



Article

DRPreter: Interpretable Anticancer Drug Response Prediction Using Knowledge-Guided Graph Neural Networks and Transformer

Jihye Shin ^{1,†} , Yinhua Piao ^{2,†} , Dongmin Bang ^{1,3} , Sun Kim ^{1,2,4,5} and Kyuri Jo ^{6,*}

¹ Interdisciplinary Program in Bioinformatics, Seoul National University, Seoul 08826, Korea

² Department of Computer Science and Engineering, Institute of Engineering Research, Seoul National University, Seoul 08826, Korea

³ AIGENDRUG Co., Ltd., Seoul 08826, Korea

⁴ Interdisciplinary Program in Artificial Intelligence, Seoul National University, Seoul 08826, Korea

⁵ MOGAM Institute for Biomedical Research, Yongin-si 16924, Korea

⁶ Department of Computer Engineering, Chungbuk National University, Cheongju 28644, Korea

* Correspondence: kyurijo@chungbuk.ac.kr

† These authors contributed equally to this work.

Abstract: Some of the recent studies on drug sensitivity prediction have applied graph neural networks to leverage prior knowledge on the drug structure or gene network, and other studies have focused on the interpretability of the model to delineate the mechanism governing the drug response. However, it is crucial to make a prediction model that is both knowledge-guided and interpretable, so that the prediction accuracy is improved and practical use of the model can be enhanced. We propose an interpretable model called DRPreter (drug response predictor and interpreter) that predicts the anticancer drug response. DRPreter learns cell line and drug information with graph neural networks; the cell-line graph is further divided into multiple subgraphs with domain knowledge on biological pathways. A type-aware transformer in DRPreter helps detect relationships between pathways and a drug, highlighting important pathways that are involved in the drug response. Extensive experiments on the GDSC (Genomics of Drug Sensitivity and Cancer) dataset demonstrate that the proposed method outperforms state-of-the-art graph-based models for drug response prediction. In addition, DRPreter detected putative key genes and pathways for specific drug–cell-line pairs with supporting evidence in the literature, implying that our model can help interpret the mechanism of action of the drug.

Keywords: transcriptomics; artificial intelligence; pharmacogenomics; human health; cancer; drug sensitivity; graph neural networks; Explainable AI; precision medicine; drug discovery



Citation: Shin, J.; Piao, Y.; Bang, D.; Kim, S.; Jo, K. DRPreter: Interpretable Anticancer Drug Response Prediction Using Knowledge-Guided Graph Neural Networks and Transformer. *Int. J. Mol. Sci.* **2022**, *23*, 13919. <https://doi.org/10.3390/ijms232213919>

Academic Editor: Pablo Minguez

Received: 30 September 2022

Accepted: 8 November 2022

Published: 11 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The advances in technology and scientific capability enable the acquisition of large amounts of personal omics data at a reduced cost [1]. Consequently, there is a growing interest in using individualized health data for precision medicine, leading to a number of data-driven healthcare models [2]. Pharmacogenomics, one of the branches of precision medicine, is the study of how a person's genetic profile influences their response to medications [3,4]. Prediction of drug response or efficacy using the omics data of patients before the actual treatment is crucial because it can help increase clinical success and minimize adverse drug effects by modifying dosages or selecting alternative medications based on predicted value for personalized chemotherapy. However, obtaining patients' tumor tissues via surgical procedure or biopsy involves safety issues [5], and performing animal experiments for clinical trials to infer human drug efficacy leads to ethical and financial concerns [6]. In addition, even though correlating the drug response and omics data can help improve understanding the drugs' mechanisms of action [7], many candidate drugs

still fail to enter clinical trials during the drug discovery process due to an incomplete understanding of the mechanisms [8,9]. In this respect, an interpretable *in silico* model for drug response prediction would be useful for numerous healthcare purposes, especially for precision medicine and drug discovery [10].

Molecular profiles of cancer cell lines and high-throughput drug sensitivity screening databases are publicly available [11–16], including CCLE (Cancer Cell Line Encyclopedia) [12] and GDSC (Genomics of Drug Sensitivity in Cancer) [13,16]. Public databases and improved computing technologies such as machine learning and deep learning have contributed to the rapid development of models for predicting anticancer drug sensitivity from cancer cell lines based on their genetic profiles.

The early studies in drug sensitivity prediction utilized machine-learning techniques [17–19] such as a random forest [20], support vector machine [21], and matrix factorization [22,23]. However, the traditional machine learning-based models can still be improved in terms of predictive performance and generalizability [3,24]. Matrix-factorization-based models leave nonlinear relationships unaddressed because they attempt to identify interactions between the drugs and cell lines using linear combinations of latent features. With the capabilities of learning complex nonlinear functions and high-dimensional representations from raw data, various deep learning techniques have been utilized for predicting drug responses [24], and the overall predictive power of drug sensitivity has been improved [25]. DeepDR [26] and MOLI [27] are drug-specific models that only use cell features such as somatic mutations, gene expression profile, and copy number variation to predict the IC₅₀ values of each sample. tCNNs [28] introduced a model to predict drug sensitivity for drug–cell pairs using SMILES (Simplified Molecular Input Line Entry System) [29] sequences as drug features in addition to the genomic profiles of cells. The models described above used vector representations in common for describing cell or drug features.

Graph-based approaches have been introduced in drug-response prediction models to take advantage of the structural information of a drug or a gene network. A drug can be represented as a molecular graph consisting of a set of atoms (nodes) and a set of bonds (edges), and the graph is transformed into a high-level representation by a neural network [30,31]. For example, GraphDRP [31] obtained drug embeddings using graph convolutional networks, and cell line embeddings were derived from binary vectors of genomic aberrations. The state of a cell line can also be characterized as a gene–gene interaction network where genes (nodes) have node features from omics data such as gene expression values [32–34]. Reference [33] proposed an end-to-end drug-response prediction model, TGDRP, with cell-line graph embedding consisting of genes that have cancer-related mutations and drug graph embedding obtained by a graph neural network. They also proposed TGSA, which updates embeddings from TGDRP with similarity information between cell lines and drugs and predicts the drug response again. Among graph-based drug-response prediction models, NIHGCN [35] constructed a cell-line–drug heterogeneous network with cell-line gene expression and drug-fingerprint vectors as node features to aggregate neighborhood interactions of drug and cell line. Then, there are two different types of GCN layer for aggregating both homogeneous and heterogeneous neighbors' information for drug response prediction. However, even though it is a graph-based model, the biochemical structures of cell line and drug are not reflected in the model.

While the recent studies described above have introduced graphs into the deep-learning models to leverage structural information and improve prediction accuracy, the models lack interpretability of the predicted results. Several methods tried to delineate the mechanisms governing the drug responses, highlighting the important genes or high-level subsystems, such as biological pathways that can cause changes in cellular phenotype. SWnet [36] explored the interactions between genetic profile and the chemical structure of drugs using self-attention and identified genes with the strongest predictive power. Reference [37] proposed a multi-layer perceptron model called pathDNN which incorporates a layer of pathway nodes and quantified the activity of each pathway to explain its

effect on the drug response. DrugCell [9] obtained binary encodings of mutational status via a visible neural network guided by a hierarchy of cell subsystems and measured the predictive performance of the subsystems. Although pathDNN and DrugCell attempted to construct an explainable model with a hierarchical structure, biological pathways were implemented as gene sets rather than gene networks, indicating that domain knowledge in gene–gene interactions is not fully reflected in the models.

According to the existing studies that suggest deep-learning models for drug response prediction, it is helpful to incorporate graph representation for both drug and cell line profiles, which enables detailed descriptions of compound structures and gene networks. Moreover, the gene network can be dissected as a set of biological pathways that include gene–gene interactions for each specific biological mechanism, which can help enhance both prediction accuracy and interpretability. However, current interpretable models for drug response prediction simply describe the network as gene–pathway layers, leaving the interaction information inside the biological pathways unused. Here, we propose a novel anticancer drug-response prediction model named DRPreter (drug response predictor and interpreter) with key features as follows:

1. *Knowledge-guided cell representation with graphs.* DRPreter constructs a cell line network as a set of subgraphs that correspond to cancer-related pathways for the detailed representation of the biological mechanism.
2. *Interpretability of drug mechanisms of action.* Using the transformer's encoder, the interactions between drugs and pathways are derived from the model, and putative key pathways for the drug mechanism can be highlighted.
3. *Enhanced performance.* DRPreter outperforms state-of-the-art drug-response prediction models, as demonstrated by comparative experiments on the GDSC drug-sensitivity dataset.

2. Results and Discussion

In this study, we developed a regression model to predict the half maximal inhibitory concentration (IC₅₀), normalized to natural logarithms, which is a representative indicator of drug sensitivity. The following is a description of the graph configuration for cell lines and drugs and the graphical abstract of DRPreter (Figure 1).

The main steps of DRPreter are: cell-line graphical representation, drug graphical representation, and drug response prediction module. As a first step in creating the cell-line graph representation, a template graph was created with all genes involved in the selected 34 cancer pathways as nodes and gene–gene interactions between the genes as edges. After that, each cancer pathway was used as a predefined subgraph of the template graph, and each pathway embedding was derived using graph attention networks (GAT). As a next step in creating drug-graph representation, we transformed SMILES-format drug structures into graphs with atoms as nodes and bonds as edges, and used a graph isomorphism network (GIN) as a graph encoder to generate drug embeddings. With the 34 pathway embeddings and drug embeddings obtained as inputs to the transformer-based cell line and drug fusion module, the embeddings were updated by reflecting inter-pathway relationships and pathway-drug relationships in the model learning process. Updated pathway embeddings were combined into graph-level embeddings for the entire cell line through concatenation, and graph-level drug embedding was obtained by combining embeddings before and after passing through the transformer encoder. In conjunction with two graph level embeddings, the IC₅₀ of a given cell-line–drug pair was predicted using a multi-layer perceptron.

