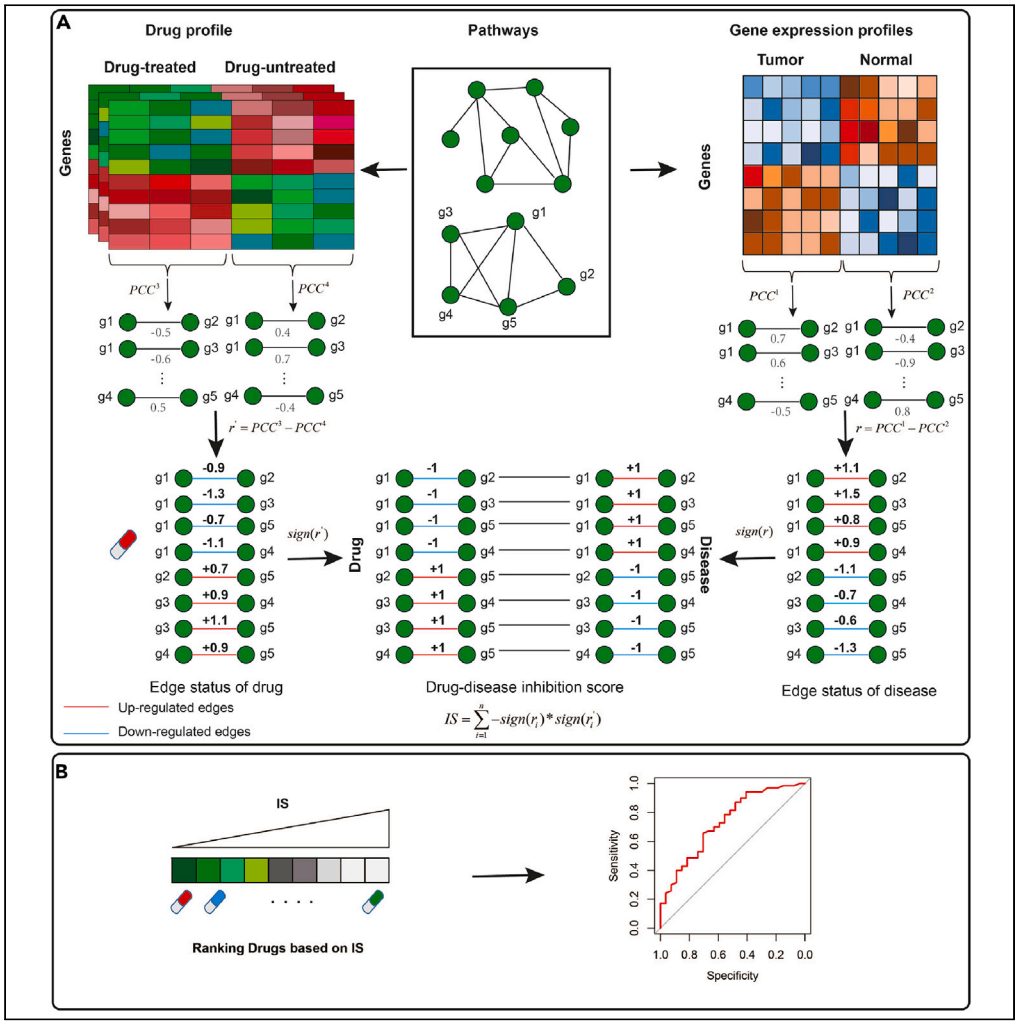
iScience

CellPress
OPEN ACCESS

Article

Integrated edge information and pathway topology for drug-disease associations



Xianbin Li, Xiangzhen Zan, Tao Liu, ..., Qizhang Li, Zhenshen Bao, Jie Lin

lixb88@gzhu.edu.cn (X.L.)
33096094@qq.com (J.L.)

Highlights

Many computational methods neglect the changes in gene interactions at the edge level

We propose a drug repurposing method based on edge information and pathway topology

Compared to the state-of-the-art approaches, our method has the superior performance

Li et al., iScience 27, 110025
July 19, 2024
https://doi.org/10.1016/j.isci.2024.110025

iScience

CellPress
OPEN ACCESS

Article

Integrated edge information and pathway topology for drug-disease associations

Xianbin Li,[1,2,7,8,*] Xiangzhen Zan,[3,7] Tao Liu,[1] Xiwei Dong,[1] Haqi Zhang,[2] Qizhang Li,[4] Zhenshen Bao,[5] and Jie Lin[6,*]

SUMMARY

Drug repurposing is a promising approach to find new therapeutic indications for approved drugs. Many computational approaches have been proposed to prioritize candidate anticancer drugs by gene or pathway level. However, these methods neglect the changes in gene interactions at the edge level. To address the limitation, we develop a computational drug repurposing method (iEdgePathDDA) based on edge information and pathway topology. First, we identify drug-induced and disease-related edges (the changes in gene interactions) within pathways by using the Pearson correlation coefficient. Next, we calculate the inhibition score between drug-induced edges and disease-related edges. Finally, we prioritize drug candidates according to the inhibition score on all disease-related edges. Case studies show that our approach successfully identifies new drug-disease pairs based on CTD database. Compared to the state-of-the-art approaches, the results demonstrate our method has the superior performance in terms of five metrics across colorectal, breast, and lung cancer datasets.

INTRODUCTION

Drug repurposing, the exploration of existing drugs for new therapeutic indications, has emerged as a promising strategy to accelerate drug development and reduce costs.[1] Traditional drug discovery processes are time-consuming and resource-intensive,[2] and often taking a decade to bring a new drug to market.[3,4] While the safety and known pharmacokinetics of approved drugs are extensively documented, these characteristics can substantially mitigate the costs associated with clinical trials.[5] Compared with the design and discovery of new drugs, drug repurposing presents notable advantages in safety, cost, and rapid results, signifying its substantial pharmaceutical significance.[6] Researchers often utilize experimental or computational methods to identify suitable drug candidates, which constitutes a key aspect of drug repurposing.[7–17]

Studies on drug repurposing have achieved significant milestones, which many researches apply computational methods to address this challenge.[18] These methods be broadly categorized into signature-based and pathway-based approaches.

Signature-based methods are predominantly devised for prioritizing drug-disease relationships by analyzing gene expression patterns. The Connectivity Map (CMAP) database, which contains gene expression profiles induced by 1309 compounds across five cancer cell lines, is a widely used resource.[19] Several computational methods have been proposed to uncover previously unknown drug-disease pairs by examining the inverse correlation of gene expression patterns between drugs and disease by using CMAP.[20–26] For example, Dudley et al.[20] systematically identified potential drugs by comparing the disease-related gene expression characteristics with drug-induced gene expression characteristics based on CMAP. Kosaka et al.[21] found an antiviral drug for prostate cancer by using drug-induced gene expression data from CMAP database. Noort et al.[22] predicted novel drugs for colorectal cancer by utilizing drug-induced gene expression profiles that invert disease profiles. Cheng et al. Pacini et al.[23] developed a framework to discover potential drugs by comparing drug-induced and disease-related gene expression profiles. Napolitano et al.[24] integrated drug targets, chemical structures, and drug-induced gene expression to facilitate drug repositioning. Chen et al.[25] proposed a drug repurposing method by leveraging tissue or species-specific transcriptome data in conjunction with drug-induced gene expression from CMAP database. Papikinos et al.[26] employed CMAP to pinpoint potential drugs for amyotrophic lateral sclerosis. Ahmed et al.[17] employed gene expression signatures of the psoriasis and compared them with perturbagen available in the CMAP to predict potentially effective drugs. However, from a system

[1]School of Computer and Big Data Science, Jiujiang University, Jiujiang, Jiangxi 332000, China
[2]Department of Digital Media Technology, Hangzhou Dianzi University, Hangzhou, Zhejiang 310018, China
[3]Institute of Computational Science and Technology, Guangzhou University, Guangzhou, Guangdong 520000, China
[4]Innovative Drug R&D Center, School of Life Sciences, Huaibei Normal University, Huaibei, Anhui 235000, China
[5]College of Information Engineering, Taizhou University, Taizhou 225300, Jiangsu, China
[6]Department of Pharmacy, the Third Affiliated Hospital of Wenzhou Medical University, Wenzhou 325200, Zhejiang Province, China
[7]These authors contributed equally
[8]Lead contact
*Correspondence: lixb88@gzhu.edu.cn (X.L.), 33096094@qq.com (J.L.)
https://doi.org/10.1016/j.isci.2024.110025

