



ACIBADEM MEHMET ALI AYDINLAR UNIVERSITY
INSTITUTE OF HEALTH SCIENCES

**DEVELOPING AN ANALYSIS PIPELINE BY INTEGRATING
MULTIOMICS DATA TO STUDY DRUG RESISTANCE
MECHANISMS IN CANCER**

SEMİH ALPSOY
PH.D. THESIS

DEPARTMENT OF BIOSTATISTICS AND BIOINFORMATICS

SUPERVISOR
Prof. Emel Timuçin

ISTANBUL-2025



ACIBADEM MEHMET ALI AYDINLAR UNIVERSITY
INSTITUTE OF HEALTH SCIENCES

**DEVELOPING AN ANALYSIS PIPELINE BY INTEGRATING
MULTIOMICS DATA TO STUDY DRUG RESISTANCE
MECHANISMS IN CANCER**

SEMİH ALPSOY
PH.D. THESIS

DEPARTMENT OF BIOSTATISTICS AND BIOINFORMATICS

SUPERVISOR
Prof. Emel Timuçin

ISTANBUL-2025

DECLARATION

I declare that this thesis work is my own work, I had no unethical behavior at any stages from the planning to the writing of the thesis, I obtained all the information in this thesis in accordance with academic and ethical rules, I cited all the information and comments that were not obtained with this thesis work, and I provided resources in the list of references. I also declare that there was no violation of any patents and copyrights during the study and writing of this thesis.

24/07/2025

Semih Alpsoy

PREFACE AND ACKNOWLEDGEMENT

This thesis would not have been accomplished without the support, guidance, and encouragement of numerous individuals, to whom I am deeply indebted.

First and foremost, I am profoundly grateful to my supervisor, Prof. Emel Timuçin, for her exceptional mentorship, insightful feedback, and constant belief in my abilities. Her expertise and thoughtful guidance have been a cornerstone throughout my academic journey.

I would also like to extend my genuine thanks to Prof. Osman Uğur Sezerman for his significant contributions to this thesis. His visionary approach, profound knowledge in the field, and guidance were invaluable. Additionally, his kindness, patience, and motivational attitude made the challenging moments of this journey much easier to navigate.

My sincere thanks go to the jury members, Prof. Şerife Esra Erdal Bağrıyanık, Assist. Prof. Aybar Can Acar, and Assist. Prof. Nazlı Keskin, for their constructive feedback and valuable insights during the evaluation of my thesis. Their contributions have greatly enhanced the quality of this thesis.

I am profoundly thankful to Prof. Orkide Coşkuner Weber, Prof. Aysu Yarman, and Assist. Prof. Dilek Göksel Duru, with whom I have the privilege of working at the Turkish-German University. Their belief in my potential, ongoing encouragement, and consistent support play an instrumental role in the successful completion of this thesis.

I am deeply indebted to my family, my mother Pembegül Alpsoy, my father Burhan Alpsoy, and my brother Eray Alpsoy, for their unconditional love, enduring encouragement, and unending patience. Their support has been my foundation, giving me the strength and resilience to achieve this milestone.

A special acknowledgment goes to my partner, Latife Yıldırım, for her constant patience, unyielding encouragement, and persistent support. Her belief in me has been a continuous source of inspiration, helping me overcome challenges and remain motivated throughout this journey.

Finally, I would like to express my sincere appreciation to TÜBİTAK (The Scientific and Technological Research Council of Türkiye) for providing Truba server support, which was indispensable for the completion of this thesis. Its technical support was a critical component in ensuring the success of my research.

TABLE OF CONTENTS

DECLARATION	iii
PREFACE AND ACKNOWLEDGEMENT.....	iv
ABBREVIATIONS AND SYMBOLS.....	viii
LIST OF FIGURES	xi
LIST OF TABLES.....	xv
ABSTRACT.....	1
ÖZET	2
1. INTRODUCTION AND AIM.....	3
2. BACKGROUND.....	6
2.1. Cancer	6
2.1.1. Treatment of cancer	7
2.1.2. Drug resistance mechanisms in cancer	12
2.2. DNNs.....	21
2.3. TL Approach.....	23
3. MATERIALS AND METHODS	24
3.1. Datasets	26
3.1.1. Pan-drug dataset with mitotic inhibitors.....	26
3.1.2. Pan-drug dataset with cytoskeletal inhibitors.....	27
3.1.3. Pan-drug dataset with DNA replication inhibitors	27
3.1.4. Pan-drug dataset with EGFR signaling inhibitors	31
3.2. Preprocessing of Multi-Omics Datasets	31
3.2.1. Preprocessing steps with gene expression data	31
3.2.2. Preprocessing steps with mutation data	33
3.2.3. Preprocessing steps with CNA data.....	35
3.3. Finding DEGs	36
3.4. Building DNNs with TL Approach for Drug Response Prediction.....	36
3.5. Explaining Drug Resistance Mechanisms with Pathway Enrichment	39
3.6. Association of Mutations and CNAs with Drug Resistance Using FET.....	40
3.7. Assessing Predictive and Prognostic Significance of Genes with TCGA ...	40
4. RESULTS	43
4.1. Predictive Performance of Pan-Drug Multi-Omics Models.....	43
4.1.1. Performance of pan-drug models tested individually in PDX or TCGA.....	43
4.1.2. Comparison of predictive power of pan-drug models with MOLI	56
4.1.3. Predictive performance of pan-drug models in the PDX and TCGA	57
4.2. Drug Resistance Mechanisms with Pathway Enrichment Analysis	64
4.2.1. Resistance mechanisms of paclitaxel.....	64

4.2.2. Resistance mechanisms of 5-FU.....	74
4.2.3. Resistant mechanisms of gemcitabine.....	78
4.2.4. Resistance mechanisms of cetuximab	80
4.3. Association of Mutation and CNA Statuses with Drug Resistance.....	84
5. DISCUSSION.....	94
6. CONCLUSION	103
7. REFERENCES	105
APPENDICES	120
APPENDIX 1: Common enriched terms for paclitaxel.....	120
APPENDIX 2: Pathway enrichment results.....	124
APPENDIX 3: Associations between CNA status and drug resistance.....	240
8. CURRICULUM VITAE.....	265

ABBREVIATIONS AND SYMBOLS

5-FU	5-fluorouracil
ABC	ATP-binding cassette
AUCPR	Area under the precision-recall curve
BRCA	Breast adenocarcinoma
CAR-T	Chimeric antigen receptor T-cell
CDC	The Centers for Disease Control and Prevention
C-index	Concordance index
CI	Confidence interval
CNA	Copy number alteration
CoxPH	Cox proportional hazards
CRC	Colorectal cancer
CSC	Cancer stem-like cell
DNN	Deep neural network
DEG	Differentially expressed gene
DTP	Drug-tolerant persister
EC	Expression-CNA integrated model
ECM	Extracellular matrix
EGFR	Epidermal growth factor receptor
EM	Expression-mutation integrated model
EMC	Expression-mutation-CNA integrated model
EMT	Epithelial-to-mesenchymal transition
ESCA	Esophageal carcinoma
FDA	Food and Drug Administration
FET	Fisher's exact test
FPKM	Fragments per kilobase of transcript per million mapped reads
GDSC	Genomics of Drug Sensitivity in Cancer
HNSC	Head and neck squamous cell carcinoma
HR	Hazard ratio

KEGG	Kyoto Encyclopedia of Genes and Genomes
KM	Kaplan-Meier
LGG	Lower grade glioma
LIHC	Liver hepatocellular carcinoma
logFC	Log fold-change
LRT	Log-rank test
LUSC	Lung squamous cell carcinoma
MAF	Mutation annotation file
MDR	Multidrug resistance
MI	Multi-omics integration
NMF	Non-negative Matrix Factorization
NSCLC	Non-small cell lung cancer
OR	Odds ratio
PAAD	Pancreatic adenocarcinoma
PCPG	Pheochromocytoma and paraganglioma
PDAC	Pancreatic ductal adenocarcinoma
PDX	Patient-derived xenograft
PPI	Protein-protein interaction
READ	Rectum adenocarcinoma
ReLU	Rectified linear unit
RMA	Robust Multiarray Average
ROS	Reactive oxygen species
RSEM	RNA-Seq expectation maximization
SARC	Sarcoma
SLC	Solute carrier
STAD	Stomach adenocarcinoma
TCGA	The Cancer Genome Atlas
TF	Transcription factor
TKI	Tyrosine kinase inhibitor
TL	Transfer learning

TME	Tumor microenvironment
TNBC	Triple-negative breast cancer
TPM	Transcripts per million

LIST OF FIGURES

Figure 1. The workflow for predicting drug responses in the PDX and TCGA using a DNN- and MI-based TL approach	25
Figure 2. Overview of the GDSC, PDX, and TCGA datasets for paclitaxel response prediction with mitotic inhibitors. (A) Number of cell lines screened with mitotic inhibitors in the GDSC, (B) List of mitotic inhibitors in the GDSC, (C) Number of cell lines screened with paclitaxel in the GDSC, (D) Number of mouse samples treated with paclitaxel in the PDX, and (E) Number of human patients treated with paclitaxel in the TCGA	28
Figure 3. Overview of the GDSC dataset for paclitaxel response prediction with cytoskeletal inhibitors. (A) The number of cell lines screened against cytoskeletal inhibitors in the GDSC, and (B) the list of cytoskeletal inhibitors included in the GDSC	29
Figure 4. Overview of datasets for response prediction with DNA replication inhibitors: (A) the number of cell lines screened against DNA replication inhibitors in the GDSC, (B) the list of DNA replication inhibitors in the GDSC, (C-D) the number of cell lines screened against 5-FU and gemcitabine in the GDSC, (E-F) mouse samples treated with 5-FU and gemcitabine in the PDX, and (G-H) human patients treated with 5-FU and gemcitabine in the TCGA.	30
Figure 5. Overview of datasets for cetuximab response prediction with EGFR signaling inhibitors: (A) the number of cell lines screened against EGFR signaling inhibitors in the GDSC, (B) the list of EGFR signaling inhibitors in the GDSC, (C) the number of cell lines screened against cetuximab in the GDSC, (D) mouse samples treated with cetuximab in the PDX, and (E) human patients treated with cetuximab in the TCGA	32
Figure 6. AUCPR predictions with mitotic inhibitors for paclitaxel response prediction using the PDX dataset.....	45
Figure 7. AUCPR predictions with cytoskeletal inhibitors for paclitaxel response prediction using the PDX dataset.....	45
Figure 8. AUCPR predictions with EGFR signaling inhibitors for erlotinib response prediction using the PDX dataset.....	46
Figure 9. AUCPR predictions with EGFR signaling inhibitors for cetuximab response prediction using the PDX dataset.....	47
Figure 10. AUCPR predictions with DNA replication inhibitors for 5-FU response prediction using the PDX dataset.....	48

Figure 11. AUCPR predictions with DNA replication inhibitors for gemcitabine response prediction using the PDX dataset.....	48
Figure 12. AUCPR predictions with mitotic inhibitors for paclitaxel response prediction using the TCGA dataset.....	49
Figure 13. AUCPR predictions with cytoskeletal inhibitors for paclitaxel response prediction using the TCGA dataset.....	50
Figure 14. AUCPR predictions with mitotic inhibitors for docetaxel response prediction using the TCGA dataset.....	50
Figure 15. AUCPR predictions with cytoskeletal inhibitors for docetaxel response prediction using the TCGA dataset.....	51
Figure 16. AUCPR predictions with EGFR signaling inhibitors for cetuximab response prediction using the TCGA dataset.....	52
Figure 17. AUCPR predictions with DNA replication inhibitors for cisplatin response prediction using the TCGA dataset.....	53
Figure 18. AUCPR predictions with DNA replication inhibitors for doxorubicin response prediction using the TCGA dataset.....	53
Figure 19. AUCPR predictions with DNA replication inhibitors for 5-FU response prediction using the TCGA dataset.....	54
Figure 20. AUCPR predictions with DNA replication inhibitors for gemcitabine response prediction using the TCGA dataset.....	55
Figure 21. AUCPR predictions with DNA replication inhibitors for temozolomide response prediction using the TCGA dataset.....	55
Figure 22. Fold enrichment plot showing the top 25 KEGG pathways for paclitaxel based on the gene features of the best performing mitotic inhibitor model	66
Figure 23. Fold enrichment plot showing the top 25 KEGG pathways for paclitaxel based on the gene features of the best performing cytoskeletal inhibitor model.....	67
Figure 24. Term-gene heatmap plot for paclitaxel based on the gene features of the best performing mitotic inhibitor model.....	68
Figure 25. Term-gene heatmap plot for paclitaxel based on the gene features of the best performing cytoskeletal inhibitor model	69
Figure 26. Term-gene graph illustrating the interactions between enriched KEGG terms and their associated genes for paclitaxel based on enrichment results from the best performing mitotic inhibitor model.....	70

Figure 27. Term-gene graph illustrating the interactions between enriched KEGG terms and their associated genes for paclitaxel based on enrichment results from the best performing cytoskeletal inhibitor model	71
Figure 28. Fold enrichment plot showing the top 25 KEGG pathways for 5-FU based on the gene features of the best performing DNA replication inhibitor model	76
Figure 29. Term-gene graph illustrating the interactions between enriched KEGG terms and their associated genes for 5-FU based on enrichment results from the best performing DNA replication inhibitor model	79
Figure 30. Fold enrichment plot showing the top 25 KEGG pathways for gemcitabine based on the gene features of the best performing DNA replication inhibitor model.....	80
Figure 31. Term-gene heatmap plot for gemcitabine based on the gene features of the best performing DNA replication inhibitor model.....	81
Figure 32. Term-gene graph illustrating the interactions between enriched KEGG terms and their associated genes for gemcitabine based on enrichment results from the best performing DNA replication inhibitor model.....	82
Figure 33. Fold enrichment plot showing the top 25 KEGG pathways for cetuximab based on the gene features of the best performing EGFR signaling inhibitor model.....	83
Figure 34. Term-gene graph illustrating the interactions between enriched KEGG terms and their associated genes for cetuximab based on enrichment results from the best performing EGFR signaling inhibitor model	84
Figure 35. C-indices from penalized univariate and multivariate CoxPH regression models for paclitaxel-unique genes (FYN, FLNB, and MYL9) within the most enriched focal adhesion pathway.	86
Figure 36. HRs and C-indices from penalized univariate CoxPH regression models for LDHB and TCF7L2, identified as significant in the pathway enrichment analysis for paclitaxel resistance, and from penalized bivariate CoxPH regression model when evaluated together	87
Figure 37. KM survival curves for paclitaxel-resistance-associated genes (TCF7L2, LDHB, MMP7, and FLNB) identified in the pathway enrichment analysis of paclitaxel.....	88
Figure 38. C-indices from penalized univariate and multivariate CoxPH regression models for 5-FU-unique genes (LYN, MYO1F, TNFSF13B, PINK1, and PREX1).....	89
Figure 39. C-indices from penalized univariate CoxPH regression models for LYN and TNFSF13B, identified as significant in the pathway enrichment analysis for 5-FU resistance.....	90

Figure 40. KM survival curves for BAMBI and CDH1, identified as resistance-associated in the pathway enrichment analysis for 5-FU, stratified by mutation status.	91
Figure 41. C-indices from penalized univariate and multivariate CoxPH regression models for gemcitabine, including genes from the top enriched pathways (GPX8, BMP4, MGST2, and CDH1)	91
Figure 42. KM survival curves for gemcitabine-resistance-associated genes (NQO1, CLIC3, TBC1D10C, and LSR).....	92

LIST OF TABLES

Table 1. Comparison of predictive performances with the MOLI results, including Elastic Net, Early Integration with NMF, Early Integration with DNNs, MOLI complete, MOLI pan-drug, and random classifier.....	56
Table 2. Predictive performance of paclitaxel with mitotic inhibitors across the four different strategies. The bolded values indicate the highest prediction results in the PDX and TCGA datasets.	58
Table 3. Predictive performance of paclitaxel with cytoskeletal inhibitors across the four different strategies. The bolded values indicate the highest prediction results in the PDX and TCGA datasets.....	59
Table 4. Predictive performance of 5-FU with DNA replication inhibitors across the four different strategies. The bolded values indicate the highest prediction results in the PDX and TCGA datasets.....	61
Table 5. Predictive performance of gemcitabine with DNA replication inhibitors across the four different strategies. The bolded values indicate the highest prediction results in the PDX and TCGA datasets.....	62
Table 6. Predictive performance of cetuximab with EGFR signaling inhibitors across the four different strategies. The bolded values indicate the highest prediction results in the PDX and TCGA datasets	65
Table 7. Common enriched terms for paclitaxel identified through enrichment analysis using gene features from pan-drug models trained with mitotic and cytoskeletal inhibitors	120
Table 8. pathfindR pathway enrichment analysis for paclitaxel utilizing the gene features of the pan-drug model trained with mitotic inhibitors (Strategy 1)	124
Table 9. pathfindR pathway enrichment analysis for paclitaxel utilizing the gene features of the pan-drug model trained with mitotic inhibitors (Strategy 2)	131
Table 10. pathfindR pathway enrichment analysis for paclitaxel utilizing the gene features of the pan-drug model trained with cytoskeletal inhibitors (Strategy 1)	137
Table 11. pathfindR pathway enrichment analysis for paclitaxel utilizing the gene features of the pan-drug model trained with cytoskeletal inhibitors (Strategy 2)	144
Table 12. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 1).....	151
Table 13. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 2).....	172

Table 14. pathfindR pathway enrichment analysis for gemcitabine utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 1) ...	201
Table 15. pathfindR pathway enrichment analysis for gemcitabine utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 2) ...	203
Table 16. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 1)....	205
Table 17. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 2)....	222
Table 18. Table of associations between CNA status and paclitaxel resistance	240
Table 19. Table of associations between CNA status and 5-FU resistance.....	241
Table 20. Table of associations between CNA status and gemcitabine resistance ...	260

ABSTRACT

Developing An Analysis Pipeline By Integrating Multi-Omics Data To Study Drug Resistance Mechanisms In Cancer

Various approaches such as surgery, radiotherapy, immunotherapy, and pharmacological treatment are employed in cancer therapy. Among these, pharmacological treatment is widely preferred due to its effectiveness across different cancer types. However, patients may develop resistance to drugs within a short period as a result of genetic mutations, epigenetic alterations, or disruptions in signaling pathways. In this thesis, the resistance mechanisms of paclitaxel, 5-fluorouracil (5-FU), gemcitabine, and cetuximab, which are commonly used in clinical practice, were investigated. For this purpose, a late multi-omics integration approach based on deep neural networks (DNNs) with transfer learning (TL) was applied. Drugs targeting mitosis, cytoskeleton, DNA replication, and epidermal growth factor receptor (EGFR) signaling pathway were independently trained to construct pan-drug training datasets for predicting drug responses. Models trained on human cell line data from the Genomics of Drug Sensitivity in Cancer (GDSC) dataset were able to predict drug responses in in vivo patient-derived xenograft (PDX) and ex vivo The Cancer Genome Atlas (TCGA) datasets. Pathway enrichment analyses based on the gene features of the best performing predictive models revealed the mechanisms of drug resistance. It was demonstrated that mutations and copy number alterations (CNAs), together with transcriptional changes, contribute to the development of resistance. Elucidating the resistance mechanisms of these four drugs through relevant pathways and molecular alterations has the potential to make a significant contribution to the development of personalized and targeted therapies.

Keywords: Multi-omics integration, Drug resistance mechanisms, Personalized oncology, Drug response prediction, Transfer learning

ÖZET

Kanserde İlaç Direnci Mekanizmalarını İncelemek İçin Çoklu-Omik Verileri Entegre Eden Bir Analiz Yöntemi Geliştirilmesi

Cerrahi, radyoterapi, immünoterapi ve farmakolojik tedavi dahil olmak üzere çeşitli yaklaşımlar kanser tedavisinde kullanılmaktadır. Bu yaklaşımlar arasında, farklı kanser türlerinde etkili olması nedeniyle farmakolojik tedavi yaygın olarak tercih edilmektedir. Ancak hastalar, genetik mutasyonlar, epigenetik değişiklikler ve ya sinyal yollarındaki bozulmalar nedeniyle kısa bir süre içinde ilaçlara karşı direnç geliştirebilmektedir. Bu tezde, klinik uygulamalarda sıklıkla kullanılan paklitaksel, 5-florourasil (5-FU), gemcitabin ve setuksimabın direnç mekanizmaları incelenmiştir. Bu amaçla, transfer öğrenmesi (TL) tabanlı derin sinir ağları (DNNs) kullanılarak çoklu-omik geç entegrasyon yaklaşımı uygulanmıştır. Mitozu, sitoskeletonu, DNA replikasyonunu ve epidermal büyüme faktörü reseptörü (EGFR) sinyal yolunu hedefleyen ilaçlar bağımsız olarak eğitilmiş ve ilaç yanıtlarını tahmin etmeye yönelik tüm-ilâç eğitim veri setleri oluşturulmuştur. Kanserde İlaç Duyarlılığı Genomiği (GDSC) veri setindeki insan hücre hattı verileri kullanılarak eğitilen modeller, in vivo hastadan türetilmiş ksenograft (PDX) ve ex vivo Kanser Genom Atlası (TCGA) veri setlerinde ilaç yanıt tahminleri yapabilmektedir. En iyi performans gösteren tahmin modellerinin gen özelliklerine dayalı yolak zenginleştirme analizleri ile ilaç direnç mekanizmaları ortaya konmuştur. Mutasyonlar ve kopya numarası değişikliklerinin (CNAs), transkripsiyonel değişikliklerle birlikte direnç gelişimine katkıda bulunduğu gösterilmiştir. Bu tezde dört ilacın direnç mekanizmalarının ilgili yollar ve moleküler değişiklikler üzerinden aydınlatılması, kişiselleştirilmiş ve hedefe yönelik tedavilerin geliştirilmesine önemli katkı sağlama potansiyeline sahiptir.

Anahtar Sözcükler: Çoklu-omik entegrasyonu, İlaç direnç mekanizmaları, Hedefe yönelik onkoloji, İlaç yanıt tahmini, Transfer öğrenmesi

1. INTRODUCTION AND AIM

Cancer is a complex, heterogeneous disease that might develop due to various genetic and environmental factors (1). There are various types of cancer, which are classified based on the part of the tissues the disease originate (2). Many treatment options are available although specific treatments are selected based on genetic profile, health status, and symptoms of the patients. Some protocols are also available in the treatment process. Especially in chemotherapy, the drugs having low effect might be administered to patients at first. After checking their effects on patients, they might be further applied or due to developing resistance, therapy might be modified with more effective drugs, or different treatment options might be applied depending on the health status of the patients (3).