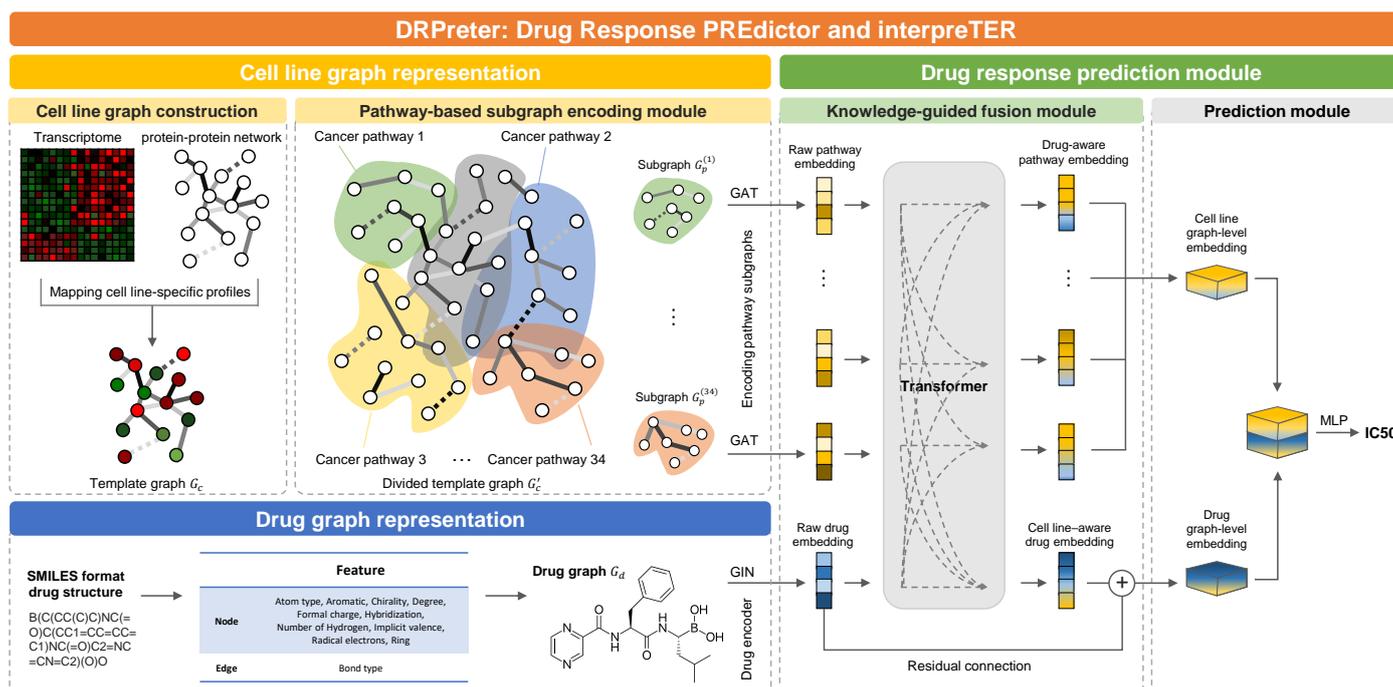


Figure 1. An overview of DRPreter. In the graph representation sections, embeddings of pathway subgraphs and drug molecule were obtained using GNN. With the obtained pathway embeddings and drug embeddings as inputs to the transformer-based cell-line and drug fusion module, the embeddings were updated by reflecting inter-pathway relationships and pathway-drug relationships in the model learning process.

2.1. Performance Comparison

2.1.1. Dataset

In the cell line template graph, the initial feature of each gene node was derived from transcriptomic data of each cell line obtained from the CCLL database version of 21Q4 [38] (<https://portals.broadinstitute.org/ccll>, accessed on 3 December 2021). The gene expression data were TPM values of the protein-coding genes for DepMap cell lines, which were inferred from RNA-seq data using the RSEM tool and were provided after \log_2 transformation, using a pseudo-count of 1; $\log_2(\text{TPM}+1)$ [38]. We assigned edges of the graph as only those interactions with high reliability scores and a combined score of at least 990 among the STRING (v11.5) [39] protein–protein interactions. The edges of the template graph and each subgraph were all STRING protein–protein interactions. Only the genes corresponding to each cancer-related pathway were obtained in KEGG, and the genes corresponding to each pathway were used as nodes in the subgraph. The STRING interactions were used as the edges connecting them. Pathways for constructing subgraphs were selected in the following manner. The non-processed pathways listed in categories 6.1 and 6.2 of the KEGG pathway database (<https://www.genome.jp/kegg/pathway.html>, accessed on 16 April 2022) were categorized according to the cancer types. These pathways include common subpathways related to cell signaling, the cell cycle, and apoptosis, which are key in various types of cancer. Consequently, if the cancer pathways provided by KEGG are used as they are, the overlap between the pathways will be excessive, and the meaning of learning for each pathway diminishes. Additionally, KEGG provides information on detailed pathways associated with each cancer type pathway. There were a total of 84 detailed pathways categorized by function. Among these pathways, we eliminated duplicate pathways, metabolic pathways, non-cancer disease pathways, viral infection pathways, and pathways with fewer than 10 genes or gene–gene interaction edges. Additionally, the focal adhesion pathway (hsa04510) was also eliminated because 91% of

the genes constituting this pathway were included in the remaining pathways. The finally selected 34 cancer-related detailed pathway list can be found in Table A1. For drug graph construction, we obtained SMILES strings from PubChem [40].

For the performance comparison experiment, we compared our method with state-of-the-art GNN-based drug-response prediction models obtaining either cell-line or drug embedding using a homogenous biochemical structure-based graph: GraphDRP, TGDRP, and TGSA. As the initial feature for each gene node, GraphDRP uses mutation (mut) and copy number variation (cnv); and TGDRP and TGSA use mut, cnv, and gene expression (exp). As GraphDRP represents cell lines as one-dimensional binary vectors, a one-dimensional CNN is used to get their embeddings. Cell lines and drugs are represented in graph format in TGDRP and TGSA, and the embeddings are obtained by an GNN. The cancer driver genes from COSMIC were selected as the genes to represent the cell lines in all baseline models [41]. The COSMIC database provides information about mutation-containing genes involved with cancer, and how these mutations can cause cancer. We selected 702 COSMIC Cancer Gene Census (<https://cancer.sanger.ac.uk/cosmic/census?tier=all>, accessed on 3 December 2021) genes; all three omics data types are provided in CCLE. We used the genes equally for the baseline model execution. Moreover, the types of cell lines and drugs used in this study were the same as in the TGDRP and TGSA. The data type used by our model differed from that of every baseline model, and ours used the most numerous omics types among them. To use only cell-line–drug pairs with three types of omics data available, intensive filtering was done on cell lines and drugs. Since all omics data had to be imported for baseline model execution, the same cell-line–drug pair was used as in the most data-intensive models. Consequently, the performance test consisted of 580 cancer cell lines that can obtain omics data from CCLE and 170 anticancer drugs provided by GDSC2. The total number of possible cell-line–drug pairs was 82,833 with log-normalized IC50 values.

In addition, as performance comparisons with deep learning, we added random forest and support vector machine (SVM) as baseline models for comparative analysis. For a fair comparison, we used the same features and preprocessed the data to feed into traditional machine-learning methods. For cell line embedding, we concatenated all gene expression vectors, resulting in a one-dimensional vector, and for drug embedding, since different drugs have different atoms, we simply the sum of each atom embedding, resulting in a one-dimensional vector. Finally, we concatenated cell line embedding and drug embedding to one-dimensional embedding and fed it into the models. The nodes constituting the cell-line graphs of the existing GNN-based drug-response prediction models were configured according to the settings in each comparison paper.

2.1.2. Experimental Setups

In the regression experiments for predicting natural log-transformed IC50 values based on drug and cancer cell line profiles, we used four standard evaluation metrics to compare the results of different models by computing the statistical correlation and accuracy of predicted and observed IC50 values. The metrics included the Pearson correlation coefficient (PCC), Spearman correlation coefficient (SCC), root absolute error (MAE), and mean-squared error (MSE). PCC measures the linear correlation of observed and predicted IC50 values. SCC is a non-parametric measure for rank correlation between observed and predicted IC50 values. MSE and MAE directly measure the difference between observed and predicted IC50 values.

2.1.3. Rediscovered Responses of Known Pairs

All possible cell-line–drug pairs were randomly divided into training, validation, and test datasets in an 8:1:1 ratio, and the experiments were conducted repeatedly on 10 random seeds. For each model, the test performance was averaged over the seeds and is reported as mean \pm standard deviation. Comparing the results of different models was based on four common evaluation indicators. The mean-squared error and mean absolute error

between the predicted IC50 and the true IC50; and the Pearson correlation coefficient and Spearman correlation coefficient of each IC50 distribution were used as evaluation criteria. Compared to the baseline model we selected above, we conducted an ablation study to examine each part of DRPreter's effectiveness (Table 1). Based on the results of the ablation study, we assume that network information from biological pathways helped improve the performance of our model. It has also been found that the ability to relate the cell line in its pre-drug treatment state to the drug through the transformer has a significant effect on performance improvement. In addition, DRPreter showed a performance improvement of about 20% in MSE compared with the next best model (Table 2). We also conducted internal validation using a 10-fold cross-validation experiment on the original dataset using CCLE and GDSC2. The validation set was selected by five random seeds (Table 3).

Table 1. Model ablation studies with different settings.

