

## Mining signaling flow to interpret mechanisms of synergy of drug combinations using deep graph neural networks

Heming Zhang<sup>1</sup>, Yixin Chen<sup>1</sup>, Philip Payne<sup>2</sup>, Fuhai Li<sup>2,3\*</sup>

<sup>1</sup>Computer Science, <sup>2</sup>Institute for Informatics (I2), Washington University School of Medicine,

<sup>3</sup>Department of Pediatrics, Washington University School of Medicine, Washington University in St. Louis, St. Louis, MO, USA. \*Correspondence: [Fuhai.Li@wustl.edu](mailto:Fuhai.Li@wustl.edu)

### Abstract

Complex signaling pathways/networks are believed to be responsible for drug resistance in cancer therapy. Drug combinations inhibiting multiple signaling targets within cancer-related signaling networks have the potential to reduce drug resistance. Deep learning models have been reported to predict drug combinations. However, these models are hard to be interpreted in terms of mechanism of synergy (MoS), and thus cannot well support the human-AI based clinical decision making. Herein, we proposed a novel computational model, *DeepSignalingFlow*, which seeks to address the preceding two challenges. Specifically, a graph convolutional network (GCN) was developed based on a core cancer signaling network consisting of 1584 genes, with gene expression and copy number data derived from 46 core cancer signaling pathways. 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. The numerical features (accumulated information due to the signaling-flows of the signaling network) of drug nodes that link to drug targets were then used to predict the synergy scores of such drug combinations. The model was evaluated using the NCI ALMANAC drug combination screening data. The evaluation results showed that the proposed

*DeepSignalingFlow* model can not only predict drug combination synergy score, but also interpret potentially interpretable MoS of drug combinations.

## 1. Introduction

Complex signaling pathways and ensuing networks are believed to be responsible for drug resistance in cancer therapy. Thus, drug combinations inhibiting multiple signaling targets in cancer-related signaling networks have the potential to be effective in terms of reduce drug resistance. However, it is challenging to discover effective and synergistic drug combinations in clinical and laboratories 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<sup>1</sup>. Further, in recent studies, 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<sup>2,3</sup>. This is notable, given that autophagy signaling was activated after treatment with RAS inhibitors<sup>2,3</sup>. 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 to ultimately improve cancer treatment outcomes.

In response to the aforementioned challenges, and to help facilitate the discovery of synergistic drug combinations for use in a variety of cancer, we sought to understand the associations between genetic biomarkers and drug response, as well as mechanism of drug synergy, using computational methods. The National Cancer Institute (NCI) has previously generated a comprehensive screening data set of ~5,232 combinations from ~100 drugs in 60 human tumor cell-lines (NCI-60), which is publicly available as the NCI-ALMANAC Drug Combination Data Set<sup>4</sup>. This dataset is valuable for developing computational models for drug combination prediction. There have been a variety of computational methods described in the scientific literature concerning the application of machine learning techniques to such data. For

example, matchmaker<sup>5</sup> was developed to analyze chemical and gene expression profiles using a deep learning framework, which is similar as the DeepSynergy<sup>6</sup> and AuDNNsynergy<sup>7</sup> models. Similarly, a simplified deep learning model, DeepSignalingSynergy, has been proposed to investigate the discriminative importance of individual signaling pathways for drug combination prediction<sup>8,9</sup>. In addition, the use of multi-task learning with transfer learning model has been proposed to prediction drug combinations<sup>10</sup>. Further, random forest (RF) and XGBoost models have been used, based on the chemical and genomics features of source data, to predict drug combinations<sup>11</sup>. These and many other computational models have also been proposed for drug combination prediction (as opposed to predicting the efficacy of single drugs). For example, the connectivity map (CMAP)<sup>12,13</sup> and associated network analysis based drug combination prediction models have been used in this context<sup>14,15</sup>. In an analogous manner, message propagation-based models have been developed based on the confluence of drug targeting and genomics data<sup>16,17</sup>. A common thread in the preceding work has been the use of semi-supervised learning models, applied to multiple pharmacogenomics datasets, for drug combination prediction<sup>18</sup>.

However, despite the promise of the existing body of work in this domain, there remain three limitations of the existing deep learning models. First, a large number of genomic and chemical features have been used in such predictive models. Given the limited amount of drug screening data currently available, the training of such models with a large number of features is intrinsically limited. Second, the signaling interactions in gene regulatory signaling networks are often not incorporated in the deep learning models, despite their criticality in terms of treatment efficacy. Third, complex mechanism of synergy (MoS) are usually not modeled in these existing constructs. As such, these models are hard to interpret and explain in order to identify effective and synergistic drug combinations for subsequent verification and validation.

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)<sup>19</sup>. 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 predication, by incorporating core cancer signaling pathways into such analyses.

## 2. Methodology

### 2.1 NCI ALMANAC drug combination screening dataset

As was introduced previously, the NCI Almanac dataset includes the combo-scores<sup>1</sup> of permutations of 104 FDA approved drugs relative to the tumor growth of NCI60 human tumor cell lines. The average combo-score of two drugs with different doses on a given tumor cell line was used as the basis for assessing the synergy score of two drugs on the given tumor cell line using a 4-element tuple:  $\langle D_A, D_B, C_C, S_{ABC} \rangle$ .

### 2.2 Gene expression and copy number data of NCI-60 Cancer Cell Lines

Multi-omics data (e.g., RNA-seq [gene expression], copy number variation, metabolomics, miRNA, and RPPA\_) for 1000 human tumor cell lines were available in the Cancer Cell Line Encyclopedia (CCLE) database<sup>20</sup>, which are of value when seeking to identify associations between genetic biomarkers and drug response. In addition, copy number data was downloaded from GDSC database<sup>21</sup>. By identifying overlapping cell lines from these two databases, we identified 43 of NCI-60 cancer cell lines that were also present in CCLE (see **Table 1**).

