signaling pathway	5.76%	35 36
Breast	MCF7	MAPK signaling pathway	7.67%	NONE
Breast	MCF7	Phospholipase D signaling pathway	5.77%	NONE
Breast	MCF7	PI3K-Akt signaling pathway	5.22%	36
Breast	MCF7	Rap1 signaling pathway	6.03%	37
Breast	MCF7	Relaxin signaling pathway	8.17%	NONE
Breast	MDA-MB-231/ATCC	Calcium signaling pathway	5.01%	38
Breast	MDA-MB-231/ATCC	IL-17 signaling pathway	5.17%	NONE
Breast	MDA-MB-231/ATCC	MAPK signaling pathway	9.35%	NONE
Breast	MDA-MB-231/ATCC	PI3K-Akt signaling pathway	9.08%	28 39
Breast	MDA-MB-231/ATCC	Rap1 signaling pathway	11.21%	NONE
Breast	MDA-MB-231/ATCC	Ras signaling pathway	9.35%	40
Breast	MDA-MB-468	Calcium signaling pathway	8.82%	41
Breast	MDA-MB-468	Glucagon signaling pathway	5.02%	NONE
Breast	MDA-MB-468	MAPK signaling pathway	9.43%	42
Breast	MDA-MB-468	PI3K-Akt signaling pathway	11.13%	43
Breast	MDA-MB-468	Rap1 signaling pathway	10.18%	NONE
Breast	MDA-MB-468	Ras signaling pathway	8.26%	NONE
Breast	T-47D	Apelin signaling pathway	5.52%	NONE
Breast	T-47D	Calcium signaling pathway	5.39%	44 45

Breast	T-47D	cAMP signaling pathway	6.86%	NONE
Breast	T-47D	GnRH signaling pathway	6.48%	NONE
Breast	T-47D	PI3K-Akt signaling pathway	10.03%	39
Breast	T-47D	Relaxin signaling pathway	5.62%	NONE
CNS	SF-268	Calcium signaling pathway	6.07%	NONE
CNS	SF-268	MAPK signaling pathway	7.61%	NONE
CNS	SF-268	NOD-like receptor signaling pathway	5.08%	NONE
CNS	SF-268	PI3K-Akt signaling pathway	9.92%	46
CNS	SF-268	Rap1 signaling pathway	6.77%	NONE
CNS	SF-295	Calcium signaling pathway	11.21%	NONE
CNS	SF-295	IL-17 signaling pathway	5.08%	NONE
CNS	SF-295	PI3K-Akt signaling pathway	13.91%	NONE
CNS	SF-295	Rap1 signaling pathway	5.45%	NONE
CNS	SF-539	GnRH signaling pathway	5.07%	NONE
CNS	SF-539	JAK-STAT signaling pathway	8.87%	NONE
CNS	SF-539	PI3K-Akt signaling pathway	10.93%	NONE
CNS	SF-539	Prolactin signaling pathway	5.40%	NONE
CNS	SNB-75	JAK-STAT signaling pathway	7.55%	NONE
CNS	SNB-75	PI3K-Akt signaling pathway	11.16%	NONE
CNS	SNB-75	Rap1 signaling pathway	5.56%	NONE
CNS	U251	MAPK signaling pathway	9.02%	47 48
CNS	U251	PI3K-Akt signaling pathway	10.55%	49 50
CNS	U251	Rap1 signaling pathway	7.81%	51
CNS	U251	Ras signaling pathway	8.82%	NONE
CNS	U251	Wnt signaling pathway	5.40%	52
Colon	HCT-116	Glucagon signaling pathway	6.12%	NONE
Colon	HCT-116	MAPK signaling pathway	8.47%	NONE
Colon	HCT-116	PI3K-Akt signaling pathway	15.45%	53 54
Colon	HCT-116	Rap1 signaling pathway	7.51%	NONE
Colon	HCT-116	Ras signaling pathway	6.17%	NONE
Colon	HCT-15	Calcium signaling pathway	5.46%	NONE
Colon	HCT-15	Glucagon signaling pathway	7.79%	NONE
Colon	HCT-15	GnRH signaling pathway	5.62%	NONE
Colon	HCT-15	PI3K-Akt signaling pathway	6.06%	NONE
Colon	HCT-15	Toll-like receptor signaling pathway	5.52%	NONE
Colon	KM12	Calcium signaling pathway	7.98%	NONE
Colon	KM12	Glucagon signaling pathway	7.41%	NONE
Colon	KM12	MAPK signaling pathway	8.29%	NONE