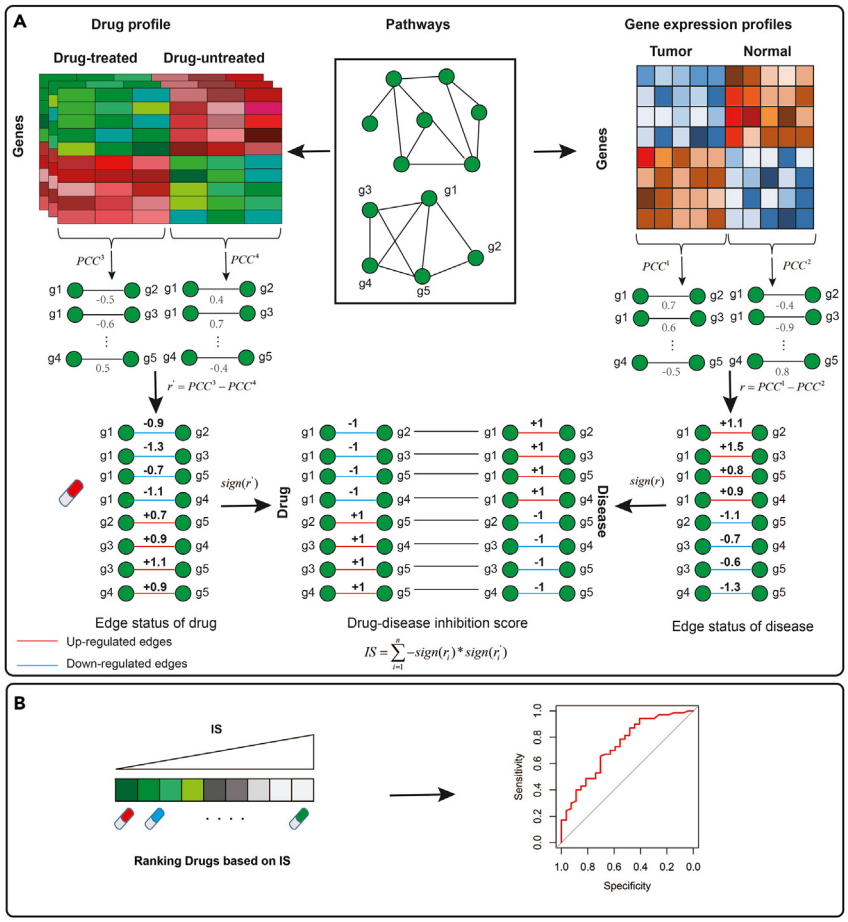


Figure 1. The flowchart of iEdgePathDDA

(A) Drug-disease inhibition score (IS).

(B) Ranking drugs based on IS.

biology perspective, drugs typically achieve their therapeutic effects on diseases by modulating biological pathways, and the signature-based approaches mainly focus on gene level.

Pathway-based methods modulate disease-associated pathways to discover drug-disease associations. By integrating comprehensive pathway data with drug-induced gene expression profiles, we aim to systematically identify drugs that perturb specific pathways implicated in diseases of interest. Recent studies have introduced that pathway-based approaches aim at identifying potential drugs for cancer treatment.[27–33] For instance, Han et al.[27] identified abnormal subpathways induced by diseases and drugs, respectively, and then evaluated the reverse correlation between drugs and diseases at the subpathway level for drug repurposing. Wu et al.[28] found candidate drugs for cancer by considering drug-induced subpathways and their crosstalk effect. Li et al.[29] introduced a drug repurposing methodology that seamlessly integrates gene expression profiles with gene regulatory networks. Lwata et al.[30] utilized molecular pathways as the therapeutic targets and prioritized potential drugs by regulating disease-related pathways. Napolitano et al.[31] proposed a drug repositioning method by incorporating drug-induced gene expression and therapeutic target genes within specific pathways. Di et al.[32] constructed a drug functional similarity network to prioritize agents by pathway activities and drug activities. Wang et al.[33] developed a drug-pathway prediction method that inferred pathway responsive to specific drugs. However, these methods primarily focused on identifying candidate drugs at the pathway level, without considering the changes of gene interactions. Gene interactions (activation or inhibition) in the pathway play a crucial role in occurrence and progression of disease. Therefore, it is imperative to develop new methods that can accurately identify cancer-related drugs based on edge level.

In this paper, we develop a drug repurposing method, which utilizes the changes of gene interactions within pathway between drug-induced state and disease-related state to identify potential candidate drugs for cancer treatment. In our framework, as shown in Figure 1, calculating inhibition score between drug and disease based on the change of gene interactions. Case studies demonstrate that our approach successfully identified new drug-disease pairs. Our findings indicate that the top-rank drug candidates have been already validated by CTD database. To evaluate the performance of iEdgePathDDA, we compare iEdgePathDDA with other existing approaches in three cancer datasets (colorectal, lung, and breast cancer). The results show that iEdgePathDDA achieves best performance in terms of AUPR, AUROC, Recall, ACC, and F1.

Table 1. The gene expression datasets of three cancers

Dataset	GEO	Case/control	Reference
Colorectal Cancer	GSE23878	19/19	Uddin et al.[34]
Breast Cancer	GSE31448	29/4	Sabatier et al.[35]
	GSE29044	73/36	Colak et al.[36]
Lung Cancer	GSE18842	46/45	Sanchez-Palencia et al.[37]
	GSE19188	91/65	Hou et al.[38]
	GSE19804	60/60	Lu et al.[39]

RESULTS

The effect of the threshold α

In this section, our focus is on scrutinizing the impact of the parameter α on the model's performance. The objective is to reveal the underlying factor α influencing its effectiveness. We use six datasets from GEO dataset. The details are outlined in Table 1.

A pivotal factor contributing to the enhancement of our model's performance is the parameter α. Therefore, the selection of an appropriate threshold α is important. Figure 2 shows the trends of five metrics (AUPR, AUROC, Recall, ACC, F1) and their corresponding average values across a range of the threshold α from 0 to 0.8 on six datasets. The general observation is that the overall trend of five metrics and the average decreases with the increment of the threshold α. As the threshold of α increases, the acquired edge information related to the disease become limited, leading to constraining the model's performance. Remarkably, the model attains its optimal overall performance on the six datasets when the threshold of α is set to 0.

Case study: Colorectal cancer

To substantiate the reliability and effectiveness of our method, we apply it to predict potential drugs for colorectal cancer. The top 5 potential anti-colorectal cancer drugs are listed in Table 2. Moreover, our method reveals some potential candidate drugs with evidence suggesting their capacity to inhibit the development of colorectal cancer. For example, dexamethasone, a glucocorticoid administered in various forms, is employed for the treating diverse inflammatory conditions, such as bronchial asthma, as well as endocrine and rheumatic disorders. Kim et al.[40] demonstrated that the inhibitory effects of dexamethasone on cell migration and invasion by suppressing epithelial-mesenchymal transition (EMT) in colon cancer cell lines under hypoxic condition.

Indomethacin, a nonsteroidal anti-inflammatory (NSAID), is employed for managing chronic musculoskeletal pain conditions and inducing closure of a hemodynamically significant patent ductus arteriosus in premature infants. Ikawa et al.[41] reported that indomethacin exhibited a potential antagonizing effect on human EP(2) receptors in LS174T human colon cancer cells.

Pentamidine, an antifungal agent, is used in the treatment of pneumocystis pneumonia in HIV-infected patients. Seguella et al.[42] introduced that pneumocystis blocked S100B activity, thereby rescuing wild-type p53 expression and determining pro-apoptotic control in colon cancer.

Colchicine is an alkaloid used to treat gout and familial Mediterranean fever as well as prevent major cardiovascular events. Kumar et al.[43] suggested that colchicine triggered apoptosis and autophagy in HCT-116 colon cancer cells.

Estradiol, an estrogenic steroid employed in treating vasomotor symptoms of vulvar and vaginal atrophy in menopause, hypoestrogenism, prevention of postmenopausal osteoporosis, treatment of breast cancer, and advanced androgen-dependent carcinoma of the prostate. Zamani et al.[44] demonstrated that estradiol could alter the migration, juxtacrine, and paracrine activities of colorectal cancer stem cells. Hsu et al.[45] found that estradiol inhibited colorectal cancer cell proliferation and migration by targeting p53.

This implies that these drugs may intervene with the disease process by modulating gene interactions (edges) and can emerge as meaningful treatment options.

Case study: Breast cancer

We then use two breast cancer datasets (GSE31448 and GSE29044) to illustrate the effectiveness of iEdgePathDDA in prioritizing candidate drugs for cancer. iEdgePathDDA identifies the top 5 potential anti-breast cancer drugs based on average rank of two breast cancer datasets (Table 3). Some candidate drugs with positive evidence have also been identified. For instance, deferoxamine is a chelating agent used to treat iron or aluminum toxicity and some blood transfusion dependent anemias. Chen et al.[46] demonstrated that deferoxamine could upregulate expression of TfR1 and DMT1, enhancing iron uptake through the activation of the IL-6/PI3K/AKT signaling pathway in aggressive triple-negative breast cancers (TNBCs).