**Table 1:** 43 NCI-60 human tumor cell lines included in CCLE database.

A498_KIDNEY	A549_LUNG	ACHN_KIDNEY	BT549_BREAST	CAKI1_KIDNEY
HOP62_LUNG	HOP92_LUNG	HS578T_BREAST	HT29_LARGE_INT ESTINE	IGROV1_OVARY
K562_HAEMATOP OIETIC_AND_LYM PHOID_TISSUE	KM12_LARGE_INT ESTINE	LOXIMVI_SKIN	MCF7_BREAST	MDAMB231_BREA ST
MDAMB468_BREA ST	NCIH226_LUNG	NCIH23_LUNG	NCIH322_LUNG	NCIH460_LUNG
NCIH522_LUNG	NIHOVCAR3_OVA RY	OVCAR4_OVARY	OVCAR8_OVARY	PC3_PROSTATE
RPMI8226_HAEMA TOPOIETIC_AND_L YMPHOID_TISSUE	SF268_CENTRAL_ NERVOUS_SYSTEM	SF295_CENTRAL_N ERVOUS_SYSTEM	SF539_CENTRAL_ NERVOUS_SYSTEM	SKMEL28_SKIN
SKMEL5_SKIN	SKOV3_OVARY	SW620_LARGE_INT ESTINE	T47D_BREAST	U251MG_CENTRA L_NERVOUS_SYS TEM
UACC257_SKIN	UACC62_SKIN	UO31_KIDNEY		

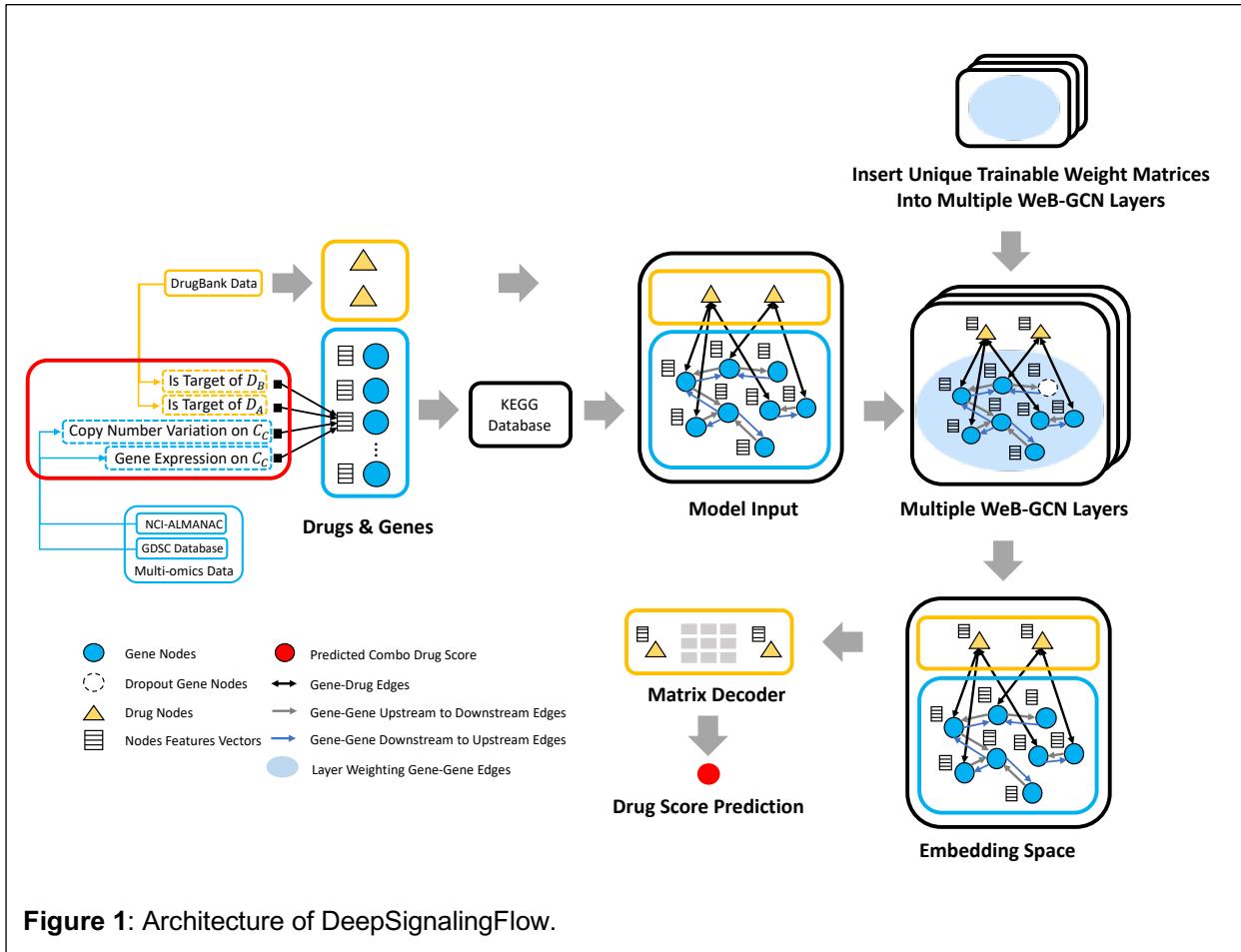
### 2.3 KEGG signaling pathways and cellular process

46 signaling pathways (45 “signaling pathways” + cell cycle)<sup>8,9</sup> were collected from the KEGG (Kyoto Encyclopedia of Genes and Genomes)<sup>22</sup> database, specifically: MAPK, FoxO, TGF-beta, T cell receptor, Adipocytokine, ErbB, Sphingolipid, VEGF, B cell receptor, Oxytocin, Ras, Phospholipase D, Apelin, Fc epsilon RI, Glucagon, Rap1, p53, Hippo, TNF, Relaxin, Calcium, mTOR, Toll-like receptor, Neurotrophin, AGE-RAGE, cGMP-PKG, PI3K-Akt, NOD-like receptor, Insulin, Cell cycle, cAMP, AMPK, RIG-I-like receptor, GnRH, Chemokine, Wnt, C-type lectin receptor, Estrogen, NF-kappa B, Notch, JAK-STAT, Prolactin, HIF-1, Hedgehog, IL-17, Thyroid hormone signaling pathways. In these 46 signaling pathways<sup>8,9</sup>, 1584 genes were identified for which gene expression (TPM) and copy number data were available in the CCLE database.