Heterogeneous nature of the disease frequently limit treatment success in most patients. Each patient may accumulate unique changes in their tumors or might have genetic predisposition for cancer. Resistance development might be rapid for such patients, which may obstacle their treatment success. Chemotherapy is mostly used with a one-size-fits-all approach without knowing a drug will work effectively in each patient. Thereby, it is mostly a trial and error approach, increasing and diversifying resistance mechanisms in cancer (4).

Various treatment options are available for cancer therapy such as surgery, radiotherapy, chemotherapy, immunotherapy, and targeted therapy. Patients mostly get successful outcomes at early stages of the treatment, but those in advanced stage may not benefit from most of the treatment options. Physicians may apply chemotherapy with single drug therapy or combination therapy, and the treatment might be supported by other treatment options such as radiotherapy and immunotherapy. Targeted therapy might be applied only when a specific cancer type with a specific genetic background. However, supportive treatment with radiotherapy or chemotherapy might be still needed. Unfortunately, even with targeted therapies, resistance could develop over time, making treatment procedure unsuccessful (5).

It is important to grasp biology of cancer in each patient in order to improve clinical outcomes. High heterogeneity is a major problem to decide effective treatment strategy. Even though a diagnosis can be made with high accuracy in patients, a standard treatment may not give high treatment success because of distinct and different mechanisms taking effect in patients. This kind of individual differences among patients is known as inter-tumor heterogeneity. It might also be observed in tumors of same patient, showing differences located in different parts of the body. That might represent another form of inter-tumor heterogeneity (6).

Individual differences might be even more complicated with intra-tumor heterogeneity. Tumors might have distinct sub-populations within the same tumor. Each might accumulate different mutations, CNAs or epigenetic changes. This might lead to differences in therapy responses. Both inter-tumor and intra-tumor heterogeneity should be considered to benefit from treatment options. However, it is not always possible to identify heterogeneity, which makes an obstacle to apply effective therapeutic strategies. For this reason, it is urgent to develop therapeutic strategies to consider heterogeneity and improve treatment success (6).

Unfortunately, the one-size-fits-all approach used in chemotherapy fails to account for heterogeneities, which can result in either a lack of response or the development of resistance in a short period. This significantly hinders the success of the treatment process. Gaining a deeper understanding of resistance mechanisms can enable more rational drug administration and facilitate the development of new therapeutic strategies. Therefore, extensive efforts have been made to elucidate resistance mechanisms in cancer (7).

In this thesis, DNN models based on a TL approach were aimed to develop robust drug response predictions and provide insights into drug resistance mechanisms. The study focused on multi-omics integration (MI), which includes gene expression, somatic mutation, and CNA data. Rather than analyzing drugs individually, drugs targeting the same pathway were aimed to be grouped to

demonstrate that combining drugs with shared mechanisms enhances the performance of drug response predictions. The GDSC was utilized as the training dataset, while in vivo PDX and ex vivo TCGA datasets were used as external datasets (8-10). Drug response predictions were conducted for erlotinib, paclitaxel, docetaxel, cetuximab, 5-FU, gemcitabine, temozolomide, cisplatin, and doxorubicin, all of which were profiled across the datasets. Multi-omics pan-drug models were aimed to apply for exploring the potential of integrating multi-omics data for improved predictive power. Gene features of the pan-drug models, which robustly predicted drug responses, were analyzed using pathfindR, which is a pathway enrichment analysis tool that identifies active subnetworks (11). This analysis aimed to investigate resistance mechanisms of paclitaxel, 5-FU, gemcitabine, and cetuximab that are shared between the PDX and TCGA, with the dual purpose of elucidating known drug resistance mechanisms and identifying novel ones. In addition, Fisher's exact test (FET) was performed to evaluate the association of mutations and CNAs with drug resistance. Finally, gene-level and pathway-level predictive performances were assessed in the TCGA using Cox proportional hazards (CoxPH) regression models. Kaplan-Meier (KM) survival curves were also generated for mutation and CNA groups with drug resistance, and statistical significance between groups was determined using the log-rank test (LRT).

2. BACKGROUND

2.1. Cancer

Cancer is a complex, high heterogeneity disease that arises from various genetic and epigenetic changes driving uncontrolled cell growth and spreading to other parts of the body. Although it was recognized in the mid-20th century that cancer cells originate from normal cells, the underlying molecular alterations have become much better understood in recent years in parallel with advances in molecular biology. The discovery that somatic mutations in driver genes cause uncontrolled cell growth and tissue invasion has led to cancer being characterized as “a disease of the genes”. However, the notion that genetic heterogeneity in cancer cell populations arises solely from fixed somatic mutations in genes has contributed to gaps in our understanding of cancer. It is now widely recognized that cancer cell populations undergo evolutionary adaptation via natural selection with diverse mechanisms such as interaction with the tumor microenvironment (TME), metabolic reprogramming, and immune evasion (12). In particular, changes in blood flow, host immune responses, and environmental conditions generate selective pressures that drive the adaptation of cancer cells, ultimately leading to resistance or sensitivity to various therapies (13).

Cancer is responsible for millions of deaths annually and ranks alongside heart disease as one of the two leading causes of death (14). According to reports from the Centers for Disease Control and Prevention (CDC), malignant neoplasms accounted for 22,7% of all deaths in the United States in 2022, making them the second leading cause of death after heart disease (15). Moreover, according to the American Cancer Society, an estimated 2,001,140 new cancer cases and 611,720 cancer-related deaths were expected in the United States in 2024. Since 1991, cancer mortality rates have declined, falling from over 4 million deaths in 1991 to fewer than 4 million per year since 2021. This decline is attributed to reductions in smoking rates, earlier detection of certain cancers, and advances in treatment options (16).

Development of cancer is driven by several hallmarks, which include sustaining proliferative signaling, evading growth suppressors, resisting programmed cell death, enabling replicative immortality, promoting angiogenesis, facilitating invasion and metastasis, deregulating cellular energetics, and avoiding immune destruction (17). Various treatment approaches such as surgery, radiation therapy, chemotherapy, targeted therapy, immunotherapy, and hormone therapy are individualized according to the type and stage of cancer, underlying molecular and genetic alterations, and the overall health status of patients (18, 19). However, the success of treatments can be limited due to the development of several drug resistance mechanisms. Increased drug efflux, drug inactivation and detoxification, changes in drug targets, enhanced DNA repair, apoptotic evasion, immune evasion, activation of alternative signaling pathways, changes in the TME, epigenetic changes, metabolic reprogramming, and epithelial-to-mesenchymal transition (EMT) are examples of such diverse and complex resistance mechanisms (20, 21).

2.1.1. Treatment of cancer

2.1.1.1. Surgery

Surgery is an effective intervention method in some cases; however, it may not be effective in removing tumors especially in the later stages of the disease. In such cases, chemotherapy is initially administered to shrink the tumor, making it more amenable to subsequent intervention. This procedure is known as neoadjuvant chemotherapy. Nevertheless, post-operative chemotherapy may still be necessary to enhance therapeutic efficacy. In this context, adjuvant chemotherapy is applied following surgical removal. Both neoadjuvant and adjuvant chemotherapy are common procedures used before and after surgical intervention (22).

2.1.1.2. Radiotherapy

Radiotherapy, which utilizes high-energy radiation, is an effective method for cancer treatment. In many cases, it is combined with surgical intervention and pharmacological therapy; however, it can also be used as a standalone treatment in certain situations. It is particularly effective for targeting localized tumors, and exerts its therapeutic effect by inducing DNA damage in cancer cells, thereby impairing their ability to replicate or triggering apoptosis (22).

2.1.1.3. Chemotherapy

Chemotherapy is a systemic treatment method that targets and eliminates rapidly dividing and uncontrollably proliferating cells. It is commonly employed when surgical resection and radiotherapy are not feasible, and when targeted therapies or hormone-based treatments are not appropriate. The major classes of chemotherapeutic agents include alkylating agents, antimetabolites, anthracyclines, mitotic inhibitors, topoisomerase inhibitors, and platinum-based compounds (23).

Alkylating agents exert their effect by adding alkyl groups to DNA, particularly during replication, which leads to DNA damage and impaired cell division. Common alkylating agents used in chemotherapy include cyclophosphamide, dacarbazine, and temozolomide (24). However, temozolomide and dacarbazine are pro-drugs that exert their effects through DNA methylation. Technically, they act like alkylating agents and are sometimes referred to as alkylating-like agents. These agents may be administered as monotherapy or in combination with other chemotherapeutic drugs depending on the cancer type, disease stage, and patient-specific clinical factors (25, 26).

Antimetabolites exert their effects by inhibiting DNA and RNA synthesis. Common examples include 5-FU, methotrexate, gemcitabine, and capecitabine. These agents are primarily used in the treatment of specific cancer types such as

pancreatic, breast, gastric, and bladder cancers. In most cases, they are administered in combination with other chemotherapeutic agents (27).

Anthracyclines are highly effective chemotherapeutic agents that kill cancer cells by interfering with DNA replication through the disruption of DNA strands. Doxorubicin, epirubicin, and daunorubicin are well-known anthracyclines commonly used in the treatment of breast cancer, lymphoma, and sarcoma (28).

Mitotic inhibitors disrupt cell division and change microtubule dynamics. They might exert their effects on cytoskeleton; thus, it may also change cytoskeletal dynamics. Paclitaxel and docetaxel are examples of mitotic inhibitors. These two drugs are also known as cytoskeletal inhibitors. These group of inhibitors are highly effective in rapidly dividing cancer cells as in the case of breast cancer (29).

Topoisomerase inhibitors block topoisomerase enzyme to show its effect by interfering in DNA replication process. Irinotecan and topotecan are two example topoisomerase inhibitors that are widely used in the treatment (30).

Platinum based drugs are used to inhibit DNA replication process and transcription by forming cross-links. They are highly effective in diverse cancer types such as ovarian and testicular cancer. Carboplatin, cisplatin, and oxaliplatin are examples of platinum-based drugs (31).

2.1.1.4. Immunotherapy

Unlike chemotherapy, which non-selectively targets all rapidly dividing cells, immunotherapy aims to eliminate cancer cells by harnessing the body's immune system. This treatment approach enables the immune system to recognize and respond to cancer cells, potentially leading to effective therapeutic outcomes. However, its efficacy is often limited to specific cancer types and depends on certain biological conditions. Among the various immunotherapeutic strategies, immune checkpoint inhibitors are the most widely used. These agents are designed to block

specific inhibitory proteins, thereby enhancing the immune system's ability to recognize and selectively attack cancer cells (32).

Another approach is the use of cancer vaccines. These vaccines, which represent an emerging form of immunotherapy, aim to stimulate the immune system in advance by enabling the recognition of specific tumor-associated antigens. Sipuleucel-T, a Food and Drug Administration (FDA)-approved vaccine for prostate cancer, is a notable example (33).

Adoptive cell therapy is another innovative immunotherapy approach with significant therapeutic potential. In this method, T-cells are genetically modified to enhance their ability to recognize and eliminate cancer cells more effectively (34).

Another cutting-edge immunotherapy with promising potential is chimeric antigen receptor T-cell (CAR-T) therapy. This approach involves engineering T-cells to express chimeric antigen receptors, thereby enhancing the immune response and promoting a more effective antitumor activity. CAR-T therapy is primarily designed for the treatment of certain leukemias and lymphomas (35).

Cytokine therapy is an immunotherapeutic approach that aims to enhance the immune response through the administration of specific cytokines. Interferon-alpha is one of the most commonly used cytokines for this purpose. By strengthening the immune system, cytokine therapy improves the body's ability to target and eliminate cancer cells. This strategy can significantly enhance treatment efficacy in certain cancer types (36).

Monoclonal antibodies represent a treatment modality that can significantly improve treatment success in specific cancer types. By specifically targeting specific antigens in the body, the immune system is given the opportunity to send a signal. Immune cells that receive this signal specifically attack only the cancer cells from which the signal come. Trastuzumab, which has been used successfully in the

treatment of HER2-positive breast cancer, is an example of a therapy using monoclonal antibody (37).

Oncolytic virus therapy uses genetically modified viruses in a dual-action approach, not only targeting cancer cells but also boosting the immune response. Such viruses infect cancer cells specifically and do not harm normal cells. While inside the cancer cell, it replicates its genetic material by replication. After a certain period of time, the cell lyses and virus fragments spread around. This in turn recruits other immune system cells to kill the remaining cancer cells (38). Talimogene laherparepvec, which has been approved by the FDA for the treatment of melanoma, is a therapeutic approach developed using this technology (39).

2.1.1.5. Targeted therapy

Unlike chemotherapy, which targets all rapidly dividing cells, targeted therapy is designed to selectively kill cancer cells based on specific molecular targets. This high level of precision offers a significant therapeutic advantage by minimizing damage to normal cells. As a result, side effects are reduced, and more effective and durable responses can be achieved in the treatment of cancer (40).

Some clinically available targeted drug examples are imatinib, erlotinib, and gefitinib. Imatinib, a tyrosine kinase inhibitor (TKI), targets the BCR-ABL fusion protein and provides high success in the treatment of chronic myeloid leukemia. Erlotinib and gefitinib, which target EGFR, are other examples that have been used effectively in lung cancer. However, targeted therapies are not usually used alone. Other supportive treatment modalities such as various forms of chemotherapy or radiotherapy are administered during or after the targeted therapy (41).

Although the development of resistance to targeted therapies is not usually as rapid as to chemotherapy, it is possible for cancer cells to develop resistance against them over prolonged treatment. Various genomic changes such as mutations and CNAs may play an important role in the failure of targeted therapy (42).

2.1.1.6. Hormone therapy

Hormone therapy requires the development of hormone-induced cancer cells before it is administered. By blocking the hormone-dependent interaction, the cancer cells are deprived of hormones and are prone to die. Drugs developed for this purpose usually bind to receptors on specific cancer cells, and either kill them or inhibit their growth signal, thus providing effective therapy against cancer. Tamoxifen, which blocks the effects of estrogen in breast cancer, is an aromatase inhibitor in this group (43).

With hormone therapy, there may be a risk of cancer recurrence. To prevent recurrence, chemotherapy is often applied even after treatment as adjuvant therapy. Chemotherapy may also be used to shrink the tumor before hormone therapy as neoadjuvant chemotherapy. However, there is a risk of resistance developing over time. Mutations in receptors may reduce the need of cancer cells for external hormones. Alternative signaling pathways are activated to increase survival and proliferation. Diversification of treatment may be important in this respect. Drug therapies that affect multiple mechanisms can be administered as combination therapy. The application of immunotherapy with hormone therapy can also lead to positive improvements in treatment success (43).

2.1.2. Drug resistance mechanisms in cancer

Drug resistance in cancer can arise through a variety of mechanisms, and understanding these mechanisms is critical for improving therapeutic efficacy. These mechanisms include altered drug transport, drug inactivation and detoxification, changes in drug targets, enhanced DNA repair, apoptotic evasion, immune evasion, activation of alternative signaling pathways, changes in the TME, epigenetic changes, metabolic reprogramming, and EMT (21, 44). In the subheadings below, drug resistance mechanisms are briefly explained.

2.1.2.1. Altered drug transport

One of the best-characterized mechanisms of drug resistance is altered drug transport resulting from the dysregulation of influx and efflux transporters (20). Resistance may develop when efflux transporters are up-regulated or influx transporters are down-regulated, leading to reduced intracellular drug accumulation. The most commonly implicated transporters in this context are the ATP-binding cassette (ABC) and solute carrier (SLC) family members (45, 46).

The ABC transporter family, composed of membrane-bound proteins, plays a key role in the efflux of drugs from cells. Numerous studies have reported that the upregulation of these transporters is associated with the development of multidrug resistance (MDR) (46). The most well-known members of the ABC transporter family include P-glycoprotein (encoded by the ABCB1 gene), MRP1 (encoded by the ABCC1 gene), and BCRP (encoded by the ABCG2 gene). Targeting these transporters has been one of the most extensively investigated strategies for overcoming the MDR in cancer therapy (47). For example, upregulation of ABCB1 has been associated with docetaxel resistance in the triple-negative breast cancer (TNBC) cell line MDA-MB-231 (47). Similarly, increased expression of ABCB1 and ABCG2 in ovarian cancer cell lines has been linked to resistance to the topoisomerase inhibitor topotecan (48). Furthermore, upregulation of ABCB1 in multiple myeloma cells has been shown to reduce the efficacy of carfilzomib, which is a proteasome inhibitor, by promoting drug efflux (49).

SLC family comprises more than 400 members, which are classified into different subfamilies based on sequence similarity. SLC transporters perform essential physiological functions by mediating the transport of nucleotides, sugars, lipids, amino acids, and peptides (45). Several SLC transporters have also been implicated in drug resistance. For instance, SLCO1A2 has been associated with resistance to imatinib and methotrexate; SLCO1B1 with cisplatin, carboplatin, oxaliplatin, regorafenib, belzutifan, SN-38, etoposide, tamoxifen, and sorafenib;

SLC22A1 with dasatinib and nintedanib; and SLC22A5 with etoposide and imatinib in various cancer types (50).

2.1.2.2. Drug inactivation and detoxification

Alterations in drug metabolism are among the key factors contributing to the development of cancer drug resistance. Similar to changes in drug transport, they influence the intracellular and extracellular concentrations of therapeutic agents. However, drug inactivation and detoxification represent distinct mechanisms. Drugs may undergo chemical modifications such as oxidation, reduction, and hydrolysis, which can result in reduced activity and contribute to cellular protection against cytotoxic effects. Low enzymatic activity may impair the activation of pro-drugs, while elevated activity can promote drug inactivation. A notable example is the development of resistance in breast cancer due to enhanced cytochrome P450 activity, which has also been specifically linked to docetaxel resistance. Conversely, reduced cytochrome P450 activity has been associated with improved treatment responses. Docetaxel is metabolized into conjugated forms such as glucuronic acid, sulfate, and glutathione, leading to a decrease in its cytotoxic activity. Studies have shown that increased glutathione levels promote docetaxel detoxification, primarily through the action of glutathione S-transferase enzymes. This detoxification mechanism is not limited to docetaxel; it also plays a significant role in resistance to various alkylating agents and platinum-based chemotherapeutics (20).

2.1.2.3. Changes in drug targets

Targeted therapies can be highly effective in certain types of cancer (51). However, factors such as point mutations, gene amplification, alternative splicing, and post-translational modifications at the drug target site can induce resistance by reducing binding affinity or disrupting molecular interactions (52, 53). As a result of these alterations, targeted drugs that were initially effective may lose their efficacy, necessitating the use of alternative therapeutic strategies (54). For example, it is well

established that the efficacy of TKIs in non-small cell lung cancer (NSCLC) is reduced due to mutations in EGFR (54). Osimertinib, which is a third-generation TKI, is specifically used to target certain EGFR mutations. However, resistance to osimertinib has been associated with the emergence of secondary mutations such as C797S and alterations in MET in some NSCLC patients (55). Another example is the association between alterations in topoisomerase II and drug resistance. A synthetic phenanthridine alkaloid (NK314) has been identified as an effective compound with this purpose. Gene deletions affecting Top2 α cleavage complexes (Top2accs) have been shown to confer resistance to NK314 in in vivo experiments (56). Furthermore, mutations in estrogen receptors in estrogen receptor-positive breast cancer have been shown to diminish the effectiveness of hormone therapy. Y537S and Y537C mutations are examples that cause resistance independently of the ligand in breast cancer cell lines (57).

2.1.2.4. Enhanced DNA repair

Various chemotherapeutic agents exert their cytotoxic effects by inducing different types of DNA damage, which contributes to their therapeutic efficacy but may also lead to the development of resistance (58). Hyperactivation of DNA repair pathways such as nucleotide excision repair, base excision repair, mismatch repair, homologous recombination, and non-homologous end joining can enable cancer cells to efficiently repair drug-induced DNA lesions (59-62). This enhanced repair capacity diminishes drug efficacy and promotes therapeutic resistance. Upregulation of key genes within these pathways has been associated with resistance to multiple drugs (63). For example, deletions in the phosphatase regulator NIPP1 have been shown to give strong resistance to chemically induced hepatocellular and skin carcinoma in mice by increasing expression of rate-limiting DNA repair proteins such as MGMT in the liver and XPD in the skin (64). In addition, upregulation of the nucleotide excision repair gene ERCC1 has been associated with cisplatin resistance in human cervical cancer cells (65). Furthermore, temozolomide resistance has been reported in mismatch repair-deficient human glioblastoma cells due to upregulation

of PARP gene by increasing DNA repair capacity. However, increased sensitivity to temozolomide has been observed when PARP inhibitors are applied, which reverses resistance caused by enhanced DNA repair mechanisms (66).

2.1.2.5. Apoptotic evasion

Apoptosis is a tightly regulated, energy-dependent process characterized by cell shrinkage, chromatin condensation, DNA fragmentation, membrane blebbing, and the formation of apoptotic bodies (67). The evasion of this programmed cell death mechanism is a fundamental strategy by which cancer cells resist the effects of therapeutic agents (68). In the apoptotic signaling pathway, upregulation of anti-apoptotic genes such as BCL-2, BCL-XL, and MCL-1, or downregulation of pro-apoptotic genes such as BAX and BAK can lead to dysfunction of apoptotic processes. As a result, cancer cells may acquire a survival advantage and evade the cytotoxic effects of therapeutic interventions (69).

Apoptosis can contribute to drug resistance through two main mechanisms: the extrinsic and intrinsic pathways. In the extrinsic pathway, ligands such as TRAIL, TWEAK, and NGF bind to pro-apoptotic death receptors, initiating downstream signaling through adaptor proteins and caspases. In the intrinsic pathway, members of the BCL-2 family such as Bmf, Bik, and Bad inhibit anti-apoptotic proteins like BCL-2, promoting mitochondrial outer membrane permeabilization. BCL-2, which can suppress p53 activity, is a well-known anti-apoptotic protein. Anti-apoptotic mediators operate in both pathways. For example, BCL-XL inhibits pro-apoptotic factors such as BAX, BAK, granzyme B, and p53 in the intrinsic pathway. In the extrinsic pathway, c-FLIP and NF- κ B are key anti-apoptotic regulators. c-FLIP inhibits the activation of caspases by preventing their recruitment to death-inducing signaling complexes, while NF- κ B promotes resistance by upregulating anti-apoptotic genes including c-FLIP. Additionally, alternative signaling pathways such as JAK/STAT, MAPK, and the expression of inhibitors of apoptosis proteins such as

survivin and XIAP also play important roles in suppressing apoptosis and promoting drug resistance (69).