Mesalazine is an aminosalicylate drug used to treat mild to moderate active ulcerative colitis and also to maintain remission once achieved. It was also used in the chemoprophylaxis of colorectal cancer associated with these conditions.[50] So far, there are no literatures reported that mesalazine can treat breast cancer. Our approach predicts that mesalazine may be effective treatment for breast cancer.

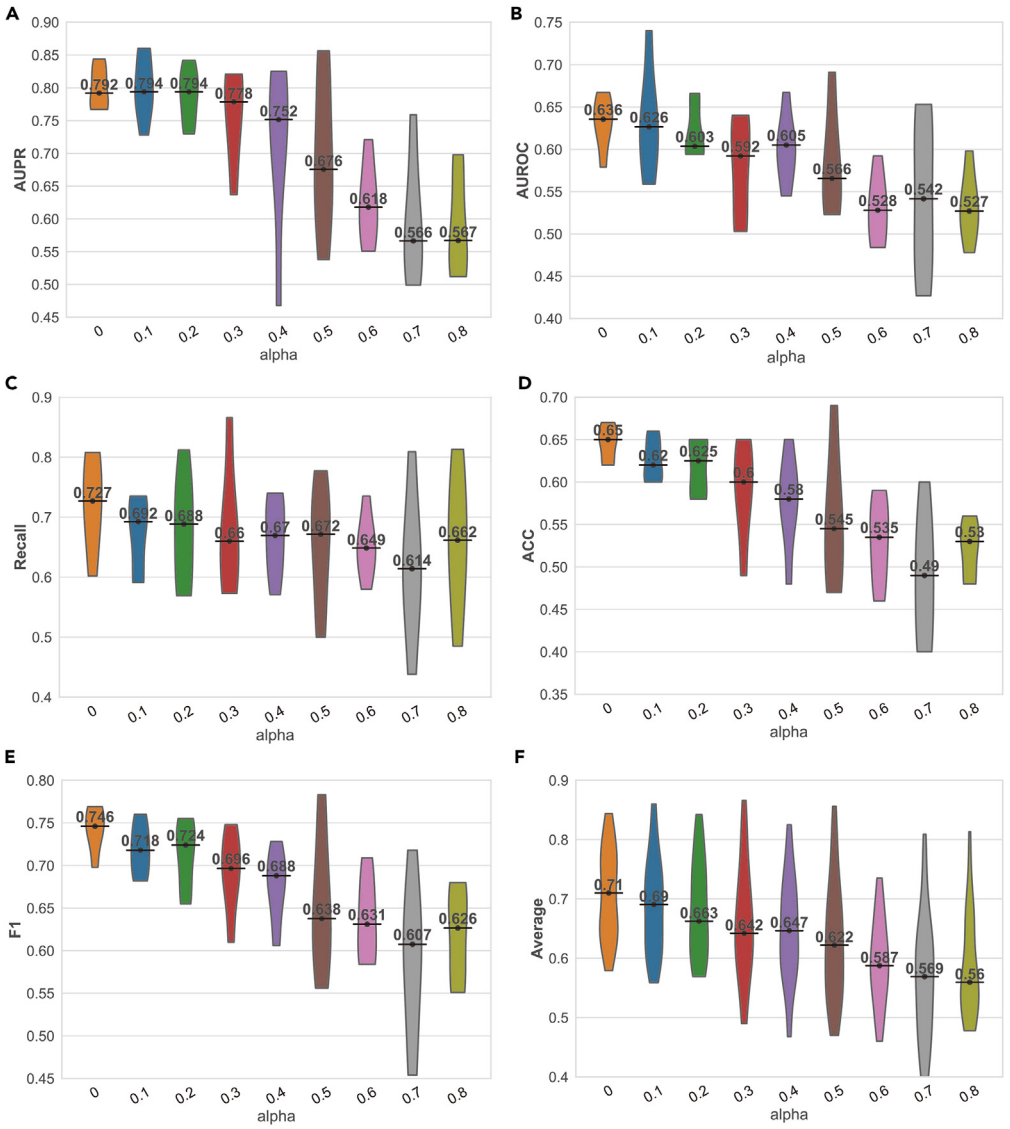


Figure 2. The performance of iEdgePathDDA of five metrics (AUPR, AUROC, Recall, ACC, F1) and the average values using six datasets with different hyperparameters

(A) AUPR, (B) AUROC, (C) Recall, (D) ACC, (E) F1, (F) Average. Average: the average of the five metrics.

Sulfasalazine is a salicylate used to treat Crohn's disease, ulcerative colitis, and rheumatoid arthritis. Wei et al.[47] found that sulfasalazine and vitamin E succinate (VES) had synergistic or antagonistic cytotoxic effects depending on VES concentration against triple-negative breast cancer cells.

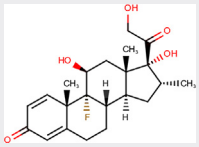
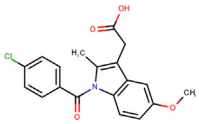
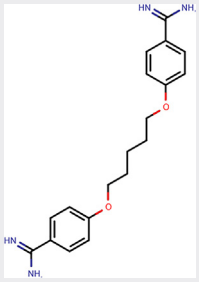
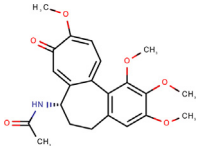
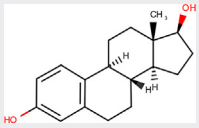
Verapamil is a non-dihydropyridine calcium channel blocker used in the treatment of angina, arrhythmia, and hypertension. Li et al.[48] reported that the combination of paclitaxel and verapamil could inhibit cell proliferation by arresting the progression of the cell cycle and promoting cell apoptosis in breast cancer cells.

Diclofenac is a nonsteroidal anti-inflammatory (NSAID) used to treat the signs and symptoms of osteoarthritis and rheumatoid arthritis. Yang et al.[49] suggested that diclofenac inhibited cell glycolysis and suppressed TNBC cell growth by decreasing GLUT1 protein expression and HK activity through the c-Myc pathway.

Case study: Lung cancer

We employ three lung cancer datasets (GSE19844, GSE18842, and GSE19804) to demonstrate the efficacy of iEdgePathDDA in prioritizing candidate drugs associated with lung cancer. iEdgePathDDA identifies the top 5 potential anti-lung cancer drugs based on the average rank

Table 2. Top 5 candidate drugs for colorectal cancer identified by iEdgePathDDA

DrugBank	Drug name	MI	MOA	Evidence
DB01234	Dexamethasone 	Bronchial asthma, endocrine and rheumatic disorders.	Inhibit neutrophil apoptosis and demargination	Kim et al.[40]
DB00328	Indomethacin 	Rheumatoid arthritis, osteoarthritis	Inhibitor of the cyclo-oxygenase enzyme or prostaglandin G/H synthase	Ikawa et al.[41]
DB00738	Pentamidine 	Pneumocystis pneumonia in patients infected with HIV	Inhibit the synthesis of DNA, RNA, phospholipids, and proteins.	Seguella et al.[42]
DB01394	Colchicine 	Prophylaxis and gout flares	Inhibitor of tubulin beta chain	Kumar et al.[43]
DB00783	Estradiol 	Osteoporosis, breast cancer, and prostate cancer	Inhibitor of ATP synthase subunit a	Zamani et al.[44]

MI, Main Indication; MOA, Mechanism of Action.

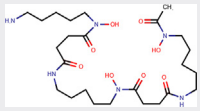
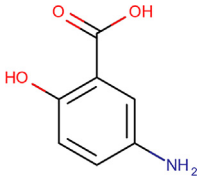
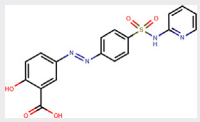
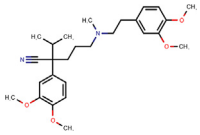
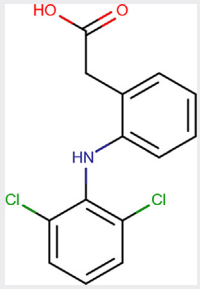
across three datasets (Table 4). Several candidate drugs with positive evidence have also been discerned. For example, Verapamil is a non-dihydropyridine calcium channel blocker used in the treatment of angina, arrhythmia, and hypertension. Huang et al.[51] found that the combination of verapamil and chemotherapeutic drugs could enhance the clinical outcomes in advanced lung cancer patients and increase the efficacy of chemotherapeutic agents.

Sulfasalazine, a salicylate prescribed for Crohn's disease, ulcerative colitis, and rheumatoid arthritis, demonstrated significant potential in enhancing chemotherapy efficacy for lung cancer according to findings by Lay et al.[55] Moreover, Bagherpoor et al.[52] reported that the combinations of sulfasalazine and disulfiram-copper displayed inhibitory effects on lung adenocarcinoma in both cell line models and mice.