### 2.3 Drug-Target interactions from DrugBank database

Drug-target information were extracted from the DrugBank<sup>23</sup> 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. Among these drug, 67 drugs were included in NCI ALMANAC with

associated targets. Further, 21 drugs with known targets corresponding to the 1584 signaling pathways previously identified were selected for use in our model, specifically: Celecoxib, Gefitinib, Quinacrine hydrochloride, Tretinoin, Cladribine, Imatinib mesylate, Romidepsin, Vinblastine sulfate (hydrate), Dasatinib, Lenalidomide, Sirolimus, Vorinostat, Docetaxel, Mitotane, Sorafenib tosylate, Thalidomide, Everolimus, Nilotinib, Tamoxifen citrate, Paclitaxel, Fulvestrant.



**Figure 1:** Architecture of DeepSignalingFlow.

## 2.4 Architecture design of DeepSignalingFlow

**Fig. 1** shows 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 adjacent matrix that links gene-gene

interactions on the signaling pathways and gene-drug interactions, and the elements of adjacent 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 in directed graph.

In the graph convolution stages of our architecture, we added the bidirectional nodes-flow with both of ‘upstream-to-downstream’ (from up-stream signaling to drug targets) and ‘downstream-to-upstream’ (from drug targets to down-stream signaling). Therefore, we call this model a Weight Bi-directional Graph Convolution Network (WeB-GCN). The weight matrices are denoted as:  $W_{up}^{(k)} \in \mathbb{R}^{n \times n}$ ,  $W_{down}^{(k)} \in \mathbb{R}^{n \times n}$ , ( $k = 1, 2, \dots, K$ ) respectively. And then the weight adjacency matrices will be denoted as:

$$A'_{up}^{(k)} = W_{up}^{(k)} \cdot A \quad (1)$$

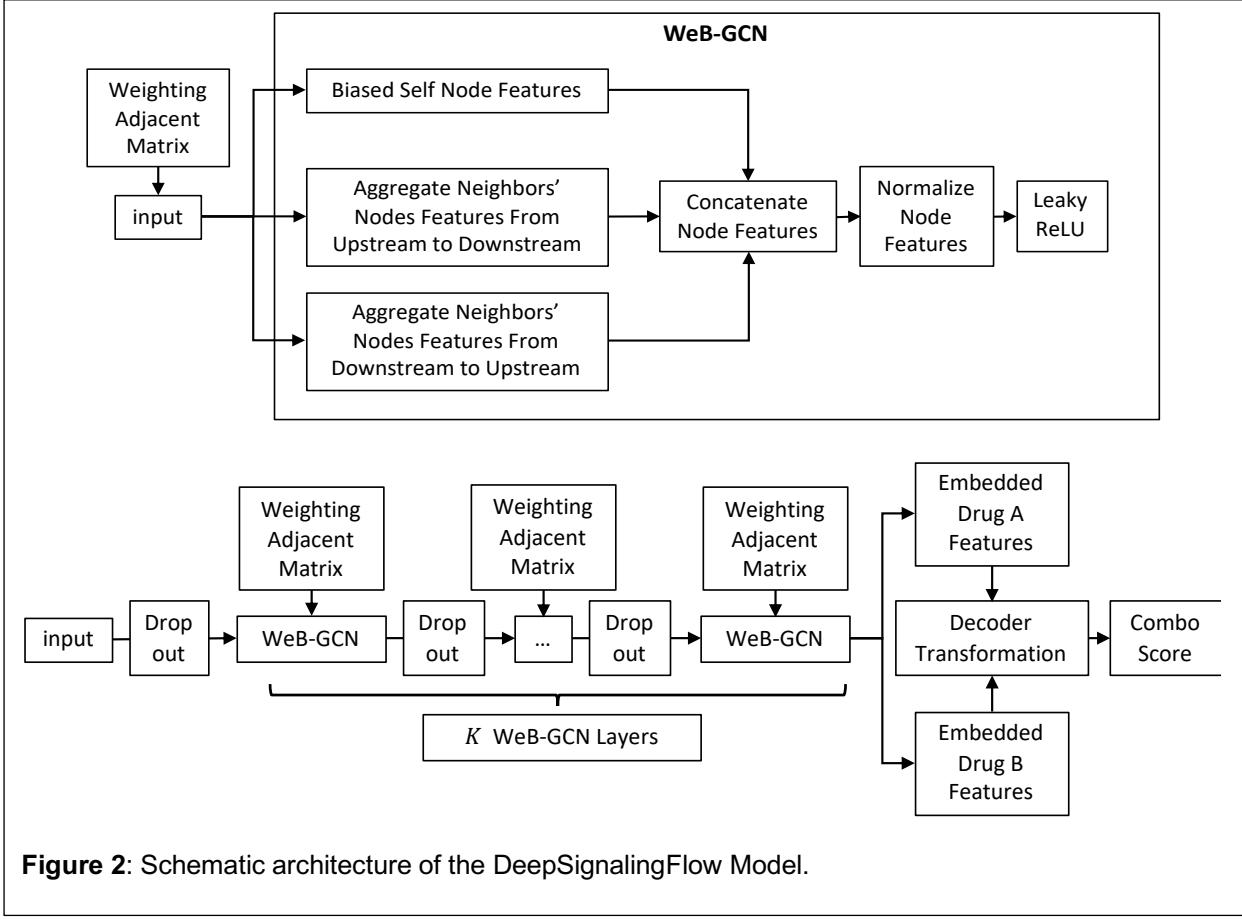
$$A'_{down}^{(k)} = W_{down}^{(k)} \cdot A \quad (2)$$

where for each  $A'_{up}^{(k)} \in \mathbb{R}^{n \times n}$ ,  $A'_{down}^{(k)} \in \mathbb{R}^{n \times n}$ , ( $k = 1, 2, \dots, K$ ), parameters for drug-gene relations values are the same with those in original adjacent matrix. Complete model parameters are specified in **Table 2**.