2.1.2.6. Immune evasion

The ability of cancer cells to evade immune surveillance and develop resistance to immunotherapies is a well-documented phenomenon. Tumor cells employ multiple strategies to escape detection and elimination by the immune system. These include downregulation of tumor-associated antigens, loss of major histocompatibility complex class I expression, secretion of immunosuppressive cytokines, and recruitment of regulatory T-cells and myeloid-derived suppressor cells. These mechanisms collectively contribute to the suppression of cytotoxic T lymphocyte and natural killer cell activation, resulting in an immunosuppressive TME that impairs effective immune responses. Moreover, the upregulation of immune checkpoint molecules such as PD-L1 on tumor cells can inhibit T-cell function and has been implicated in resistance to immune checkpoint inhibitors. Therefore, overcoming immune evasion remains a major challenge, and continued research is essential to enhance the efficacy of immunotherapies (70).

2.1.2.7. Activation of alternative signaling pathways

The activation of alternative signaling pathways is a well-recognized mechanism underlying resistance to targeted therapies. In response to the inhibition of primary oncogenic drivers, cancer cells can engage compensatory signaling networks to sustain survival, proliferation, invasion, and metastatic capabilities. This adaptive mechanism is frequently observed across various cancer types and therapeutic contexts, emphasizing the necessity of developing effective combination strategies that simultaneously target both primary and alternative pathways. Addressing this mode of resistance is crucial for enhancing the long-term efficacy of anti-cancer treatments (71, 72).

There are well-documented examples in the literature of resistance mechanisms involving the activation of alternative signaling pathways. One such example is resistance to EGFR-targeted inhibitors such as erlotinib and gefitinib in the NSCLC, where MET amplification maintains downstream signaling via the PI3K/Akt pathway independently of EGFR (73, 74). Similarly, HER2 amplification can sustain MAPK/PI3K pathway activity despite EGFR inhibition (75). In melanoma, resistance to BRAF inhibitors such as vemurafenib has been linked to activation of MEK and/or PI3K/Akt signaling. This may occur through MEK upregulation, which lies downstream of BRAF, or through loss of PTEN, leading to PI3K/Akt pathway activation (76). Another example involves resistance to ALK inhibitors, where compensatory activation of EGFR and/or IGF-1R signaling reactivates PI3K/Akt and MAPK signaling pathways, promoting tumor survival and drug resistance (77).

2.1.2.8. Changes in the TME

Understanding the alterations within the TME is essential for elucidating the mechanisms underlying cancer drug resistance. Disruptions in the dynamic and heterogeneous interactions between cancer cells and various TME components such as fibroblasts, endothelial cells, immune cells, the extracellular matrix (ECM), signaling molecules, and cytokines have been strongly associated with resistance to therapy (78). Several TME-related changes can contribute to drug resistance through distinct ways. First, cancer-associated fibroblasts secrete growth factors, cytokines, and ECM components that impair drug penetration, promote tumor cell survival, and enhance immune evasion (79). Second, the formation of a hypoxic environment within the TME activates hypoxia-inducible factors, which in turn upregulate genes involved in angiogenesis, glycolysis, autophagy, and drug efflux, which can reduce drug efficacy (80). Third, the recruitment of immunosuppressive cells such as tumor-associated macrophages myeloid-derived suppressor cells, and regulatory T-cells creates an immunosuppressive environment, which can diminish the effectiveness of immunotherapies (81). Fourth, ECM remodeling and increased tissue stiffness can further restrict drug diffusion. In such conditions, activation of integrin-mediated

signaling has been shown to promote resistance to apoptosis and contribute to therapeutic failure (82). These TME-mediated alterations create a protective niche for tumor cells, representing a major challenge to effective cancer drug treatment. Therefore, targeting TME components is critical for overcoming drug resistance and improving therapeutic outcomes.

2.1.2.9. Epigenetic changes

Epigenetic changes such as DNA methylation, histone modifications, chromatin-remodeling complexes, and non-coding RNAs regulate gene expression through reversible mechanisms without altering the DNA sequence (83). These modifications play a critical role in the development of drug resistance in cancer (84). Epigenetic alterations can lead to the downregulation of tumor suppressor genes and the upregulation of oncogenes, thereby disrupting apoptotic pathways and enhancing the survival potential of cancer cells (85). For instance, hypermethylation of promoter CpG islands in tumor suppressor genes such as CDKN2A disrupt apoptosis and contribute to drug resistance (86). Conversely, hypomethylation of enhancer regions can up-regulate ABC transporter genes, which are responsible for promoting drug efflux and reducing intracellular drug accumulation (87). Histone deacetylation, which is mediated by histone deacetylases, can repress pro-apoptotic genes such as BAX and PUMA, thereby lead to drug resistance (88). Enhanced activity of histone methyltransferases like EZH2 has been shown to silence immune-related genes, promoting immune evasion and resistance to immune checkpoint inhibitors (89). In addition, non-coding RNAs can contribute to drug resistance via epigenetic regulation. For example, upregulation of miR-21 can down-regulate PTEN, leading to activation of the PI3K/Akt signaling pathway. These changes have been linked to drug resistance in cancer (90). Similarly, HOTAIR can interact with the PRC2 complex to suppress HOXD gene expression, which has been reported to contribute to cancer drug resistance (91).

2.1.2.10. Metabolic reprogramming

Cancer cells can gain a survival advantage by reprogramming their metabolic activities. By efficiently managing energy production, they adapt to extreme environmental conditions, which can lead to the development of drug resistance. Even in the presence of oxygen, they may preferentially utilize glycolysis over oxidative phosphorylation through a process known as the Warburg effect (92). Under hypoxic conditions with limited oxygen availability, they can use alternative metabolites such as lactate as an energy source (93). Cancer cells also heavily rely on glutamine metabolism to support their metabolic demands. By upregulating glutaminase expression, they enhance the stabilization of reactive oxygen species (ROS), thereby improving their capacity to neutralize ROS generated by chemotherapeutic agents (94). Fatty acid metabolism may also play an important role in drug resistance as it can be exploited by cancer cells. To evade oxidative stress, cancer cells may rely on fatty acids as an alternative energy source (95). Moreover, they can alter their mitochondrial metabolism to enhance their ability to withstand therapeutic stress by dynamically modulating oxidative phosphorylation and glycolysis according to changing metabolic demands (96).

2.1.2.11. EMT

EMT is a dynamic biological process in which cancer cells transition from an epithelial phenotype to a mesenchymal phenotype and is recognized as a key mechanism in the development of drug resistance. During EMT, epithelial cells lose their cell-cell adhesion properties and acquire enhanced motility and invasive capabilities. This transition is marked by the downregulation of epithelial markers such as E-cadherin and the upregulation of mesenchymal markers including N-cadherin, vimentin, and transcription factors (TFs) such as ZEB1, SNAIL (Snail), SNAIL2 (Slug), and TWIST1 (97). During EMT, cancer cells gain a survival advantage by promoting anti-apoptotic signaling pathways, increasing drug efflux pumps, and decreasing proliferation, which allow them to evade the cytotoxic effects

of therapeutic agents (98). Many studies have demonstrated a strong association between EMT and the development of drug resistance in various cancer types (99). For example, it has been reported that the upregulation of SNAI1 and SNAI2 induces resistance to gefitinib in EGFR-mutant lung cancer cell lines (100). Similarly, PAX6 has been shown to bind and regulate the promoter region of ZEB2, another well-known EMT TF, thereby promoting cisplatin resistance in the NSCLC via the PI3K/Akt signaling pathway during the EMT process (101). Moreover, AXL, a receptor tyrosine kinase associated with EMT, is up-regulated and confers resistance to osimertinib in EGFR-mutant NSCLC cells (102). Finally, the upregulation of SNAI1 has been observed in basal-like breast cancer MDA-MB-231 cells and is associated with resistance to gemcitabine and doxorubicin (103).

2.2. DNNs

DNNs are a type of artificial neural networks with several layers. In its simplest form, there must be at least one input layer, one output layer and at least two hidden layers between them (104). The more hidden layers there are, the more complex a neural network can be, resulting in increased computational power. However, it provides an advantage in handling more complex problems. With fewer hidden layers, the ability of model handling complex problems may be limited although the computational need decreases and rapid modeling is achievable (105). Adding nonlinearity is necessary to understand complex relationships (106). For this purpose, various activation functions are used in the layers. The most commonly used activation function is the rectified linear unit (ReLU). However, alternative activation functions such as sigmoid and tanh are also available (107). The use of activation functions is indispensable for understanding complex data relationships.

During model training, optimization is performed to find the optimal weights and bias values. For this purpose, the predictive performance of the model is evaluated using a loss function that measures the difference between predicted and actual outputs. The choice of the loss function depends on the type of problem. For

binary classification, binary cross-entropy loss is usually used, while cross-entropy loss is preferred for multi-class problems. For regression tasks, the mean square error is a standard choice. Optimization algorithms such as stochastic gradient descent, Adam, Adagrad, and RMSProp update the parameters of the model to minimize the loss function. DNNs use back-propagation for this process. It enables iterative updates and loss minimization to update parameters (108).

DNNs have trainable parameters such as weights and biases that are learned during training. These parameters can be further fine-tuned to improve the performance of the model. Hyperparameters that are tuned during training include learning rate, batch size, epoch, dropout rate, and weight decay. The learning rate determines the range for updating the weights. High rates might miss the optimal solution, while low rates might reach convergence slowly. Batch size divides the dataset into smaller subsets, standardizing training and increasing generalizability. Epochs represent the exact number of passes through the dataset, and early stopping can be applied to terminate training when performance stops improving. Dropout prevents overfitting by randomly deactivating a fraction of neurons during training and provides various weight updates. Weight decay regularizes the model by penalizing large weights to prevent overfitting and improve generalization (109).

DNNs have gained significant popularity in drug response prediction and various other biological applications involving multi-omics data integration. They are favored for their ability to effectively capture complex relationships and show high performance. However, DNNs require a significant number of samples to perform optimally and their performance can be suboptimal when sample size is limited. The sample size of omics data is unfortunately not large. Despite the presence of tens of thousands of gene features, the sample size is often significantly smaller. In fact, working with samples with drug response information leads to even smaller sample sizes. Therefore, predicting drug response robustly with DNNs at low sample size is a great challenge (110).

2.3. TL Approach

TL is a machine learning approach where knowledge gained from large, diverse, and generalizable models is transferred to other relevant domains (111). This approach is particularly useful when the target dataset is limited in size. However, training large datasets is both computationally demanding and time-consuming. TL has a great advantage in this regard. It does not need to train large datasets over several times. Instead, with fine-tuning, the trainable weight and bias parameters can be updated and the model can be easily used for the data to be tested (112) .

Another advantage of TL is its ability to handle distribution shifts between training and test datasets, a challenge that is often difficult to address with traditional methods. TL provides a highly generalizable approach. It enables robust performance across different datasets. This eliminates the need for dataset-specific training and testing (113).

DNNs are particularly effective when combined with a TL approach. They are excellent in handling high-dimensional data; thus, it makes them ideal for models integrating multi-omics data. Unlike traditional machine learning methods, DNNs can manage sparse data types such as mutations. In the mean time, they can effectively model complex relationships. Moreover, DNNs do not need feature selection methods because they are good at feature representation. This eliminates the need for dataset-specific feature selection as in the case of other machine learning methods (114).

3. MATERIALS AND METHODS

In this thesis, drug response prediction was performed by integrating multi-omics data with DNNs using a TL approach. Additionally, the resistance mechanisms of the drugs were investigated. For this purpose, the GDSC was used as the training dataset, while the PDX and TCGA were used as test datasets. The GDSC dataset contains gene expression, mutation, and CNA data from human cancer cell lines representing various cancer types. In the modeling approach, three different integrations were applied: expression-mutation integrated model (EM), expression-CNA integrated model (EC), and expression-mutation-CNA integrated model (EMC). Binary drug response data from the GDSC was used as input, and model performances were evaluated on the PDX and TCGA datasets (8-10).

When integrating response profiles of drugs, drugs targeting the same pathway were grouped to build pan-drug models. Instead of using all genes, differentially expressed genes (DEGs) were identified between resistant and sensitive samples for each drug. DEGs from each drug in the inhibitor group were unionized. In this way, it was possible to represent at least one DEG per drug. The EM, EC, and EMC pan-drug models were trained separately with drug groups including mitotic, cytoskeletal, DNA replication, and EGFR signaling pathway inhibitors. In other words, for each inhibitor group, all the three pan-drug models were built. Their performance in external datasets was evaluated using the area under the precision-recall curve (AUCPR) metric. Predictions were made for paclitaxel, 5-FU, gemcitabine, and cetuximab, which are shared between external datasets. For drug response prediction, four different strategies were followed with these common drugs, and robust predictions were performed. Afterwards, the performance of pan-drug models was compared in detail across strategies for assessing predictive power. The workflow outlining the steps involved in modeling and drug response prediction is presented in Figure 1.

Enrichment analyses using pathfindR were performed with the gene features of the strongest models predicting the TCGA data in order to obtain clinically relevant resistance mechanisms (11). By comparing the enrichment results of the top-performing pan-drug models across strategies, both general and drug-specific resistance mechanisms were explored. The shared enriched terms across strategies were regarded as general resistance mechanisms; on the other hand, unique terms were regarded as drug-specific resistance mechanisms. In addition, FET was applied to associate resistance with mutation and CNA statuses of genes identified within enriched terms. While pathfindR associates resistance primarily with transcriptional changes, FET can reveal potential associations with mutations or CNAs. This might show genomic-level changes complement transcriptional changes in resistance mechanisms. The detailed information for all stages is provided under the subheadings below.

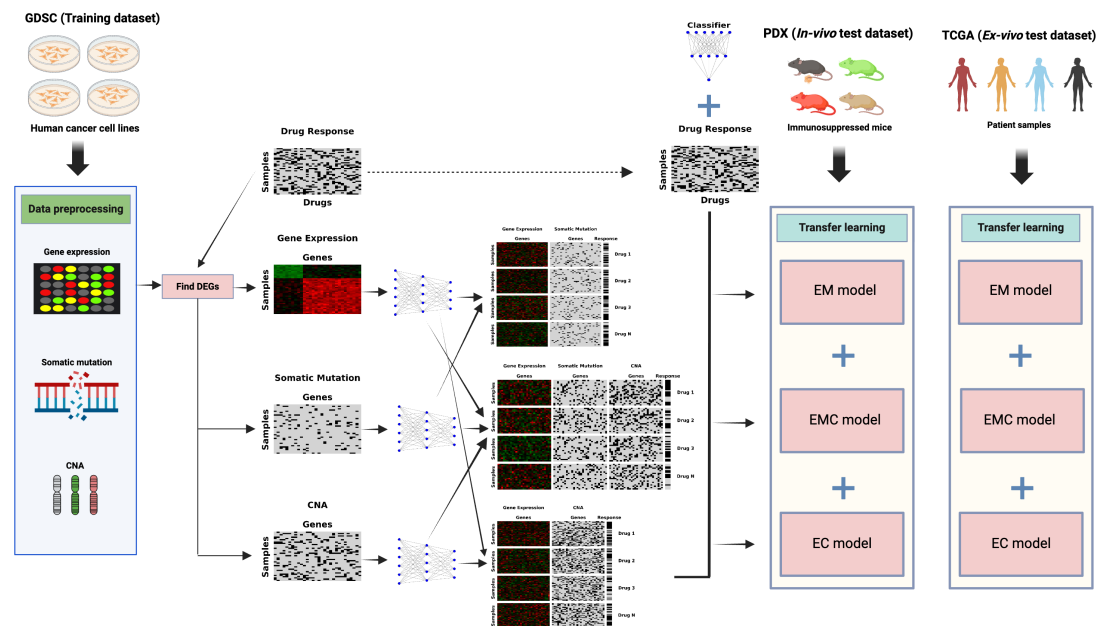


Figure 1. The workflow for predicting drug responses in the PDX and TCGA using a DNN- and MI-based TL approach

3.1. Datasets

Separate training datasets were constructed for the four distinct inhibitor groups (mitotic, cytoskeletal, DNA replication, and EGFR signaling inhibitors) based on their pathways in the GDSC dataset. Due to inclusion of paclitaxel in both the mitotic and cytoskeletal inhibitor groups, drug response of paclitaxel was evaluated in both contexts. Sample sizes of the GDSC, PDX, and TCGA datasets are given as to inhibitor groups below.

3.1.1. Pan-drug dataset with mitotic inhibitors

Multi-omics data (gene expression, somatic mutation, and CNA) from 19 different tissue cell lines were utilized in models along with response data from 11 drugs in the mitotic inhibitor group. Paclitaxel, which is regarded as both mitotic and cytoskeletal inhibitor, was included in the test dataset (115). Figure 2 illustrates the tissue origins of the cell lines in the GDSC, and lists the drugs with their corresponding numbers.

GDSC consists of 6360 samples, including 911 samples belonging to the NSCLC and breast adenocarcinoma (BRCA) cohorts. There are 602 sensitive and 5758 resistant samples in the dataset which shows a significant class imbalance. Only 9% of the samples are sensitive. In addition, 14% of the samples are sensitive for both cohorts, which were used in test datasets for paclitaxel response prediction (Figure 2A).

GDSC consists of 370 cell lines treated with paclitaxel (Figure 2B). Only 31 samples correspond to the NSCLC (20 samples) and the BRCA (11 samples), accounting for 8% of the total samples. 26 samples are paclitaxel-sensitive, whereas 344 are paclitaxel-resistant, showing only 7% of samples are sensitive (Figure 2C).

PDX includes 43 samples; however, only 5 of the samples are paclitaxel-sensitive, and 38 samples are paclitaxel-resistant. All sensitive samples belong to the

BRCA cohort, and the NSCLC cohort are full of with resistant samples. In total, 56% of the samples belong to the BRCA cohort, comprising 24 samples (Figure 2D).

The TCGA dataset is more diverse, including samples from five different cohorts. The BRCA is the dominant cohort, accounting for 77% of the 35 TCGA paclitaxel treated samples. There are 25 paclitaxel-sensitive and 10 paclitaxel-resistant samples in the TCGA. 22 of the samples are sensitive, and five of the samples are resistant in the BRCA cohort (Figure 2E). However, there is a significant inconsistency in sensitivity ratios between the PDX and TCGA datasets. The inconsistency is particularly evident within the BRCA cohort.

3.1.2. Pan-drug dataset with cytoskeletal inhibitors

There are 7561 samples treated with cytoskeletal inhibitors including 11 drugs in GDSC. 800 samples belong to the NSCLC cohort, while 405 samples belong to the BRCA cohort, which accounts for 16% of the dataset (Figure 3A). High class imbalance is also present with cytoskeletal inhibitors. 866 sensitive and 6695 resistant samples are present among cytoskeleton treated samples. In addition, only 11% of the samples is sensitive. The names and counts of cytoskeletal inhibitors are presented in Figure 3B.

3.1.3. Pan-drug dataset with DNA replication inhibitors

There are 13 different drugs among DNA replication inhibitors. Total sample size is 9435. Since both 5-FU and gemcitabine are included in the test datasets, tissue types common to pan-drug models in the training set are presented in separate graphs for each drug. Figure 4A shows the names and numbers of cell lines treated with DNA replication inhibitors categorized by tissue type. Common cancer types in the test sets for 5-FU account for 10% of all cell line samples. Overall, 9% of samples among DNA replication inhibitors are classified as sensitive. The names and counts of inhibitors used, including 5-FU and gemcitabine colored as red, are listed in Figure 4B. 11% of the 764 samples treated with 5-FU are sensitive, accounting for

82 sensitive samples and 682 resistant samples (Figure 4C). On the other hand, gemcitabine samples consist of 761 samples with 7% sensitivity ratio, indicating class imbalance is high (Figure 4D).

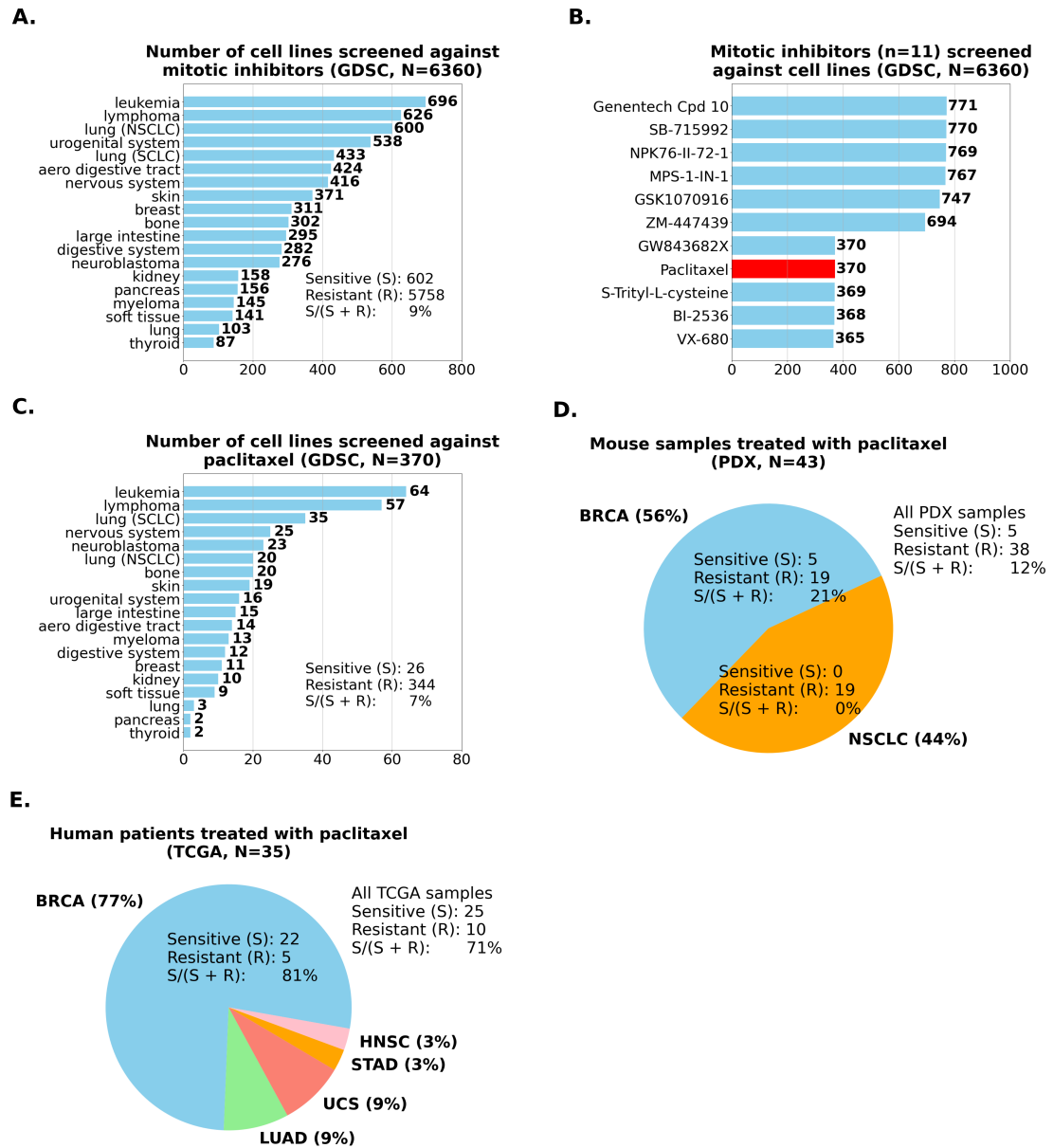


Figure 2. Overview of the GDSC, PDX, and TCGA datasets for paclitaxel response prediction with mitotic inhibitors. (A) Number of cell lines screened with mitotic inhibitors in the GDSC, (B) List of mitotic inhibitors in the GDSC, (C) Number of cell lines screened with paclitaxel in the GDSC, (D) Number of mouse samples treated with paclitaxel in the PDX, and (E) Number of human patients treated with paclitaxel in the TCGA

PDX samples belong to colorectal cancer (CRC) and pancreatic ductal adenocarcinoma (PDAC) cohorts for 5-FU and gemcitabine, respectively. The 5-FU PDX samples consists only one sensitive and 22 resistant samples, indicating 4% sensitivity ratio (Figure 4E). On the other hand, the gemcitabine PDX samples are with more sensitive samples. 28% of the samples are sensitive, accounting for 7 samples (Figure 4F).