Colon	KM12	PI3K-Akt signaling pathway	16.25%	55
Colon	SW-620	Glucagon signaling pathway	6.15%	NONE
Colon	SW-620	MAPK signaling pathway	8.70%	NONE
Colon	SW-620	PI3K-Akt signaling pathway	6.92%	NONE
Colon	SW-620	Rap1 signaling pathway	7.53%	NONE
Colon	SW-620	Ras signaling pathway	6.88%	56
Leukemia	K-562	Glucagon signaling pathway	6.21%	NONE
Leukemia	K-562	PI3K-Akt signaling pathway	13.76%	57 58 59
Leukemia	K-562	Rap1 signaling pathway	7.47%	60
Leukemia	RPMI-8226	Calcium signaling pathway	5.38%	61
Leukemia	RPMI-8226	Chemokine signaling pathway	8.56%	NONE
Leukemia	RPMI-8226	GnRH signaling pathway	6.17%	62
Leukemia	RPMI-8226	PI3K-Akt signaling pathway	7.08%	63
Leukemia	RPMI-8226	Rap1 signaling pathway	5.77%	NONE
Leukemia	SR	Apelin signaling pathway	7.09%	NONE
Leukemia	SR	PI3K-Akt signaling pathway	11.36%	NONE
Leukemia	SR	Relaxin signaling pathway	6.22%	NONE
Lung	A549/ATCC	Calcium signaling pathway	8.68%	64
Lung	A549/ATCC	JAK-STAT signaling pathway	5.15%	60
Lung	A549/ATCC	MAPK signaling pathway	8.08%	65 66 67
Lung	A549/ATCC	PI3K-Akt signaling pathway	16.51%	68 69 70
Lung	A549/ATCC	Rap1 signaling pathway	8.23%	71
Lung	A549/ATCC	Ras signaling pathway	6.94%	NONE
Lung	EKVX	JAK-STAT signaling pathway	5.43%	NONE
Lung	EKVX	PI3K-Akt signaling pathway	14.58%	NONE
Lung	EKVX	Rap1 signaling pathway	5.42%	NONE
Lung	HOP-62	MAPK signaling pathway	8.98%	72
Lung	HOP-62	PI3K-Akt signaling pathway	18.72%	NONE
Lung	HOP-62	Rap1 signaling pathway	9.31%	NONE
Lung	HOP-62	Ras signaling pathway	7.92%	NONE
Lung	HOP-92	Calcium signaling pathway	10.26%	NONE
Lung	HOP-92	GnRH signaling pathway	5.08%	NONE
Lung	HOP-92	IL-17 signaling pathway	6.17%	NONE
Lung	HOP-92	PI3K-Akt signaling pathway	10.78%	NONE
Lung	HOP-92	Rap1 signaling pathway	5.68%	NONE
Lung	HOP-92	Relaxin signaling pathway	5.71%	NONE
Lung	NCI-H23	MAPK signaling pathway	10.27%	73

Lung	NCI-H23	PI3K-Akt signaling pathway	13.73%	73 74
Lung	NCI-H23	Rap1 signaling pathway	10.45%	NONE
Lung	NCI-H23	Ras signaling pathway	8.20%	NONE
Lung	NCI-H460	Calcium signaling pathway	6.93%	NONE
Lung	NCI-H460	GnRH signaling pathway	5.92%	75
Lung	NCI-H460	Neurotrophin signaling pathway	5.01%	76
Lung	NCI-H460	Rap1 signaling pathway	6.09%	NONE
Lung	NCI-H460	Relaxin signaling pathway	5.11%	NONE
Lung	NCI-H460	Toll-like receptor signaling pathway	5.09%	NONE
Lung	NCI-H522	Apelin signaling pathway	5.34%	NONE
Lung	NCI-H522	Glucagon signaling pathway	5.24%	NONE
Lung	NCI-H522	PI3K-Akt signaling pathway	9.34%	77
Lung	NCI-H522	Wnt signaling pathway	8.02%	78
Melanoma	LOX IMVI	Calcium signaling pathway	5.10%	NONE
Melanoma	LOX IMVI	IL-17 signaling pathway	5.34%	NONE
Melanoma	LOX IMVI	PI3K-Akt signaling pathway	13.54%	79
Melanoma	LOX IMVI	Rap1 signaling pathway	5.32%	NONE
Melanoma	SK-MEL-2	Apelin signaling pathway	6.02%	NONE
Melanoma	SK-MEL-2	MAPK signaling pathway	9.54%	80
Melanoma	SK-MEL-2	PI3K-Akt signaling pathway	15.75%	81
Melanoma	SK-MEL-2	Relaxin signaling pathway	5.10%	NONE
Melanoma	SK-MEL-28	Calcium signaling pathway	5.54%	NONE
Melanoma	SK-MEL-28	Oxytocin signaling pathway	5.36%	NONE
Melanoma	SK-MEL-28	PI3K-Akt signaling pathway	8.72%	82 83
Melanoma	SK-MEL-28	Rap1 signaling pathway	7.08%	NONE
Melanoma	SK-MEL-5	Calcium signaling pathway	5.88%	84
Melanoma	SK-MEL-5	GnRH signaling pathway	5.52%	NONE
Melanoma	SK-MEL-5	Rap1 signaling pathway	6.70%	NONE
Melanoma	SK-MEL-5	Wnt signaling pathway	6.37%	NONE
Melanoma	UACC-257	Calcium signaling pathway	6.35%	85
Melanoma	UACC-257	IL-17 signaling pathway	6.56%	NONE
Melanoma	UACC-257	PI3K-Akt signaling pathway	6.94%	NONE
Melanoma	UACC-257	Rap1 signaling pathway	6.12%	NONE
Melanoma	UACC-257	Toll-like receptor signaling pathway	5.97%	NONE
Melanoma	UACC-62	cAMP signaling pathway	6.23%	86
Melanoma	UACC-62	GnRH signaling pathway	5.06%	NONE