**Table 2:** Model parameters and notations.

Input Matrices	When $k = 2, 3, \dots, K$	When $k = 1$
Nodes Features Matrix	$H^{(k-1)} \in \mathbb{R}^{n \times 4d^{(k-1)}}$	$H^{(0)} \in \mathbb{R}^{n \times d^{(0)}}$
New Nodes Adjacent Matrices	$A'_{up}^{(k)} \in \mathbb{R}^{n \times n}, A'_{down}^{(k)} \in \mathbb{R}^{n \times n}$	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	$M_{up}^{(k)} \in \mathbb{R}^{4d^{(k-1)} \times d^{(k)}}$ $M_{down}^{(k)} \in \mathbb{R}^{4d^{(k-1)} \times d^{(k)}}$	$M_{up}^{(1)} \in \mathbb{R}^{d^{(0)} \times d^{(1)}}$ $M_{down}^{(1)} \in \mathbb{R}^{d^{(0)} \times d^{(1)}}$
Trainable Biased Feature Transformation Matrix	$B^{(k)} \in \mathbb{R}^{4d^{(k-1)} \times d^{(k)}}$	$B^{(1)} \in \mathbb{R}^{d^{(0)} \times d^{(1)}}$



**Figure 2:** Schematic architecture of the DeepSignalingFlow Model.

**Fig. 2** shows the structural details of our model. First, the mean aggregation of each node neighbors' features from upstream to downstream is achieved by:

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

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

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

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

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

For  $k = 1, 2, \dots, K$ , the model uses normalization function  $N: \mathbb{R}^{n \times 4d^{(k)}} \rightarrow \mathbb{R}^{n \times 4d^{(k)}}$  and LeakyReLU function with parameter  $\alpha$  to map concatenated nodes features  $H_{concat}^{(k)} = (H_{up\_neigh}^{(k)} | H_{down\_neigh}^{(k)} | H_{self}^{(k)}) \in \mathbb{R}^{n \times 4d^{(k)}}$  to  $H^{(k)} \in \mathbb{R}^{n \times 4d^{(k)}}$  with the equation:

$$H^{(k)} = \text{LeakyReLU}(N(H_{concat}^{(k)})) \quad (6)$$

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

$$v'_{ij} = \frac{v_{ij}}{\sqrt{\sum_{j=1}^q v_{ij}^2}} \quad (7)$$

where  $v'_{ij}$  is the element of new matrix  $V' \in \mathbb{R}^{p \times q}$ . Therefore, the output of  $k$ -th layer WeB-GCN is  $H^{(k)}$ . After obtaining the embedded nodes features with  $H^{(K)} \in \mathbb{R}^{n \times 4d^{(K)}}$ , the embedded nodes features of drug\_A are obtained:  $H_A^{(K)} \in \mathbb{R}^{4d^{(K)}}$ , and the drug\_B embedded nodes features are obtained as:  $H_B^{(K)} \in \mathbb{R}^{4d^{(K)}}$ , which are just the 1585<sup>th</sup> and 1586<sup>th</sup> rows in  $H^{(K)}$ . Based on the idea of decagon decoder<sup>24</sup>, the prediction of combo score will be calculated in following equation:

$$g(H_A^{(K)}, H_B^{(K)}) = H_A^{(K)} D Q D^T H_B^{(K)} \quad (8)$$

where  $D \in \mathbb{R}^{4d^{(K)} \times N}$  and  $Q \in \mathbb{R}^{N \times N}$  are trainable decoder matrices.

## Drug combination response interpretation using the directional signaling flow

As aforementioned, we added directional signaling interaction weights to indicate the potential signaling flow from up-stream signaling to drug targets, and from drug targets to down-stream signaling. The trainable directional weights can potentially indicate and interpret the signaling flow

on the signaling network to affect the drug combination response. Specifically, the signaling flow derived from the directional weighted matrices are defined as follows. Each such model had  $k=3$  layers, and each layer has upstream weight adjacent parameters and downstream weight adjacent parameters, which are  $A'^{(1)}_{up}, A'^{(2)}_{up}, A'^{(3)}_{up}$  and  $A'^{(1)}_{down}, A'^{(2)}_{down}, A'^{(3)}_{down}$ . Through aggregating 3 layers upstream and downstream to weight adjacent matrices, we arrive at an aggregated upstream, downstream and bind-stream for  $k$ th 5-fold cross valuation model with  $k$ th model's  $A'^{(1)}_{up}, A'^{(2)}_{up}, A'^{(3)}_{up}$  and  $A'^{(1)}_{down}, A'^{(2)}_{down}, A'^{(3)}_{down}$  using the formula:

$$A'^{(k)}_{upagg} = \frac{1}{3} (A'^{(1)}_{up} + A'^{(2)}_{up} + A'^{(3)}_{up}) \quad (9)$$

$$A'^{(k)}_{downagg} = \frac{1}{3} (A'^{(1)}_{down} + A'^{(2)}_{down} + A'^{(3)}_{down}) \quad (10)$$

$$A'^{(k)}_{bindagg} = \frac{1}{3} \left[ \frac{1}{2} (A'^{(1)}_{up} + A'^{(1)}_{down}) + \frac{1}{2} (A'^{(2)}_{up} + A'^{(2)}_{down}) + \frac{1}{2} (A'^{(3)}_{up} + A'^{(3)}_{down}) \right] \quad (11)$$