Bucladesine, recognized for its role as a cell-permeable cAMP analog, holds versatile applications in research due to its ability to mimics cAMP and can induce normal physiological responses under experimental conditions. Zhang et al.[53] highlighted the inhibitory effects of dibutyryl cAMP (bucladesine) on the growth and differentiation of metastatic human lung cancer cells.

Chlorpromazine, a phenothiazine antipsychotic with applications ranging from treating nausea and preoperative anxiety to managing severe behavioral problems, schizophrenia and bipolar disorder, is proposed by Fujiwara et al.[54] as holding therapeutic potential for non-small cell lung cancer (NSCLC) with mutated EGFR. Importantly, this potential is attributed to a novel mechanism distinct from conventional tyrosine kinase inhibitors (TKLs) either alone or in combination with other agents.

Table 3. Top 5 candidate drugs for breast cancer identified by iEdgePathDDA

DrugBank	Drug name	MI	MOA	Evidence
DB00746	Deferoxamine 	Acute iron or aluminum toxicity	Binding trivalent iron	Chen et al.[46]
DB00244	Mesalazine 	ulcerative colitis	Inhibitor of nuclear factor kappa-B kinase subunit beta	–
DB00795	Sulfasalazine 	Crohn's disease, ulcerative colitis, and rheumatoid arthritis	Inhibitor of prostaglandin G/H synthase 1/2	Wei et al.[47]
DB00661	Verapamil  0	angina, arrhythmia, and hypertension	Inhibit L-type calcium channels	Li et al.[48]
DB00586	Diclofenac  1	Osteoarthritis and rheumatoid arthritis	Inhibitor of prostaglandin G/H synthase 1/2	Yang et al.[49]

MI, Main Indication; MOA, Mechanism of Action.

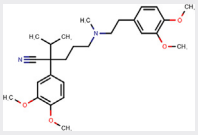
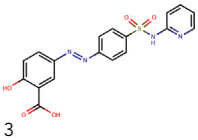
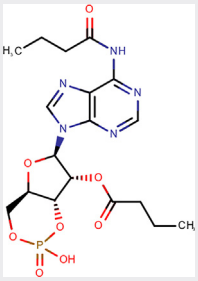
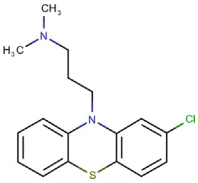
Alpha-estradiol, recognized as both an estrogen and 5 alpha-reductase inhibitor primarily employed in topical drug treatment of hair loss,[56] lacks existing literature supporting its role in lung cancer treatment. Nevertheless, our approach predicts a potential therapeutic effect of alpha-estradiol for lung cancer.

These findings underscore the capability of iEdgePathDDA to present precise and diverse treatment options for patients, potentially expediting drug discovery. The results open new avenues for research and clinical applications, offering new prospects and opportunities for further investigation.

Compare with other drug repurposing methods

To assess the efficacy of our method, we conduct a comprehensive comparison with four existing approaches using five metrics (AUPR, AURPC, Recall, ACC, and F1). As depicted in Table 5, iEdgePathDDA outperforms other drug repurposing methods across all five metrics on six datasets. Notably, it achieves exceptional performance, achieving an AUPR, AUROC, Recall, ACC and F1 score of 0.799, 0.631, 0.722, 0.645 and 0.743, respectively. Furthermore, a comparison with suboptimal methods reveals substantial improvements in

Table 4. Top 5 candidate drugs for lung cancer identified by iEdgePathDDA

DrugBank	Drug name	MI	MOA	Evidence
DB00661	Verapamil  2	angina, arrhythmia, and hypertension	Inhibit L-type calcium channels	Huang et al.[51]
DB00795	Sulfasalazine  3	Crohn's disease, ulcerative colitis, and rheumatoid arthritis	Inhibitor of prostaglandin G/H synthase 1/2	Lay et al.[52]
DB13242	Bucladesine 	–	Plasminogen activator inhibitor 1	Zhang et al.[53]
DB00477	Chlorpromazine 	Schizophrenia, nausea, vomiting, preoperative anxiety	Antagonist on dopamine D1/D2 receptor	Fujiwara et al.[54]
–	Alpha-estradiol	Hair loss	–	–

MI, Main Indication; MOA, Mechanism of Action.

iEdgePathDDA, with increases of 14.4% in AUPR, 9.3% in AUROC, 8.5% in Recall, 5.4% in ACC, and 18.3% in F1. These enhancements underscore a significant improvement in the model's capacity to detect relevant associations while maintaining accuracy. The results affirm that the improvements achieved by iEdgePathDDA are both comprehensive and substantial, indicating its practical significance in identifying drug-disease associations. Therefore, the incorporation of gene interactions (edges) within pathways contributes to the enhanced predictive performance of iEdgePathDDA, highlighting its practical importance in improving prediction accuracy and advancing the identification of drug-disease associations.

Table 5. Performance comparison in different methods based on six datasets

Method	AUPR	AUROC	Recall	ACC	F1
iEdgePathDDA	**0.799** ± 0.032	**0.631** ± 0.030	**0.722** ± 0.078	**0.645** ± 0.020	**0.743** ± 0.026
CMAP	0.443 ± 0.043	0.428 ± 0.040	0.315 ± 0.081	0.591 ± 0.031	0.436 ± 0.090
NP	0.655 ± 0.072	0.505 ± 0.040	0.380 ± 0.098	0.446 ± 0.053	0.469 ± 0.088
SubtypeDrug	0.579 ± 0.131	0.538 ± 0.034	0.602 ± 0.092	0.481 ± 0.021	0.560 ± 0.049
DRviaSPCN	0.497 ± 0.068	0.520 ± 0.013	0.637 ± 0.071	0.488 ± 0.033	0.537 ± 0.045

NP, Network proximity.

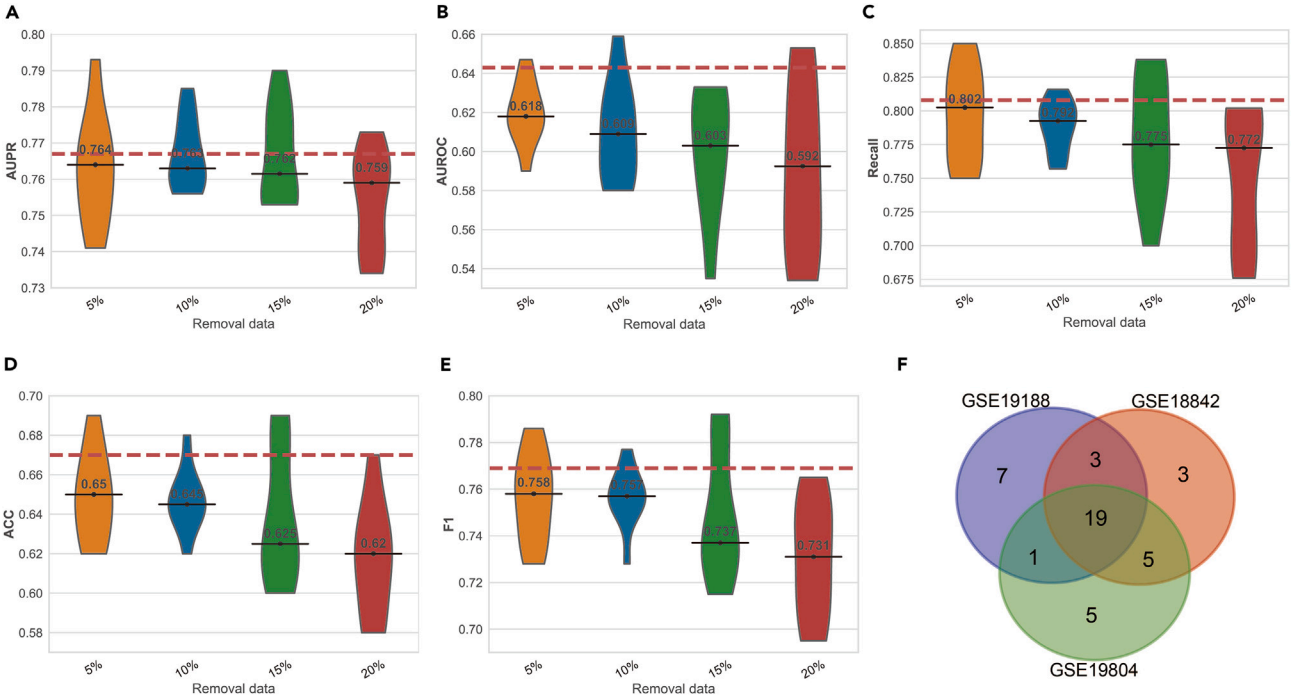


Figure 3. Performance of partially removing edges based on GSE19188 using our proposed method

(A) AUPR (0.767), (B)AUROC (0.643), (C) Recall (0.808), (D)ACC (0.67), (E) F1 (0.769), (F) Venn diagram shows the overlapped drugs among three datasets. Bracket represents value of the original data in red line.