Melanoma	UACC-62	IL-17 signaling pathway	5.74%	NONE
Melanoma	UACC-62	Oxytocin signaling pathway	6.43%	NONE
Melanoma	UACC-62	Phospholipase D signaling pathway	6.47%	NONE
Melanoma	UACC-62	Rap1 signaling pathway	6.64%	NONE
Melanoma	UACC-62	Relaxin signaling pathway	9.14%	NONE
Ovarian	IGROV1	Calcium signaling pathway	7.10%	87
Ovarian	IGROV1	cAMP signaling pathway	5.35%	88
Ovarian	IGROV1	GnRH signaling pathway	6.45%	NONE
Ovarian	IGROV1	Rap1 signaling pathway	6.80%	NONE
Ovarian	OVCAR-3	MAPK signaling pathway	9.14%	89
Ovarian	OVCAR-3	PI3K-Akt signaling pathway	19.19%	90 91
Ovarian	OVCAR-3	Rap1 signaling pathway	10.20%	NONE
Ovarian	OVCAR-3	Ras signaling pathway	7.29%	NONE
Ovarian	OVCAR-3	Relaxin signaling pathway	5.69%	NONE
Ovarian	OVCAR-4	Apelin signaling pathway	8.11%	NONE
Ovarian	OVCAR-4	cAMP signaling pathway	6.16%	92
Ovarian	OVCAR-4	PI3K-Akt signaling pathway	19.03%	93 91
Ovarian	OVCAR-4	Rap1 signaling pathway	5.62%	NONE
Ovarian	OVCAR-8	IL-17 signaling pathway	5.42%	NONE
Ovarian	OVCAR-8	PI3K-Akt signaling pathway	5.99%	94
Ovarian	OVCAR-8	Toll-like receptor signaling pathway	5.01%	NONE
Ovarian	SK-OV-3	Calcium signaling pathway	6.10%	95
Ovarian	SK-OV-3	MAPK signaling pathway	10.32%	96
Ovarian	SK-OV-3	PI3K-Akt signaling pathway	22.45%	97
Ovarian	SK-OV-3	Rap1 signaling pathway	7.86%	98
Ovarian	SK-OV-3	Ras signaling pathway	7.71%	99
Prostate	DU-145	IL-17 signaling pathway	7.07%	23 100 101 102 103 104
Prostate	DU-145	PI3K-Akt signaling pathway	17.79%	105 106 107 108 109 110
Prostate	PC-3	MAPK signaling pathway	9.10%	111 112
Prostate	PC-3	PI3K-Akt signaling pathway	21.74%	113 114
Prostate	PC-3	Rap1 signaling pathway	10.06%	115
Prostate	PC-3	Ras signaling pathway	8.62%	116 117

Renal	A498	Calcium signaling pathway	8.05%	118 119
Renal	A498	MAPK signaling pathway	8.61%	120
Renal	A498	PI3K-Akt signaling pathway	20.69%	121 122
Renal	A498	Rap1 signaling pathway	9.75%	NONE
Renal	A498	Ras signaling pathway	7.13%	NONE
Renal	ACHN	Calcium signaling pathway	9.62%	123
Renal	ACHN	JAK-STAT signaling pathway	7.10%	124
Renal	ACHN	MAPK signaling pathway	10.19%	125
Renal	ACHN	PI3K-Akt signaling pathway	8.34%	126
Renal	ACHN	Rap1 signaling pathway	10.39%	127
Renal	ACHN	Ras signaling pathway	7.81%	NONE
Renal	CAKI-1	Apelin signaling pathway	6.02%	NONE
Renal	CAKI-1	Calcium signaling pathway	5.94%	NONE
Renal	CAKI-1	JAK-STAT signaling pathway	6.04%	128
Renal	CAKI-1	PI3K-Akt signaling pathway	16.94%	129
Renal	CAKI-1	Rap1 signaling pathway	5.15%	130
Renal	UO-31	Apelin signaling pathway	6.85%	NONE
Renal	UO-31	PI3K-Akt signaling pathway	19.27%	131
Renal	UO-31	Relaxin signaling pathway	5.35%	NONE

References

1. TCGA. *https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga*.
2. Barretina, J. *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–7 (2012).
3. Yang, W. *et al.* Genomics of Drug Sensitivity in Cancer (GDSC): A resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res* **41**, D955–D961 (2013).
4. Li, F. *et al.* Weakly activated core inflammation pathways were identified as a central signaling mechanism contributing to the chronic neurodegeneration in Alzheimer's disease. *bioRxiv* 2021.08.30.458295 (2021) doi:10.1101/2021.08.30.458295.
5. Li, F. *et al.* Loss of estrogen unleashing neuro-inflammation increases the risk of Alzheimer's disease in women. *bioRxiv* 2022.09.19.508592 (2022) doi:10.1101/2022.09.19.508592.
6. Zhou, G., Li, S. & Xia, J. Network-Based Approaches for Multi-omics Integration. in *Computational Methods and Data Analysis for Metabolomics* (ed. Li, S.) 469–487 (Springer US, 2020). doi:10.1007/978-1-0716-0239-3_23.
7. Subramanian, I., Verma, S., Kumar, S., Jere, A. & Anamika, K. Multi-omics Data Integration, Interpretation, and Its Application. *Bioinform Biol Insights* **14**, 1177932219899051 (2020).
8. Zhang, T., Zhang, L., Payne, P. & Li, F. Synergistic Drug Combination Prediction by Integrating Multi-omics Data in Deep Learning Models. *arXiv preprint arXiv:1811.07054* (2018).
9. Li, G., Muller, M., Thabet, A. & Ghanem, B. DeepGCNs: Can GCNs go as deep as CNNs? *Proceedings of the IEEE International Conference on Computer Vision* **2019-Octob**, 9266–9275 (2019).
10. Cai, C. & Wang, Y. A Note on Over-Smoothing for Graph Neural Networks. (2020).
11. Holbeck, S. L. *et al.* The National Cancer Institute ALMANAC: A comprehensive screening resource for the detection of anticancer drug pairs with enhanced therapeutic activity. *Cancer Research* **77**, 3564–3576 (2017).
12. Van Der Meer, D. *et al.* Cell Model Passports - a hub for clinical, genetic and functional datasets of preclinical cancer models. *Nucleic Acids Research* **47**, D923–D929 (2019).
13. Barretina, J. *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–7 (2012).
14. Ogata, H. *et al.* KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Research* 28 Preprint at https://doi.org/10.1093/nar/27.1.29 (1999).
15. Wishart, D. S. *et al.* DrugBank 5.0: A major update to the DrugBank database for 2018. *Nucleic Acids Research* **46**, D1074–D1082 (2018).