Structural Settings of DRPreter	Data	MSE (\downarrow)	MAE (\downarrow)	PCC (\uparrow)	SCC (\uparrow)
Template graph	COSMIC ¹	0.8926 \pm 0.0363	0.6909 \pm 0.0146	0.9423 \pm 0.0027	0.9196 \pm 0.0034
Template graph	Pathway ²	0.8536 \pm 0.0420	0.6759 \pm 0.0161	0.9449 \pm 0.0032	0.9224 \pm 0.0035
Pathway	Pathway ²	0.8645 \pm 0.0277	0.6791 \pm 0.0113	0.9446 \pm 0.0014	0.9233 \pm 0.0008
Pathway + Transformer	Pathway ²	0.8302 \pm 0.0156	0.6676 \pm 0.0051	0.9465 \pm 0.0015	0.9242 \pm 0.0015
Pathway + Transformer + Similarity	Pathway ²	0.8251 \pm 0.0122	0.6682 \pm 0.0047	0.9467 \pm 0.0013	0.9248 \pm 0.0014

¹ COSMIC: 702 COSMIC genes. ² Pathway: 2369 genes of 34 cancer-related pathways. The best performance is shown in bold for each metric.

Table 2. Performance comparison with baseline models.

Model	Cell Encoder	Data	MSE (\downarrow)	MAE (\downarrow)	PCC (\uparrow)	SCC (\uparrow)
SVM ¹	-	Pathway	8.5780 \pm 2.0615	2.2976 \pm 0.3005	0.5282 \pm 0.0355	0.4471 \pm 0.0476
RF ²	-	Pathway	1.6711 \pm 0.0422	0.9608 \pm 0.0100	0.8887 \pm 0.0021	0.8497 \pm 0.0034
GraphDRP	1D CNN	COSMIC	1.0110 \pm 0.0157	0.7618 \pm 0.0083	0.9386 \pm 0.0018	0.9151 \pm 0.0021
TGDRP	GNN	COSMIC	0.9004 \pm 0.0341	0.6933 \pm 0.0148	0.9417 \pm 0.0026	0.9188 \pm 0.0040
TGSA	GNN	COSMIC	0.8955 \pm 0.0536	0.6913 \pm 0.0238	0.9425 \pm 0.0043	0.9201 \pm 0.0051
DRPreter	Knowledge-guided GNN	Pathway	0.8251 \pm 0.0122	0.6682 \pm 0.0047	0.9467 \pm 0.0013	0.9248 \pm 0.0014

¹ SVM: support vector machine. ² RF: random forest. The best performance is shown in bold for each metric.

Table 3. Internal validation using 10-fold cross-validation on 5 random seeds.

Comparison Models	Data	MSE (\downarrow)	MAE (\downarrow)	PCC (\uparrow)	SCC (\uparrow)
TGDRP	COSMIC ¹	1.9398 \pm 0.0231	1.0435 \pm 0.0058	0.8665 \pm 0.0026	0.8164 \pm 0.0074
DRPreter Template graph	COSMIC ¹	1.9665 \pm 0.0323	1.0435 \pm 0.0089	0.8685 \pm 0.0018	0.8232 \pm 0.0022
DRPreter Template graph	Pathway ²	1.9276 \pm 0.0495	1.0351 \pm 0.0130	0.8711 \pm 0.0034	0.8270 \pm 0.0042
DRPreter w/o Trans ³ and Similarity	Pathway ²	1.8536 \pm 0.0548	1.0085 \pm 0.0123	0.8820 \pm 0.0049	0.8445 \pm 0.0094
DRPreter w/o similarity	Pathway ²	1.8317 \pm 0.0276	1.0076 \pm 0.0067	0.8778 \pm 0.0018	0.8356 \pm 0.0022

¹ COSMIC: 702 COSMIC genes. ² Pathway: 2369 genes of 34 cancer-related pathways. ³ Trans: transformer-based cell-line-drug fusion module.

2.2. Case Study

2.2.1. Interpolation of Unknown Values

The method of missing values prediction has been widely used in drug-response-prediction studies [28,30,31,33] to identify whether the model is capable of inductive prediction. For evaluating the inductive predictability of our model, we trained with all the known cell-line–drug pairs and predicted values without experimental results of pairs in the GDSC2 database. There were a total of 98,600 pairs using 580 cancer cell lines and 170 drugs, but 15,767 cell lines were not covered by our data due to filtering because of a lack of omics data or due to the absence of drug response experiments in GDSC. The model with the highest performance was used to predict missing drug response values.

We illustrate the distributions of known IC₅₀ values in GDSC2 and the predicted values of our model (Figure 2). The box plots are grouped by drugs, and each box represents the distribution of the IC₅₀ values within a cell line. We displayed the drugs with the top 10 highest and top 10 lowest median IC₅₀ value. After conducting Mann–Whitney Wilcoxon test for each drug distribution, 18 drugs among the 20 selected drugs showed no significant difference between the GDSC2 and predicted unknown IC₅₀ value distribution. The result implies the predicted missing IC₅₀ values follow the measured value distribution.

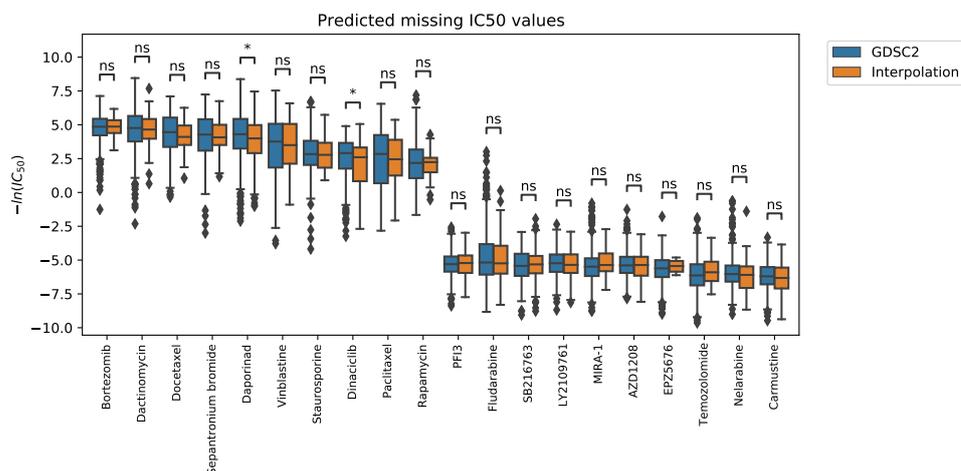


Figure 2. Box plot of drug-specific IC₅₀ distributions of cell lines. The distribution of GDSC2 data (blue) compared with predicted missing IC₅₀ values (orange). The 10 drugs with the highest median IC₅₀ values and the 10 drugs with the lowest median were selected. Among the 20 drugs, IC₅₀ value distributions of 18 drugs showed no significant differences through the Mann–Whitney Wilcoxon Test. ns: not significant, *: 0.01 < *p*-value < 0.05.

The total predicted missing values using our model can be found in Table S1 of the Supplementary Data.

Not knowing the actual values for these missing pairs, we conducted literature searches to assess our predictions. Bortezomib had the smallest overall IC₅₀ distribution, and the most sensitive cell-line pair was LP-1 in our model. LP-1 is a cell line derived from the peripheral blood of a multiple myeloma patient. Bortezomib is a proteasome inhibitor that is widely used in patients with multiple myeloma [42,43]. Rapamycin was not included among the top 10 sensitive drugs in the known GDSC data but in our predicted values, so we analyzed it further. In our study, rapamycin was most sensitive to the MV-4-11 cell line. The MV-4-11 cells are macrophages that were isolated from the blast cells of a biphenotypic B myelomonocytic leukemia patient. Rapamycin can inhibit leukemic activity in acute myeloid leukemia by mTOR inhibition through the blockade in G₀/G₁ phase of the cell cycle [44].

Based on the biological processes at the cellular and molecular level of cancer cells and drugs, DRPreter can make inductive predictions for cell lines and drugs when there are no known responses and seems to have the potential to select candidates for drug treatment.

2.2.2. Gradient-Weighted Gene Nodes Interpretation

It is essential for drug-response prediction methods to capture significant biological implications and to make accurate predictions. A gene-level analysis was performed first to determine whether the model was taking into account genes that are known as drug targets, involved in target pathways, or biomarkers of disease. We prioritized genes from an input drug and cancer–cell-line pair by scoring each gene with a gradient-weighted extent to check whether it is drug-target-related. The importance score of each gene node was determined by GradCAM, which is a widely utilized technique to produce explanations of model decisions [45], and we considered the score as the extent of its contribution. In our model, GradCAM determined the influence of input gene nodes on the label by tracing back the gradient backpropagation process of the model for predicting IC50 value. Table 4 shows the top five most significant genes of each cell-line–drug pair in the test dataset.

Table 4. Gradient-based gene importance analysis.

Drug	Cell Line	Disease	Top 5 Significant Genes	ln(IC50)	
				True	Predicted
Afatinib	GMS-10	Glioblastoma	<i>ACTR3B</i> , <i>PRR5</i> , <i>PRKCZ</i> , <i>ERBB2</i> , <i>LTBR</i>	0.5372	0.5324
Vinblastine	NCI-H1792	NSCLC	<i>CYP7A1</i> , <i>GTF2H2</i> , <i>DVL2</i> , <i>RAB5B</i> , <i>TP53</i>	−5.9258	−5.27633
Docetaxel	PANC0327	Pancreatic cancer	<i>CLDN18</i> , <i>SOX17</i> , <i>FGF19</i> , <i>WNT7A</i> , <i>CDH5</i>	−3.7668	−3.8204
Rapamycin	IGR1	Melanoma	<i>TYRP1</i> , <i>DCT</i> , <i>TYR</i> , <i>FRZB</i> , <i>CDK2</i>	−1.6747	−1.7651
Bortezomib	EBC-1	Lung squamous cell carcinoma Derived from metastatic site: Skin	<i>SHC4</i> , <i>TNR</i> , <i>IL17RA</i> , <i>MAPK12</i> , <i>SMURF1</i>	−5.7714	−6.0714

Genes in bold are direct targets of drugs, are involved in target pathways, or are biomarkers of disease.