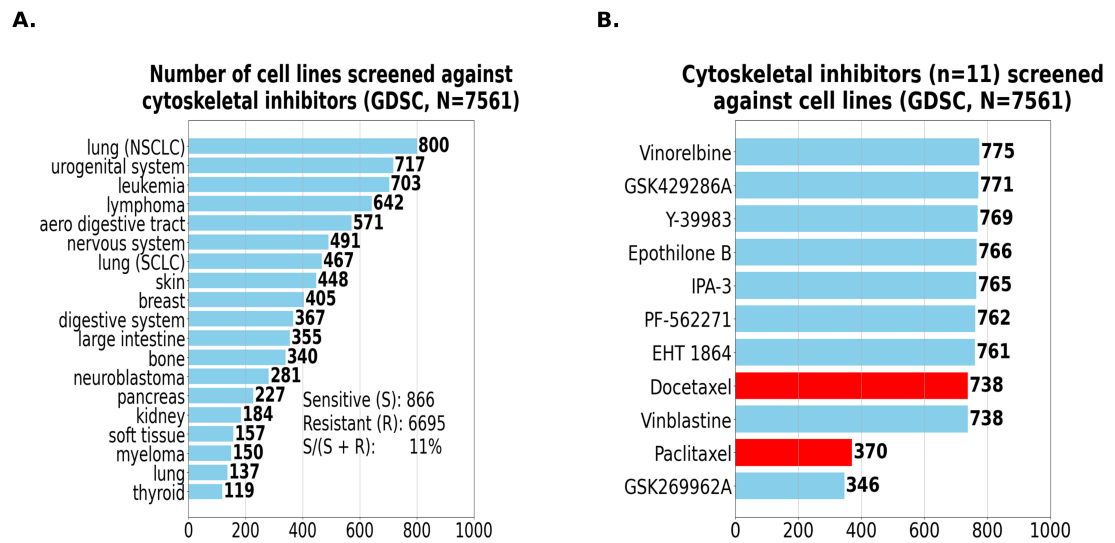


Figure 3. Overview of the GDSC dataset for paclitaxel response prediction with cytoskeletal inhibitors. (A) The number of cell lines screened against cytoskeletal inhibitors in the GDSC, and (B) the list of cytoskeletal inhibitors included in the GDSC.

TCGA samples are more diverse. There are four different cohorts for 5-FU and stomach adenocarcinoma (STAD) is the dominant cohort with 18 samples, five of which are resistant. In addition, there are six pancreatic adenocarcinoma (PAAD) samples; however, only two of them are sensitive. Esophageal carcinoma (ESCA) and rectum adenocarcinoma (READ) cohorts are all sensitive (Figure 4G). For gemcitabine, PAAD consists of 87% of TCGA samples, accounting for 48 samples with 20 sensitive and 28 resistant profile. Sarcoma (SARC), liver hepatocellular carcinoma (LIHC), lung squamous cell carcinoma (LUSC), and pheochromocytoma and paraganglioma (PCPG) cohorts are with sensitive samples. There are 55 patients treated with gemcitabine in the TCGA, and sensitivity ratio is 36% (Figure 4H).

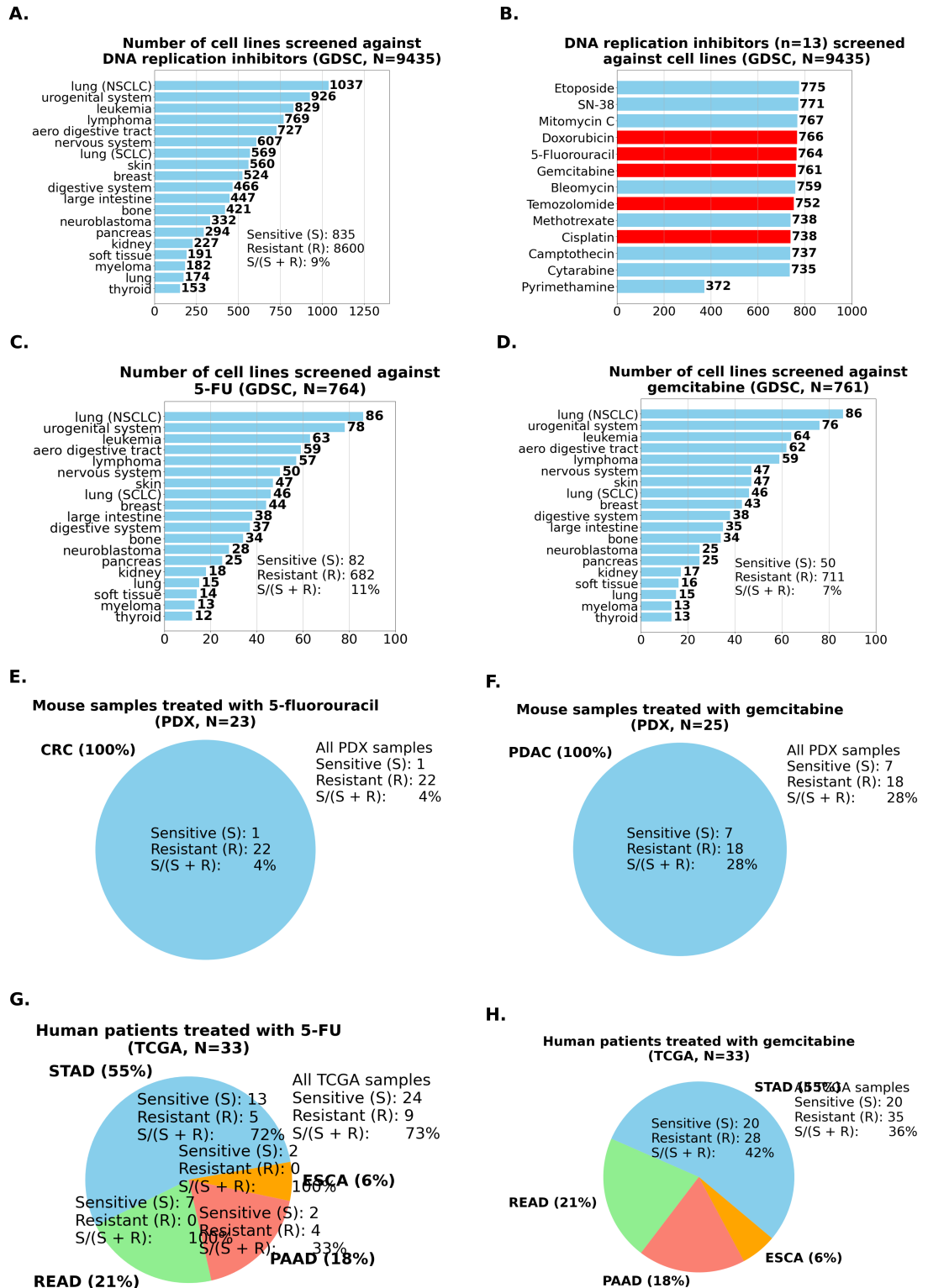


Figure 4. Overview of datasets for response prediction with DNA replication inhibitors: (A) the number of cell lines screened against DNA replication inhibitors in the GDSC, (B) the list of DNA replication inhibitors in the GDSC, (C–D) the number of cell lines screened against 5-FU and gemcitabine in the GDSC, (E–F) mouse samples treated with 5-FU and gemcitabine in the PDX, and (G–H) human patients treated with 5-FU and gemcitabine in the TCGA.

3.1.4. Pan-drug dataset with EGFR signaling inhibitors

EGFR signaling inhibitors include a total of 5633 samples, 16% of which are common to the cancer types treated in the test datasets. The test datasets include the NSCLC, the CRC, and head and neck squamous cell carcinoma (HNSC) cohorts. 794 of the samples are sensitive, representing 14% sensitivity ratio (Figure 5A). The training dataset includes eight different EGFR signaling inhibitors. However, afatinib is used twice in the dataset since the GDSC consists of duplicated drug response for afatinib. Both response data was used in training dataset. 725 of the samples are treated with cetuximab (Figure 5B). 725 cetuximab treated cell lines are present, consisting of 99 sensitive and 626 resistant samples (Figure 5C).

In the PDX dataset, the CRC and NSCLC are the only two cohorts. While only one sensitive samples in the NSCLC cohort is present in a total of 24 samples, four sensitive and 32 resistant CRC samples are present in the dataset. In addition, all sensitive samples represent 8% of the total samples, indicating a high class imbalance (Figure 5D). In the TCGA, only the HNSC cohort samples consisting of nine samples are available. Six of the samples are sensitive and three samples are resistant (Figure 5E).

3.2. Preprocessing of Multi-Omics Datasets

The GDSC, PDX, and TCGA datasets use different platforms for gene expression, mutation, and CNA analyses. Several preprocessing steps were applied to handle the differences in platforms to eliminate batch effects. In this way, the datasets were made comparable. Preprocessing steps were performed exactly as described in the MOLI (116).

3.2.1. Preprocessing steps with gene expression data

Affymetrix Human Genome U219 Array was used for expression profiling in the GDSC (8). CEL-format files of all cancer cell lines were downloaded from the

ArrayExpress database (117). The affy package in R was used to preprocess the raw data (118). Robust Multiarray Average (RMA) was chosen as the normalization method (119). Background correction, quantile normalization, and log transformation of probe intensities were then applied to preprocess the gene expression data (120).

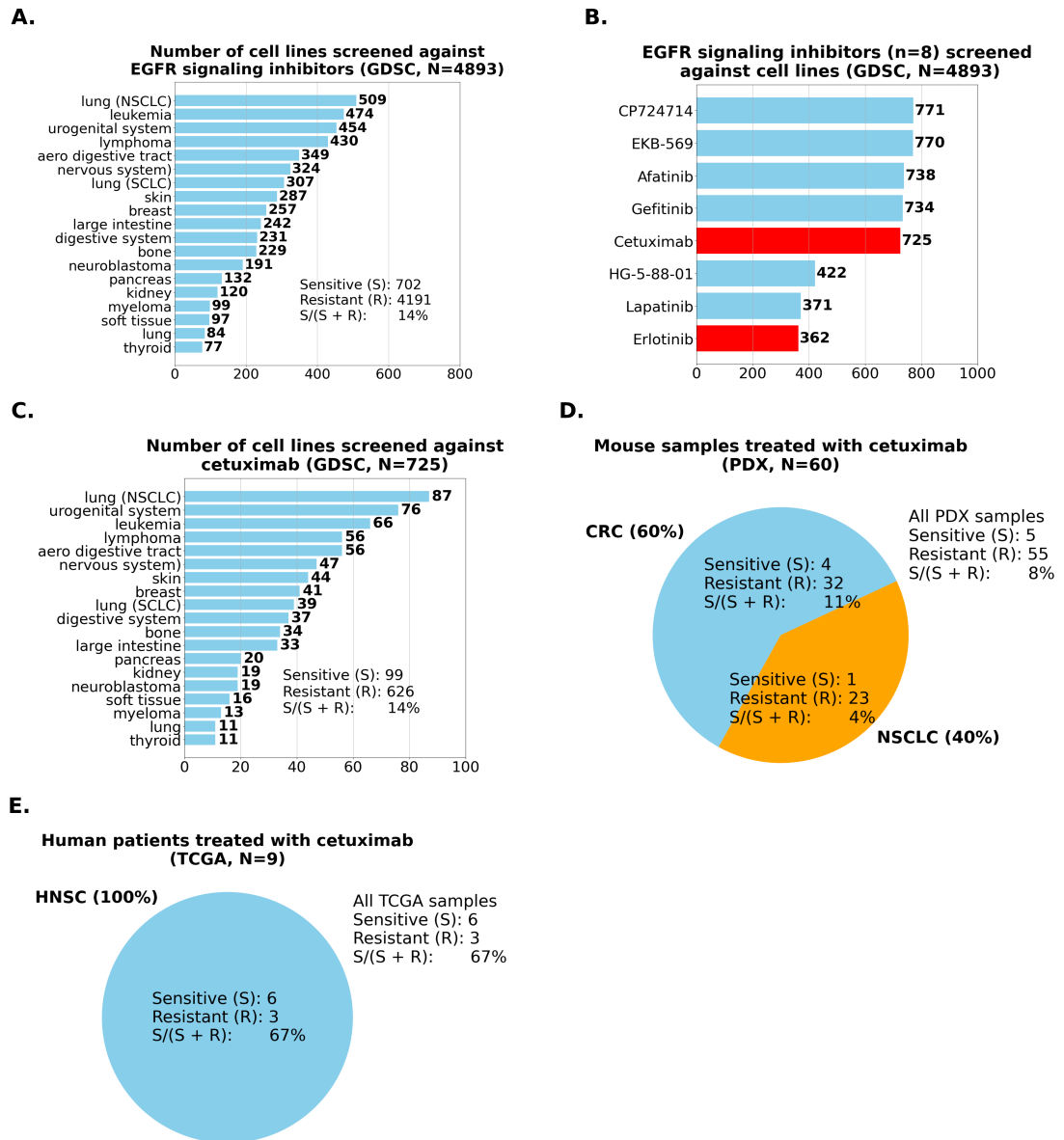


Figure 5. Overview of datasets for cetuximab response prediction with EGFR signaling inhibitors: (A) the number of cell lines screened against EGFR signaling inhibitors in the GDSC, (B) the list of EGFR signaling inhibitors in the GDSC, (C) the number of cell lines screened against cetuximab in the GDSC, (D) mouse samples treated with cetuximab in the PDX, and (E) human patients treated with cetuximab in the TCGA.

Chip description library files and Affymetrix platform specific probe set descriptions were downloaded from BrainArray (121). After this step, the Entrez Gene IDs of the probe sets obtained by the normalization were found for the probe sets. Using the WGCNA R package, intensity values were calculated using the “Average” method for those matching more than one probe set (122). Those matching multiple Entrez Gene IDs were excluded from the analysis as they were considered non-specific.

Unlike GDSC, expression data in the PDX and TCGA were analyzed with RNA-Seq. However, there is no raw data in both datasets. Therefore, preprocessed expression data were used. RNA-Seq data of the PDX samples was accessed from the supplementary file of the manuscript of Gao et al. as fragments per kilobase of transcript per million mapped reads (FPKM) format (9). Estimated fractions of transcripts in the TCGA calculated with RNA-Seq expectation maximization (RSEM) method are available on Firehose Broad GDAC website, from which the files of all cohorts in the TCGA were downloaded (123). To make the two different RNA-Seq data comparable, both were converted to transcripts per million (TPM) normalization format as described in the MOLI (116). Then both were log₂-transformed to make all datasets comparable. However, for the expression data to be comparable between datasets, the batch effects between them had to be removed. For this reason, pairwise homogenization between the datasets was performed with the sva R package (124). An important point here is that only common genes were used in the homogenization step between the datasets.

3.2.2. Preprocessing steps with mutation data

Somatic mutation data in the GDSC was accessed from the website https://www.cancerrxgene.org/downloads/bulk_download and the “WES_variants.xlsx” file was used in the analysis (8). For the PDX, the excel file in the supplementary table in Gao et al.'s manuscript was used (9). The excel file contains all the omics-data, but somatic mutation data is in the “pdx_mut_and_cn2” tab. For the TCGA mutation

data, mutation annotation file (MAF) files in Firehose Broad GDAC were downloaded (123). After accessing the MAF files of all cohorts, amplifications and deletions were removed. Only protein-coding non-synonymous mutations such as missense, nonsense, and frameshift variants were retained for analysis. Silent mutations were also removed. In the mutation dataset, the genes with non-synonymous mutations were binarized as 1, while non-mutated genes as 0. All gene IDs were mapped to Entrez Gene IDs as in the gene expression dataset.

The OncoKB database was used to identify tumor suppressor genes and oncogenes (125). They were classified according to the list in this database and used in the mutation data with their functional impact indicated. In tumor suppressor genes, -1 was given for loss of function, whereas for oncogenes, 1 was given for gain of function mutations. Driver genes were also taken into account, and mutations near driver genes were also scaled according to their distance with the method described in the MOLI (116). However, as mentioned in the previous paragraph, the values were binarized to be used in the DNN.

The genes listed in the database as both tumor suppressor genes and oncogenes were considered for truncating mutations and given a value of -1 if truncating mutations were found. The genes without truncating mutations were directly excluded from the analysis. In addition, the oncogenes activated only by upregulation or amplification were also excluded from the analysis because somatic mutations in these genes can not be realistically assessed for functional impact (116).

cBioPortal database was also used for one-dimensional and three-dimensional hotspot mutation data (126). In addition, hotspot information from both COSMIC database and the manuscript of Bailey et al. were used (127, 128). Driver mutations could be annotated robustly with information obtained from multiple datasets in this way.

3.2.3. Preprocessing steps with CNA data

For the TCGA dataset, the probe intensity values of the samples were normalized. For this purpose, normalization was performed by comparing the intensities of the most similar normal samples in the HapMap (International HapMap3) database (129). Log₂-transformation was performed afterwards. The obtained log-ratio values were then segmented using the circular binary segmentation algorithm as described in the MOLI (116, 130). Thus, the regions with consistent log-ratio values were identified.

The regions likely to be germ-line segments and the regions with a small number of probes were removed from the tumor CNA profile since these segments might be noisy (116). The remaining segments were matched with gene annotation using the GRCh37/hg19 assembly in NCBI (131). The log-ratio values were assigned to the regions with overlapping segments. In cases with more than one matching segment, the most extreme log-ratio score was given. Segments with no matching segments or segments with low $|\log\text{-ratio}|$ scores were considered as copy-neutral regions (116). After this step, the regions with amplification or deletion were assigned a value of 1 to be used in DNN, while copy-neutral regions were assigned a value of 0. The reason for the binary coding is that DNNs can better understand the relationships in this way (116). Deletions could have been assigned -1, amplifications 1, others 0. However, this coding method was not preferred.

Unlike TCGA, GDSC and PDX have predicted copy number information at the gene level. To make them comparable to the CNAs in the TCGA, the logarithm of the copy number of the genes was calculated by dividing the logarithm by the ploidy of the copy-neutral state in the sample as described in the manuscript of Sharifi-Noghabi et al. (116). In these datasets, the copy-neutral state was found by estimating the distribution of copy number estimates at the gene level. In doing so, it was considered that the mode closest to the median of the copy-neutral state could represent copy number estimates. As in the TCGA, those with low $|\log\text{-ratio}|$ values

were considered copy-neutral. The high ones were coded as 1 for amplification and -1 for deletion (116). However, since binary representation was used in the DNN, regions with amplification or deletion were coded as 1 and copy-neutral as 0 as in the TCGA.

3.3. Finding DEGs

For each drug within the inhibitor groups, the DEGs were identified from the GDSC dataset by comparing sensitive and resistant cell line samples. The limma R package was utilized for this analysis, applying criteria of an adjusted p-value $< 0,05$ and $|\log_2FC| > 1$ (132). The DEGs meeting these criteria were determined individually for each drug, and subsequently, the DEGs from all drugs within the same inhibitor group were combined into a union set. This ensured that they represent all the inhibitor group. These unionized DEGs were then filtered across gene expression, mutation, and CNA datasets within the inhibitor group and used as features for downstream analyses.

3.4. Building DNNs with TL Approach for Drug Response Prediction

Multi-omics data integration was performed using DNNs to predict drug response in the PDX and TCGA datasets. Similar to the MOLI framework, auto-encoders, which are a type of feed-forward network, were independently trained for gene expression, mutation, and CNA datasets (116). Each dataset was optimized separately, and the outputs were then integrated into the DNN. At this stage, drug response data were incorporated to validate the overall predictive performance of the model on two independent datasets. These steps are visualized in the workflow in Figure 1.

MI approach utilizing DNNs with TL was applied to predict the responses of unique and shared drugs in the PDX and TCGA datasets. Drugs targeting the mitotic, cytoskeletal, DNA replication, and EGFR signaling pathway were trained separately using the GDSC data to develop the EM, EC, and EMC models. In all models, only

features common across datasets were used to ensure consistency between training and test datasets. This approach enabled response prediction for each drug in external datasets.

Although the training datasets included response data for the drugs being tested, this did not result in data leakage, as no samples from the external datasets were used in training. Instead, only the response data of drugs shared between datasets were utilized. While this may introduce potential bias favoring the drugs being predicted, the same strategy as MOLI was adopted to enable fair performance comparisons.