16. O'Neil, J. *et al.* An unbiased oncology compound screen to identify novel combination strategies. *Molecular Cancer Therapeutics* **15**, 1155–1162 (2016).
17. Zagidullin, B. *et al.* DrugComb: An integrative cancer drug combination data portal. *Nucleic Acids Research* **47**, W43–W51 (2019).
18. Zitnik, M., Agrawal, M. & Leskovec, J. Modeling polypharmacy side effects with graph convolutional networks. *Bioinformatics* **34**, i457–i466 (2018).
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. & Li, F. Synergistic Drug Combination Prediction by Integrating Multi-omics Data in Deep Learning Models. *arXiv preprint arXiv:1811.07054* (2018).
21. Kipf, T. N. & Welling, M. Semi-supervised classification with graph convolutional networks. *5th International Conference on Learning Representations, ICLR 2017 - Conference Track Proceedings* 1–14 (2017).
22. Veličković, P. *et al.* Graph attention networks. *6th International Conference on Learning Representations, ICLR 2018 - Conference Track Proceedings* 1–12 (2018).
23. Guo, N. *et al.* Interleukin-17 promotes migration and invasion of human cancer cells through upregulation of MTA1 expression. *Frontiers in Oncology* **9**, 1–11 (2019).
24. Huang, S., New, L., Pan, Z., Han, J. & Nemerow, G. R. Urokinase plasminogen activator/urokinase-specific surface receptor expression and matrix invasion by breast cancer cells requires constitutive p38α mitogen-activated protein kinase activity. *Journal of Biological Chemistry* **275**, 12266–12272 (2000).
25. Whyte, J., Bergin, O., Bianchi, A., McNally, S. & Martin, F. Key signalling nodes in mammary gland development and cancer. Mitogen-activated protein kinase signalling in experimental models of breast cancer progression and in mammary gland development. *Breast Cancer Research* **11**, 1–14 (2009).
26. Choi, A. H. *et al.* Endogenous Akt Activity Promotes Virus Entry and Predicts Efficacy of Novel Chimeric Orthopoxvirus in Triple-Negative Breast Cancer. *Molecular Therapy - Oncolytics* **9**, 22–29 (2018).
27. Ma, L. & Gan, L. AR Artificial Antagonist HC-1119 Regulated TNBC BT549 Cells Proliferation and Metastasis via PI3K / Akt / mTOR Signaling Pathway. 1–17.
28. Cheng, W., Liu, D., Guo, M., Li, H. & Wang, Q. Sophoraflavanone G suppresses the progression of triple-negative breast cancer via the inactivation of EGFR–PI3K–AKT signaling. *Drug Development Research* (2022).
29. Shahi Thakuri, P. *et al.* Phytochemicals inhibit migration of triple negative breast cancer cells by targeting kinase signaling. *BMC Cancer* **20**, 1–14 (2020).
30. Meng, L. *et al.* ZEB1-mediated transcriptional upregulation of circWWC3 promotes breast cancer progression through activating ras signaling pathway. *Molecular Therapy-Nucleic Acids* **22**, 124–137 (2020).
31. Kim, J. H. *et al.* Calsequestrin 2 overexpression in breast cancer increases tumorigenesis and metastasis by modulating the tumor microenvironment. *Molecular oncology* **16**, 466–484 (2022).
32. Makena, M. R. & Rao, R. Subtype specific targeting of calcium signaling in breast cancer. *Cell calcium* **85**, 102109 (2020).

33. White, C. The regulation of tumor cell invasion and metastasis by endoplasmic reticulum-to-mitochondrial Ca2+ transfer. *Frontiers in Oncology* **7**, 171 (2017).
34. Mumin, N. H. *et al.* Overcoming acquired resistance to HSP90 inhibition by targeting JAK-STAT signalling in triple-negative breast cancer. *BMC cancer* **19**, 1–14 (2019).
35. Manerba, M. *et al.* Metabolic activation triggered by cAMP in MCF-7 cells generates lethal vulnerability to combined oxamate/etomoxir. *Biochimica et Biophysica Acta (BBA)-General Subjects* **1863**, 1177–1186 (2019).
36. Zibara, K. *et al.* Signaling pathways activated by PACAP in MCF-7 breast cancer cells. *Cellular Signalling* **50**, 37–47 (2018).
37. Zhang, Y.-L., Wang, R.-C., Cheng, K., Ring, B. Z. & Su, L. Roles of Rap1 signaling in tumor cell migration and invasion. *Cancer biology & medicine* **14**, 90 (2017).
38. Curry, M. C., Luk, N. A., Kenny, P. A., Roberts-Thomson, S. J. & Monteith, G. R. Distinct regulation of cytoplasmic calcium signals and cell death pathways by different plasma membrane calcium ATPase isoforms in MDA-MB-231 breast cancer cells. *Journal of Biological Chemistry* **287**, 28598–28608 (2012).
39. Ma, B. *et al.* BRCA1 subcellular localization regulated by PI3K signaling pathway in triple-negative breast cancer MDA-MB-231 cells and hormone-sensitive T47D cells. *Open Life Sciences* **15**, 501–510 (2020).
40. Wagner, K.-U. Know thy cells: commonly used triple-negative human breast cancer cell lines carry mutations in RAS and effectors. *Breast Cancer Research* **24**, 1–2 (2022).
41. Bong, A. H. L. & Monteith, G. R. Calcium signaling and the therapeutic targeting of cancer cells. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* **1865**, 1786–1794 (2018).
42. Au, C. C. *et al.* Three-dimensional growth of breast cancer cells potentiates the anti-tumor effects of unacylated ghrelin and AZP-531. *Elife* **9**, e56913 (2020).
43. Niu, L. *et al.* Sericin inhibits MDA-MB-468 cell proliferation via the PI3K/Akt pathway in triple-negative breast cancer. *Molecular Medicine Reports* **23**, 1 (2021).
44. Peters, A. A. *et al.* Calcium Channel TRPV6 as a Potential Therapeutic Target in Estrogen Receptor–Negative Breast CancerCharacterization of TRPV6 in Breast Cancer. *Molecular cancer therapeutics* **11**, 2158–2168 (2012).
45. Bolanz, K. A., Hediger, M. A. & Landowski, C. P. The role of TRPV6 in breast carcinogenesis. *Molecular cancer therapeutics* **7**, 271–279 (2008).
46. Yuan, G. *et al.* Serum-derived hepatitis C virus can infect human glioblastoma cell line SF268 and activate the PI3K-Akt pathway. *Molecular Medicine Reports* **19**, 4441–4448 (2019).
47. Wu, M.-H. *et al.* Blockage of U251 cells in G0/G1 through MAPK signaling pathway by LRRC4. *Zhong nan da xue xue bao. Yi xue ban= Journal of Central South University. Medical Sciences* **32**, 226–230 (2007).
48. Wang, Y. *et al.* Maslinic Acid Inhibits the Growth of Malignant Gliomas by Inducing Apoptosis via MAPK Signaling. *Journal of Oncology* **2022**, (2022).
49. Jiang, H. *et al.* Resveratrol downregulates PI3K/Akt/mTOR signaling pathways in human U251 glioma cells. *Journal of experimental therapeutics & oncology* **8**, 25 (2009).