As verified by literature searches, the bold genes in Table 4 are the target genes or genes associated with the target pathway for each drug–cell-line pair. The targets were obtained from DrugBank [46] and GDSC, and the genes corresponding to the target pathway were obtained from GeneCards [47] and Harmonizome [48]. Afatinib is a irreversible ErbB family blocker [49] that targets *EGFR* and *ERBB2*, and its target pathway is the EGFR signaling pathway. Our model found *ERBB2* as a significant gene of the afatinib pair. Among the other top five genes, *LTBR* was found to be related with tumor treatment for its potential in triggering apoptosis of tumor cells or anti-tumor immune response [50]. As with the majority of cancers, *TP53* is the most common mutated gene, showing a predominant clonal expression in Non-Small-Cell Lung Cancer (NSCLC) [51]. Additionally, microtubule-active drugs, including vinblastine, are known to induce apoptosis through inducing expression of p53 protein [52]. It is known to be possible to use *CLDN18* as an early-stage indicator of pancreatic ductal carcinogenesis and to study *CLDN18*'s regulatory mechanisms for uncovering key pathways such as the PKC pathway of pancreatic cancer [53]. *WNT7A*, the fourth-ranked gene, shows relation with docetaxel for Wnt signaling, playing a role in docetaxel resistance [54]. *CDK2* corresponds to the mTOR signaling pathway, which is the target pathway of rapamycin. Additionally, an in vivo experiment reported that upregulation of *TYRP1* and *TYR* proteins may explain the melanogenesis of rapamycin-treated melanoma cells [55]. The use of bortezomib and paclitaxel suggests the potential for rationally designed treatments for solid tumors with MAPK pathway activation [56].

2.2.3. Pathway-Level Interpretation Using the Transformer

We examined which pathways were stimulated in various cancer types that are sensitive to drugs and some that are not, and whether our model could capture such meaningful pathways. The self-attention score from the transformer-based structure (Figure 3) was investigated for a drug that is sensitive only to specific cell lines. All the GDSC data with known IC₅₀ values were observed in the same way as Figure 2a, and dasatinib was selected as having the widest IC₅₀ distributions. The wide distribution of its IC₅₀ means that the drug exhibits the greatest differences in efficacy based on the type of cell line. We compared the self-attention score of the transformer on MEG-01, the cell line judged to be sensitive by having the smallest IC₅₀ value, and BT-483, the most insensitive cell line with the largest IC₅₀ value, among the 548 cell line pairs with dasatinib (Figure 4). The MEG-01 cell line was derived from the hematopoietic and lymphoid tissue of a leukemia patient, and the BT-483 cell line was derived from the breast tissue of a breast-cancer patient.

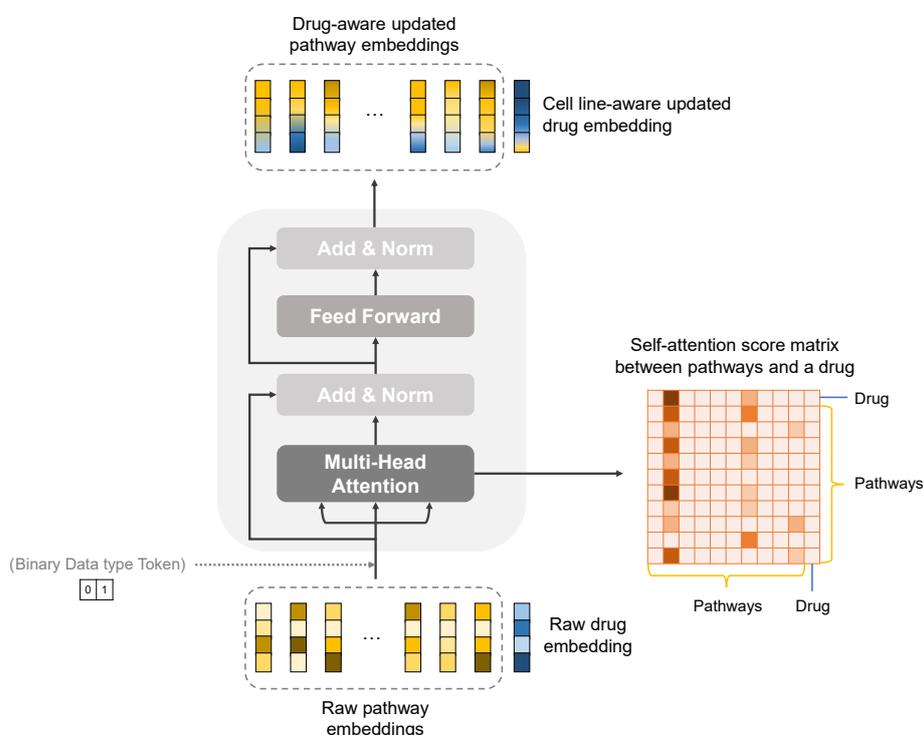


Figure 3. A detailed structure of type-aware transformer encoder reflecting interactions and relationships between pathways and a drug. We extracted drug-pathway interaction information from the modified encoder of the Transformer module and identified putative key pathways for the drug's mechanism of action using a matrix of self-attention scores between pathways and the drug.

The TGF- β signaling pathway (hsa04350) was the pathway with the highest attention score for the MEG-01 cell line, which is the most sensitive to dasatinib. The second-most-important pathway, ubiquitin-mediated proteolysis (hsa04120), involves the covalent binding of ubiquitin to the target protein and its degradation. It is known that ubiquitin-mediated degradation can regulate the TGF- β signaling pathway [57]. The TGF- β signaling pathway suppresses tumors in normal and premalignant cells, yet promotes oncogenesis in advanced cancer cells, and its components are regulated by ubiquitin-modifying enzymes; abnormalities of the enzymes can cause malfunctioning of the pathway, which can cause cancer, tissue fibrosis, and metastasis [58–60]. In this regard, the ubiquitin-modifying enzymes in the pathway and their counterparts are increasingly being explored as potential drug targets [59]. Dasatinib is a tyrosine kinase inhibitor that can be a treatment for chronic myeloid leukemia [61]. Dasatinib functions by binding to the ATP site of the active conformation of BCR-Abl [62]. As a signal-transduction inhibitor, dasatinib inhibits the

proliferation of tumor cells by inhibiting tyrosine kinase action, especially blocking transcriptional and promigratory responses to TGF- β through inhibition of Smad signaling [63]. The ubiquitin pathway can regulate the basal level of Smads, and altered Smad proteins can cause a malfunction in responding to the incoming signals due to their importance in transducing TGF- β signals [57]. From the ubiquitin to the TGF- β pathway, our model captures the mechanism of action of the drug.

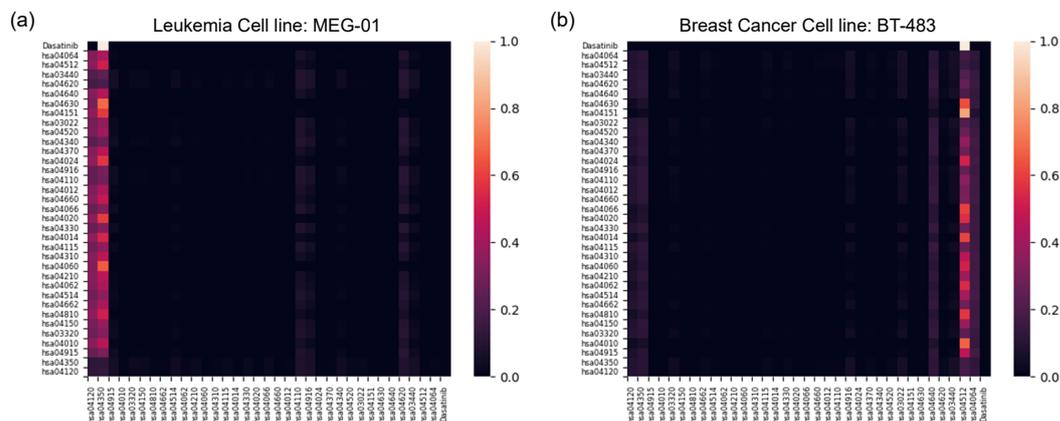


Figure 4. Visualization of all-pairwise self-attention scores from the transformer. (a) Dasatinib and leukemia cell line MEG-01 pair. (b) Dasatinib and breast cancer cell line BT-483. The figures show the y-axis as the query of the transformer and the x-axis as the key. On each axis, there is a drug and 34 pathways which start with “hsa”, indicating KEGG pathway identifiers.

Moreover, the ECM–receptor interaction pathway (hsa04512) was found to be most important in the breast-cancer cell line, which is the most insensitive to dasatinib. The ECM–receptor interaction pathway has been shown to be possibly useful as a biomarker for breast cancer [64], but it does not relate to dasatinib’s mechanism of action. Hence, our model identified the pathways related to the drug’s mechanism of action for drug-sensitive carcinoma and focused on the biomarker for carcinoma without drug efficacy.

3. Materials and Methods

3.1. Graph Neural Networks

A graph neural network (GNN) is a type of neural network that operates on graph-structured data. GNN uses the topology of the graph to learn the relationships between the input features. It can perform more effectively than other representation learning methods on input data with topological information. In this study, we represent a graph as $G = (V, E)$ where $V = \{v_1, \dots, v_n\}$ is the set of n nodes and $E \subseteq V \times V$ is the set of edges. The node v_i has node feature $x_i \in \mathbb{R}^d$, where d is a dimension of the feature. The node feature matrix of the graph can be represented as $X \in \mathbb{R}^{n \times d}$, where n is the number of nodes in the graph. Adjacency matrix $A \in \mathbb{R}^{n \times n}$ indicates the total connectivity of nodes in the graph, where $A_{i,j} = 1$ means nodes, v_i and v_j are linked, and $W^{(l)}$ represents the parameters of the l -th layer of the graph (Table 5).