To ensure robust response predictions for drugs shared between both the PDX and TCGA datasets, four different strategies were implemented. In the first strategy, data for all cancer types and response data for all drugs were included in the pan-drug training datasets. In the second strategy, response data for the tested drugs were excluded. The third strategy removed common cancer types from the training dataset but retained response data for the tested drug. The fourth strategy excluded both response data for the tested drug and the common cancer types. To address class imbalance in drug responses, a weighted random sampler was applied during training, ensuring the minority-sensitive class was sampled more frequently. The PyTorch framework in Python was used for the sampler and all modeling steps (133). Model performance was evaluated using the AUCPR metric, which is particularly suitable for datasets with significant class imbalance (134).

Separate auto-encoders were trained for each omic dataset to be integrated into the DNN models. Each auto-encoder consisted of a single linear layer, batch normalization, a ReLU activation function, and a dropout parameter. Batch normalization was employed to accelerate the training process and improve performance by normalizing the data to a mean of 0 and a standard deviation of 1. The ReLU activation function introduced non-linearity, improving performance and enabling the learning of complex patterns. Dropout was applied to prevent overfitting. All auto-encoders used the Adagrad optimizer, with weight updates

controlled by the weight decay parameter to shrink large weights and enhance the generalizability of the models (116).

The outputs of the auto-encoders were concatenated, and L2 regularization was applied. This step added the squared magnitude of all weights as a penalty term in the loss function to prevent overfitting and improve the model's ability to generalize to unseen data. For binary classification, an auto-encoder with a single linear layer, dropout parameter, and sigmoid activation function was used in the final layer. Binary cross-entropy were used as the loss function, classifying output values above 0,5 as 1 (sensitive) and those below 0,5 as 0 (resistant) (116).

The hyperparameters used during model training were optimized through random search over 50 iterations. To prevent overfitting and ensure generalizability, 5-fold cross-validation was applied (134). The hyperparameter ranges were as follows:

Mini-batch size: [8, 16, 32, 64]

Input layer size: [1024, 512, 256, 128, 64, 32, 16]

Learning rate: [0,5, 0,1, 0,05, 0,01, 0,001, 0,005, 0,0005, 0,0001, 0,00005, 0,00001]

Epoch: [20, 50, 10, 15, 30, 40, 60, 70, 80, 90, 100]

Dropout rate: [0,3, 0,4, 0,5, 0,6, 0,7, 0,8]

Weight decay: [0,01, 0,001, 0,1, 0,0001]

Following initial drug response prediction, model fine-tuning was conducted using 5-fold cross-validation over 50 iterations, employing the same hyperparameters described above. The pan-drug models initially trained on the GDSC data were subsequently adapted to the PDX and TCGA datasets in a drug-specific manner after freezing gene expression, mutation and CNA layers. Model performance was evaluated across iterations, and the reported AUCPR scores reflect those of the highest-performing models obtained during the fine-tuning process.

3.5. Explaining Drug Resistance Mechanisms with Pathway Enrichment

The pathfindR R package was used for enrichment analyses (11). It is a pathway enrichment tool that can be used to identify drug resistance mechanisms. It creates a protein-protein interaction (PPI) network using an active sub-network approach. pathfindR uses various public PPI networks such as Kyoto Encyclopedia of Genes and Genomes (KEGG) and Reactome for this purpose (135, 136). Starting from the most significant gene, it calculates a significance score for neighboring genes around it. If a neighboring gene increases the score, it is included in the sub-network. By repeating this procedure iteratively, it keeps genes that are consistently in the active-subnetwork together in the network. With this approach, genes are not considered as acting in isolation; on the contrary, they are regarded to be in activity through interactions. pathfindR uses DEGs and p-values to determine which genes in the active sub-network are important for biological activity. Log fold-change (logFC) values are also optionally available but are important for depth of analysis and visualization. By default it uses the greedy approach. It also offers options such as simulated annealing and genetic algorithm. However, these can be slower (11).

In this thesis, the greedy approach was used. The KEGG database was utilized for the construction of the PPI network. With 25 iterations, active sub-networks were found. Resistance-related genes used in the modeling for inhibitors were entered as input with logFC values in the enrichment analyses. pathfindR analyses were performed under the first and second strategies. Drug-specific resistance mechanisms were examined through comparative analyses. General resistance mechanisms were determined by identifying common enriched terms, whereas drug-specific resistance mechanisms were identified from unique terms using the first strategy. Since three different models were tested, the analyses were compared based on the gene features of the best performing models on the TCGA data. The aim was to link these clinically relevant findings to underlying resistance mechanisms.

3.6. Association of Mutations and CNAs with Drug Resistance Using FET

Mutations and CNAs associated with drug resistance between sensitive and resistant samples were analyzed using FET with the scipy python package (137). Mutation and CNA associations with drug resistance for the four drugs common to both the PDX and TCGA datasets (paclitaxel, 5-FU, gemcitabine, and cetuximab) were assessed using the GDSC. In addition, it was examined whether genes showing significant mutation and CNA associations were included in the pathfindR enrichment analyses. Due to the small sample size and class imbalance in the PDX and TCGA datasets, FET was performed only for the GDSC datasets, with 1000 bootstrapping iterations. The median odds ratio (OR) and p-value were used for evaluation.

3.7. Assessing Predictive and Prognostic Significance of Genes with TCGA

The univariate CoxPH regression models were applied using mutation, CNA, and resistance information to investigate how genomic features collectively influence survival outcomes and drug resistance in the TCGA datasets. Instead of analyzing the independent effects of CNAs and mutations on resistance, univariate CoxPH regression models with interaction terms were used to identify synergistic or antagonistic combined effects. This provided a more comprehensive understanding of complex resistance mechanisms.

Two separate univariate penalized CoxPH regression models with ridge regularization were used to estimate hazard ratios (HRs) for resistance-mutation and resistance-CNA interaction terms as shown in Equations (1) and (2), respectively. Model performances were optimized using 5-fold cross-validation, and 1000 bootstrapping iterations were performed to compute HRs with 95% confidence intervals (CIs) for assessing the strength and reliability of each covariate's association with survival, as well as the concordance index (C-index).

$$h(t) = h_0(t) \cdot \exp(\beta_1 \cdot \text{expression} + \beta_2 \cdot \text{cna} + \beta_3 \cdot \text{response} + \beta_4 \cdot (\text{cna} \times \text{response})) \quad (1)$$

$$h(t) = h_0(t) \cdot \exp(\beta_1 \cdot \text{expression} + \beta_2 \cdot \text{mutation} + \beta_3 \cdot \text{response} + \beta_4 \cdot (\text{mutation} \times \text{response})) \quad (2)$$

In both equations, $h(t)$ denotes the hazard function, representing the instantaneous risk of the event at time t ; expression is the gene expression level; and response indicates drug response status (0 or 1). In Equation (1), cna represents CNA status (0 or 1) with the interaction term $\text{cna} \times \text{response}$ capturing whether the effect of CNA on survival depends on drug response. In Equation (2), mutation denotes mutation status (0 or 1) with the interaction term $\text{mutation} \times \text{response}$ capturing whether the effect of mutation on survival depends on drug response. In both models, β_1 - β_4 are the regression coefficients estimating the effect of each variable on the hazard.

To incorporate ridge regularization into the CoxPH regression models, the coefficient estimates were obtained by maximizing the penalized partial log-likelihood as shown in Equation (3).

$$\text{Maximize: } \ell(\beta) - \lambda \sum_{j=1}^p \beta_j^2 \quad (3)$$

where:

- $\ell(\beta)$ is the Cox partial log-likelihood,
- $\lambda \geq 0$ is the ridge penalty parameter,
- p is the number of predictors.

For multiple gene combinations, two separate multivariate CoxPH regression models with ridge regularization were constructed, as shown in Equations (4) and (5).

$$h(t) = h_0(t) \cdot \exp\left(\sum_{i=1}^n \beta_{1i} \cdot \text{expression}_i + \sum_{i=1}^n \beta_{2i} \cdot \text{cna}_i + \beta_3 \cdot \text{response} + \sum_{i=1}^n \beta_{4i} \cdot (\text{cna}_i \times \text{response})\right) \quad (4)$$

$$h(t) = h_0(t) \cdot \exp \left(\sum_{i=1}^n \beta_{1i} \cdot \text{expression}_i + \sum_{i=1}^n \beta_{2i} \cdot \text{mutation}_i + \beta_3 \cdot \text{response} + \sum_{i=1}^n \beta_{4i} \cdot (\text{mutation}_i \times \text{response}) \right) \quad (5)$$

The C-indices of genes identified in the pathfindR enrichment analyses were compared to the multi C-indices of all genes within their respective pathways. Additionally, bar graphs were generated to visualize the C-index for evaluating model performances. KM plots were also visualized to assess survival differences based on CNA, mutation, and resistance status with statistical significance determined using the LRT.

4. RESULTS

The predictive performances of the multi-omics-integrated pan-drug models, developed using DNNs with a TL approach for drugs targeting the same pathways in the GDSC training datasets, are presented separately for the PDX and TCGA datasets in the following sections. Drug response prediction performances for drugs common to both datasets are then compared across four different strategies, followed by a comparison with MOLI results (116). Drug resistance mechanisms for each drug are identified based on pathfindR enrichment results from the models with demonstrated predictive power. Finally, FET results for mutation and CNA statuses are examined.

4.1. Predictive Performance of Pan-Drug Multi-Omics Models

For drugs with response data available either in the PDX or TCGA datasets, response predictions were performed separately for each dataset since there are some drugs having response data in one dataset but not in other dataset. Additionally, predictions for drugs common to both datasets were made using the four robust strategies described as in method section. By comparing the performance of the response predictions in the MOLI with the pan-drug models trained in this thesis, it was demonstrated that the constructed pan-drug models outperformed the MOLI predictions, achieving stronger predictive performance across all the six drugs (116).

4.1.1. Performance of pan-drug models tested individually in PDX or TCGA

Drug response predictions were performed for paclitaxel, erlotinib, cetuximab, gemcitabine, and 5-FU in the PDX dataset. In the TCGA dataset, drug response predictions were performed for cetuximab, paclitaxel, cisplatin, doxorubicin, 5-FU, gemcitabine, temozolomide, paclitaxel, and docetaxel.

4.1.1.1. Performance of pan-drug models tested with PDX dataset

The following subheadings present detailed results of the predictions for paclitaxel, erlotinib, cetuximab, 5-FU, and gemcitabine in that order.

4.1.1.1.1. Predictive performance of pan-drug models for paclitaxel response prediction

Predictive performance of paclitaxel was evaluated using both mitotic and cytoskeletal inhibitor classes with multi-omics pan-drug models. All samples from the PDX and TCGA datasets, as well as the BRCA cohort, which constitutes the majority of the samples, were used for prediction separately. Therefore, a total of six different models were trained to build the EM, EC, and EMC models.

First, mitotic inhibitors were trained for response prediction of paclitaxel using the EM, EC, and EMC configurations. All models were trained with 6360 samples. However, the feature sizes differed due to variations in common features across datasets during integration. The performances of the models in the GDSC training dataset are visualized in Figure 6. The EC model yielded the strongest predictive performance across all 43 samples with AUCPR = 0,33 and AUCPR = 0,56 after fine-tuning. When the models were tested on the BRCA cohort, consisting of 24 samples, the EMC model showed the strongest performance with AUCPR = 0,42 and AUCPR = 0,61 after fine-tuning, significantly outperforming random prediction, which had AUCPR = 0,21 (Figure 6).

Predictions made with cytoskeletal inhibitors showed slightly different performance compared to that of mitotic inhibitors when all samples were considered. In the GDSC training dataset, the strongest model was with the EC. However, when all samples were used, the EMC model achieved the superior predictive performance with AUCPR = 0,42 with cytoskeletal inhibitors. The best model was still the EMC after fine-tuning though the AUCPR did not change. However, for predictions made with the BRCA cohort, the best performance was

achieved with the EMC model, similar to the mitotic inhibitors with AUCPR = 0,69 and AUCPR = 0,72 after fine-tuning (Figure 7).

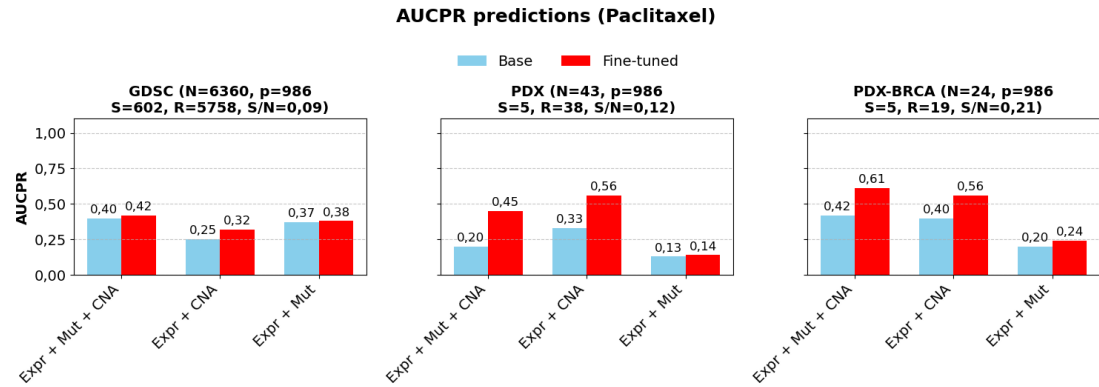


Figure 6. AUCPR predictions with mitotic inhibitors for paclitaxel response prediction using the PDX dataset

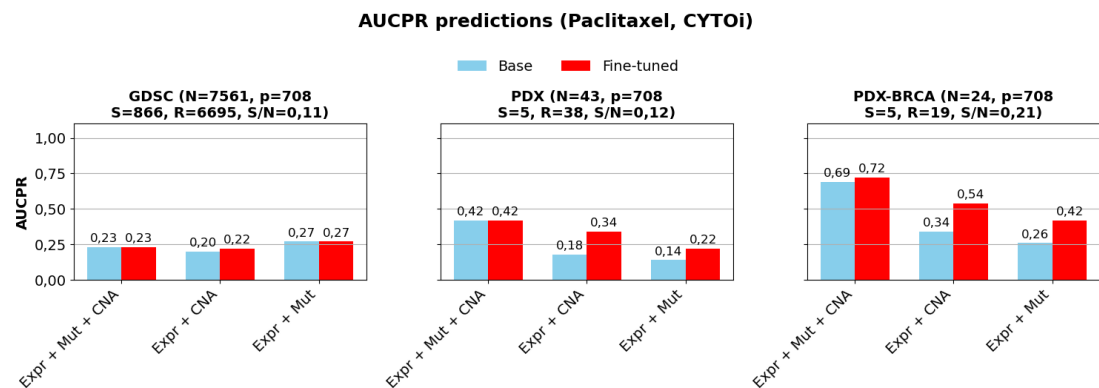


Figure 7. AUCPR predictions with cytoskeletal inhibitors for paclitaxel response prediction using the PDX dataset

4.1.1.1.2. Predictive performance of pan-drug models for erlotinib and cetuximab response prediction

Erlotinib and cetuximab are EGFR signaling inhibitors; thereby, the performance of the models is presented together for both drugs. Although a single training model could have been used for predicting the response for both drugs, separate training datasets were created for each. The purpose of this was to compare whether predictions made with two different models would yield similar results. Similar

prediction outcomes were obtained; however, the results presented here are those from the models with the best response performance.

In the training dataset, the strongest prediction for erlotinib was achieved with the EM model, yielding an AUCPR = 0,50. However, in the PDX dataset, the EC model demonstrated the best performance with AUCPR = 0,48 and the EM model demonstrated the best performance with AUCPR = 0,81 after fine-tuning. Since only the NSCLC cohort samples were present in the PDX dataset, strongest models yielded same AUCPR scores (Figure 8).

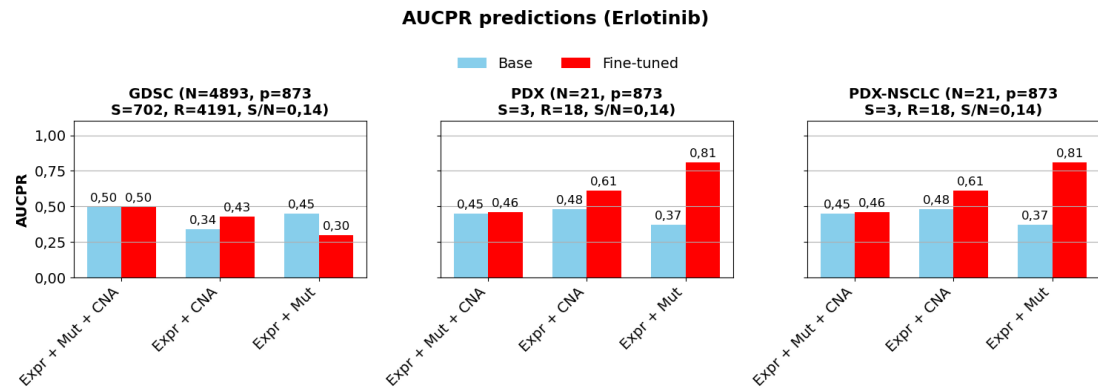


Figure 8. AUCPR predictions with EGFR signaling inhibitors for erlotinib response prediction using the PDX dataset

For the training models built for cetuximab, the EMC model exhibited the best performance. On the other hand, in predictions made with all 60 samples in the PDX dataset, the EC model was the strongest. The presence of only five sensitive samples highlighted a significant class imbalance. Similarly, for the CRC cohort with 36 samples, the EMC model achieved the strongest prediction with AUCPR = 0,33 and the EC model achieved the strongest prediction with AUCPR = 0,40 after fine-tuning.¹ (Figure 9). For the NSCLC cohort with 24 samples but only one sensitive sample, the EC model performed best. Nevertheless, due to the presence of only one sensitive sample and the high class imbalance, making a definitive evaluation of the prediction performance is unrealistic.

4.1.1.1.3. Predictive performance of pan-drug models for 5-FU and gemcitabine response prediction

5-FU and gemcitabine are inhibitors of DNA replication and were therefore tested using DNA replication models. The EM, EC, and EMC models were trained for both drugs, following the same approach used for EGFR signaling inhibitors.

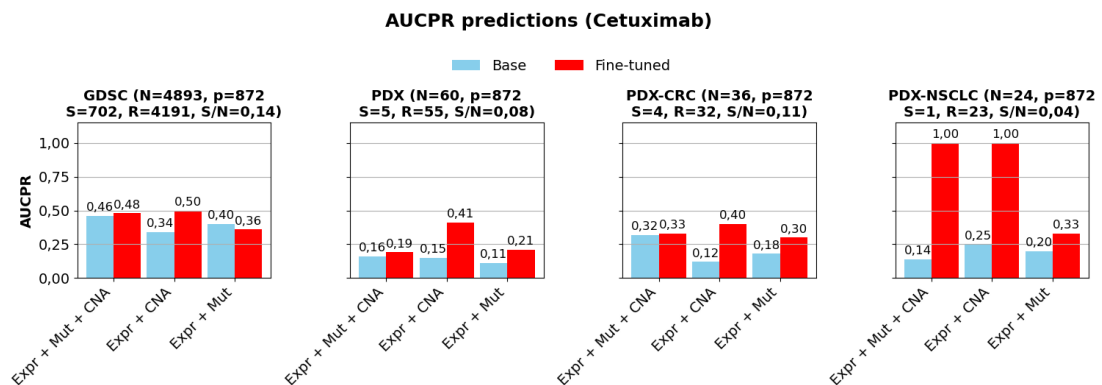


Figure 9. AUCPR predictions with EGFR signaling inhibitors for cetuximab response prediction using the PDX dataset

For 5-FU, the EMC model demonstrated the strongest performance in the training dataset. In the PDX dataset, which consisted solely of the CRC cohort, the EM model achieved the best performance with AUCPR = 1,00 (Figure 10). Figure 10 shows the same prediction in the CRC cohort, which has the same number of samples in all 5-FU samples in the PDX. However, given that there was only one sensitive sample, a realistic evaluation of the prediction is challenging.

Similarly, the EMC model showed the strongest performance in the training dataset for gemcitabine. In the PDX test dataset, which included 25 samples from the PDAC cohort, the EMC model achieved the best prediction with AUCPR = 0,47 and with AUCPR = 0,67 after fine-tuning, compared to a random prediction of AUCPR = 0,28. Since only PDAC samples were in the PDX samples, same predictions were obtained (Figure 11).

4.1.1.2. Performance of pan-drug models tested with TCGA dataset

The TCGA dataset contains sufficient data to perform drug response predictions for paclitaxel, docetaxel, cetuximab, cisplatin, doxorubicin, gemcitabine, 5-FU, and temozolomide. Predictions were carried out for each of these drugs.

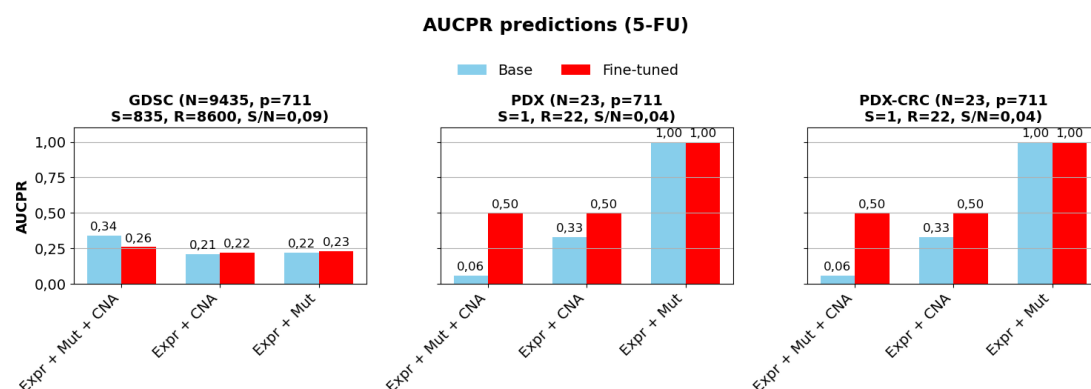


Figure 10. AUCPR predictions with DNA replication inhibitors for 5-FU response prediction using the PDX dataset

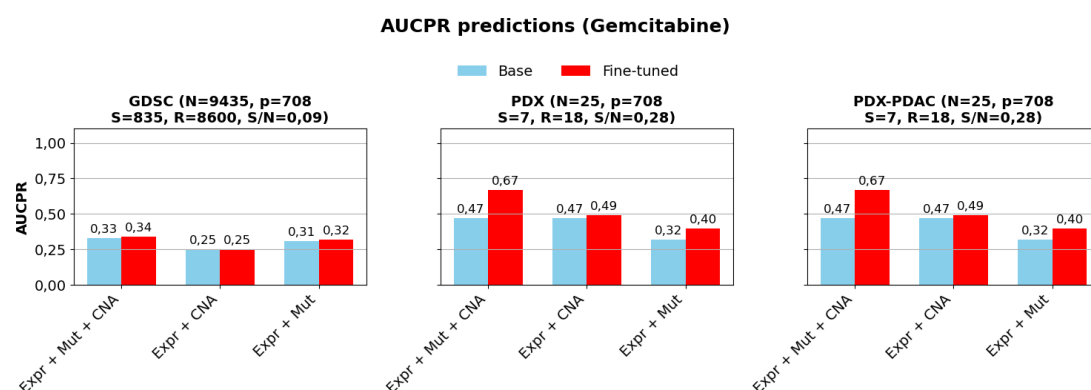


Figure 11. AUCPR predictions with DNA replication inhibitors for gemcitabine response prediction using the PDX dataset

4.1.1.2.1. Predictive performance of pan-drug models for paclitaxel and docetaxel response prediction

For paclitaxel and docetaxel, both mitotic and cytoskeletal models were used for response prediction, resulting in a total of six models.