Robustness analysis

To validate the stability and robustness of the iEdgePathDDA method, we perform the data removal tests using the GSE19188 dataset. We remove 5%, 10%, 15%, and 20% of the edges and repeat the iEdgePathDDA method 10 times for each removal. For each removal, we derive the candidate drugs by IS and then recalculate the AUPR, AUROC, Recall, ACC and F1 based on the CTD database.[57] The results show that the AUPR, AUROC, Recall, ACC and F1 decrease slowly compared to the original data, and the median AUPR, AUROC, Recall, ACC and F1 are still above 0.759, 0.592, 0.772, 0.62, and 0.731 even removing 20% of the edges, indicating that the iEdgePathDDA is robust to data removal (Figures 3A–3E).

Furthermore, we test the reproducibility of our method using two additional gene expression datasets (GSE18842 and GSE19804). Employing the iEdgePathDDA approach on these datasets and comparing the top 30 drugs for each ranked drug list, we observe that 63.3% (19/30) of drugs are consistently present in all three datasets (Figure 3F). This consistency across datasets underscores the remarkable robustness and stability of our approach in diverse datasets.

Conclusion

With the escalating availability of high-throughput sequencing technologies and advancements in bioinformatics, computational methodologies for drug repurposing have gained prominence over traditional experimental techniques. This study presents a drug repurposing method, which leverages the changes of drug-induced or disease-related gene interactions within pathway to identify potential candidate drugs for cancer treatment. Through a comprehensive case study, we showcase the practical predictive capability of iEdgePathDDA, revealing previously undiscovered drug-disease associations. To further underscore the predictive power of our approach, we compare it with other drug repurposing approaches in six datasets. The results show that our approach could achieve excellent predictive performance in six datasets. Ultimately, the results of the robustness analysis confirm the reliability and credibility of iEdgePathDDA and provide a solid foundation for its application in future research.

Limitations of the study

While iEdgePathDDA exhibits superior performance, there are still some limitations. First, numerous diseases are multifactorial, involving complex biological pathways, computational models may oversimplify these complexities. Second, the computational methods lack the indispensable support of experimental validation for the identified drug candidates. Essential experimental studies, encompassing *in vitro* or *in vivo* assays, are imperative to validate the efficacy of the predicted drug.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- METHOD DETAILS
  - Gene expression profiles
  - All pathways from KEGG database
  - Drug-induced gene expression profiles
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Calculating inhibition score between drug-induced edges with disease-related edges
  - Identifying candidate anticancer drugs
  - Performance evaluation
  - Baseline methods
- QUANTIFICATION AND STATISTICAL ANALYSIS

ACKNOWLEDGMENTS

This work was supported in part by the National Natural Science Foundation of China (grant no. 62202115, 62102104), the Fundamental Research Funds for the Provincial Universities of Zhejiang GK239909299001-019, and the Research Foundation of Hangzhou Dianzi University (KYS335622091; KYH333122029M). The authors acknowledge the Supercomputing Center of Hangzhou Dianzi University for providing computing resources.

AUTHOR CONTRIBUTIONS

Participated in research design: X.L., X.Z., T.L., X.D., H.Z., Q.L., Z.B., and J.L. Performed data analysis: X.L. and J.L. Wrote or contributed to the writing of the manuscript: X.L., X.Z., T.L., X.D., H.Z., Q.L., Z.B., and J.L. Overall supervision of the project: X.L., X.Z., T.L., X.D., H.Z., Q.L., Z.B., and J.L.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: January 31, 2024
Revised: April 6, 2024
Accepted: May 15, 2024
Published: May 18, 2024

REFERENCES

1. Ashburn, T.T., and Thor, K.B. (2004). Drug repositioning: identifying and developing new uses for existing drugs. Nat. Rev. Drug Discov. *3*, 673–683. https://doi.org/10.1038/nrd1468.
2. DiMasi, J.A., Hansen, R.W., and Grabowski, H.G. (2003). The price of innovation: new estimates of drug development costs. J. Health Econ. *22*, 151–185. https://doi.org/10.1016/s0167-6296(02)00126-1.
3. Adams, C.P., and Brantner, V.V. (2006). Estimating the cost of new drug development: is it really 802 million dollars? Health Aff. *25*, 420–428. https://doi.org/10.1377/hlthaff.25.2.420.
4. Yella, J.K., Yaddanapudi, S., Wang, Y., and Jegga, A.G. (2018). Changing Trends in Computational Drug Repositioning. Pharmaceuticals *11*, 57. https://doi.org/10.3390/ph11020057.
5. Pushpakom, S., Iorio, F., Eyers, P.A., Escott, K.J., Hopper, S., Wells, A., Doig, A., Guilliams, T., Latimer, J., McNamee, C., et al. (2019). Drug repurposing: progress, challenges and recommendations. Nat. Rev. Drug Discov. *18*, 41–58. https://doi.org/10.1038/nrd.2018.168.
6. Parvathaneni, V., Kulkarni, N.S., Muth, A., and Gupta, V. (2019). Drug repurposing: a promising tool to accelerate the drug discovery process. Drug Discov. Today *24*, 2076–2085. https://doi.org/10.1016/j.drudis.2019.06.014.
7. Panda, S., Kumari, L., Badwaik, H.R., and Shanmugarajan, D. (2022). Computational approaches for drug repositioning and repurposing to combat SARS-CoV-2 infection. In Computational Approaches for Novel Therapeutic and Diagnostic Designing to Mitigate SARS-CoV-2 Infection. Copyright © 2022, A. Parihar, R. Khan, A. Kumar, A.K. Kaushik, and H. Gohel, eds. (Elsevier Inc. All rights reserved), pp. 247–265. https://doi.org/10.1016/b978-0-323-91172-6.00008-x.
8. Ahmed, F., Samantasinghar, A., Soomro, A.M., Kim, S., and Choi, K.H. (2023). A systematic review of computational approaches to understand cancer biology for informed drug repurposing. J. Biomed. Inform. *142*, 104373. https://doi.org/10.1016/j.jbi.2023.104373.
9. Ahmed, F., Kang, I.S., Kim, K.H., Asif, A., Rahim, C.S.A., Samantasinghar, A., Memon, F.H., and Choi, K.H. (2023). Drug repurposing for viral cancers: A paradigm of machine learning, deep learning, and virtual screening-based approaches. J. Med. Virol. *95*, e28693. https://doi.org/10.1002/jmv.28693.
10. Ahmed, F., Soomro, A.M., Chethikkattuveli Salih, A.R., Samantasinghar, A., Asif, A., Kang, I.S., and Choi, K.H. (2022). A comprehensive review of artificial intelligence and network based approaches to drug repurposing in Covid-19. Biomed. Pharmacother. *153*,

113350. https://doi.org/10.1016/j.biopha.2022.113350.