Table 5. Notation of graph neural networks used in this paper.

Notation	Description
G	A graph.
V	Set of nodes of a graph.
v	A node included in V .
i, j	Indexes of the nodes.
l	Index of the layer of a graph.
v_i	i -th node in V .
x_i	Node feature of node v_i
$N(i)$	Set of neighbor nodes of a node v_i
E	Set of edges of a graph.
A	Adjacency matrix between nodes.
$W^{(l)}$	Trainable parameter matrix of l -th layer.
$X^{(l)}$	Node feature matrix of l -th layer.
σ	Nonlinear activation function softmax.
ϵ	Learnable parameter.

In each GNN layer, a key mechanism, called message passing, updates the node representation by using the node features of the previous layer and the topology of the graph [65]. The message passing mechanism involves aggregating the information of neighboring nodes and updating the hidden state of each node by combining the node representation from the previous layer and the aggregated messages. For every node in each layer, a transformed feature vector is generated capturing the structural information of the k -hop neighbor nodes. The GNN can update the i -th node representation in the l -th layer as in the following Equation [66,67], where $N(i)$ is the set of neighbor nodes linked to the target node i . For a given node, the AGGREGATE step applies a permutation invariant function to its neighboring nodes to produce the aggregated node feature of neighbors, and the COMBINE step delivers the aggregated node feature to the learnable layer to produce updated node embedding by integrating the existing embedding and the aggregated neighbor embedding.

$$x_i^{(l)} = \text{COMBINE}^{(l)} \left(x_i^{(l-1)}, \text{AGGREGATE}^{(l-1)} (x_j^{(l-1)} : j \in N(i)) \right) \quad (1)$$

3.2. Cell-Line Graph Representation

3.2.1. Cell-Line Graph Construction

We used a biological template network to represent cell lines to simulate gene–gene interactions in actual cells. In the cell-line graph G_c , genes are represented by nodes and edges represent the relationships between genes. This template graph contained 2369 genes selected using the pathway selection method described in the next section.

It is known that drugs do not have a universal effect throughout all cellular components, but tend to have distinct effects on specific genes or pathway targets. In this way, cancer cells undergo phenotypic changes as a result of drug molecules inhibiting or activating their target pathways. Motivated by this point, instead of representing the cell line as a homogeneous large-scale graph that contains the entire genes, we divided the template network G_c into pathway subgraphs G_p according to the biological domain knowledge inspired by [68] and learned graph embeddings from the selected subgraph units. Finally, the divided cell-line graph G_c' was represented as a heterogeneous graph containing multiple subgraphs. We selected pathways that can be targeted by drugs, as they are associated with cancer from the KEGG pathway database [69], and used these pathways as pre-defined subgraphs of the cell-line template.

In the case of template graph G_c , the i -th selected pathway subgraph can be described as $G_p^{(i)} = (V_p^{(i)}, E_p^{(i)})$, where $V_p^{(i)}$ refers to a set of nodes and $E_p^{(i)}$ refers to a set of edges of the pathway. Thus, the template graph G_c is extended as a union of disjoint graphs,

with overlaps between the pathways in the form of $\mathbf{G}'_c = \{\mathbf{G}_p^{(1)}, \dots, \mathbf{G}_p^{(34)}\}$. In the template cell-line graph \mathbf{G}_c , gene sets included in 34 pathways were represented by 2369 nodes and 7954 edges. A divided template graph \mathbf{G}'_c with pathways as subgraphs had 4646 nodes and 12,004 edges after combining the data from all pathways. The types of constituting genes remained the same, but the numbers of nodes and edges increased when the template network was divided into subnetworks due to the overlap of functions.

3.2.2. Cell-Line Graph Encoder on Pathway Subgraphs

Transcriptomic features of nodes and biological network topology were captured within each subgraph using a graph attention network (GAT) [70]. Using the self-attention mechanism, GAT calculates a normalized attention score α_{ij} , indicating the importance of the features of the neighbor nodes for a target node i , where $j \in N(i)$. A subsequent step in the message passing process is for each node to reflect the importance of the neighboring nodes' information in accordance with the previously obtained attention scores. Details and graphical overview of GAT can be found in the original paper [70].

$$X^{(l)} = \sigma \left(\sum_{j \in N(i)} \alpha_{ij}^{(l-1)} W^{(l-1)} X^{(l-1)} \right) \quad (2)$$

If template graph \mathbf{G}_c is used as it is, edges connected to one gene include interactions from multi pathways, which can be noise. Node representations were updated through GAT on the cell-line graph constructed in the previous subsection. The cell-line graph consists of pathway-based subgraphs; thus, the updated node representation can reflect the intra-pathway gene–gene interaction information. To pool the cell-line graph-level embedding, we initially used simple hierarchical permutation-invariant graph-pooling strategies [71–73]. However, the graph-pooling strategies we employed resulted in slight performance degradation. We assumed that this was due to the relatively large size of the cell-line graph, and simply pooling the nodes into a vector of the same dimension may lose the information of the nodes in a cell line. As a result, the embeddings of each node learned through GAT were concatenated to form a graph-level embedding for each pathway.

3.3. Drug Graph Representation

3.3.1. Drug Graph Construction

We used a graph neural network to learn the drug representation by reflecting the relationships between atoms connected by bonds and the overall molecular structural information. A drug can be represented as a graph in which atoms are nodes and bonds are edges. We used RDKit [74] to transform SMILES [29], a one-dimensional string format drug structure, into a graph format that can reflect structural information of an actual drug. The ten initial features of atomic nodes were imported from previous research [30,33], which predicted drug sensitivity from GNN-based embeddings of drug structures. The details of atomic and bond features can be found in Table A2.

3.3.2. Drug Graph Encoder

We used the graph isomorphism network (GIN) [75] to learn the features of the atomic nodes within the drug graph. GIN applies a neighborhood aggregation method similar to the Weisfeiler–Lehman test [76] and updates the i -th node feature of the l -th layer as follows.

$$x_i^{(l)} = MLP^{(l)} \left((1 + \epsilon^{(l)}) \cdot x_i^{(l-1)} + \sum_{j \in N(i)} x_j^{(l-1)} \right) \quad (3)$$

Details and graphical overview of GIN can be found in the original paper [75]. The graph encoder was chosen following the results of GraphDRP, which involved a comparison of different types of graph neural networks—GIN, GAT, and GCN+GAT—in order to analyze the effectiveness of each graph encoder in predicting drug response. In addition, GIN was widely used for the embeddings of drug graphs in various drug-response predic-

tion models [31,33,34,77]. All embeddings of each atom node were updated through GIN, then they were concatenated to create raw drug embeddings before the pathway affected them.

3.4. Drug Response Prediction Module

3.4.1. Knowledge-Guided Cell-Line–Drug Fusion Module Using Transformer

The transformer model tracks relationships in sequential data, such as the words in a sentence, to discover context and meaning from the components [78]. We used a transformer-based module to reflect not only inter-pathway interactions but also the interactions between the pathways and each drug, which would allow exploring the pharmacological mechanisms of action at the pathway level during a therapeutic process (Figure 3).

The embedding of 34 pathways and that of a drug obtained from graph representation modules were updated in the transformer-based cell line and drug fusion module. The encoder is based on the transformer encoder, and does not use positional encoding, as the order of the embeddings does not matter. Instead, we conducted two experiments: first with binary-data-type tokens indicating whether the embedding is a pathway or a drug, and then without positional encoding or data-type tokens. All-pairwise self-attention scores can be obtained between pathways and drug embeddings via multi-head attention. As a result of the encoder-based transformer model, the raw pathway embeddings were updated to drug-aware pathway embeddings according to the effects of the drugs, and the raw drug embedding was updated based on the relationship with each pathway.

Our structure has a single encoder-based layer taking pathway embeddings $X_p^{(l)}$ ($l = 1, \dots, 34$) and a drug embedding X_d derived from knowledge-guided GNNs as input values. Inputs in a typical transformer's encoder are constructed by adding positional encoding to embeddings of source sequences. Unlike translation, where an order of words in a sentence is important, the pathway embeddings entering our encoder are not affected by the order in which they are encoded, so a transformer's encoder structure other than positional encoding was used for this study. As an alternative, we added a type-encoded token that indicates whether the embedding is a drug or a pathway. In an element-wise manner, type-encoded binary tokens are added to the input feature matrix with the same number of dimensions before input embeddings are fed into the module.

On the fusion pathway and drug embeddings, self-attention was performed several times through multi-head attention, and the average of each trial was used as the final attention score. After the encoder completed its execution, the encoder produced drug-aware updated pathway embeddings $X_p^{(l)'}$ and pathways' transcriptome-aware updated drug embedding X_d' reflecting interaction information. These drug-aware pathway embeddings facilitate interpreting the medication's mechanism of action, since they can reflect both the drug–pathway interaction information and the interactions between the pathways. Drugs have a large structural variation when compared to cell-line graphs which are composed of the same genes and are structurally equal but have different node feature values. Therefore, it is possible that the variation of the drug embedding may be blurred because the new drug embedding updated as a result of the transformer is affected by the cell line embedding. Hence, we connect the raw drug embedding obtained through GNN prior to the transformer with the updated drug embedding obtained after the transformer using residual connection [79]. By residual connection, it is possible to preserve the original drug structure information and utilize the cell-line–drug interaction information using the updated drug embedding which recognizes the transcriptomic information of each pathway. We concatenated the resulting 34 subgraph embeddings in order to prevent information loss, thereby embedding the entire cell line.