For paclitaxel, the most effective model in the training dataset was the EMC model. In the PDX dataset, which consisted of 35 samples, the EMC model again yielded the strongest prediction with AUCPR = 0,88. However, fine-tuning did not improve AUCPR score. The random prediction was AUCPR = 0,71. In the BRCA cohort, the EMC model was again the most powerful model. It achieved AUCPR = 0,96 closely followed by the EC model with AUCPR = 0,94. Fine-tuning yielded the same AUCPR score. Random prediction for this cohort resulted AUCPR = 0,81. Both models performed exceptionally well (Figure 12).

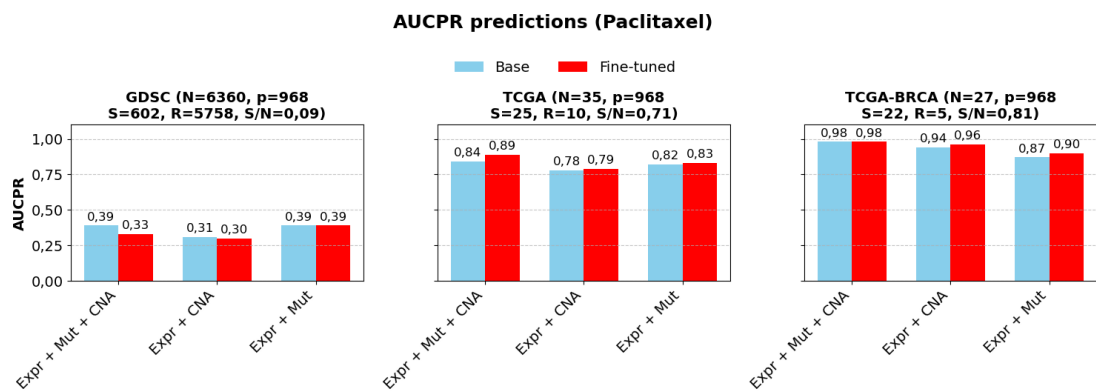


Figure 12. AUCPR predictions with mitotic inhibitors for paclitaxel response prediction using the TCGA dataset

For cytoskeletal models, the EMC model demonstrated the strongest prediction in the training dataset with AUCPR = 0,31. In the TCGA dataset, the EC model outperformed others for all samples with AUCPR = 0,78 and AUCPR = 0,87 after fine-tuning. The EMC and EM models also exhibited AUCPR scores greater than 0,70, which indicated all three models showed almost same performance. In the BRCA cohort, the EC model again showed the highest performance with AUCPR = 0,95 and AUCPR = 0,96 after fine-tuning (Figure 13). However, unlike the mitotic inhibitor models where all models outperformed random predictions, only the EC model achieved strong predictions in this case.

When mitotic models were applied to docetaxel, the EMC model showed the strongest performance in the training dataset. In the TCGA dataset, consisting of 16

samples, the EMC remained the strongest model with AUCPR = 0,78 and the EC was the strongest model with AUCPR = 0,85 after fine-tuning. Notably, docetaxel was the only drug with balanced classes, as the dataset included eight sensitive and eight resistant samples. All models performed better than random prediction in this dataset. In the BRCA cohort, consisting of six sensitive and two resistant samples, the EMC model yielded AUCPR = 0,90 while the EC model yielded AUCPR = 0,98 after fine-tuning (Figure 14).

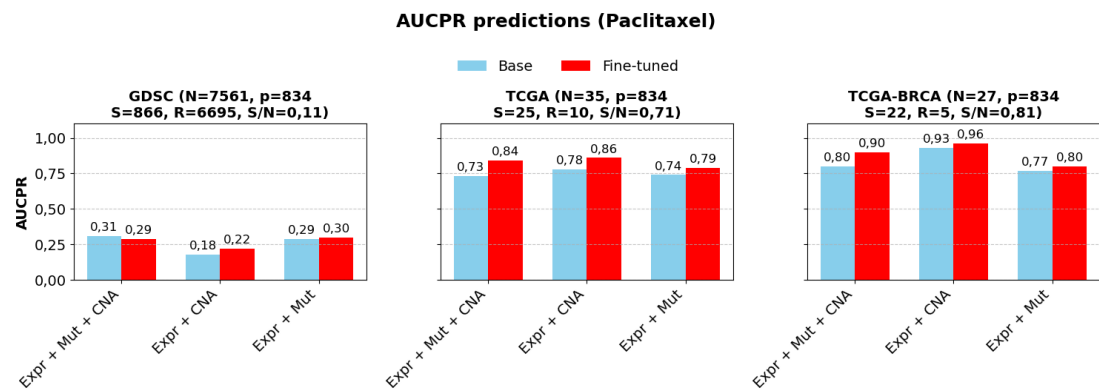


Figure 13. AUCPR predictions with cytoskeletal inhibitors for paclitaxel response prediction using the TCGA dataset

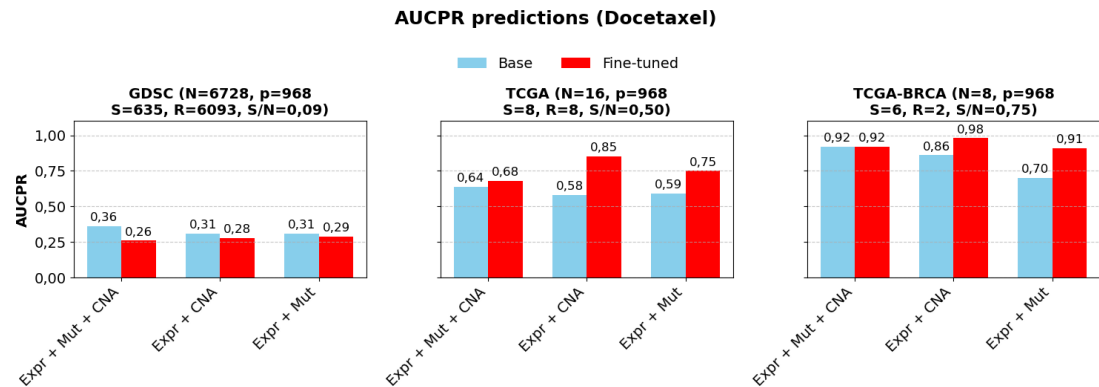


Figure 14. AUCPR predictions with mitotic inhibitors for docetaxel response prediction using the TCGA dataset

In predictions using cytoskeletal models, the EM and EC models achieved the strongest performance in the GDSC training dataset with AUCPR = 0,30. In the TCGA dataset, the EMC model did not outperform random predictions; however, the

EC model demonstrated the strongest performance with AUCPR = 0,64 and AUCPR = 0,70 after fine-tuning. The EM model also predicted response with AUCPR = 0,57, which was higher than random AUCPR. In the BRCA cohort, the EM model emerged as the strongest, achieving AUCPR = 0,96, but fine-tuning did not improve AUCPR score further. Similarly, the EMC model achieved an AUCPR of 0,93, which improved to 1,00 after fine-tuning (Figure 15).

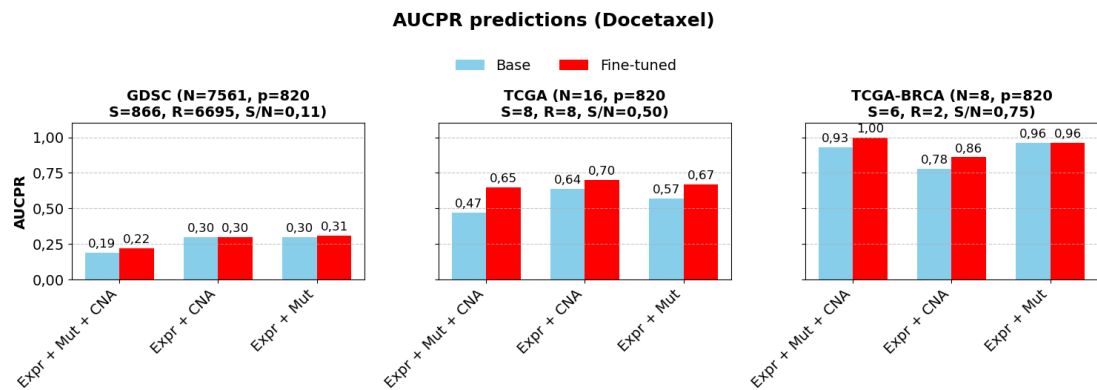


Figure 15. AUCPR predictions with cytoskeletal inhibitors for docetaxel response prediction using the TCGA dataset

4.1.1.2.2. Predictive performance of pan-drug models for cetuximab response prediction

Cetuximab was the only drug targeting the EGFR signaling pathway in the TCGA. Therefore, response predictions were exclusively made for this drug. The strongest model in the training dataset was the EMC model with AUCPR = 0,43. In the TCGA dataset, response predictions were performed for the nine samples. EMC model was the strongest with AUCPR = 0,72 and AUCPR = 0,95 after fine-tuning. The EC model had the same AUCPR score, but fine-tuning improved AUCPR to 0,93. All the nine samples belong to the HNSC cohort. Six of the samples were sensitive and three of the samples were resistant, indicating random prediction with AUCPR = 0,67 (Figure 16).

4.1.1.2.3. Predictive performance of pan-drug models for cisplatin, doxorubicin, 5-FU, gemcitabine, and temozolomide response prediction

All the five drugs belong to DNA replication inhibitors. Separate models were trained for each drug, and the metrics for the best performing models were reported. The training dataset for DNA replication inhibitors included 9435 samples.

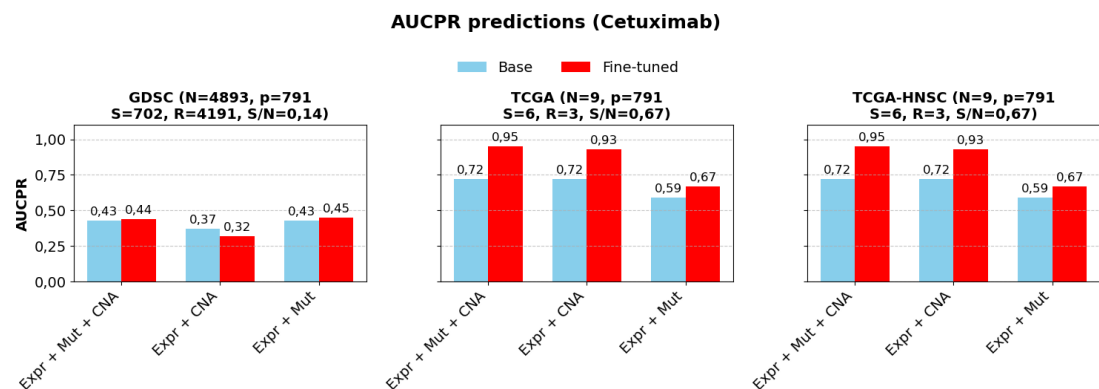


Figure 16. AUCPR predictions with EGFR signaling inhibitors for cetuximab response prediction using the TCGA dataset

In the training dataset with cisplatin, the strongest performance was observed with the EMC model with AUCPR = 0,36. In the test dataset consisting of 64 TCGA samples, both the EC and EMC models demonstrated strong performance, each achieving AUCPR = 0,96. In contrast, the EM model did not outperform random prediction, which had AUCPR = 0,87. Fine-tuning improved AUC scores of each model, but the highest score was achieved with the EC model with AUCPR = 0,98. Within the HNSC cohort, which included 37 samples, the EC model was the strongest, with AUCPR = 0,94 closely followed by the EMC model. Fine-tuning improved prediction for the EC model and yielded AUCPR = 0,97. For the EMC model, prediction also improved with AUCPR = 0,95. However, the EM model once again performed poorly compared to random prediction even after fine-tuning with AUCPR = 0,84 (Figure 17).

In the training dataset with doxorubicin, the best performance was achieved by the EM model with AUCPR = 0,31. However, in the TCGA dataset, the EMC model performed significantly better than all others, achieving AUCPR = 0,67, where random AUCPR = 0,22. Fine-tuning further improved prediction and yielded AUCPR = 1,00. Similarly, the model surpassed random prediction with AUCPR = 0,41 and AUCPR = 0,67 after fine-tuning. In the SARC cohort, all pan-drug models surpassed random predictions even before fine-tuning. The EMC model achieved the highest predictive performance with AUCPR = 0,67 and AUCPR = 1,00 after fine-tuning. The SARC cohort consisted of 15 samples, two of which were sensitive, resulting in a random AUCPR = 0,13. Thus, the performance of the EMC model was notably strong (Figure 18).

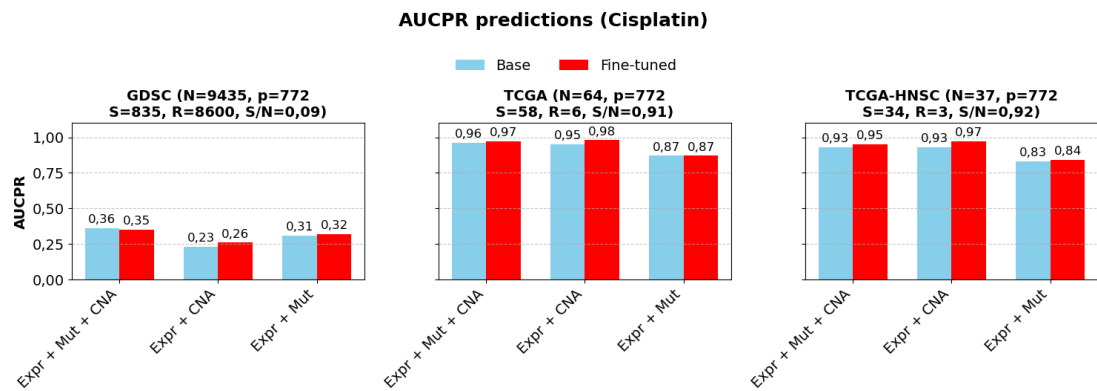


Figure 17. AUCPR predictions with DNA replication inhibitors for cisplatin response prediction using the TCGA dataset

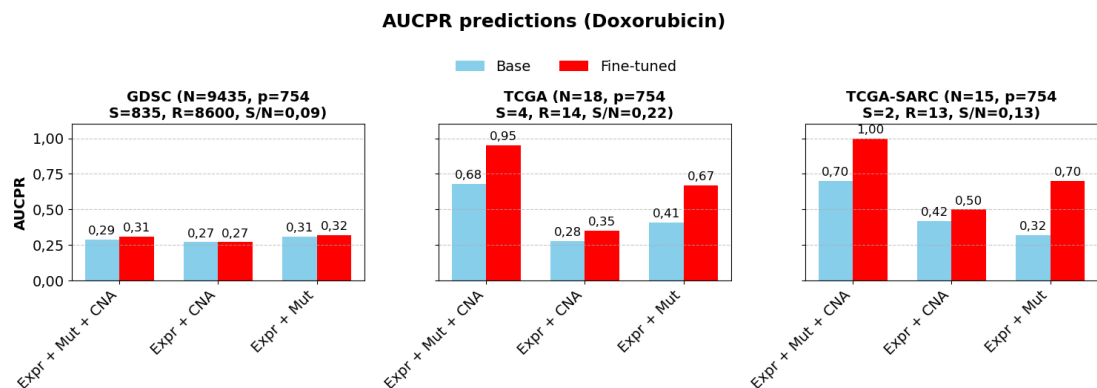


Figure 18. AUCPR predictions with DNA replication inhibitors for doxorubicin response prediction using the TCGA dataset

The strongest model in the training dataset for 5-FU prediction was the EMC model with AUCPR = 0,33. Similarly, in the TCGA dataset, the EMC model achieved the highest performance with AUCPR = 0,81 and AUCPR = 0,84 after fine-tuning. There were 33 samples, nine of which were resistant. Thus, random AUCPR was 0,73. In the STAD cohort, which included 18 samples, the prediction was exceptional, with AUCPR = 0,93 compared to a random prediction of AUCPR = 0,72. Fine-tuning further improved prediction and yielded AUCPR = 0,96. Moreover, in the PAAD cohort, consisting of only six samples with two sensitive samples, the EM model demonstrated the strongest prediction, achieving AUCPR = 0,67 and AUCPR = 0,75 after fine-tuning (Figure 19).

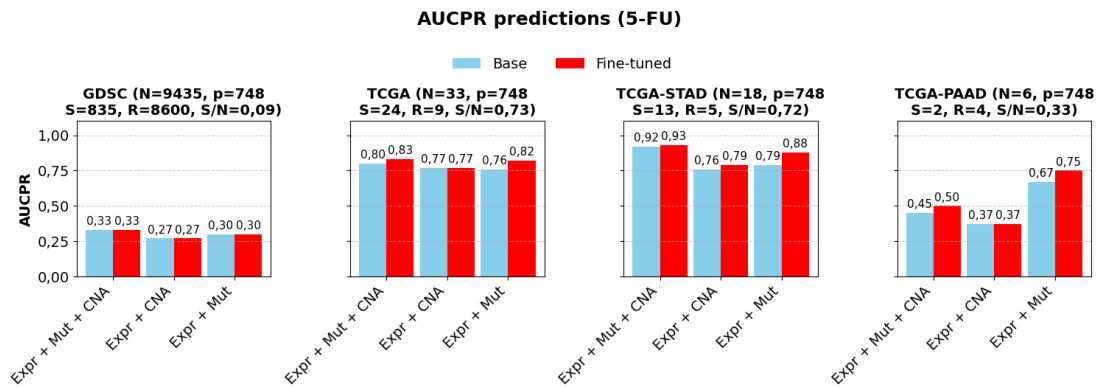


Figure 19. AUCPR predictions with DNA replication inhibitors for 5-FU response prediction using the TCGA dataset

For gemcitabine, the EMC model was also the strongest in the training dataset with AUCPR = 0,36. However, in the TCGA dataset, although the predictive performances were close, the EM model achieved the highest performance with AUCPR = 0,43 and AUCPR = 0,47 after fine-tuning. All the three models had higher AUCPR scores than random prediction (AUCPR = 0,36). Out of 55 samples, 48 belonged to the PAAD cohort. For predictions specifically within the PAAD cohort, the strongest model was the EC model with AUCPR = 0,63, outperforming the random prediction with AUCPR = 0,42. Fine-tuning improved prediction further with AUCPR = 0,69. In addition, the performances of all models exceeded that of random predictions even before fine-tuning (Figure 20).

The final drug among the DNA replication inhibitors for drug response prediction was temozolomide. In the training dataset, all models achieved nearly identical performance around AUCPR = 0,35. In the TCGA dataset, which included 90 samples (11 sensitive and 79 resistant), the EC model achieved the highest predictive performance with AUCPR = 0,22 and AUCPR = 0,25 after fine-tuning. The random AUCPR was 0,12. All models demonstrated stronger prediction performance compared to random prediction even before fine-tuning. All TCGA samples belonged to the lower grade glioma (LGG) cohort. Thus, the same AUCPR was achieved with the models (Figure 21).

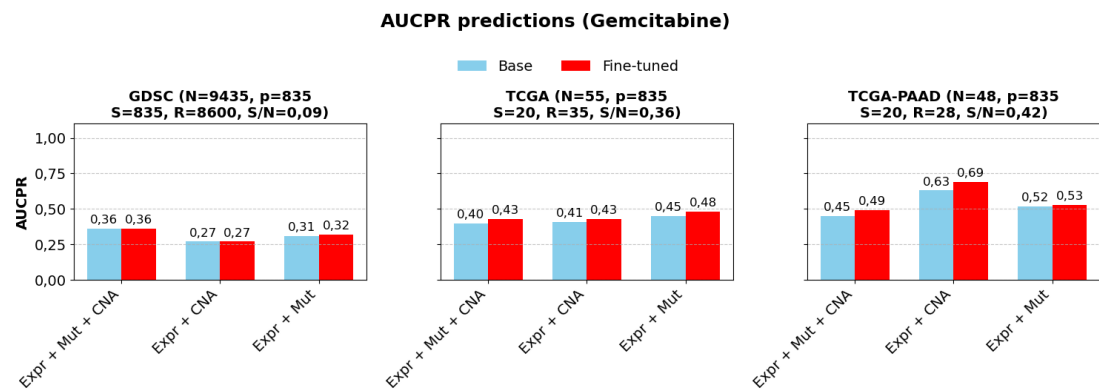


Figure 20. AUCPR predictions with DNA replication inhibitors for gemcitabine response prediction using the TCGA dataset

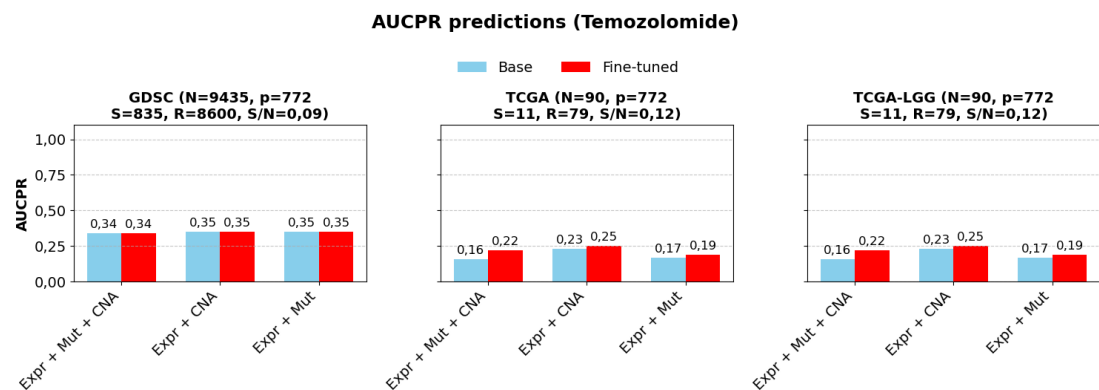


Figure 21. AUCPR predictions with DNA replication inhibitors for temozolomide response prediction using the TCGA dataset

4.1.2. Comparison of predictive power of pan-drug models with MOLI

The predictive performance of multi-omics pan-drug models was evaluated against the MOLI prediction models using the PDX and TCGA datasets as test data. The pan-drug models consistently outperformed the MOLI and other methods in predicting the response to erlotinib, cetuximab, docetaxel, paclitaxel, and cisplatin. Significant improvements in AUCPR were particularly evident for erlotinib, cetuximab, docetaxel, and paclitaxel (Table 1).