11. Ahmed, F., Lee, J.W., Samantasinghar, A., Kim, Y.S., Kim, K.H., Kang, I.S., Memon, F.H., Lim, J.H., and Choi, K.H. (2022). SperoPredictor: An Integrated Machine Learning and Molecular Docking-Based Drug Repurposing Framework With Use Case of COVID-19. Front. Public Health *10*, 902123. https://doi.org/10.3389/fpubh.2022.902123.
12. Ahmed, F., Yang, Y.J., Samantasinghar, A., Kim, Y.W., Ko, J.B., and Choi, K.H. (2023). Network-based drug repurposing for HPV-associated cervical cancer. Comput. Struct. Biotechnol. J. *21*, 5186–5200. https://doi.org/10.1016/j.csbj.2023.10.038.
13. Ahmed, F., Samantasinghar, A., Ali, W., and Choi, K.H. (2024). Network-based drug repurposing identifies small molecule drugs as immune checkpoint inhibitors for endometrial cancer. Mol. Divers. *21*, 5186. https://doi.org/10.1007/s11030-023-10784-7.
14. Samantasinghar, A., Ahmed, F., Rahim, C.S.A., Kim, K.H., Kim, S., and Choi, K.H. (2023). Artificial intelligence-assisted repurposing of lubiprostone alleviates tubulointerstitial fibrosis. Transl. Res. *262*, 75–88. https://doi.org/10.1016/j.trsl.2023.07.010.
15. Samantasinghar, A., Sunildutt, N.P., Ahmed, F., Soomro, A.M., Salih, A.R.C., Parihar, P., Memon, F.H., Kim, K.H., Kang, I.S., and Choi, K.H. (2023). A comprehensive review of key factors affecting the efficacy of antibody drug conjugate. Biomed. Pharmacother. *161*, 114408. https://doi.org/10.1016/j.biopha.2023.114408.
16. Asif, A., Park, S.H., Manzoor Soomro, A., Khalid, M.A.U., Salih, A.R.C., Kang, B., Ahmed, F., Kim, K.H., and Choi, K.H. (2021). Microphysiological system with continuous analysis of albumin for hepatotoxicity modeling and drug screening. J. Ind. Eng. Chem. *98*, 318–326. https://doi.org/10.1016/j.jiec.2021.03.035.
17. Ahmed, F., Ho, S.G., Samantasinghar, A., Memon, F.H., Rahim, C.S.A., Soomro, A.M., Choi, K.H., Sunildutt, N., Sunildutt, N., and Kim, K.H. (2022). Drug repurposing in psoriasis, performed by reversal of disease-associated gene expression profiles. Comput. Struct. Biotechnol. J. *20*, 6097–6107. https://doi.org/10.1016/j.csbj.2022.10.046.
18. Lotfi Shahreza, M., Ghadiri, N., Mousavi, S.R., Varshosaz, J., and Green, J.R. (2018). A review of network-based approaches to drug repositioning. Brief. Bioinform. *19*, 878–892. https://doi.org/10.1093/bib/bbx017.
19. Lamb, J., Crawford, E.D., Peck, D., Modell, J.W., Blat, I.C., Wrobel, M.J., Lerner, J., Brunet, J.P., Subramanian, A., Ross, K.N., et al. (2006). The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. Science (New York, N.Y.) *313*, 1929–1935. https://doi.org/10.1126/science.1132939.
20. Dudley, J.T., Sirota, M., Shenoy, M., Pai, R.K., Roedder, S., Chiang, A.P., Morgan, A.A., Sarwal, M.M., Pasricha, P.J., and Butte, A.J. (2011). Computational repositioning of the anticonvulsant topiramate for inflammatory bowel disease. Sci. Transl. Med. *3*, 96ra76. https://doi.org/10.1126/scitranslmed.3002648.
21. Kosaka, T., Nagamatsu, G., Saito, S., Oya, M., Suda, T., and Horimoto, K. (2013). Identification of drug candidate against prostate cancer from the aspect of somatic cell reprogramming. Cancer Sci. *104*, 1017–1026. https://doi.org/10.1111/cas.12183.
22. van Noort, V., Schölch, S., Iskar, M., Zeller, G., Ostertag, K., Schweitzer, C., Werner, K., Weitz, J., Koch, M., and Bork, P. (2014). Novel drug candidates for the treatment of metastatic colorectal cancer through global inverse gene-expression profiling. Cancer Res. *74*, 5690–5699. https://doi.org/10.1158/0008-5472.can-13-3540.
23. Pacini, C., Iorio, F., Gonçalves, E., Iskar, M., Klabunde, T., Bork, P., and Saez-Rodriguez, J. (2013). DvD: An R/Cytoscape pipeline for drug repurposing using public repositories of gene expression data. Bioinformatics *29*, 132–134. https://doi.org/10.1093/bioinformatics/bts656.
24. Napolitano, F., Zhao, Y., Moreira, V.M., Tagliaferri, R., Kere, J., D'Amato, M., and Greco, D. (2013). Drug repositioning: a machine-learning approach through data integration. J. Cheminform. *5*, 30. https://doi.org/10.1186/1758-2946-5-30.
25. Chen, Y.W., Diamante, G., Ding, J., Nghiem, T.X., Yang, J., Ha, S.M., Cohn, P., Arneson, D., Blencowe, M., Garcia, J., et al. (2022). PharmOmics: A species- and tissue-specific drug signature database and gene-network-based drug repositioning tool. iScience *25*, 104052. https://doi.org/10.1016/j.isci.2022.104052.
26. Papikinos, T., Krokidis, M.G., Vrahatis, A., Vlamos, P., and Exarchos, T.P. (2023). Signature-Based Computational Drug Repurposing for Amyotrophic Lateral Sclerosis. Adv. Exp. Med. Biol. *1424*, 201–211. https://doi.org/10.1007/978-3-031-31982-2_22.
27. Han, X., Kong, Q., Liu, C., Cheng, L., and Han, J. (2021). SubtypeDrug: a software package for prioritization of candidate cancer subtype-specific drugs. Bioinformatics *37*, 2491–2493. https://doi.org/10.1093/bioinformatics/btab011.
28. Wu, J., Li, X., Wang, Q., and Han, J. (2022). DRviaSPCN: a software package for drug repurposing in cancer via a subpathway crosstalk network. Bioinformatics *38*, 4975–4977. https://doi.org/10.1093/bioinformatics/btac611.
29. Li, X., Liao, M., Wang, B., Zan, X., Huo, Y., Liu, Y., Bao, Z., Xu, P., and Liu, W. (2023). A drug repurposing method based on inhibition effect on gene regulatory network. Comput. Struct. Biotechnol. J. *21*, 4446–4455. https://doi.org/10.1016/j.csbj.2023.09.007.
30. Iwata, M., Hirose, L., Kohara, H., Liao, J., Sawada, R., Akiyoshi, S., Tani, K., and Yamanishi, Y. (2018). Pathway-Based Drug Repositioning for Cancers: Computational Prediction and Experimental Validation. J. Med. Chem. *61*, 9583–9595. https://doi.org/10.1021/acs.jmedchem.8b01044.
31. Napolitano, F., Carrella, D., Mandriani, B., Pisonero-Vaquero, S., Sirci, F., Medina, D.L., Brunetti-Pierri, N., and di Bernardo, D. (2018). gene2drug: a computational tool for pathway-based rational drug repositioning. Bioinformatics *34*, 1498–1505. https://doi.org/10.1093/bioinformatics/btx800.
32. Di, J., Zheng, B., Kong, Q., Jiang, Y., Liu, S., Yang, Y., Han, X., Sheng, Y., Zhang, Y., Cheng, L., et al. (2019). Prioritization of candidate cancer drugs based on a drug functional similarity network constructed by integrating pathway activities and drug activities. Mol. Oncol. *13*, 2259–2277. https://doi.org/10.1002/1878-0261.12564.
33. Wang, Y., Qi, J., and Ai, D. (2023). DPADM: a novel algorithm for detecting drug-pathway associations based on high-throughput transcriptional response to compounds. Brief. Bioinform. *24*, bbac517. https://doi.org/10.1093/bib/bbac517.
34. Uddin, S., Ahmed, M., Hussain, A., Abubaker, J., Al-Sanea, N., AbdulJabbar, A., Ashari, L.H., Alhomoud, S., Al-Dayel, F., Jehan, Z., et al. (2011). Genome-wide expression analysis of Middle Eastern colorectal cancer reveals FOXM1 as a novel target for cancer therapy. Am. J. Pathol. *178*, 537–547. https://doi.org/10.1016/j.ajpath.2010.10.020.
35. Sabatier, R., Finetti, P., Adelaide, J., Guille, A., Borg, J.P., Chaffanet, M., Lane, L., Birnbaum, D., and Bertucci, F. (2011). Down-regulation of ECRG4, a candidate tumor suppressor gene, in human breast cancer. PLoS One *6*, e27656. https://doi.org/10.1371/journal.pone.0027656.
36. Colak, D., Nofal, A., Albakheet, A., Nirmal, M., Jeprel, H., Eldali, A., Al-Tweigeri, T., Tulbah, A., Ajarim, D., Malik, O.A., et al. (2013). Age-specific gene expression signatures for breast tumors and cross-species conserved potential cancer progression markers in young women. PLoS One *8*, e63204. https://doi.org/10.1371/journal.pone.0063204.
37. Sanchez-Palencia, A., Gomez-Morales, M., Gomez-Capilla, J.A., Pedraza, V., Boyero, L., Rosell, R., and Fárez-Vidal, M.E. (2011). Gene expression profiling reveals novel biomarkers in nonsmall cell lung cancer. Int. J. Cancer *129*, 355–364. https://doi.org/10.1002/ijc.25704.
38. Hou, J., Aerts, J., den Hamer, B., van Ijcken, W., den Bakker, M., Riegman, P., van der Leest, C., van der Spek, P., Foekens, J.A., Hoogsteden, H.C., et al. (2010). Gene expression-based classification of non-small cell lung carcinomas and survival prediction. PLoS One *5*, e10312. https://doi.org/10.1371/journal.pone.0010312.
39. Lu, T.P., Hsiao, C.K., Lai, L.C., Tsai, M.H., Hsu, C.P., Lee, J.M., and Chuang, E.Y. (2015). Identification of regulatory SNPs associated with genetic modifications in lung adenocarcinoma. BMC Res. Notes *8*, 92. https://doi.org/10.1186/s13104-015-1053-8.
40. Kim, J.H., Hwang, Y.J., Han, S.H., Lee, Y.E., Kim, S., Kim, Y.J., Cho, J.H., Kwon, K.A., Kim, J.H., and Kim, S.H. (2015). Dexamethasone inhibits hypoxia-induced epithelial-mesenchymal transition in colon cancer. World J. Gastroenterol. *21*, 9887–9899. https://doi.org/10.3748/wjg.v21.i34.9887.
41. Ikawa, Y., Fujino, H., Otake, S., and Murayama, T. (2012). Indomethacin antagonizes EP(2) prostanoid receptor activation in LS174T human colon cancer cells. Eur. J. Pharmacol. *680*, 16–21. https://doi.org/10.1016/j.ejphar.2012.01.033.
42. Seguella, L., Rinaldi, F., Marianecci, C., Capuano, R., Pesce, M., Annunziata, G., Casano, F., Bassotti, G., Sidoni, A., Milone, M., et al. (2020). Pentamidine niosomes thwart S100B effects in human colon carcinoma biopsies favouring wtp53 rescue. J. Cell Mol. Med. *24*, 3053–3063. https://doi.org/10.1111/jcmm.14943.
43. Kumar, A., Singh, B., Sharma, P.R., Bharate, S.B., Saxena, A.K., and Mondhe, D.M. (2016). A novel microtubule depolymerizing colchicine analogue triggers apoptosis and autophagy in HCT-116 colon cancer cells. Cell Biochem. Funct. *34*, 69–81. https://doi.org/10.1002/cbf.3166.