3.4.2. Improving Predictive Performance Using a Similarity Graph

Based on the idea that similar drugs and similar cell lines exhibit interchangeable drug response behaviors, some drug-response prediction models use prior knowledge of drug

and cell-line similarity to minimize differences between drugs and cell lines in the latent space. Reference [22] applied regularization terms based on chemical-structure similarities between drugs and similarities between cell lines based on gene expression profiles to improve prediction accuracy and prevent overfitting.

We followed the similarity-based embedding updating strategy of [33]. From the completed end-to-end model up to Section 3.3, embeddings of all 580 cell lines and 170 drugs were created. Then, we constructed two homogeneous graphs, each consisting of cell lines and drug nodes, with the initial feature of each node set having the resulting embeddings of the previous step. Using GraphSAGE [80], we updated the embeddings of each homogeneous cell line and drug graph. After that, we updated embeddings of each cell-line–drug pair from two homogeneous graphs. We concatenated two embeddings into the one-dimensional vector and used a multi-layer perceptron to predict final IC₅₀ values.

4. Conclusions

In this paper, we proposed an interpretable drug-response prediction model called DRPreter which integrates biological and chemical-domain knowledge with cutting-edge deep learning technologies to deliver outstanding predictive performance and interpretability. We introduced cancer-related pathways and constructed the cell line network as a set of subgraphs to represent and interpret biological mechanisms in detail. We extracted drug–pathway interaction information from the modified encoder of the transformer module and obtained putative key pathways for the drug’s mechanism. Ablation studies verified the effectiveness of each component of the model, and performance comparison experiments showed DRPreter has enhanced predictive power compared to the state-of-the-art graph-based drug-response prediction models which obtain either the cell line or drug embedding using a homogeneous biochemical structure-based graph.

To properly apply the drug response predicted by the model for clinical use or drug discovery, it is essential to understand the process and mechanism from which it was derived due to safety and reliability issues. Accordingly, we implemented gene and pathway-level analysis via DRPreter, and it has been shown that DRPreter predicts drug sensitivity based on known drug mechanisms of action and target-related factors. We also identified the cell line that would act most sensitively for each drug in the absence of experimented data through a case study and confirmed that it is widely used for each drug currently in the clinical situation. By doing so, patients who have shown resistance to a specific drug may be able to select a drug candidate group that would replace the unsuitable drug. It will be remarkably efficient to have comprehensive public databases of drug targets and predictive models that can interpret pharmacological mechanisms for personalized medicine and drug discovery.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms232213919/s1>.

Author Contributions: Conceptualization, J.S., Y.P., S.K., and K.J.; methodology, J.S. and Y.P.; software, J.S. and Y.P.; validation, J.S. and Y.P.; investigation, J.S., D.B. and K.J.; data curation, J.S.; writing—original draft preparation, J.S.; writing—review and editing, J.S., Y.P., D.B., S.K., and K.J.; visualization, J.S. and D.B.; supervision, S.K. and K.J.; funding acquisition, S.K. and K.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Bio and Medical Technology Development Program of the National Research Foundation (NRF) funded by the Ministry of Science and ICT (NRF-2022M3E5F3085677); by a grant (number DY0002259501) from the Ministry of food and Drug Safety; by an Institute of Information and communications Technology Planning and Evaluation (IITP) grant funded by the Korea government (MSIT) (number 2021-0-01343, Artificial Intelligence Graduate School Program (Seoul National University)); by a National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (number NRF-2020R1G1A1003558); and by the Collaborative Genome Program for Fostering New Post Genome Industry of the National Research Foundation (NRF) funded by the Ministry of Science and ICT (MSIT) (NRF2014M3C9A3063541).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data and code used in this study are available at <https://github.com/babaling/DRPreter>, (accessed on 30 September 2022).

Acknowledgments: We would like to show our gratitude to Sangseon Lee and Dohoon Lee (Bio and Health Informatics Lab, Seoul National University) for providing insight and expertise that greatly assisted the research.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

mut	Mutation of gene
exp	Gene expression
cnv	Copy number variation
CCL	Cancer Cell Line Encyclopedia
CNN	Convolutional Neural Network
COSMIC	Catalogue of Somatic Mutations in Cancer
GAT	Graph Attention Network
GCN	Graph Convolutional Network
GDSC	Genomics of Drug Sensitivity in Cancer
GIN	Graph Isomorphism Network
GNN	Graph Neural Network
GradCAM	Gradient-weighted Class Activation Mapping
IC50	half maximal inhibitory concentration
MLP	Multi-Layer Perceptron
RSEM	RNA-Seq by Expectation Maximization
TGF	Transforming Growth Factor
TPM	Transcripts Per Kilobase Million

Appendix A

The number of genes and edges in Table A1 is the result of filtering according to the threshold of protein–protein interactions' combined score.

Table A1. A list of cancer-related pathways used as subgraphs in the cell-line template graph.

Pathway Name	KEGG Identifier	Number of Genes	Number of Edges
Ubiquitin mediated proteolysis	hsa04120	142	534
TGF- β signaling pathway	hsa04350	94	228
Estrogen signaling pathway	hsa04915	137	222
MAPK signaling pathway	hsa04010	294	692
PPAR signaling pathway	hsa03320	74	28
mTOR signaling pathway	hsa04150	155	688
Regulation of actin cytoskeleton	hsa04810	218	552
B cell receptor signaling pathway	hsa04662	79	208
Cell adhesion molecules	hsa04514	146	150
Chemokine signaling pathway	hsa04062	190	514
Apoptosis	hsa04210	136	424
Cytokine-cytokine receptor interaction	hsa04060	293	588
Wnt signaling pathway	hsa04310	167	384
p53 signaling pathway	hsa04115	73	180

Table A1. Cont.

Pathway Name	KEGG Identifier	Number of Genes	Number of Edges
Ras signaling pathway	hsa04014	232	600
Notch signaling pathway	hsa04330	59	76
Calcium signaling pathway	hsa04020	239	218
HIF-1 signaling pathway	hsa04066	109	204
T cell receptor signaling pathway	hsa04660	104	336
ErbB signaling pathway	hsa04012	85	326
Cell cycle	hsa04110	126	1076
Melanogenesis	hsa04916	101	110
cAMP signaling pathway	hsa04024	221	222
VEGF signaling pathway	hsa04370	59	102
Hedgehog signaling pathway	hsa04340	56	80
Adherens junction	hsa04520	71	172
Basal transcription factors	hsa03022	44	470
PI3K-Akt signaling pathway	hsa04151	351	1030
JAK-STAT signaling pathway	hsa04630	162	508
Hematopoietic cell lineage	hsa04640	96	102
Toll-like receptor signaling pathway	hsa04620	102	328
Homologous recombination	hsa03440	41	140
ECM-receptor interaction	hsa04512	88	120
NF- κ B signaling pathway	hsa04064	102	392

Table A2. Atomic and bond features of the drug graph.

	Feature	Size	Description
Node	Atom type	43	[B, C, N, O, F, ...] (one-hot)
	Aromatic	1	Whether the atom is in aromatic system (binary)
	Chirality	2	[R, S] (one-hot or null)
	Degree	11	[0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10] (one-hot)
	Formal charge	1	electric charge (integer)
	Hybridization	5	[sp , sp^2 , sp^3 , sp^3d , sp^3d^2] (one-hot or null)
	Number of Hydrogens	5	[0, 1, 2, 3, 4] (one-hot)
	Implicit valence	7	[0, 1, 2, 3, 4, 5, 6] (one-hot)
	Radical electrons	1	Number of radical electrons (integer)
	Ring	1	Whether the atom is in ring (binary)
Edge	Bond type	4	[single, double, triple, aromatic] (one-hot)

References

1. Kellogg, R.A.; Dunn, J.; Snyder, M.P. Personal omics for precision health. *Circ. Res.* **2018**, *122*, 1169–1171. [[CrossRef](#)] [[PubMed](#)]
2. Ahmed, Z. Practicing precision medicine with intelligently integrative clinical and multi-omics data analysis. *Hum. Genom.* **2020**, *14*, 1–5. [[CrossRef](#)] [[PubMed](#)]
3. Kalamara, A.; Tobalina, L.; Saez-Rodriguez, J. How to find the right drug for each patient? Advances and challenges in pharmacogenomics. *Curr. Opin. Syst. Biol.* **2018**, *10*, 53–62. [[CrossRef](#)] [[PubMed](#)]
4. Singh, D.B. The impact of pharmacogenomics in personalized medicine. *Curr. Appl. Pharm. Biotechnol.* **2019**, *171*, 369–394.