50. Li, F., Jin, D., Tang, C. & Gao, D. CEP55 promotes cell proliferation and inhibits apoptosis via the PI3K/Akt/p21 signaling pathway in human glioma U251 cells. *Oncology letters* **15**, 4789–4796 (2018).
51. Nie, J.-H. *et al.* Differential exosomic proteomic patterns and their influence in resveratrol sensitivities of glioblastoma cells. *International Journal of Molecular Sciences* **20**, 191 (2019).
52. Wang, H. *et al.* Downregulation of Pygopus 2 inhibits vascular mimicry in glioma U251 cells by suppressing the canonical Wnt signaling pathway. *Oncology Letters* **11**, 678–684 (2016).
53. Liu, L. *et al.* Catalpol promotes cellular apoptosis in human HCT116 colorectal cancer cells via microRNA-200 and the downregulation of PI3K-Akt signaling pathway. *Oncology Letters* **14**, 3741–3747 (2017).
54. Zhang, H., Sun, J., Ma, R. & Zhao, S. Role of Episamarcandin in Promoting the Apoptosis of Human Colon Cancer HCT116 Cells through the PI3K-Akt Signaling Pathway. *Evidence-Based Complementary and Alternative Medicine* **2021**, (2021).
55. Soo, H.-C. *et al.* Cudraflavone C induces tumor-specific apoptosis in colorectal cancer cells through inhibition of the phosphoinositide 3-kinase (PI3K)-AKT pathway. *PLoS One* **12**, e0170551 (2017).
56. Lin, G., Tang, Z., Ye, Y.-B. & Chen, Q. NF-κB activity is downregulated by KRAS knockdown in SW620 cells via the RAS-ERK-IκBα pathway. *Oncology reports* **27**, 1527–1534 (2012).
57. Kang, Q. & Yan, S. Piperlongumine reverses doxorubicin resistance through the PI3K/Akt signaling pathway in K562/A02 human leukemia cells. *Experimental and therapeutic medicine* **9**, 1345–1350 (2015).
58. Wang, H. *et al.* HOXB4 knockdown reverses multidrug resistance of human myelogenous leukemia K562/ADM cells by downregulating P-gp, MRP1 and BCRP expression via PI3K/Akt signaling pathway. *International Journal of Oncology* **49**, 2529–2537 (2016).
59. Al-Rawashde, F. A. *et al.* Thymoquinone Inhibits JAK/STAT and PI3K/Akt/mTOR Signaling Pathways in MV4-11 and K562 Myeloid Leukemia Cells. *Pharmaceuticals* **15**, 1123 (2022).
60. Ma, X. *et al.* MiR-486-5p-directed MAGI1/Rap1/RASSF5 signaling pathway contributes to hydroquinone-induced inhibition of erythroid differentiation in K562 cells. *Toxicology in Vitro* **66**, 104830 (2020).
61. Pick, T. *et al.* Remodelling of Ca2+ homeostasis is linked to enlarged endoplasmic reticulum in secretory cells. *Cell Calcium* **99**, 102473 (2021).
62. Lamharzi, N., Schally, A. V & Koppán, M. Luteinizing hormone–releasing hormone (LH–RH) antagonist Cetrorelix inhibits growth of DU-145 human androgen-independent prostate carcinoma in nude mice and suppresses the levels and mRNA expression of IGF-II in tumors. *Regulatory peptides* **77**, 185–192 (1998).
63. Zhang, H., Chen, J., Zeng, Z., Que, W. & Zhou, L. Knockdown of DEPTOR induces apoptosis, increases chemosensitivity to doxorubicin and suppresses autophagy in RPMI-8226 human multiple myeloma cells in vitro. *International Journal of Molecular Medicine* **31**, 1127–1134 (2013).
64. Tatur, S., Groulx, N., Orlov, S. N. & Grygorczyk, R. Ca2+-dependent ATP release from A549 cells involves synergistic autocrine stimulation by coreleased uridine nucleotides. *The Journal of physiology* **584**, 419–435 (2007).