With upstream, downstream and bind-stream weight adjacent matrices parameters, we then aggregated weight adjacent matrices of those 5 models by the following formula:

$$A'_{up\_final} = \frac{1}{5} (A'^{(1)}_{upagg} + A'^{(2)}_{upagg} + A'^{(3)}_{upagg} + A'^{(4)}_{upagg} + A'^{(5)}_{upagg}) \quad (12)$$

$$A'_{down\_final} = \frac{1}{5} (A'^{(1)}_{downagg} + A'^{(2)}_{downagg} + A'^{(3)}_{downagg} + A'^{(4)}_{downagg} + A'^{(5)}_{downagg}) \quad (13)$$

$$A'_{bind\_final} = \frac{1}{5} (A'^{(1)}_{bindagg} + A'^{(2)}_{bindagg} + A'^{(3)}_{bindagg} + A'^{(4)}_{bindagg} + A'^{(5)}_{bindagg}) \quad (14)$$

With the final upstream, downstream and bind-stream weight adjacent matrices parameters for 1584 genes of  $A'_{up\_final}, A'_{down\_final}, A'_{bind\_final}$ , we then were able to study the signaling flow on the signaling network.

### 3. Results

#### 3.1 Experimental Setup

Model input: <drug\_A, drug\_B, cellline name, combo score> were used to construct each point in an input dataset, which was comprised of 5445 such points; among these data, using a 5-fold cross validation with 5 splits of 1089 nodes, we created a training dataset. For each point, the cell line name contained the RNA sequence of the 1584 genes. Given all possible 2-drug permutations, each point created a graph with 1586 nodes, and each such graph had 3 initial features. For those 1584 genes, their 3 features indicate the RNA sequence number, and whether they have connection spanning drug\_A and drug\_B. For each 2-drug dyad, their 3 features were initiated with zeros. Furthermore, to indicate connections between nodes, adjacent matrices were also created. Those matrices were formed from file 'Selected\_Kegg\_Pathways\_edges\_1', 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 adjacent matrices, those elements are symmetric. At the same time, in-degree matrices are formed according to said adjacent matrices.

#### 3.2 Hyperparameters Setting

Subsequently, a model was developed by using pytorch. Using the Adam optimizer, learning rate started with 0.01 and reduced equally within each batch of first 50 epochs. We empirically set the training epochs as 75, which allows for optimal validation results.  $K = 3$  WeB-GCN layers were used, and the initial input feature dimension was:  $d^{(0)} = 3$ , with the feature dimensions varying at the different layers, and denoted by:  $(d^{(1)}, d^{(2)}, \dots, d^{(k)}, \dots, d^{(K)})$ ,  $k = 1, 2, 3$ . WeB-GCN concatenate biased last layer node features and transformed nodes features, with the final concatenated

dimensions of the WeB-GCN layers for the output dims being  $3 \times d^{(k)}$ . Output dims of the last layer served as the input dims for this layer, as follows:

- (1) First layer (input dims, output dims):  $(d^{(0)}, 4d^{(1)})$
- (2) Second layer (input dims, output dims):  $(4d^{(1)}, 4d^{(2)})$
- (3) Third layer (input dims, output dims):  $(4d^{(2)}, 4d^{(3)})$

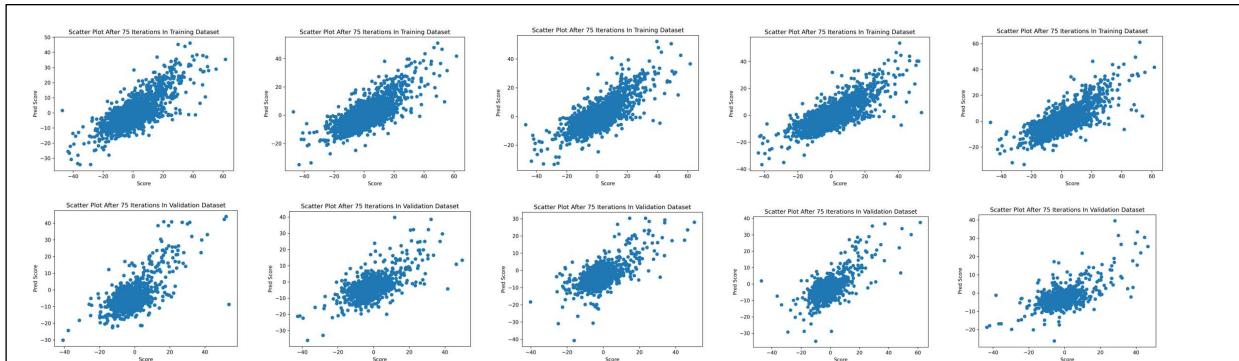
The final embedded drug nodes dims were  $4d^{(3)}$  (see **Table 2**). Decoder trainable transformation matrix dims:  $D \in \mathbb{R}^{2d^{(K)} \times N}$  and  $Q \in \mathbb{R}^{N \times N}$  were used as trainable decoder matrices, with changeable parameters of  $N$  to adapt model performance. As shown in **Table 2**, the model uses  $N = 40$ . As for the LeakyReLU function, the parameter  $\alpha = 0.1$ .

**Table 2:** 5-fold Cross Validation on 2 Sets of Hyper-parameters

Number of Dataset	Input Dims, Output Dims			Decoder dims	Dropout	Training MSE Loss	Training Pearson Correlation	Test MSE Loss	Test Pearson Correlation
	1 <sup>st</sup> layer	2 <sup>nd</sup> layer	3 <sup>rd</sup> layer						
1 <sup>st</sup> dataset	(4, 12)	(12, 12)	(12, 36)	40	0.01	34.11	74.64%	53.34	68.20%
2 <sup>nd</sup> dataset	(4, 12)	(12, 12)	(12, 36)	40	0.01	33.48	74.40%	51.91	64.21%
3 <sup>rd</sup> dataset	(4, 12)	(12, 12)	(12, 36)	40	0.01	33.01	75.75%	44.59	66.64%
4 <sup>th</sup> dataset	(4, 12)	(12, 12)	(12, 36)	40	0.01	32.78	75.71%	42.19	67.93%
5 <sup>th</sup> dataset	(4, 12)	(12, 12)	(12, 36)	40	0.01	35.38	73.39%	48.01	62.48%
Average	(4, 12)	(12, 12)	(12, 36)	40	0.01	<b>33.76</b>	<b>74.78%</b>	<b>48.01</b>	<b>65.89%</b>

### 3.3 Synergistic score prediction evaluation

To evaluate the model performance in terms of synergy score prediction for drug combinations, we conducted 5-fold validation. As seen, the average prediction (using the Pearson correlation coefficient), was about 66% accuracy using the testing data. These prediction results are comparable with existing deep learning models<sup>6,7</sup>. The results indicated the feasibility of drug combination synergy score prediction using a graph neural network with a small set of core signaling pathways genes, compared with the existing complex deep neural network (DNN) models using a large number of genomics and chemical structure features.