Table 1. Comparison of predictive performances with the MOLI results, including Elastic Net, Early Integration with NMF, Early Integration with DNNs, MOLI complete, MOLI pan-drug, and random classifier

Test Datasets	Drug Classes	Drugs	Integration	AUCPR	AUCPR (Fine-tuned)	Improvement	Methods					
							Elastic Net	Early Integration via NMF	Early Integration via DNNs	MOLI complete	MOLI pan-drug	Random classifier
PDX	Mitosis	Paclitaxel	EMC	0,20	0,45	+0,21	0,10	0,21	-	0,24	-	0,12
			EC	0,33	0,56	+0,32						
			EM	0,14	0,22	-0,02						
	Cytoskeleton	Paclitaxel	EMC	0,42	0,42	+0,18	0,10	0,21	-	0,24	-	0,12
			EC	0,18	0,34	+0,10						
			EM	0,14	0,22	-0,02						
	DNA replication	Gemcitabine	EMC	0,47	0,67	+0,18	0,28	0,35	0,35	0,49	-	0,28
			EC	0,47	0,49	0,00						
			EM	0,32	0,40	-0,09						
	EGFR signaling	Cetuximab	EMC	0,15	0,19	-0,01	0,06	0,07	-	0,11	0,20	0,08
			EC	0,15	0,41	+0,21						
			EM	0,11	0,21	+0,01						
	EGFR signaling	Erlotinib	EMC	0,45	0,46	+0,18	0,11	0,28	-	0,33	0,28	0,14
			EC	0,48	0,61	+0,33						
			EM	0,36	0,61	+0,53						
TCGA	Mitosis	Docetaxel	EMC	0,78	0,80	+0,29	0,51	0,51	0,45	0,49	-	0,50
			EC	0,57	0,85	+0,34						
			EM	0,58	0,75	+0,24						
	Cytoskeleton	Docetaxel	EMC	0,50	0,77	+0,26	0,51	0,51	0,45	0,49	-	0,50
			EC	0,67	0,72	+0,21						
			EM	0,57	0,67	+0,16						
	DNA replication	Cisplatin	EMC	0,96	0,97	+0,04	0,85	0,93	-	0,93	-	0,91
			EC	0,96	0,98	+0,05						
			EM	0,87	0,87	-0,06						
	DNA replication	Gemcitabine	EMC	0,40	0,43	-0,02	0,38	0,37	0,46	0,45	-	0,37
			EC	0,41	0,43	-0,02						
			EM	0,43	0,47	+0,02						

For cisplatin, although the MOLI models demonstrated strong performance in the TCGA with high AUCPR scores, the EC model achieved even better prediction (AUCPR = 0,98). While the response prediction for gemcitabine in the both PDX and TCGA datasets was comparable to the MOLI, the pan-drug models still yielded superior AUCPR scores compared to all other methods (122). The most remarkable

improvements were observed for erlotinib and cetuximab. For erlotinib, the EM, EC, and EMC models achieved substantially higher AUCPR scores while for cetuximab, the EC model demonstrated the strongest performance (Table 1).

4.1.3. Predictive performance of pan-drug models in the PDX and TCGA

Drug response predictions for paclitaxel, 5-FU, gemcitabine, and cetuximab, which are shared between both datasets, were conducted robustly by following the four strategies outlined in the methods section.

4.1.3.1. Predictive performance of pan-drug models for paclitaxel response prediction across the four different strategies

Drug response predictions for paclitaxel was performed using mitotic and cytoskeletal inhibitors separately. Response predictions for both were described in separate paragraphs. However, the drug response prediction performances with mitotic inhibitors were evaluated first.

In the prediction of paclitaxel response using all 35 samples from the TCGA dataset and 27 samples from the BRCA cohort, AUCPR scores consistently outperformed random prediction across all the four strategies with the EM, EC, and EMC models. However, the PDX dataset did not follow the same trend. While AUCPR scores consistently exceeded random predictions under the first strategy and second strategy, some models fell below random predictions under the third and fourth strategies. The models that most effectively predicted paclitaxel response using all TCGA samples achieved AUCPR values ranging from 0,86 to 0,92 compared to a random AUCPR = 0,71 although the strongest models varied across strategies. Similarly, in the TCGA BRCA cohort, the strongest models varied across strategies, with AUCPR values ranging from 0,96 to 0,99 compared to random AUCPR = 0,81. In the TCGA dataset, response predictions demonstrated a more stable range while slightly greater variation was observed in the PDX dataset. For all 43 PDX samples, the strongest models varied across strategies, achieving AUCPR

values between 0,27 and 0,46 compared to a random AUCPR = 0,12. In the PDX BRCA cohort, consisting of 24 samples, even greater variation was observed, with AUCPR values ranging from 0,54 to 0,74 compared to a random AUCPR = 0,21 (Table 2).

Table 2. Predictive performance of paclitaxel with mitotic inhibitors across the four different strategies. The bolded values indicate the highest prediction results in the PDX and TCGA datasets

					AUCPR						
Strategies	Drug Classes (Cohorts)	Size (Drug Classes)	Size (Drugs)	Integration	GDSC	GDSC (PDX, Fine-tuned)	GDSC (TCGA, Fine-tuned)	PDX	PDX (Fine-tuned)	TCGA	TCGA (Fine-tuned)
STRATEGY 1 (All cancer types included & Paclitaxel included)	MITOSISi	6360	(PDX:43, TCGA:35)	EMC	0,43	0,41	0,43	0,23	0,42	0,84	0,90
				EC	0,31	0,31	0,31	0,15	0,14	0,84	0,84
				EM	0,37	0,37	0,36	0,21	0,22	0,74	0,74
	MITOSISi (BRCA)		(PDX:24, TCGA:27)	EMC	0,43	0,40	0,41	0,60	0,74	0,97	0,99
				EC	0,31	0,31	0,33	0,22	0,21	0,94	0,95
				EM	0,37	0,37	0,37	0,27	0,32	0,85	0,85
STRATEGY 2 (All cancer types included & Paclitaxel excluded)	MITOSISi	5990	(PDX:43, TCGA:35)	EMC	0,31	0,35	0,40	0,21	0,25	0,82	0,84
				EC	0,44	0,45	0,45	0,25	0,39	0,82	0,84
				EM	0,45	0,46	0,46	0,21	0,24	0,86	0,86
	MITOSISi (BRCA)		(PDX:24, TCGA:27)	EMC	0,31	0,35	0,40	0,34	0,54	0,87	0,92
				EC	0,44	0,45	0,44	0,38	0,50	0,95	0,96
				EM	0,45	0,46	0,43	0,33	0,42	0,93	0,95
STRATEGY 3 (Common cancer types excluded & Paclitaxel included)	MITOSISi	5086	(PDX:43, TCGA:35)	EMC	0,38	0,40	0,39	0,11	0,14	0,84	0,89
				EC	0,41	0,24	0,24	0,16	0,27	0,73	0,76
				EM	0,41	0,38	0,41	0,10	0,12	0,83	0,86
	MITOSISi (BRCA)		(PDX:24, TCGA:27)	EMC	0,38	0,41	0,40	0,20	0,22	0,96	0,97
				EC	0,41	0,24	0,42	0,28	0,64	0,85	0,87
				EM	0,41	0,41	0,39	0,17	0,24	0,87	0,89
STRATEGY 4 (Common cancer types excluded & Paclitaxel excluded)	MITOSISi	4758	(PDX:43, TCGA:35)	EMC	0,48	0,47	0,48	0,10	0,12	0,89	0,92
				EC	0,36	0,40	0,35	0,37	0,46	0,76	0,89
				EM	0,48	0,48	0,48	0,10	0,10	0,76	0,77
	MITOSISi (BRCA)		(PDX:24, TCGA:27)	EMC	0,48	0,47	0,47	0,18	0,21	0,96	0,96
				EC	0,36	0,37	0,35	0,61	0,67	0,89	0,96
				EM	0,48	0,48	0,48	0,17	0,20	0,84	0,85

Unlike the trend observed with mitotic models, paclitaxel response predictions using cytoskeletal models fell sometimes below random predictions in the both PDX and TCGA datasets across all strategies. However, for each strategy, there were models that outperformed the random AUCPR. The AUCPR predictions across strategies using all TCGA samples showed slight variation, ranging from 0,82 to 0,87. In the TCGA BRCA cohort, AUCPR scores changed between 0,88 and 0,94 across strategies. Similarly, in the PDX dataset, the variation across all samples was evident in AUCPR scores ranging from 0,23 to 0,41, whereas the PDX BRCA cohort

exhibited greater variation, with AUCPR scores ranging from 0,36 to 0,61 (Table 3).

Table 3. Predictive performance of paclitaxel with cytoskeletal inhibitors across the four different strategies. The bolded values indicate the highest prediction results in the PDX and TCGA datasets

					AUCPR						
Strategies	Drug Classes (Cohorts)	Size (Drug Classes)	Size (Drugs)	Integration	GDSC	GDSC (PDX, Fine-tuned)	GDSC (TCGA, Fine-tuned)	PDX	PDX (Fine-tuned)	TCGA	TCGA (Fine-tuned)
STRATEGY 1 (All cancer types included & Paclitaxel included)	CYTOi	7561	(PDX:43, TCGA:35)	EMC	0,25	0,29	0,23	0,24	0,40	0,79	0,87
				EC	0,25	0,26	0,25	0,29	0,41	0,79	0,79
				EM	0,28	0,30	0,29	0,11	0,13	0,68	0,82
	CYTOi (BRCA)		(PDX:24, TCGA:27)	EMC	0,25	0,29	0,23	0,42	0,54	0,93	0,94
				EC	0,25	0,26	0,25	0,36	0,46	0,85	0,86
				EM	0,28	0,30	0,31	0,21	0,30	0,83	0,90
STRATEGY 2 (All cancer types included & Paclitaxel excluded)	CYTOi	7191	(PDX:43, TCGA:35)	EMC	0,32	0,33	0,33	0,10	0,16	0,76	0,77
				EC	0,32	0,33	0,32	0,29	0,34	0,78	0,82
				EM	0,33	0,33	0,33	0,12	0,13	0,68	0,76
	CYTOi (BRCA)		(PDX:24, TCGA:27)	EMC	0,32	0,33	0,33	0,19	0,25	0,86	0,88
				EC	0,32	0,32	0,32	0,41	0,61	0,86	0,92
				EM	0,33	0,33	0,34	0,28	0,32	0,87	0,88
STRATEGY 3 (Common cancer types excluded & Paclitaxel included)	CYTOi	5853	(PDX:43, TCGA:35)	EMC	0,30	0,30	0,33	0,12	0,23	0,82	0,83
				EC	0,30	0,31	0,26	0,15	0,21	0,79	0,86
				EM	0,31	0,32	0,25	0,09	0,10	0,63	0,76
	CYTOi (BRCA)		(PDX:24, TCGA:27)	EMC	0,30	0,30	0,32	0,17	0,29	0,87	0,88
				EC	0,30	0,26	0,26	0,23	0,42	0,85	0,87
				EM	0,31	0,32	0,25	0,19	0,22	0,80	0,86
STRATEGY 4 (Common cancer types excluded & Paclitaxel excluded)	CYTOi	5525	(PDX:43, TCGA:35)	EMC	0,32	0,33	0,30	0,13	0,19	0,71	0,81
				EC	0,32	0,33	0,30	0,14	0,30	0,75	0,85
				EM	0,31	0,22	0,29	0,11	0,21	0,73	0,82
	CYTOi (BRCA)		(PDX:24, TCGA:27)	EMC	0,32	0,33	0,29	0,20	0,25	0,83	0,89
				EC	0,32	0,27	0,32	0,24	0,36	0,86	0,93
				EM	0,31	0,33	0,29	0,19	0,30	0,83	0,92

An important observation emerged in the response predictions. Between the first and second strategies, both the PDX and TCGA demonstrated nearly identical performance with mitotic and cytoskeletal models across all samples. However, a decline in AUCPR was observed in the PDX BRCA cohort with mitotic inhibitors and in the TCGA BRCA cohort with cytoskeletal inhibitors. While the decrease in the TCGA was modest, the decline in the PDX was considerably more profound.

4.1.3.2. Predictive performance of pan-drug models for 5-FU response prediction across the four different strategies

For 5-FU, differences were present in the cohorts of samples between the PDX and TCGA datasets. In the PDX dataset, all 23 samples were from the CRC cohort,

with only one sample classified as sensitive. In the TCGA dataset, while samples were derived from various cohorts, response predictions were performed specifically for the STAD cohort of 18 samples and the PAAD cohort of six samples. Additionally, the READ cohort comprised seven samples; however, since all of them were sensitive, response prediction was not conducted for this cohort (Figure 4G).

All AUCPR values across the four strategies in the PDX dataset were higher than the random AUCPR. However, the AUCPR values remained low due to the presence of only one sensitive sample. In the TCGA dataset, while all models for all samples exceeded the random AUCPR values across the strategies, certain models fell below random predictions in the PAAD cohort under second strategy and in the STAD cohort under fourth strategy. Despite this, each strategy consistently included at least one strong-performing model (Table 4).

Across the four strategies, the AUCPR for predictions with PDX samples ranged from 0,08 to 0,17, while the random AUCPR was 0,04. In contrast, predictions with TCGA samples showed an AUCPR range from 0,82 to 0,93. The random AUCPR was 0,73. The highest AUCPR = 0,93 was achieved under the third strategy. Similarly, the AUCPR ranged from 0,82 to 0,95 in the STAD cohort and the third strategy yielded the highest AUCPR = 0,95. In both cases, these predictions were made with the EM model. Finally, in the TCGA PAAD cohort, the AUCPR ranged from 0,75 to 1,00, which showed significant variation due to the small sample size. The EM model again achieved an excellent AUCPR in this cohort (Table 4).

4.1.3.3. Predictive performance of pan-drug models for gemcitabine response prediction across the four different strategies

For gemcitabine response prediction, the PDX dataset included only 28 samples, all from the PDAC cohort (Figure 4). On the other hand, the TCGA dataset comprised samples from various cohorts, with 48 of the 55 samples belonging to the PAAD cohort (Figure 4). Only under the third strategy with the TCGA some models

performed below the random AUCPR. Similarly, some models fell below random AUCPR in both the first and third strategies in the PDX dataset. (Table 5).

Table 4. Predictive performance of 5-FU with DNA replication inhibitors across the four different strategies. The bolded values indicate the highest prediction results in the PDX and TCGA datasets

					AUCPR						
Strategies	Drug Classes (Cohorts)	Size (Drug Classes)	Size (Drugs)	Integration	GDSC	GDSC (PDX, Fine-tuned)	GDSC (TCGA, Fine-tuned)	PDX	PDX (Fine-tuned)	TCGA	TCGA (Fine-tuned)
STRATEGY 1 (All cancer types included & 5-FU included)	DNA_REPi	9435	(PDX:23, TCGA:33)	EMC	0,29	0,29	0,30	0,14	0,17	0,81	0,83
				EC	0,21	0,27	0,27	0,08	0,14	0,76	0,82
				EM	0,31	0,31	0,31	0,07	0,08	0,83	0,84
	DNA_REPi (CRC/STAD)		(PDX:23, TCGA:18)	EMC	0,29	0,29	0,32	0,14	0,17	0,86	0,88
				EC	0,21	0,27	0,28	0,08	0,14	0,73	0,78
				EM	0,31	0,31	0,31	0,07	0,08	0,85	0,89
	DNA_REPi (CRC/PAAD)		(PDX:23, TCGA:6)	EMC	0,29	0,29	0,30	0,14	0,17	0,70	0,70
				EC	0,21	0,27	0,20	0,08	0,14	0,50	1,00
				EM	0,31	0,31	0,31	0,07	0,08	0,67	0,75
STRATEGY 2 (All cancer types included & 5-FU excluded)	DNA_REPi	8671	(PDX:23, TCGA:33)	EMC	0,34	0,34	0,33	0,05	0,05	0,74	0,78
				EC	0,33	0,33	0,33	0,05	0,05	0,75	0,75
				EM	0,33	0,33	0,33	0,05	0,07	0,82	0,85
	DNA_REPi (CRC/STAD)		(PDX:23, TCGA:18)	EMC	0,34	0,34	0,32	0,05	0,05	0,89	0,90
				EC	0,33	0,33	0,33	0,05	0,05	0,81	0,81
				EM	0,33	0,33	0,33	0,05	0,07	0,91	0,92
	DNA_REPi (CRC/PAAD)		(PDX:23, TCGA:6)	EMC	0,34	0,34	0,32	0,05	0,05	0,32	0,45
				EC	0,33	0,33	0,33	0,05	0,05	0,27	0,27
				EM	0,33	0,33	0,33	0,05	0,07	0,45	0,75
STRATEGY 3 (Common cancer types excluded & 5-FU included)	DNA_REPi	8202	(PDX:23, TCGA:33)	EMC	0,35	0,34	0,35	0,07	0,10	0,79	0,81
				EC	0,33	0,33	0,33	0,08	0,09	0,87	0,88
				EM	0,33	0,34	0,33	0,05	0,10	0,93	0,93
	DNA_REPi (CRC/STAD)		(PDX:23, TCGA:18)	EMC	0,35	0,34	0,35	0,07	0,10	0,90	0,92
				EC	0,33	0,33	0,33	0,08	0,09	0,87	0,90
				EM	0,33	0,34	0,33	0,05	0,10	0,95	0,95
	DNA_REPi (CRC/PAAD)		(PDX:23, TCGA:6)	EMC	0,35	0,34	0,35	0,07	0,10	0,42	0,45
				EC	0,33	0,33	0,33	0,08	0,09	0,45	0,70
				EM	0,33	0,34	0,34	0,05	0,10	1,00	1,00
STRATEGY 4 (Common cancer types excluded & 5-FU excluded)	DNA_REPi	7540	(PDX:23, TCGA:33)	EMC	0,34	0,33	0,34	0,05	0,05	0,74	0,78
				EC	0,33	0,33	0,33	0,07	0,08	0,79	0,82
				EM	0,34	0,34	0,34	0,06	0,07	0,77	0,78
	DNA_REPi (CRC/STAD)		(PDX:23, TCGA:18)	EMC	0,34	0,33	0,34	0,05	0,05	0,71	0,74
				EC	0,33	0,33	0,33	0,07	0,08	0,79	0,82
				EM	0,34	0,34	0,34	0,06	0,07	0,81	0,82
	DNA_REPi (CRC/PAAD)		(PDX:23, TCGA:6)	EMC	0,34	0,33	0,33	0,05	0,05	0,83	0,83
				EC	0,33	0,33	0,33	0,07	0,08	0,58	0,83
				EM	0,34	0,34	0,34	0,06	0,07	0,58	0,58

AUCPR values for all TCGA samples ranged from 0,45 to 0,50, with a random AUCPR = 0,36. Variations were observed in the strongest performing models across strategies. In the TCGA PAAD cohort, the AUCPR ranged from 0,52 to 0,65, where the random AUCPR was 0,42. Notably, the EC model demonstrated the highest performance across all strategies with the exception of the third strategy in the PAAD

cohort. In the PDX dataset, the AUCPR ranged from 0,40 to 0,58 with a random AUCPR = 0,28. The EC model was the strongest under the first and second strategies, while the EM model emerged as the strongest under the third strategy. Nevertheless, all the three models showed the same predictive power with AUCPR = 0,40 under the fourth strategy (Table 5).

Table 5. Predictive performance of gemcitabine with DNA replication inhibitors across the four different strategies. The bolded values indicate the highest prediction results in the PDX and TCGA datasets

					AUCPR						
Strategies	Drug Classes (Cohorts)	Size (Drug Classes)	Size (Drugs)	Integration	GDSC	GDSC (PDX, Fine-tuned)	GDSC (TCGA, Fine-tuned)	PDX	PDX (Fine-tuned)	TCGA	TCGA (Fine-tuned)
STRATEGY 1 (All cancer types included & Gemcitabine included)	DNA_REPi	9435	(PDX:25, TCGA:55)	EMC	0,34	0,34	0,33	0,50	0,57	0,40	0,42
				EC	0,32	0,33	0,33	0,23	0,26	0,41	0,43
				EM	0,33	0,34	0,34	0,29	0,35	0,44	0,46
	DNA_REPi (PDAC/PAAD)		(PDX:25, TCGA:48)	EMC	0,34	0,34	0,34	0,50	0,57	0,45	0,50
				EC	0,32	0,33	0,33	0,23	0,26	0,56	0,62
				EM	0,33	0,34	0,34	0,29	0,35	0,48	0,49
STRATEGY 2 (All cancer types included & Gemcitabine excluded)	DNA_REPi	8674	(PDX:25, TCGA:55)	EMC	0,37	0,36	0,36	0,43	0,58	0,47	0,43
				EC	0,34	0,34	0,34	0,48	0,52	0,48	0,50
				EM	0,35	0,36	0,36	0,43	0,50	0,44	0,48
	DNA_REPi (PDAC/PAAD)		(PDX:25, TCGA:48)	EMC	0,37	0,36	0,36	0,43	0,58	0,47	0,51
				EC	0,34	0,34	0,34	0,48	0,52	0,59	0,64
				EM	0,35	0,36	0,36	0,43	0,50	0,48	0,54
STRATEGY 3 (Common cancer types excluded & Gemcitabine included)	DNA_REPi	8843	(PDX:25, TCGA:55)	EMC	0,28	0,33	0,29	0,34	0,40	0,44	0,47
				EC	0,32	0,32	0,32	0,33	0,36	0,42	0,44
				EM	0,32	0,34	0,34	0,32	0,57	0,43	0,40
	DNA_REPi (PDAC/PAAD)		(PDX:25, TCGA:48)	EMC	0,28	0,33	0,30	0,34	0,40	0,47	0,52
				EC	0,32	0,32	0,32	0,33	0,36	0,48	0,51
				EM	0,32	0,34	0,33	0,32	0,57	0,49	0,44
STRATEGY 4 (Common cancer types excluded & Gemcitabine excluded)	DNA_REPi	8132	(PDX:25, TCGA:55)	EMC	0,32	0,32	0,30	0,22	0,40	0,39	0,44
				EC	0,34	0,33	0,34	0,35	0,40	0,41	0,43
				EM	0,33	0,33	0,33	0,32	0,40	0,42	0,45
	DNA_REPi (PDAC/PAAD)		(PDX:25, TCGA:48)	EMC	0,32	0,32	0,30	0,22	0,40	0,46	0,52
				EC	0,34	0,33	0,34	0,35	0,40	0,62	0,65
				EM	0,33	0,33	0,33	0,32	0,40	0,49	0,53

When the first and second strategies were compared, the AUCPR values of the strongest models were similar. This indicated that gemcitabine response could be robustly predicted even without response data included in the training dataset. Furthermore, the AUCPR value of the best performing model for predicting PDX samples remained almost stable under the third strategy. Similarly, there was slight decrease in AUCPR score under the third strategy for all TCGA samples. These results highlighted robust prediction in both the PDX PDAC cohort and all TCGA

samples regardless of the inclusion of common cancer types. However, a decrease in AUCPR was observed in the TCGA PAAD cohort. This suggested that the removal of the common cancer type from the training dataset resulted in the loss of some cohort-specific features. The fourth strategy, however, showed an opposite trend. The AUCPR of the best performing model for the PDX PDAC cohort decreased, whereas the TCGA PAAD cohort achieved the highest AUCPR with the best model. For all TCGA samples, there was little variation under the fourth strategy. This suggested that removing common cancer types and gemcitabine response data from the training dataset, which could introduce bias and noise, improved predictive performance specifically in the TCGA PAAD cohort (Table 5).