65. Sun, D.-Z., Song, C.-Q., Xu, Y.-M. & Dong, X.-S. Role of the MAPK pathway in human lung epithelial-like A549 cells apoptosis induced by paraquat. *Genetics and Molecular Biology* **43**, (2020).
66. Yao, Q. *et al.* Curcumin induces the apoptosis of A549 cells via oxidative stress and MAPK signaling pathways. *International journal of molecular medicine* **36**, 1118–1126 (2015).
67. Fan, Q. *et al.* Optimizing component formula suppresses lung cancer by blocking DTL-mediated PDCD4 ubiquitination to regulate the MAPK/JNK pathway. *Journal of Ethnopharmacology* **299**, 115546 (2022).
68. Yuan, S. *et al.* Tumor-associated macrophages affect the biological behavior of lung adenocarcinoma A549 cells through the PI3K/AKT signaling pathway. *Oncology Letters* **18**, 1840–1846 (2019).
69. Kang, C., Wang, L., Wang, D., Zhang, X. & Chen, J. Lung cancer A549 cells suppressed with overexpressed HNF1B or PCDHA13 inhibited PI3K/AKT phosphorylation. *Translational Cancer Research* **9**, 3819 (2020).
70. Chen, J. *et al.* Polyphenols in Ilex latifolia Thunb. inhibit human lung cancer cell line A549 by regulation of the PI3K-Akt signaling pathway. *BMC Complementary Medicine and Therapies* **22**, 1–11 (2022).
71. Xiao, L. *et al.* Cytoplasmic RAP1 mediates cisplatin resistance of non-small cell lung cancer. *Cell death & disease* **8**, e2803–e2803 (2017).
72. Galan-Moya, E. M. *et al.* Balance between MKK6 and MKK3 mediates p38 MAPK associated resistance to cisplatin in NSCLC. *PloS one* **6**, e28406 (2011).
73. Wan Mohd Tajuddin, W. N. B., Abas, F., Othman, I. & Naidu, R. Molecular Mechanisms of Antiproliferative and Apoptosis Activity by 1, 5-Bis (4-Hydroxy-3-Methoxyphenyl) 1, 4-Pentadiene-3-one (MS13) on Human Non-Small Cell Lung Cancer Cells. *International Journal of Molecular Sciences* **22**, 7424 (2021).
74. Yu, H. *et al.* Knockdown of SNORA47 inhibits the tumorigenesis of NSCLC via mediation of PI3K/Akt signaling pathway. *Frontiers in Oncology* **11**, 620213 (2021).
75. Cai, T., Long, J., Wang, H., Liu, W. & Zhang, Y. Identification and characterization of miR-96, a potential biomarker of NSCLC, through bioinformatic analysis. *Oncology Reports* **38**, 1213–1223 (2017).
76. Odate, S. *et al.* TrkB/BDNF signaling pathway is a potential therapeutic target for pulmonary large cell neuroendocrine carcinoma. *Lung Cancer* **79**, 205–214 (2013).
77. Gao, T. *et al.* LncRNA HAND2-AS1 inhibits proliferation and promotes apoptosis of non-small cell lung cancer cells by inactivating PI3K/Akt pathway. *Bioscience Reports* **40**, (2020).
78. Busch, A. M. *et al.* Evidence for tankyrases as antineoplastic targets in lung cancer. *BMC cancer* **13**, 1–15 (2013).
79. Fattore, L. *et al.* Activation of an early feedback survival loop involving phospho-ErbB3 is a general response of melanoma cells to RAF/MEK inhibition and is abrogated by anti-ErbB3 antibodies. *Journal of translational medicine* **11**, 1–11 (2013).
80. Hao, H., Dong, Y. Bin, Bowling, M. T., Zhou, H. S. & McMASTERS, K. M. Alteration of gene expression in melanoma cells following combined treatment with E2F-1 and doxorubicin. *Anticancer research* **26**, 1947–1956 (2006).

81. Kwon, S.-J. *et al.* Isoegomaketone induces apoptosis in SK-MEL-2 human melanoma cells through mitochondrial apoptotic pathway via activating the PI3K/Akt pathway. *International journal of oncology* **45**, 1969–1976 (2014).
82. Irvine, M. *et al.* Oncogenic PI3K/AKT promotes the step-wise evolution of combination BRAF/MEK inhibitor resistance in melanoma. *Oncogenesis* **7**, 1–11 (2018).
83. Liu, W. *et al.* Taraxasterol attenuates melanoma progression via inactivation of reactive oxygen species-mediated PI3K/Akt signaling pathway. *Human & Experimental Toxicology* **41**, 09603271211069034 (2022).
84. Webster, M. R. *et al.* Bisphosphonamidate Clodronate Prodrug Exhibits Selective Cytotoxic Activity against Melanoma Cell LinesClodronate Prodrug Is Cytotoxic against Melanoma. *Molecular cancer therapeutics* **13**, 297–306 (2014).
85. Stewart, T. A., Yapa, K. T. D. S. & Monteith, G. R. Altered calcium signaling in cancer cells. *Biochimica et Biophysica Acta (BBA)-Biomembranes* **1848**, 2502–2511 (2015).
86. Khaled, M., Levy, C. & Fisher, D. E. Control of melanocyte differentiation by a MITF–PDE4D3 homeostatic circuit. *Genes & Development* **24**, 2276–2281 (2010).
87. Huang, H.-K. *et al.* Chemoresistant ovarian cancer enhances its migration abilities by increasing store-operated Ca2+ entry-mediated turnover of focal adhesions. *Journal of biomedical science* **27**, 1–14 (2020).
88. Yue, W. *et al.* The apoptotic resistance of BRCA1-deficient ovarian cancer cells is mediated by cAMP. *Frontiers in Cell and Developmental Biology* **10**, (2022).
89. Lee, K., Kim, S. & Lee, H. Orostachys japonicus induce p53-dependent cell cycle arrest through the MAPK signaling pathway in OVCAR-3 human ovarian cancer cells. *Food Science & Nutrition* **6**, 2395–2401 (2018).
90. Baghal-Sadriforoush, S., Bagheri, M., Rad, I. A. & Sotoodehnejadnematalahi, F. Melatonin Sensitizes OVCAR-3 Cells to Cisplatin through Suppression of PI3K/Akt Pathway. *Cellular and Molecular Biology* **68**, 158–169 (2022).
91. Wu, Y.-H., Huang, Y.-F., Chen, C.-C. & Chou, C.-Y. Akt inhibitor SC66 promotes cell sensitivity to cisplatin in chemoresistant ovarian cancer cells through inhibition of COL11A1 expression. *Cell death & disease* **10**, 1–16 (2019).
92. Huang, W.-L. *et al.* Thyrostimulin-TSHR signaling promotes the proliferation of NIH: OVCAR-3 ovarian cancer cells via trans-regulation of the EGFR pathway. *Scientific reports* **6**, 1–13 (2016).
93. Lau, M.-T., So, W.-K. & Leung, P. C. K. Fibroblast growth factor 2 induces E-cadherin down-regulation via PI3K/Akt/mTOR and MAPK/ERK signaling in ovarian cancer cells. *PloS one* **8**, e59083 (2013).
94. Cao, S. *et al.* Silencing of type Iγ phosphatidylinositol phosphate kinase suppresses ovarian cancer cell proliferation, migration and invasion Corrigendum in/10.3892/or. 2021.7982. *Oncology reports* **38**, 253–262 (2017).
95. Predescu, D.-V. *et al.* G protein-coupled receptors (GPCRs)-mediated calcium signaling in ovarian cancer: Focus on GPCRs activated by neurotransmitters and inflammation-associated molecules. *International journal of molecular sciences* **20**, 5568 (2019).
96. Zhang, L. *et al.* Aloesin suppresses cell growth and metastasis in ovarian cancer SKOV3 cells through the inhibition of the MAPK signaling pathway. *Analytical Cellular Pathology* **2017**, (2017).