44. Zamani, A.R.N., Avci, Ç.B., Ahmadi, M., Pouyafar, A., Bagheri, H.S., Fathi, F., Heidarzadeh, M., Rezaie, J., Mirhosseini, Y., Saberianpour, S., et al. (2020). Estradiol modulated colorectal cancer stem cells bioactivity and interaction with endothelial cells. Life Sci. *257*, 118078. https://doi.org/10.1016/j.lfs.2020.118078.
45. Hsu, H.H., Kuo, W.W., Ju, D.T., Yeh, Y.L., Tu, C.C., Tsai, Y.L., Shen, C.Y., Chang, S.H., Chung, L.C., and Huang, C.Y. (2014). Estradiol agonists inhibit human LoVo colorectal-cancer cell proliferation and migration through p53. World J. Gastroenterol. *20*, 16665–16673. https://doi.org/10.3748/wjg.v20.i44.16665.
46. Chen, C., Liu, P., Duan, X., Cheng, M., and Xu, L.X. (2019). Deferoxamine-induced high expression of TfR1 and DMT1 enhanced iron uptake in triple-negative breast cancer cells by activating IL-6/PI3K/AKT pathway. OncoTargets Ther. *12*, 4359–4377. https://doi.org/10.2147/ott.s193507.
47. Wei, C.W., Yu, Y.L., Lu, J.Y., Hung, Y.T., Liu, H.C., and Yiang, G.T. (2019). Anti-Cancer Effects of Sulfasalazine and Vitamin E Succinate in MDA-MB 231 Triple-Negative Breast Cancer Cells. Int. J. Med. Sci. *16*, 494–500. https://doi.org/10.7150/ijms.30380.
48. Li, P., Zhong, D., and Gong, P.Y. (2019). Synergistic effect of paclitaxel and verapamil to overcome multi-drug resistance in breast cancer cells. Biochem. Biophys. Res. Commun. *516*, 183–188. https://doi.org/10.1016/j.bbrc.2019.05.189.
49. Yang, L., Li, J., Li, Y., Zhou, Y., Wang, Z., Zhang, D., Liu, J., and Zhang, X. (2021). Diclofenac impairs the proliferation and glucose metabolism of triple-negative breast cancer cells by targeting the c-Myc pathway. Exp. Ther. Med. *21*, 584. https://doi.org/10.3892/etm.2021.10016.
50. Słoka, J., Madej, M., and Strzalka-Mrozik, B. (2023). Molecular Mechanisms of the Antitumor Effects of Mesalazine and Its Preventive Potential in Colorectal Cancer. Molecules *28*, 5081. https://doi.org/10.3390/molecules28135081.
51. Huang, J., Zhang, T., Ma, K., Fan, P., Liu, Y., Weng, C., Fan, G., Duan, Q., and Zhu, X. (2013). Clinical evaluation of targeted arterial perfusion of verapamil and chemotherapeutic drugs in interventional therapy of advanced lung cancer. Cancer Chemother. Pharmacol. *72*, 889–896. https://doi.org/10.1007/s00280-013-2271-1.
52. Bagherpoor, A.J., Shameem, M., Luo, X., Seelig, D., and Kassie, F. (2023). Inhibition of lung adenocarcinoma by combinations of sulfasalazine (SAS) and disulfiram-copper (DSF-Cu) in cell line models and mice. Carcinogenesis *44*, 291–303. https://doi.org/10.1093/carcin/bgad020.
53. Zhang, J., Fang, W., and Zheng, J. (1997). [Regulatory effects of cAMP analogs on growth and differentiation of metastatic human lung cancer cells]. Zhonghua Bing Li Xue Za Zhi *26*, 16–18.
54. Fujiwara, R., Taniguchi, Y., Rai, S., Iwata, Y., Fujii, A., Fujimoto, K., Kumode, T., Serizawa, K., Morita, Y., Espinoza, J.L., et al. (2022). Chlorpromazine cooperatively induces apoptosis with tyrosine kinase inhibitors in EGFR-mutated lung cancer cell lines and restores the sensitivity to gefitinib in T790M-harboring resistant cells. Biochem. Biophys. Res. Commun. *626*, 156–166. https://doi.org/10.1016/j.bbrc.2022.08.010.
55. Lay, J.D., Hong, C.C., Huang, J.S., Yang, Y.Y., Pao, C.Y., Liu, C.H., Lai, Y.P., Lai, G.M., Cheng, A.L., Su, I.J., et al. (2007). Sulfasalazine suppresses drug resistance and invasiveness of lung adenocarcinoma cells expressing AXL. Cancer Res. *67*, 3878–3887. https://doi.org/10.1158/0008-5472.can-06-3191.
56. Kim, J.H., Lee, S.Y., Lee, H.J., Yoon, N.Y., and Lee, W.S. (2012). The Efficacy and Safety of 17α-Estradiol (Ell-Cranell® alpha 0.025%) Solution on Female Pattern Hair Loss: Single Center, Open-Label, Non-Comparative, Phase IV Study. Ann. Dermatol. *24*, 295–305. https://doi.org/10.5021/ad.2012.24.3.295.
57. Davis, A.P., Grondin, C.J., Lennon-Hopkins, K., Saraceni-Richards, C., Sciaky, D., King, B.L., Wiegers, T.C., and Mattingly, C.J. (2015). The Comparative Toxicogenomics Database's 10th year anniversary: update 2015. Nucleic Acids Res. *43*, D914–D920. https://doi.org/10.1093/nar/gku935.
58. Ogata, H., Goto, S., Sato, K., Fujibuchi, W., Bono, H., and Kanehisa, M. (1999). KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. *27*, 29–34.
59. Geistlinger, L., Csaba, G., and Zimmer, R. (2016). Bioconductor's EnrichmentBrowser: seamless navigation through combined results of set- & network-based enrichment analysis. BMC Bioinf. *17*, 45. https://doi.org/10.1186/s12859-016-0884-1.
60. Tarca, A.L., Draghici, S., Bhatti, G., and Romero, R. (2012). Down-weighting overlapping genes improves gene set analysis. BMC Bioinf. *13*, 136. https://doi.org/10.1186/1471-2105-13-136.
61. Peyvandipour, A., Saberian, N., Shafi, A., Donato, M., and Draghici, S. (2018). A novel computational approach for drug repurposing using systems biology. Bioinformatics *34*, 2817–2825. https://doi.org/10.1093/bioinformatics/bty133.
62. Duan, Y., Evans, D.S., Miller, R.A., Schork, N.J., Cummings, S.R., and Girke, T. (2020). signatureSearch: environment for gene expression signature searching and functional interpretation. Nucleic Acids Res. *48*, e124. https://doi.org/10.1093/nar/gkaa878.
63. Debrabant, B., and Soerensen, M. (2014). Gene set analysis for GWAS: assessing the use of modified Kolmogorov-Smirnov statistics. Stat. Appl. Genet. Mol. Biol. *13*, 553–566. https://doi.org/10.1515/sagmb-2013-0015.
64. Cheng, F., Desai, R.J., Handy, D.E., Wang, R., Schneeweiss, S., Barabási, A.L., and Loscalzo, J. (2018). Network-based approach to prediction and population-based validation of in silico drug repurposing. Nat. Commun. *9*, 2691. https://doi.org/10.1038/s41467-018-05116-5.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Source code	This study	https://github.com/eshinesimida/iEdgePathDDA
Colorectal Cancer	Uddin et al.[34]	GSE23878
Breast Cancer	Sabatier et al.[35]	GSE31448
Breast Cancer	Colak et al.[36]	GSE29044
Lung Cancer	Sanchez-Palencia et al.[37]	GSE18842
Lung Cancer	Hou et al.[38]	GSE19188
Lung Cancer	Lu et al.[39]	GSE19804
KEGG pathway database	Ogata et al.[58]	https://www.kegg.jp/
CMAP database	Lamb et al.[19]	https://www.broadinstitute.org/connectivity-map-cmap
CTD database	Davis et al.[57]	https://ctdbase.org/
Software and algorithms		
R programming language V4.1.2	R core	https://www.r-project.org/
EnrichmentBrowser	Geistlinger et al.[59]	https://bioconductor.org/packages/EnrichmentBrowser/
KEGGdzPathwaysGEO	Tarca et al.[60]	http://bioconductor.org/packages/KEGGdzPathwaysGEO/
DrugDiseaseNet	Peyvandipour et al.[61]	https://github.com/azampvd/DrugDiseaseNet
signatureSearchData	Duan et al.[62]	https://bioconductor.org/packages/signatureSearchData/
iEdgePathDDA	This study	https://github.com/eshinesimida/iEdgePathDDA

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Xianbin Li (lixb88@gzhu.edu.cn).