**Figure 3:** Scatterplots of the predicted and experimental synergy scores on the 5-fold training and validation dataset with dropout rate 0.01

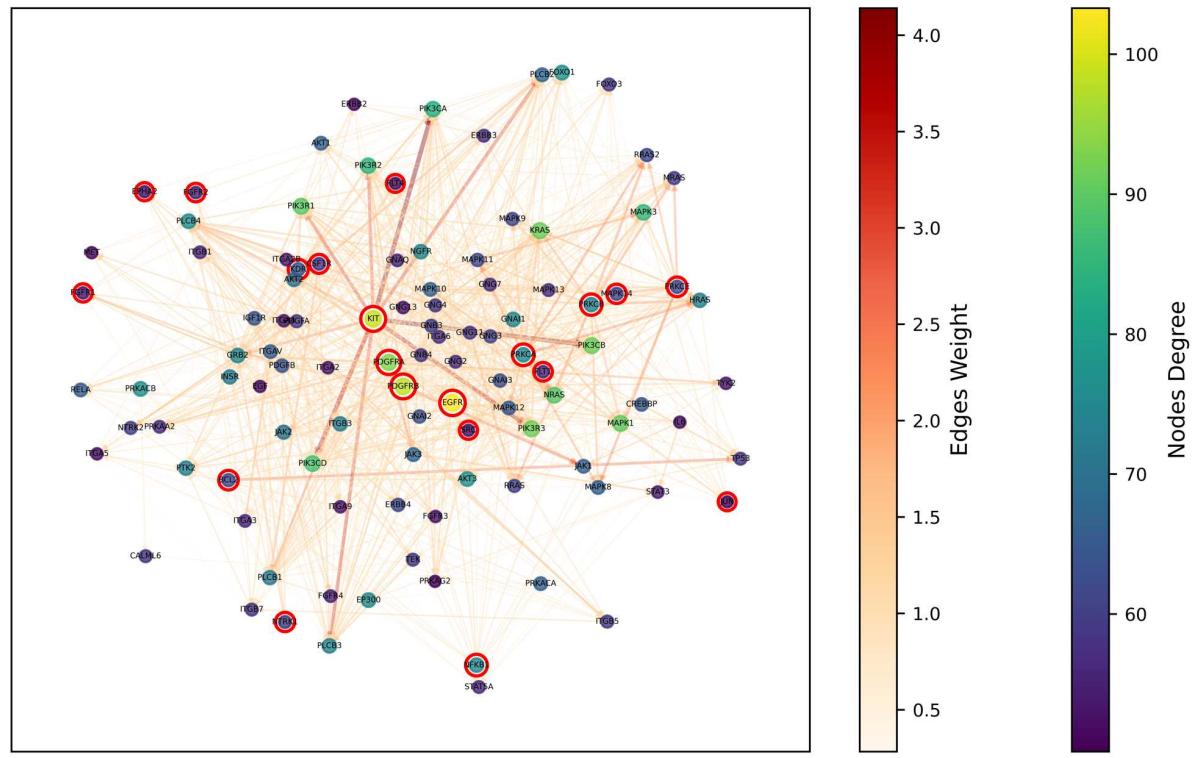
### 3.3 Signaling flow interpreting the mechanism of synergy

As aforementioned, we added directional signaling interaction weights to indicate the potential signaling flow from up-stream signaling to drug targets, and from drug targets to down-stream signaling. In this sub-aim, we identified a set of important sub-network, including important signaling interactions, to show the feasibility of model interpretation. As an example, first, the signaling interactions (edges) were filtered, based on the trained signaling interaction weights, by selecting the signaling interactions with weights more than a given threshed, which indicates the important signaling flow on the signaling network. Then network nodes with degrees larger than