97. Cao, S. *et al.* Silencing of type Iγ phosphatidylinositol phosphate kinase suppresses ovarian cancer cell proliferation, migration and invasion Corrigendum in/10.3892/or. 2021.7982. *Oncology reports* **38**, 253–262 (2017).
98. Zwartkruis, F. J. T., Wolthuis, R. M. F., Nabben, N. M. J. M., Franke, B. & Bos, J. L. Extracellular signal-regulated activation of Rap1 fails to interfere in Ras effector signalling. *The EMBO Journal* **17**, 5905–5912 (1998).
99. Young, T. W. *et al.* Up-regulation of tumor susceptibility gene 101 protein in ovarian carcinomas revealed by proteomics analyses. *Molecular & Cellular Proteomics* **6**, 294–304 (2007).
100. You, Z. *et al.* Interleukin-17 induces expression of chemokines and cytokines in prostatic epithelial cells but does not stimulate cell growth in vitro. *International journal of medical and biological frontiers* **18**, 629 (2012).
101. Zhang, Q. *et al.* Interleukin-17 promotes prostate cancer via MMP7-induced epithelial-to-mesenchymal transition. *Oncogene* **36**, 687–699 (2017).
102. Vitiello, G. A. & Miller, G. Targeting the interleukin-17 immune axis for cancer immunotherapy. *Journal of Experimental Medicine* **217**, (2020).
103. McGeachy, M. J., Cua, D. J. & Gaffen, S. L. The IL-17 family of cytokines in health and disease. *Immunity* **50**, 892–906 (2019).
104. Wu, N. *et al.* Cathepsin K regulates the tumor growth and metastasis by IL-17/CTSK/EMT axis and mediates M2 macrophage polarization in castration-resistant prostate cancer. *Cell Death & Disease* **13**, 1–9 (2022).
105. Li, X. *et al.* Inhibition of prostate cancer DU-145 cells proliferation by anthopleura anjunae oligopeptide (YVPGP) via PI3K/AKT/mTOR signaling pathway. *Marine Drugs* **16**, 325 (2018).
106. Shin, D. Y. *et al.* Apoptosis induction of human prostate carcinoma DU145 cells by diallyl disulfide via modulation of JNK and PI3K/AKT signaling pathways. *International journal of molecular sciences* **13**, 14158–14171 (2012).
107. Zeng, Y. & Yang, Y. Piperine depresses the migration progression via downregulating the Akt/mTOR/MMP-9 signaling pathway in DU145 cells. *Molecular medicine reports* **17**, 6363–6370 (2018).
108. Yang, T. & Wang, S. Novel PI3K inhibitor 5d inhibits prostate cancer cell growth in vitro and in vivo. 1–11 (2022).
109. Choo, G.-S., Lee, H.-N., Shin, S.-A., Kim, H.-J. & Jung, J.-Y. Anticancer effect of fucoidan on DU-145 prostate cancer cells through inhibition of PI3K/Akt and MAPK pathway expression. *Marine Drugs* **14**, 126 (2016).
110. Park, J. S. *et al.* Farnesol induces apoptosis of DU145 prostate cancer cells through the PI3K/Akt and MAPK pathways. *International journal of molecular medicine* **33**, 1169–1176 (2014).
111. Gopalakrishnan, A. *et al.* Modulation of activator protein-1 (AP-1) and MAPK pathway by flavonoids in human prostate cancer PC3 cells. *Archives of pharmacal research* **29**, 633–644 (2006).
112. Kim, S.-O. Effect of Indole-3-Carbinol on Inhibition of MMP Activity via MAPK Signaling Pathway in Human Prostate Cancer Cell Line, PC3 Cells. *Journal of Nutrition and Health* **41**, 224–231 (2008).