5. Cho, S.Y. Patient-derived xenografts as compatible models for precision oncology. *Lab. Anim. Res.* **2020**, *36*, 1–11. [[CrossRef](#)] [[PubMed](#)]
6. Singh, V.P.; Pratap, K.; Sinha, J.; Desiraju, K.; Bahal, D.; Kukreti, R. Critical evaluation of challenges and future use of animals in experimentation for biomedical research. *Int. J. Immunopathol. Pharmacol.* **2016**, *29*, 551–561. [[CrossRef](#)]
7. Rees, M.G.; Seashore-Ludlow, B.; Cheah, J.H.; Adams, D.J.; Price, E.V.; Gill, S.; Javaid, S.; Coletti, M.E.; Jones, V.L.; Bodycombe, N.E.; et al. Correlating chemical sensitivity and basal gene expression reveals mechanism of action. *Nat. Chem. Biol.* **2016**, *12*, 109–116. [[CrossRef](#)]
8. Seyhan, A.A. Lost in translation: The valley of death across preclinical and clinical divide—identification of problems and overcoming obstacles. *Transl. Med. Commun.* **2019**, *4*, 1–19. [[CrossRef](#)]
9. Kuenzi, B.M.; Park, J.; Fong, S.H.; Sanchez, K.S.; Lee, J.; Kreisberg, J.F.; Ma, J.; Ideker, T. Predicting drug response and synergy using a deep learning model of human cancer cells. *Cancer Cell* **2020**, *38*, 672–684. [[CrossRef](#)]
10. Savage, N. Tapping into the drug discovery potential of AI. *Biopharma Deal* **2021**. [[CrossRef](#)]
11. Shoemaker, R.H. The NCI60 human tumour cell line anticancer drug screen. *Nat. Rev. Cancer* **2006**, *6*, 813–823. [[CrossRef](#)]
12. Barretina, J.; Caponigro, G.; Stransky, N.; Venkatesan, K.; Margolin, A.A.; Kim, S.; Wilson, C.J.; Lehár, J.; Kryukov, G.V.; Sonkin, D.; et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **2012**, *483*, 603–607. [[CrossRef](#)]
13. Yang, W.; Soares, J.; Greninger, P.; Edelman, E.J.; Lightfoot, H.; Forbes, S.; Bindal, N.; Beare, D.; Smith, J.A.; Thompson, I.R.; et al. Genomics of Drug Sensitivity in Cancer (GDSC): A resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res.* **2012**, *41*, D955–D961. [[CrossRef](#)]
14. Basu, A.; Bodycombe, N.E.; Cheah, J.H.; Price, E.V.; Liu, K.; Schaefer, G.I.; Ebright, R.Y.; Stewart, M.L.; Ito, D.; Wang, S.; et al. An interactive resource to identify cancer genetic and lineage dependencies targeted by small molecules. *Cell* **2013**, *154*, 1151–1161. [[CrossRef](#)]
15. Seashore-Ludlow, B.; Rees, M.G.; Cheah, J.H.; Cokol, M.; Price, E.V.; Coletti, M.E.; Jones, V.; Bodycombe, N.E.; Soule, C.K.; Gould, J.; et al. Harnessing Connectivity in a Large-Scale Small-Molecule Sensitivity Dataset. *Cancer Discov.* **2015**, *5*, 1210–1223. [[CrossRef](#)]
16. Iorio, F.; Knijnenburg, T.A.; Vis, D.J.; Bignell, G.R.; Menden, M.P.; Schubert, M.; Aben, N.; Gonçalves, E.; Barthorpe, S.; Lightfoot, H.; et al. A landscape of pharmacogenomic interactions in cancer. *Cell* **2016**, *166*, 740–754. [[CrossRef](#)]
17. Güvenç Paltun, B.; Mamitsuka, H.; Kaski, S. Improving drug response prediction by integrating multiple data sources: Matrix factorization, kernel and network-based approaches. *Brief. Bioinform.* **2021**, *22*, 346–359. [[CrossRef](#)]
18. Adam, G.; Rampášek, L.; Safikhani, Z.; Smirnov, P.; Haibe-Kains, B.; Goldenberg, A. Machine learning approaches to drug response prediction: Challenges and recent progress. *NPJ Precis. Oncol.* **2020**, *4*, 1–10. [[CrossRef](#)]
19. Firoozbakht, F.; Yousefi, B.; Schwikowski, B. An overview of machine learning methods for monotherapy drug response prediction. *Brief. Bioinform.* **2022**, *23*, bbab408. [[CrossRef](#)]
20. Riddick, G.; Song, H.; Ahn, S.; Walling, J.; Borges-Rivera, D.; Zhang, W.; Fine, H.A. Predicting in vitro drug sensitivity using Random Forests. *Bioinformatics* **2011**, *27*, 220–224. [[CrossRef](#)]
21. Dong, Z.; Zhang, N.; Li, C.; Wang, H.; Fang, Y.; Wang, J.; Zheng, X. Anticancer drug sensitivity prediction in cell lines from baseline gene expression through recursive feature selection. *BMC Cancer* **2015**, *15*, 1–12. [[CrossRef](#)] [[PubMed](#)]
22. Wang, L.; Li, X.; Zhang, L.; Gao, Q. Improved anticancer drug response prediction in cell lines using matrix factorization with similarity regularization. *BMC Cancer* **2017**, *17*, 1–12. [[CrossRef](#)] [[PubMed](#)]
23. Guan, N.N.; Zhao, Y.; Wang, C.C.; Li, J.Q.; Chen, X.; Piao, X. Anticancer drug response prediction in cell lines using weighted graph regularized matrix factorization. *Mol. Ther. Nucleic Acids* **2019**, *17*, 164–174. [[CrossRef](#)] [[PubMed](#)]
24. Baptista, D.; Ferreira, P.G.; Rocha, M. Deep learning for drug response prediction in cancer. *Brief. Bioinform.* **2021**, *22*, 360–379. [[CrossRef](#)] [[PubMed](#)]
25. Sakellaropoulos, T.; Vougas, K.; Narang, S.; Koinis, F.; Kotsinas, A.; Polyzos, A.; Moss, T.J.; Piha-Paul, S.; Zhou, H.; Kardala, E.; et al. A deep learning framework for predicting response to therapy in cancer. *Cell Rep.* **2019**, *29*, 3367–3373. [[CrossRef](#)]
26. Chiu, Y.C.; Chen, H.I.H.; Zhang, T.; Zhang, S.; Gorthi, A.; Wang, L.J.; Huang, Y.; Chen, Y. Predicting drug response of tumors from integrated genomic profiles by deep neural networks. *BMC Med. Genom.* **2019**, *12*, 143–155.
27. Sharifi-Noghabi, H.; Zolotareva, O.; Collins, C.C.; Ester, M. MOLI: Multi-omics late integration with deep neural networks for drug response prediction. *Bioinformatics* **2019**, *35*, i501–i509. [[CrossRef](#)]
28. Liu, P.; Li, H.; Li, S.; Leung, K.S. Improving prediction of phenotypic drug response on cancer cell lines using deep convolutional network. *BMC Bioinform.* **2019**, *20*, 1–14. [[CrossRef](#)]
29. Weininger, D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J. Chem. Inf. Comput. Sci.* **1988**, *28*, 31–36. [[CrossRef](#)]
30. Liu, Q.; Hu, Z.; Jiang, R.; Zhou, M. DeepCDR: A hybrid graph convolutional network for predicting cancer drug response. *Bioinformatics* **2020**, *36*, i911–i918. [[CrossRef](#)]
31. Nguyen, T.; Nguyen, G.T.; Nguyen, T.; Le, D.H. Graph convolutional networks for drug response prediction. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2021**, *19*, 146–154. [[CrossRef](#)]
32. Kim, S.; Bae, S.; Piao, Y.; Jo, K. Graph convolutional network for drug response prediction using gene expression data. *Mathematics* **2021**, *9*, 772. [[CrossRef](#)]