Materials availability

This study did not generate new unique materials.

Data and code availability

- The data mentioned in this paper are publicly available, and are listed in key resources table with accessibility
- All the codes are available online at Github and is fully publicly available as of the date of publication.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

METHOD DETAILS

Gene expression profiles

We download three cancer datasets from the Gene Expression Omnibus (GEO) database, covering colorectal, lung, and breast cancer. Each set of gene expression data contains cancer and control samples. The expression values of each gene are standardized to a normal distribution using the quantile normalization method across all samples. The gene expression profiles are derived from the Affymetrix Human Genome U133 Plus 2.0 array.

All pathways from KEGG database

We download 237 pathways from the KEGG database[58] in the XML format. All pathways are converted into a gene-gene network, where nodes represent genes and edges represent the signals transmitted between these genes. The gene-gene network contains all 5593 genes and 28371 edges present in the pathways extracted from the KEGG database.

Drug-induced gene expression profiles

We download drug-induced gene expression profiles from the CMAP database,[19] including 6,100 instances associated with 1,309 drugs. These instances are measured across five human cancer cell lines, covering breast cancer epithelial cell lines (MCF7, ssMCF7), a prostate cancer epithelial cell line (PC3), a nonepithelial leukemia cell line (HL60), and melanoma cell line (SKMEL5). However, due to the limited number of instances available for ssMCF7 and SKMEL5, our subsequent analysis focuses on MCF7, PC3, and HL60 cell lines.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

For a given drug at a specific concentration, we obtain the edges within the pathways influenced by the drug. We then use the PCC method to identify the edges modulated by the drug (Figure 1A). An edge with a significantly higher positive or negative impact score indicates a stronger activation or inhibition (up or down-regulation) of that edge by the drug. These edges are subsequently employed in the drug identification process. Pearson correlation coefficient (PCC) between two genes based on gene expression profiles. As we know, such a correlation can be calculated for two genes with a group of samples as:

$$PCC(x_i, x_j) = \frac{C(x_i, x_j)}{\sqrt{V(x_i)V(x_j)}} \quad \text{(Equation 1)}$$

Where, x_i and x_j represent two genes' expression profiles, their covariance for a group of samples is $C(x_i,x_j) = E((x_i - \mu_i)(x_j - \mu_j))$; and their expression variance is $V(x_i) = E((x_i - \mu_i))$. Here, $E(x)$ is the operation of expectation for variable x over a group of samples.

To obtain the disease-related edges from KEGG pathways, we utilize the PCC method to calculate the changes of edges (gene interactions) between normal and tumor samples.

$$r = \begin{cases} PCC^1(A,B) - PCC^2(A,B) \geq \alpha \\ PCC^1(A,B) - PCC^2(A,B) \leq -\alpha \end{cases} \quad \text{(Equation 2)}$$

Where PCC^1 represents the Pearson correlation coefficient between gene A and gene B in tumor samples. PCC^2 represents the Pearson correlation coefficient between gene A and gene B in normal samples. r represents the change of PCC from normal to tumor state between gene A and gene B. And α represents the threshold, which screen important and critical edges from the pathway.

To acquire the drug-induced edges. We use the PCC method to calculate the changes of gene interactions from untreated to treated state.

$$r' = \begin{cases} PCC^3(A,B) - PCC^4(A,B) \geq \alpha \\ PCC^3(A,B) - PCC^4(A,B) \leq -\alpha \end{cases} \quad \text{(Equation 3)}$$

Where PCC^3 represents the Pearson correlation coefficient between gene A and gene B in drug-treated samples. PCC^4 represents the Pearson correlation coefficient between gene A and gene B in untreated samples. r' represents the change of the PCC from untreated to treated state between gene A and gene B.

Calculating inhibition score between drug-induced edges with disease-related edges

To verify the potential effectiveness of a drug in treating a disease, we introduce an inhibition score (IS) for identifying drug-disease interactions. This score serves as a measure of the drug's therapeutic effect at the edge level. Specifically, for a drug j, the edges are screened by using PCC method (Figure 1A). For a given disease, we utilize the same approach to calculate the PCC in the context of disease gene expression profiles. Then, we respectively map the up- or down-regulated edges induced by the disease to the edges list induced by the drug to calculate the *IS*. The inhibition score (IS) is defined as:

$$IS = \sum_{i=1}^{n} -sign(r_i) * sign(r_i') \quad \text{(Equation 4)}$$

Where n represents the number of disease-related edges based on all pathways. We use the PCC to calculate the status (up- or down-regulated) of the drug-induced edges. $sign(r_i)$ represents the status (up- or down-regulated) of edge i by disease, $sign(r_i')$ represents the status (up- or down-regulated) of edge i by drug.

Identifying candidate anticancer drugs

We measure the inhibition score on each drug-disease pair and derive an IS (Figure 1B). A negative score indicates that the expression pattern of drug is similar to that of the disease, suggesting a potential antagonistic effect of the drug on the disease. Conversely, a positive score indicates that the expression pattern of the drug contrasts with that of the disease, hinting at its potential as a treatment option for the disease. In our study, each drug is utilized in various instances involving different concentrations, cancer cell lines, or durations. Consequently, we compute the IS for each drug. Subsequently, we construct an ascending ranked list of instances based on their respective IS values (Figure 1B).

Performance evaluation

Drug-disease interactions are acquired from the Comparative Toxicogenomics Database (CTD,[57]), which offers meticulously curated data covering cross-species chemical-gene/protein interactions and gene-disease associations. In our study, drug-disease associations documented in the CTD are regarded as positive samples, whereas associations not found in CTD are treated as negative samples.

To ensure a robust performance assessment, we establish CMAP, NP, SubtypeDrug, and DRviaSPCN as benchmarks. The effectiveness of these methods is evaluated using five metrics, including AUPR, AUROC, Recall, ACC, and F1.

$$Recall = \frac{TP}{TP+FN} \quad \text{(Equation 5)}$$

$$ACC = \frac{TP+TN}{TP+TN+FP+FN} \quad \text{(Equation 6)}$$

$$F1 = \frac{2TP}{2TP+FP+FN} \quad \text{(Equation 7)}$$

Where TP is the count of true positive associations, FP is false positive associations, TN is true negative associations, and FN is false negative associations. AUROC is the area under the receiver operating characteristic curve, a graph of the true positive rate against the false positive rate. AUPR is the area under the precision-recall curve, which is constructed by plotting precision against recall.

Baseline methods

To evaluate the performance of iEdgePathDDA, we compare it to other four methods, which contain signature-based method (CMAP) and pathway-based methods (NP, SubtypeDrug and DRviaSPCN). The detailed approaches are as follows:

CMAP

Detect drugs capable of reversing disease-associated gene expression signatures through the utilization of the Kolmogorov-Smirnov (KS) statistical method.[63] This method involves assessing the statistical distribution differences between drug-induced genes and disease-related genes.[19]

Network proximity (NP)

Cheng et al.[64] introduced an approach for identifying potential drugs by computing the average shortest distance between genes perturbed by the drug and those associated with the disease within the PPI network.

SubtypeDrug

Han et al.[27] provided a framework (SubtypeDrug) to discover the subtype-specific drugs at the subpathway level.

DRviaSPCN

Wu et al.[28] developed an R-based software package (DRviaSPCN) to prioritize candidate drugs for cancer via a subpathway crosstalk network.

QUANTIFICATION AND STATISTICAL ANALYSIS

Drug-disease associations documented in the CTD are regarded as positive samples, whereas associations not found in CTD are treated as negative samples. We evaluate model performance by using five metrics, including AUPR, AUROC, Recall, ACC, and F1.