113. Arisan, E. D. *et al.* The role of the PI3K/AKT/mTOR signaling axis in the decision of the celastrol-induced cell death mechanism related to the lipid regulatory pathway in prostate cancer cells. *Phytochemistry Letters* **39**, 73–83 (2020).
114. Wang, Z. *et al.* DT-13 inhibits proliferation and metastasis of human prostate cancer cells through blocking PI3K/Akt pathway. *Frontiers in pharmacology* **9**, 1450 (2018).
115. Bailey, C. L., Kelly, P. & Casey, P. J. Activation of Rap1 promotes prostate cancer metastasis. *Cancer research* **69**, 4962–4968 (2009).
116. Pan, W., Bodempudi, V., Esfandyari, T. & Farassati, F. Utilizing ras signaling pathway to direct selective replication of herpes simplex virus-1. *PLoS One* **4**, e6514 (2009).
117. Chen, W.-Y. *et al.* MicroRNA-34a regulates WNT/TCF7 signaling and inhibits bone metastasis in Ras-activated prostate cancer. *Oncotarget* **6**, 441 (2015).
118. Elzamzamy, O. M. *Targeting Calcium Homeostasis for the Treatment of Multiple Myeloma*. (West Virginia University, 2021).
119. Zhu, Z. *et al.* STING Suppresses Mitochondrial VDAC2 to Govern RCC Growth Independent of Innate Immunity. *Advanced Science* 2203718 (2022).
120. Li, W. *et al.* YTHDC1 is downregulated by the YY1/HDAC2 complex and controls the sensitivity of ccRCC to sunitinib by targeting the ANXA1-MAPK pathway. *Journal of Experimental & Clinical Cancer Research* **41**, 1–18 (2022).
121. Zhu, Y., Rao, Q., Zhang, X. & Zhou, X. Galangin induced antitumor effects in human kidney tumor cells mediated via mitochondrial mediated apoptosis, inhibition of cell migration and invasion and targeting PI3K/AKT/mTOR signalling pathway. *J. BUON* **23**, 795–799 (2018).
122. Wang, X. *et al.* Chlorogenic acid inhibits proliferation and induces apoptosis in A498 human kidney cancer cells via inactivating PI3K/Akt/mTOR signalling pathway. *Journal of Pharmacy and Pharmacology* **71**, 1100–1109 (2019).
123. Kim, J. *et al.* WNK1 promotes renal tumor progression by activating TRPC6-NFAT pathway. *The FASEB Journal* **33**, 8588–8599 (2019).
124. Xiao, C. *et al.* RNF7 inhibits apoptosis and sunitinib sensitivity and promotes glycolysis in renal cell carcinoma via the SOCS1/JAK/STAT3 feedback loop. *Cellular & Molecular Biology Letters* **27**, 1–22 (2022).
125. He, Z. H., Guo, F., Hu, X. X., Luo, Z. Y. & Yi, J. W. Knockdown of HMGB2 inhibits proliferation and invasion of renal tumor cells via the p-38MAPK pathway. *Eur Rev Med Pharmacol Sci* **24**, 4729–4737 (2020).
126. Sun, P. *et al.* MicroRNA-195 targets VEGFR2 and has a tumor suppressive role in ACHN cells via PI3K/Akt and Raf/MEK/ERK signaling pathways. *International Journal of Oncology* **49**, 1155–1163 (2016).
127. Khan, M. I. *et al.* Comparative gene expression profiling of primary and metastatic renal cell carcinoma stem cell-like cancer cells. *PLoS One* **11**, e0165718 (2016).
128. Jung, J. E. *et al.* Caffeic acid and its synthetic derivative CADPE suppress tumor angiogenesis by blocking STAT3-mediated VEGF expression in human renal carcinoma cells. *Carcinogenesis* **28**, 1780–1787 (2007).
129. Sourbier, C. *et al.* The phosphoinositide 3-kinase/Akt pathway: a new target in human renal cell carcinoma therapy. *Cancer research* **66**, 5130–5142 (2006).

130. Dong, X. *et al.* RAP1GAP inhibits cytoskeletal remodeling and motility in thyroid cancer cells. *Endocrine-related cancer* **19**, 575 (2012).
131. Mygland, L. *et al.* Identification of response signatures for tankyrase inhibitor treatment in tumor cell lines. *Iscience* **24**, 102807 (2021).

Appendix A

Algorithm: $\boldsymbol{k}$-th hop adjacency matrix

Input:	initial adjacency matrix $A \in \mathbb{R}^{n \times n}$
Step 1:	$A_{sum} = A$
Step 2:	For k in $2, 3, \ldots, K$:
Step 3:	$A^{(k)} = A^k$
Step 4:	For i in $1, 2, 3, \ldots, n$:
Step 5:	For j in $1, 2, 3, \ldots, n$:
Step 6:	If $A_{ij}^{(k)} > 1$:
Step 7:	$A_{ij}^{(k)} = 1$
Step 8:	End If
Step 9:	If $i = j$:
Step 10:	$A_{ij}^{(k)} = A_{ij}^{(k)} - (A_{sum})_{ij} - 1$
Step 11:	Else:
Step 12:	$A_{ij}^{(k)} = A_{ij}^{(k)} - (A_{sum})_{ij}$
Step 13:	End if
Step 14:	If $A_{ij}^{(k)} < 0$:
Step 15:	$A_{ij}^{(k)} = 0$
Step 16:	End If
Step 17:	End For
Step 18:	End For
Step 19:	$A_{sum} = A_{sum} + A^{(k)}$
Step 20:	End For
Output:	$A^{(2)}, A^{(3)}, \ldots, A^{(K)}$

Appendix B

Following is the video link for our visualization tool ***NetFlowVis***:

https://www.youtube.com/watch?v=qL6nS770NMI&ab_channel=HemingZhang