33. Zhu, Y.; Ouyang, Z.; Chen, W.; Feng, R.; Chen, D.Z.; Cao, J.; Wu, J. TGSA: Protein–protein association-based twin graph neural networks for drug response prediction with similarity augmentation. *Bioinformatics* **2022**, *38*, 461–468. [[CrossRef](#)]
34. Feng, R.; Xie, Y.; Lai, M.; Chen, D.Z.; Cao, J.; Wu, J. AGMI: Attention-Guided Multi-omics Integration for Drug Response Prediction with Graph Neural Networks. In Proceedings of the 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Houston, TX, USA, 9–12 December 2021; pp. 1295–1298.
35. Peng, W.; Liu, H.; Dai, W.; Yu, N.; Wang, J. Predicting cancer drug response using parallel heterogeneous graph convolutional networks with neighborhood interactions. *Bioinformatics* **2022**, *38*, 4546–4553. [[CrossRef](#)]
36. Zuo, Z.; Wang, P.; Chen, X.; Tian, L.; Ge, H.; Qian, D. SWnet: A deep learning model for drug response prediction from cancer genomic signatures and compound chemical structures. *BMC Bioinform.* **2021**, *22*, 1–16. [[CrossRef](#)]
37. Deng, L.; Cai, Y.; Zhang, W.; Yang, W.; Gao, B.; Liu, H. Pathway-guided deep neural network toward interpretable and predictive modeling of drug sensitivity. *J. Chem. Inf. Model.* **2020**, *60*, 4497–4505. [[CrossRef](#)]
38. DepMap Broad. DepMap 21Q4 Public. Figshare. Dataset, 2021. Available online: <https://portals.broadinstitute.org/ccl> (accessed on 3 December 2021).
39. Szklarczyk, D.; Gable, A.L.; Lyon, D.; Junge, A.; Wyder, S.; Huerta-Cepas, J.; Simonovic, M.; Doncheva, N.T.; Morris, J.H.; Bork, P.; et al. STRING v11: Protein–Protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **2019**, *47*, D607–D613. [[CrossRef](#)]
40. Wang, Y.; Xiao, J.; Suzek, T.O.; Zhang, J.; Wang, J.; Bryant, S.H. PubChem: A public information system for analyzing bioactivities of small molecules. *Nucleic Acids Res.* **2009**, *37*, W623–W633. [[CrossRef](#)]
41. Sondka, Z.; Bamford, S.; Cole, C.G.; Ward, S.A.; Dunham, I.; Forbes, S.A. The COSMIC Cancer Gene Census: Describing genetic dysfunction across all human cancers. *Nat. Rev. Cancer* **2018**, *18*, 696–705. [[CrossRef](#)]
42. Field-Smith, A.; Morgan, G.J.; Davies, F.E. Bortezomib (Velcade™) in the treatment of multiple myeloma. *Ther. Clin. Risk Manag.* **2006**, *2*, 271. [[CrossRef](#)]
43. Kouroukis, T.; Baldassarre, F.; Haynes, A.; Imrie, K.; Reece, D.; Cheung, M. Bortezomib in multiple myeloma: Systematic review and clinical considerations. *Curr. Oncol.* **2014**, *21*, 573–603. [[CrossRef](#)] [[PubMed](#)]
44. Récher, C.; Beyne-Rauzy, O.; Demur, C.; Chicanne, G.; Dos Santos, C.; Mas, V.M.D.; Benzaquen, D.; Laurent, G.; Huguet, F.; Payrastre, B. Antileukemic activity of rapamycin in acute myeloid leukemia. *Blood* **2005**, *105*, 2527–2534. [[CrossRef](#)] [[PubMed](#)]
45. Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-cam: Visual explanations from deep networks via gradient-based localization. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 618–626.
46. Wishart, D.S.; Knox, C.; Guo, A.C.; Cheng, D.; Shrivastava, S.; Tzur, D.; Gautam, B.; Hassanali, M. DrugBank: A knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Res.* **2008**, *36*, D901–D906. [[CrossRef](#)] [[PubMed](#)]
47. Safran, M.; Dalah, I.; Alexander, J.; Rosen, N.; Iny Stein, T.; Shmoish, M.; Nativ, N.; Bahir, I.; Doniger, T.; Krug, H.; et al. GeneCards Version 3: The human gene integrator. *Database* **2010**, *2010*, baq020. [[CrossRef](#)] [[PubMed](#)]
48. Rouillard, A.D.; Gundersen, G.W.; Fernandez, N.F.; Wang, Z.; Monteiro, C.D.; McDermott, M.G.; Ma’ayan, A. The harmonizome: A collection of processed datasets gathered to serve and mine knowledge about genes and proteins. *Database* **2016**, *2016*, baw100. [[CrossRef](#)]
49. Ioannou, N.; Dalglish, A.; Seddon, A.; Mackintosh, D.; Guertler, U.; Solca, F.; Modjtahedi, H. Anti-tumour activity of afatinib, an irreversible ErbB family blocker, in human pancreatic tumour cells. *Br. J. Cancer* **2011**, *105*, 1554–1562. [[CrossRef](#)]
50. Fernandes, M.T.; Dejardin, E.; dos Santos, N.R. Context-dependent roles for lymphotoxin- β receptor signaling in cancer development. *Biochim. Biophys. Acta BBA Rev. Cancer* **2016**, *1865*, 204–219. [[CrossRef](#)]
51. Canale, M.; Andrikou, K.; Priano, I.; Cravero, P.; Pasini, L.; Urbini, M.; Delmonte, A.; Crinò, L.; Bronte, G.; Ulivi, P. The Role of TP53 Mutations in EGFR-Mutated Non-Small-Cell Lung Cancer: Clinical Significance and Implications for Therapy. *Cancers* **2022**, *14*, 1143. [[CrossRef](#)]
52. Tishler, R.B.; Lamppu, D.M.; Park, S.; Price, B.D. Microtubule-active drugs taxol, vinblastine, and nocodazole increase the levels of transcriptionally active p53. *Cancer Res.* **1995**, *55*, 6021–6025.
53. Tanaka, M.; Shibahara, J.; Fukushima, N.; Shinozaki, A.; Umeda, M.; Ishikawa, S.; Kokudo, N.; Fukayama, M. Claudin-18 is an early-stage marker of pancreatic carcinogenesis. *J. Histochem. Cytochem.* **2011**, *59*, 942–952. [[CrossRef](#)]
54. Stewart, D.J. Wnt signaling pathway in non-small cell lung cancer. *J. Natl. Cancer Inst.* **2014**, *106*, djt356. [[CrossRef](#)]
55. Hah, Y.S.; Cho, H.Y.; Lim, T.Y.; Park, D.H.; Kim, H.M.; Yoon, J.; Kim, J.G.; Kim, C.Y.; Yoon, T.J. Induction of melanogenesis by rapamycin in human MNT-1 melanoma cells. *Ann. Dermatol.* **2012**, *24*, 151–157. [[CrossRef](#)]
56. Mehnert, J.M.; Tan, A.R.; Moss, R.; Poplin, E.; Stein, M.N.; Sovak, M.; Levinson, K.; Lin, H.; Kane, M.; Gounder, M.; et al. Rationally Designed Treatment for Solid Tumors with MAPK Pathway Activation: A Phase I Study of Paclitaxel and Bortezomib Using an Adaptive Dose-Finding Approach Paclitaxel and Bortezomib for Tumors with MAPK Activation. *Mol. Cancer Ther.* **2011**, *10*, 1509–1519. [[CrossRef](#)]
57. Izzi, L.; Attisano, L. Regulation of the TGF β signalling pathway by ubiquitin-mediated degradation. *Oncogene* **2004**, *23*, 2071–2078. [[CrossRef](#)]
58. Huang, F.; Chen, Y.G. Regulation of TGF- β receptor activity. *Cell Biosci.* **2012**, *2*, 1–10. [[CrossRef](#)]
59. Iyengar, P.V. Regulation of Ubiquitin Enzymes in the TGF- β Pathway. *Int. J. Mol. Sci.* **2017**, *18*, 877. [[CrossRef](#)]

60. Seoane, J.; Gomis, R.R. TGF- β family signaling in tumor suppression and cancer progression. *Cold Spring Harb. Perspect. Biol.* **2017**, *9*, a022277. [[CrossRef](#)]
61. Keskin, D.; Sadri, S.; Eskazan, A.E. Dasatinib for the treatment of chronic myeloid leukemia: Patient selection and special considerations. *Drug Des. Dev. Ther.* **2016**, *10*, 3355. [[CrossRef](#)]
62. Sun, H.; Kapuria, V.; Peterson, L.F.; Fang, D.; Bornmann, W.G.; Bartholomeusz, G.; Talpaz, M.; Donato, N.J. Bcr-Abl ubiquitination and Usp9x inhibition block kinase signaling and promote CML cell apoptosis. *Blood J. Am. Soc. Hematol.* **2011**, *117*, 3151–3162. [[CrossRef](#)]
63. Bartscht, T.; Rosien, B.; Rades, D.; Kaufmann, R.; Biersack, H.; Lehnert, H.; Gieseler, F.; Ungefroren, H. Dasatinib blocks transcriptional and promigratory responses to transforming growth factor-beta in pancreatic adenocarcinoma cells through inhibition of Smad signalling: Implications for in vivo mode of action. *Mol. Cancer* **2015**, *14*, 1–12. [[CrossRef](#)]
64. Bao, Y.; Wang, L.; Shi, L.; Yun, F.; Liu, X.; Chen, Y.; Chen, C.; Ren, Y.; Jia, Y. Transcriptome profiling revealed multiple genes and ECM-receptor interaction pathways that may be associated with breast cancer. *Cell. Mol. Biol. Lett.* **2019**, *24*, 1–20. [[CrossRef](#)] [[PubMed](#)]
65. Gilmer, J.; Schoenholz, S.S.; Riley, P.F.; Vinyals, O.; Dahl, G.E. Neural message passing for quantum chemistry. In Proceedings of the International Conference on Machine Learning (PMLR), Sydney, Australia, 6–11 August 2017; pp. 1263–1272.
66. Li, M.M.; Huang, K.; Zitnik, M. Graph Representation Learning in Biomedicine. *arXiv* **2021**, arXiv:2104.04883.
67. Dai, E.; Zhao, T.; Zhu, H.; Xu, J.; Guo, Z.; Liu, H.; Tang, J.; Wang, S. A Comprehensive Survey on Trustworthy Graph Neural Networks: Privacy, Robustness, Fairness, and Explainability. *arXiv* **2022**, arXiv:2204.08570.
68. Lee, S.; Lim, S.; Lee, T.; Sung, I.; Kim, S. Cancer subtype classification and modeling by pathway attention and propagation. *Bioinformatics* **2020**, *36*, 3818–3824. [[CrossRef](#)] [[PubMed](#)]
69. Kanehisa, M.; Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **2000**, *28*, 27–30. [[CrossRef](#)]
70. Veličković, P.; Cucurull, G.; Casanova, A.; Romero, A.; Lio, P.; Bengio, Y. Graph attention networks. *arXiv* **2017**, arXiv:1710.10903.
71. Zhang, M.; Cui, Z.; Neumann, M.; Chen, Y. An end-to-end deep learning architecture for graph classification. In Proceedings of the AAAI conference on Artificial Intelligence, New Orleans, LA, USA, 2 February 2018; Volume 32.
72. Gao, H.; Ji, S. Graph u-nets. In Proceedings of the International Conference On Machine Learning (PMLR), Long Beach, CA, USA, 9–15 June 2019; pp. 2083–2092.
73. Lee, J.; Lee, I.; Kang, J. Self-attention graph pooling. In Proceedings of the International Conference on Machine Learning (PMLR), Long Beach, CA, USA, 9–15 June 2019; pp. 3734–3743.
74. Landrum, G. RDKit: A Software Suite for Cheminformatics, Computational Chemistry, and Predictive Modeling. 2013. Available online: <http://rdkit.sourceforge.net> (accessed on 16 September 2021).
75. Xu, K.; Hu, W.; Leskovec, J.; Jegelka, S. How powerful are graph neural networks? *arXiv* **2018**, arXiv:1810.00826.
76. Weisfeiler, B.; Leman, A. The reduction of a graph to canonical form and the algebra which appears therein. *NTI Ser.* **1968**, *2*, 12–16.
77. Zheng, K.; Zhao, H.; Zhao, Q.; Wang, B.; Gao, X.; Wang, J. NASMDR: A framework for miRNA-drug resistance prediction using efficient neural architecture search and graph isomorphism networks. *Brief. Bioinform.* **2022**, *20*, bbac338. [[CrossRef](#)]
78. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 6000–6010.
79. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 770–778.
80. Hamilton, W.; Ying, Z.; Leskovec, J. Inductive representation learning on large graphs. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 1025–1035.