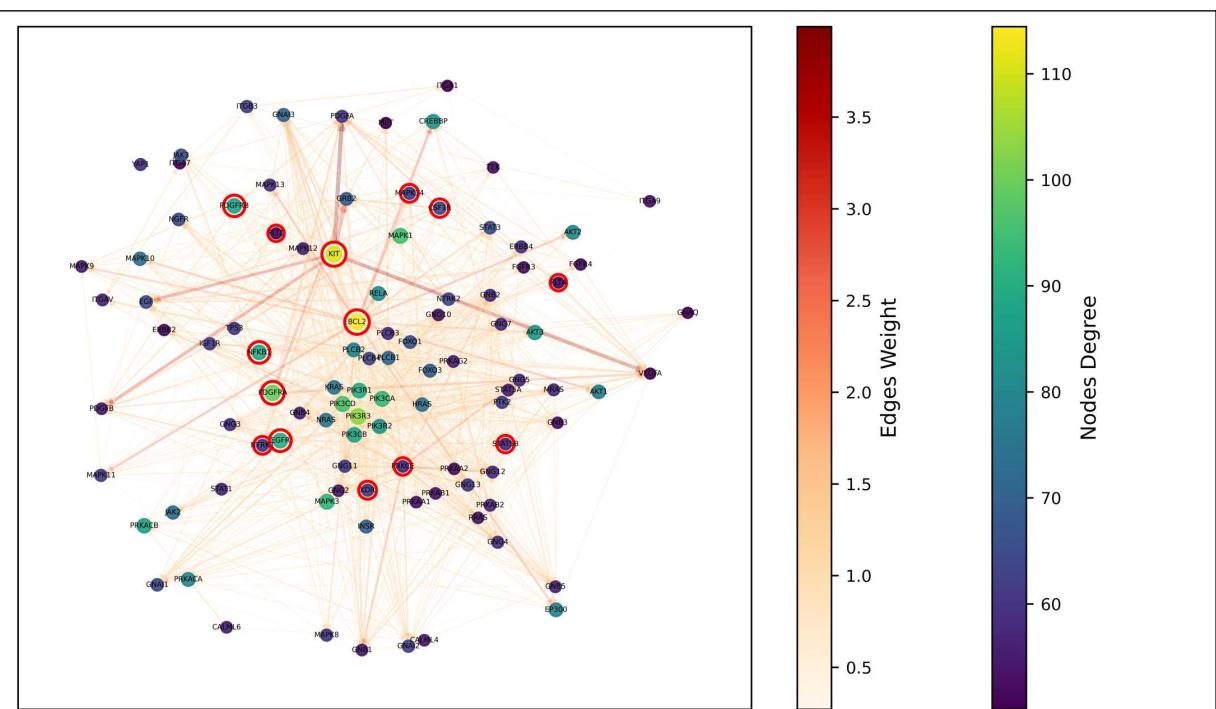
50 were selected to generate signaling flow network to partially explain the mechanism of drug combination synergy (see **Table 3**). Different thresholds of edges and nodes could generate different signaling flow subnetworks. As seen in **Figs. 4,5,6**, the core up-stream, down-stream and integrated (up- and down-stream) signaling flow networks were identified from the large signaling network based on the average trained directional weight matrices on the 5 training datasets. This sub-network module has the potential to uncover the potential mechanism of synergy of the involved drugs.

**Table 3:** Filtering results for upstream

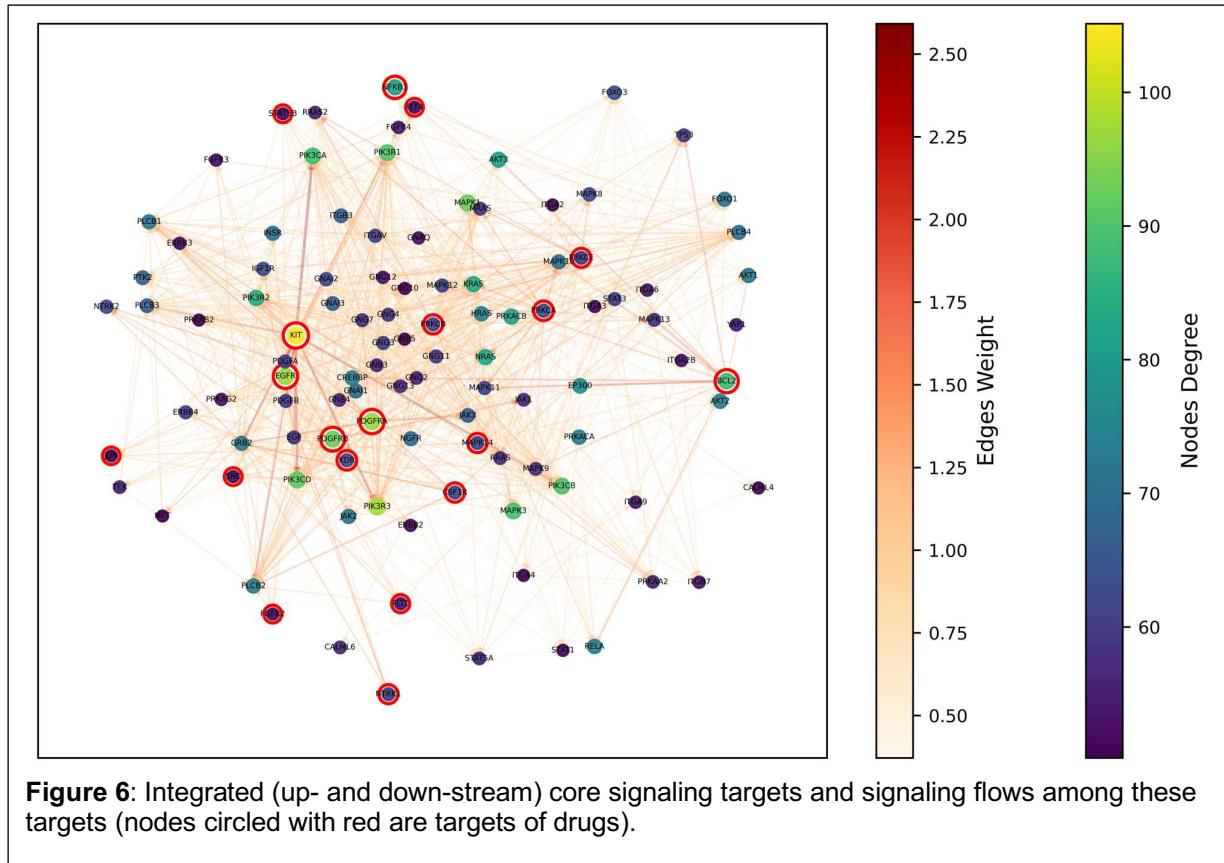
Network	Filtering Threshold	Nodes Number	Edges Number	Target Genes Number
Upstream	Nodes Degree > 50	105	1034	20
Downstream	Nodes Degree > 50	100	1101	14
Bindstream	Nodes Degree > 50	106	1095	19



**Figure 4:** Up-stream core signaling targets and signaling flows among these targets (nodes circled with red are targets of drugs).



**Figure 5:** Down-stream core signaling targets and signaling flows among these targets (nodes circled with red are targets of drugs).



#### 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, heterogenous 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 *DeepSignalingFlow* model. The model is built on a core cancer signaling network consisting of 1584 genes, with gene expression and copy number data derived from 46 core cancer signaling pathways. 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. The selected signaling interaction edges with large weights indicate 1) the important biomarker genes; and 2) the signaling cascades among them, which can potentially interpret the model prediction results. In conclusion, the first-time proposed up-stream and down-stream signaling flow design is novel and can potentially interpret the mechanism of drug combination response.

This is still an exploratory study to prediction 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. Moreover, more omics data, like mutation, methylation, miRNA, should be integrated and 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 prediction. We will investigate these challenges in future work.

## Reference

1. Jia J, Zhu F, Ma X, Cao ZW, Li YX, Chen YZ. Mechanisms of drug combinations:

Interaction and network perspectives. *Nat Rev Drug Discov.* 2009. doi:10.1038/nrd2683

2. Kinsey CG, Camolotto SA, Boespflug AM, et al. Protective autophagy elicited by RAF→MEK→ERK inhibition suggests a treatment strategy for RAS-driven cancers. *Nat Med.* 2019. doi:10.1038/s41591-019-0367-9
3. Bryant KL, Stalnecker CA, Zeitouni D, et al. Combination of ERK and autophagy inhibition as a treatment approach for pancreatic cancer. *Nat Med.* 2019. doi:10.1038/s41591-019-0368-8
4. Holbeck SL, Camalier R, Crowell JA, et al. The National Cancer Institute ALMANAC: A comprehensive screening resource for the detection of anticancer drug pairs with enhanced therapeutic activity. *Cancer Res.* 2017;77(13):3564-3576. doi:10.1158/0008-5472.CAN-17-0489
5. Kuru HI, Tastan O, Cicek AE. MatchMaker: A Deep Learning Framework for Drug Synergy Prediction. *bioRxiv.* January 2020:2020.05.24.113241. doi:10.1101/2020.05.24.113241
6. Preuer K, Lewis RPI, Hochreiter S, Bender A, Bulusu KC, Klambauer G. DeepSynergy: Predicting anti-cancer drug synergy with Deep Learning. *Bioinformatics.* 2018. doi:10.1093/bioinformatics/btx806
7. Zhang T, Zhang L, Payne P, Li F. Synergistic Drug Combination Prediction by Integrating Multi-omics Data in Deep Learning Models. *arXiv Prepr arXiv181107054.* 2018.
8. 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 Annu Symp.* 2020. doi:10.1101/2020.04.10.036491
9. Feng J, Zhang H, Li F. Investigate the relevance of major signaling pathways in cancer survival using a biologically meaningful deep learning model. *bioRxiv.* January 2020:2020.04.13.039487. doi:10.1101/2020.04.13.039487
10. Gilvary C, Dry JR, Elemento O. Multi-task learning predicts drug combination synergy in

cells and in the clinic. *bioRxiv*. January 2019:576017. doi:10.1101/576017

11. Sidorov P, Naulaerts S, Ariey-Bonnet J, Pasquier E, Ballester PJ. Predicting Synergism of Cancer Drug Combinations Using NCI-ALMANAC Data. *Front Chem*. 2019. doi:10.3389/fchem.2019.00509
12. Lamb J, Crawford ED, Peck D, et al. The connectivity map: Using gene-expression signatures to connect small molecules, genes, and disease. *Science (80-)*. 2006;313(5795):1929-1935. doi:10.1126/science.1132939
13. Subramanian A, Narayan R, Corsello SM, et al. A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. *Cell*. 2017;171(6):1437-1452. doi:10.1016/j.cell.2017.10.049
14. Regan KE, Payne PRO, Li F. Integrative network and transcriptomics-based approach predicts genotype- specific drug combinations for melanoma. *AMIA Jt Summits Transl Sci proceedings AMIA Jt Summits Transl Sci*. 2017:247-256.
15. Huang L, Li F, Sheng J, et al. DrugComboRanker: Drug combination discovery based on target network analysis. *Bioinformatics*. 2014;30(12):i228–i236. doi:10.1093/bioinformatics/btu278
16. Xu J, Regan-Fendt K, Deng S, Carson WE, Payne PRO, Li F. Diffusion mapping of drug targets on disease signaling network elements reveals drug combination strategies. In: *Proceedings of Tpacific Symposium on Biocomputing*. ; 2018:92-103. doi:10.1142/9789813235533\_0009
17. Li H, Li T, Quang D, Guan Y. Network propagation predicts drug synergy in cancers. *Cancer Res*. 2018. doi:10.1158/0008-5472.CAN-18-0740
18. Chen X, Ren B, Chen M, Wang Q, Zhang L, Yan G. NLLSS: Predicting Synergistic Drug Combinations Based on Semi-supervised Learning. *PLoS Comput Biol*. 2016. doi:10.1371/journal.pcbi.1004975
19. Sanchez-Vega F, Mina M, Armenia J, et al. Oncogenic Signaling Pathways in The

Cancer Genome Atlas. *Cell*. 2018;173(2):321-337.e10.  
doi:<https://doi.org/10.1016/j.cell.2018.03.035>

20. Barretina J, Caponigro G, Stransky N, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature*. 2012;483(7391):603-607.  
doi:10.1038/nature11003

21. Yang W, Soares J, Greninger P, et al. Genomics of Drug Sensitivity in Cancer (GDSC): A resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res*. 2013;41(D1):D955–D961. doi:10.1093/nar/gks1111

22. Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 1999;28. doi:10.1093/nar/27.1.29

23. Wishart DS, Feunang YD, Guo AC, et al. DrugBank 5.0: A major update to the DrugBank database for 2018. *Nucleic Acids Res*. 2018;46(D1):D1074-D1082.  
doi:10.1093/nar/gkx1037

24. Zitnik M, Agrawal M, Leskovec J. Modeling polypharmacy side effects with graph convolutional networks. *Bioinformatics*. 2018;34(13):i457-i466.  
doi:10.1093/bioinformatics/bty294