4.1.3.4. Predictive performance of pan-drug models for cetuximab response prediction across the four different strategies

Different cohorts were utilized in PDX and TCGA for predicting cetuximab response. The PDX dataset included samples from the CRC and NSCLC cohorts (Figure 5). On the other hand, the TCGA dataset consisted exclusively of samples from the HNSC cohort (Figure 5).

In both PDX and TCGA samples, AUCPR values were generally higher than the random AUCPR scores across all strategies (Table 6). In the TCGA HNSC cohort samples, AUCPR values ranged from 0,86 to 1,00, whereas the random AUCPR was 0,67. The small sample size of only nine samples likely contributed to the observed variation. The first and third strategies yielded the highest prediction with the EM model with AUCPR = 1,00. Similarly, the AUCPR showed variations between 0,19 and 0,44 for all PDX samples, and the strongest models varied across strategies. The random AUCPR in this case was 0,08. The EC model outperformed other models across all strategies except the first strategy in the PDX CRC cohort. The AUCPR values ranged from 0,28 to 0,42, while random AUCPR was 0,11. Lastly, AUCPR values ranged between 0,50 and 1,00 in the PDX NSCLC cohort. However, this can not be considered a reliable evaluation due to the presence of only one sensitive

sample. The random AUCPR was 0,04, while high AUCPR values across strategies were achieved with different models (Table 6).

In cetuximab response prediction across four different strategies, no significant decrease in predictive performance was observed for the TCGA HNSC samples compared to the first strategy. However, the PDX NSCLC and CRC cohorts showed variable performance across strategies (Table 6). The highest prediction was with the EMC model for the PDX CRC cohort samples under the first strategy with AUCPR = 0,42. However, a notable decrease in AUCPR was observed under the second and third strategies for the PDX CRC cohort. The fourth strategy demonstrated slight decrease in AUCPR compared to the first strategy with AUCPR = 0,37. This slight decrease was likely due to the removal of bias and noise resulting from both cetuximab response data and common cancer types in the training dataset.

4.2. Drug Resistance Mechanisms with Pathway Enrichment Analysis

For drugs shared between PDX and TCGA, enrichment analyses were conducted using a pathway-level, gene expression-based active subnetwork approach with pathfindR tool (11). The genes from the most powerful models in TCGA predictions across strategies were used as inputs for the enrichment analyses since TCGA captures the most clinically relevant features.

4.2.1. Resistance mechanisms of paclitaxel

Pathway enrichment analysis under the first strategy identified 79 common enriched terms between enrichment results obtained with the gene features of the best performing mitotic and cytoskeletal models. Spearman rank correlation test revealed a strong correlation in the ranking of the enriched terms ($\rho = 0,93$, $p\text{-value} = 6,92 \times 10^{-33}$). “Focal adhesion”, “ECM-receptor interaction”, and “p53 signaling pathway” were the top three shared enriched terms (Table 7 in Appendix 1).

Table 6. Predictive performance of cetuximab with EGFR signaling inhibitors across the four different strategies. The bolded values indicate the highest prediction results in the PDX and TCGA datasets

					AUCPR						
Strategies	Drug Classes (Cohorts)	Size (Drug Classes)	Size (Drugs)	Integration	GDSC	GDSC (PDX, Fine-tuned)	GDSC (TCGA, Fine-tuned)	PDX	PDX (Fine-tuned)	TCGA	TCGA (Fine-tuned)
STRATEGY 1 (All cancer types included & Cetuximab included)	EGFRi	5633	(PDX:60, TCGA:9)	EMC	0,43	0,41	0,39	0,09	0,44	0,61	0,86
				EC	0,48	0,46	0,48	0,20	0,28	0,73	0,87
				EM	0,50	0,33	0,33	0,09	0,13	0,76	1,00
	EGFRi (CRC/HNSC)		(PDX:36, TCGA:9)	EMC	0,43	0,45	0,39	0,13	0,42	0,61	0,86
				EC	0,48	0,48	0,48	0,31	0,38	0,73	0,87
				EM	0,50	0,33	0,33	0,12	0,19	0,76	1,00
	EGFRi (NSCLC/HNSC)		(PDX:24, TCGA:9)	EMC	0,43	0,43	0,39	0,08	1,00	0,61	0,86
				EC	0,48	0,47	0,48	0,50	1,00	0,73	0,87
				EM	0,50	0,49	0,33	0,09	0,25	0,76	1,00
STRATEGY 2 (All cancer types included & Cetuximab excluded)	EGFRi	4908	(PDX:60, TCGA:9)	EMC	0,47	0,48	0,47	0,09	0,10	0,80	0,82
				EC	0,47	0,47	0,47	0,10	0,11	0,78	0,80
				EM	0,46	0,46	0,47	0,16	0,19	0,84	0,94
	EGFRi (CRC/HNSC)		(PDX:36, TCGA:9)	EMC	0,47	0,46	0,47	0,17	0,19	0,80	0,82
				EC	0,47	0,47	0,47	0,20	0,30	0,78	0,80
				EM	0,46	0,47	0,47	0,17	0,18	0,84	0,94
	EGFRi (NSCLC/HNSC)		(PDX:24, TCGA:9)	EMC	0,47	0,47	0,47	0,06	0,14	0,80	0,82
				EC	0,47	0,47	0,47	0,07	0,08	0,78	0,80
				EM	0,46	0,46	0,47	0,50	0,50	0,84	0,94
STRATEGY 3 (Common cancer types excluded & Cetuximab included)	EGFRi	4352	(PDX:60, TCGA:9)	EMC	0,40	0,40	0,41	0,13	0,17	0,62	0,91
				EC	0,40	0,39	0,38	0,16	0,28	0,79	0,92
				EM	0,38	0,38	0,37	0,12	0,14	0,98	1,00
	EGFRi (CRC/HNSC)		(PDX:36, TCGA:9)	EMC	0,40	0,41	0,41	0,10	0,14	0,62	0,91
				EC	0,40	0,40	0,38	0,19	0,28	0,79	0,92
				EM	0,38	0,38	0,37	0,12	0,15	0,98	1,00
	EGFRi (NSCLC/HNSC)		(PDX:24, TCGA:9)	EMC	0,40	0,40	0,41	0,50	1,00	0,62	0,91
				EC	0,40	0,39	0,38	0,50	1,00	0,79	0,92
				EM	0,38	0,36	0,37	0,17	0,50	0,98	1,00
STRATEGY 4 (Common cancer types excluded & Cetuximab excluded)	EGFRi	3803	(PDX:60, TCGA:9)	EMC	0,41	0,41	0,41	0,18	0,20	0,79	0,84
				EC	0,39	0,40	0,40	0,18	0,20	0,62	0,76
				EM	0,40	0,40	0,38	0,14	0,14	0,87	0,86
	EGFRi (CRC/HNSC)		(PDX:36, TCGA:9)	EMC	0,41	0,39	0,41	0,23	0,3	0,79	0,84
				EC	0,39	0,40	0,40	0,37	0,37	0,62	0,76
				EM	0,40	0,39	0,38	0,16	0,16	0,87	0,86
	EGFRi (NSCLC/HNSC)		(PDX:24, TCGA:9)	EMC	0,41	0,41	0,41	1,00	1,00	0,79	0,84
				EC	0,39	0,40	0,40	0,17	0,17	0,62	0,76
				EM	0,40	0,39	0,38	0,25	0,50	0,87	0,86

In addition to the three top shared enriched terms, many cancer related signaling pathways as well as terms related to cell-cell interaction, cellular adhesion, cancer, metabolism, cytoskeleton, immune response, and TME were enriched. All these enriched terms obtained with the enrichment analyses of mitotic and cytoskeletal inhibitor models are given in Tables 8 and 10 in Appendix 2. The enrichment results of the models with high performance in the TCGA under the second strategy are also provided in Tables 9 and 11 in Appendix 2.

The top 25 of the enriched terms were specifically selected to plot fold enrichment graphs. Figure 22 and Figure 23 present the fold enrichment results for these terms for mitotic and cytoskeletal inhibitor models used under the first strategy, respectively. A comparison of the two figures reveals that the enriched terms are indeed shared and ranked similarly.

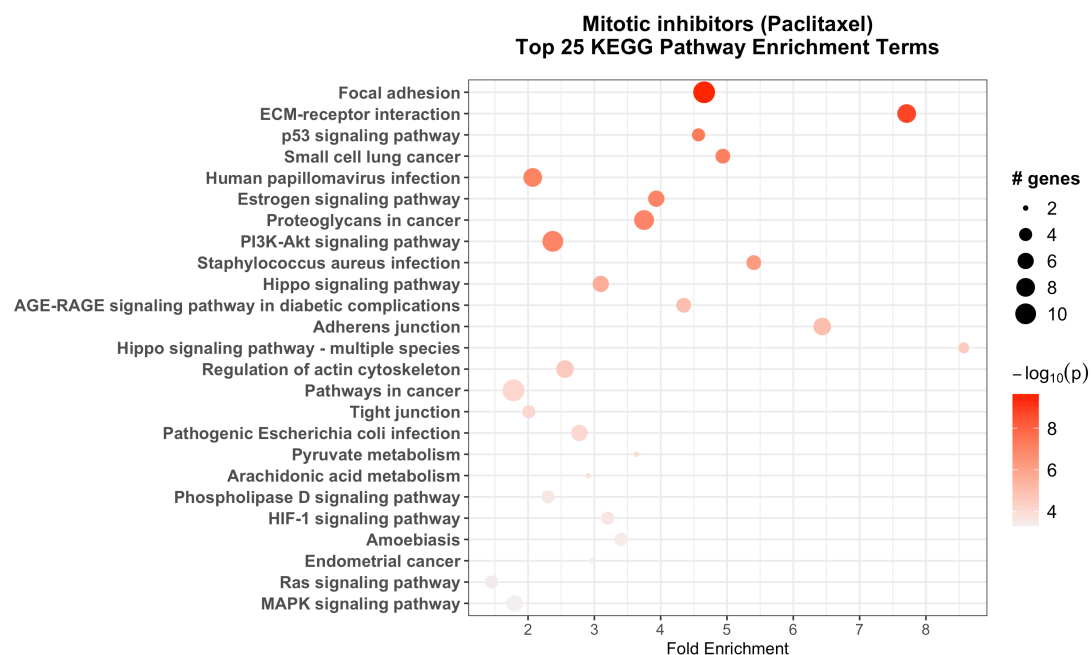


Figure 22. Fold enrichment plot showing the top 25 KEGG pathways for paclitaxel based on the gene features of the best performing mitotic inhibitor model

Term-gene heatmap plots displaying upregulated and downregulated genes within the top 25 enriched terms are presented in Figure 24 and Figure 25, respectively. These graphs provide insights into active and suppressed pathways. They also highlight genes that contribute to multiple pathways. For instance, TCF7L2 was upregulated in seven of the top 25 enriched terms.

Term-gene plots for the top 25 terms were generated based on the enrichment analyses using features from the models of the two inhibitor groups (Figures 26-27). These plots illustrate the interactions between the enriched terms and their associated genes. Notably, only a few genes were upregulated in these plots, including FYN,

LDHB, TCF7L2, CXCR4, and CYFIP2. In addition, these genes had multiple connections to other pathways since they were involved in multiple enriched terms.

When evaluating paclitaxel resistance through signaling pathways, it becomes evident that most pathways were suppressed while compensatory signaling pathways were activated. PI3K/Akt, MAPK, NF- κ B, JAK/STAT, Ras, and FoxO signaling pathways, which are most commonly associated with cancer related survival pathways, were all suppressed (138). However, alternative survival pathways such as Wnt and Hippo were active only with TCF7L2 upregulation (Figures 24-25, Tables 8-11 in Appendix 2). Dysregulation of the Wnt and Hippo signaling pathways is closely linked to drug resistance in cancer (139). Moreover, “Focal adhesion”, the most significant term and highly associated with resistance in cancer with cytoskeletal reorganization, ECM remodeling, EMT, cell migration and metastasis, was only active with FYN upregulation (140). “Adherens junction”, which was active only by upregulation of FYN and TCF7L2, was also significantly enriched (Figures 24-25). This term is regarded as critical in drug resistance in cancer due to its role in cytoskeletal reorganization, EMT, Wnt/ β -catenin signaling pathway, microenvironment, tumor heterogeneity, stem-like properties, and cancer cell survival (141).

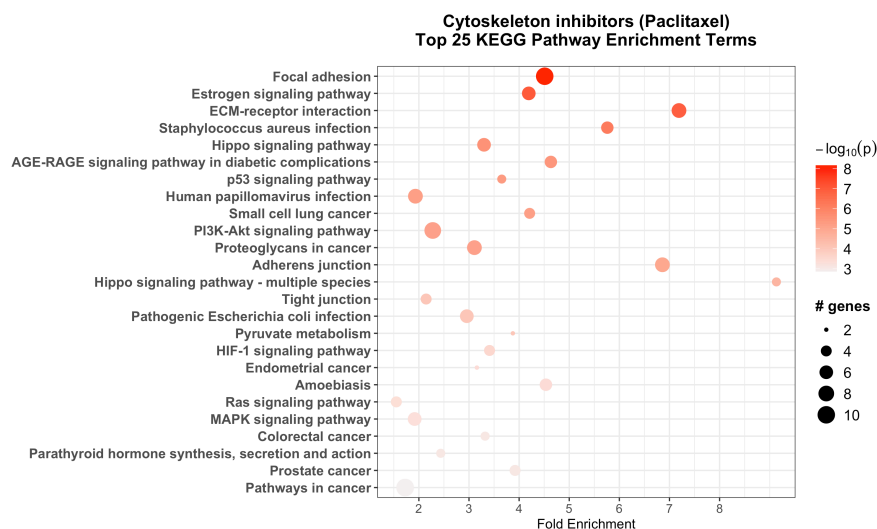


Figure 23. Fold enrichment plot showing the top 25 KEGG pathways for paclitaxel based on the gene features of the best performing cytoskeletal inhibitor model

There were various cancer-related terms such as endometrial, colorectal and breast cancer, in which TCF7L2 was exclusively upregulated (Figures 24-25, Tables 8-11 in Appendix 2). It is a key TF in the Wnt/ β -catenin pathway. In addition to activating the Wnt/ β -catenin pathway, it might contribute to drug resistance through activities such as cancer stem cell maintenance, metabolic adaptation, and cross-talk with other pathways (138).

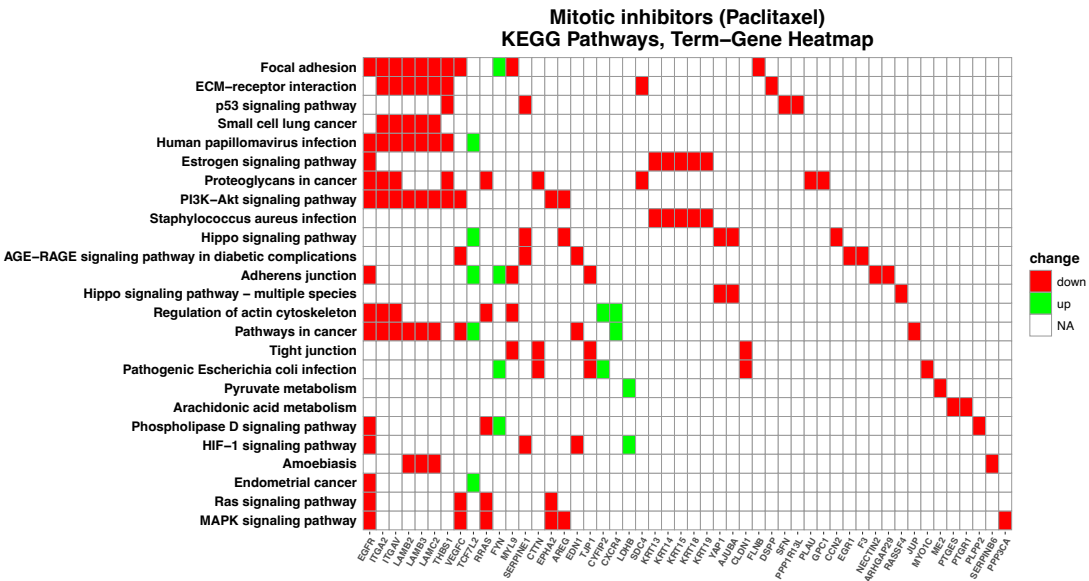


Figure 24. Fold enrichment plot showing the top 25 KEGG pathways for paclitaxel based on the gene features of the best performing mitotic inhibitor model

“Regulation of actin cytoskeleton”, which is related to the cytoskeleton that paclitaxel specifically targets by binding to microtubules, was active only with CYFIP2 and CXCR4 upregulation (Figures 24-25) (142). CYFIP2 is associated with actin remodeling, while CXCR4 is associated with cell migration, metastasis, cancer cell survival, and drug resistance in cancer (143, 144). CXCR4, along with FYN, was also upregulated in the “Axon guidance” (Tables 8-11 in Appendix 2). The relevance of this term to drug resistance in cancer is critical. It has been shown to contribute to the development of drug resistance through cytoskeletal remodeling, anti-apoptotic signaling, TME interactions, cancer stemness and plasticity, and drug efflux (145, 146). Furthermore, CXCR4 alone was upregulated in the “Calcium signaling pathway”, which has been found to contribute to drug resistance in cancer through

evasion of apoptosis, enhanced cellular repair, drug efflux, EMT, and modulation of the TME (Tables 8-11 in Appendix 2) (147).

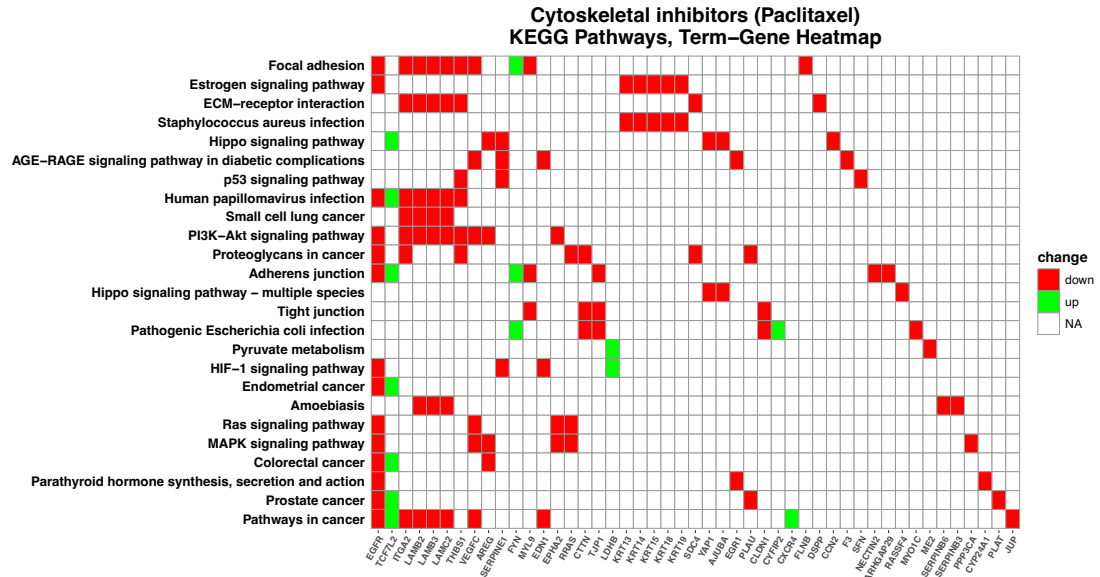


Figure 25. Term-gene heatmap plot for paclitaxel based on the gene features of the best performing cytoskeletal inhibitor model

“Transcriptional misregulation in cancer” was inactive, which suggests that resistance mechanisms of paclitaxel develop from mechanisms other than transcriptional dysregulation. Furthermore, “Tight junction”, associated with epithelial integrity, cancer cell survival, and drug efflux, and “Gap junction”, associated with apoptosis, pro-survival microenvironment, and tumor cell heterogeneity, were inactive (Figures 24-25, Tables 8-11 in Appendix 2) (148, 149). TJP1 was downregulated in both. It was also downregulated in the “Adherens junction” (Figures 24-25, Tables 8-11 in Appendix 2). Down-regulation of TJP1 has been found in the literature to correlate with EMT-like features. Its downregulation is also associated with resistance in cancer due to its link to reduced barrier function, altered polarity signaling, and drug efflux (150).

“Phagosome” was another inactive term (Tables 8-11 in Appendix 2). Down-regulation of the phagosome has critical roles in drug resistance as it increases

immune evasion and survival, metabolic flexibility, drug inactivation, inhibition of drug delivery, and efflux pumps (151).

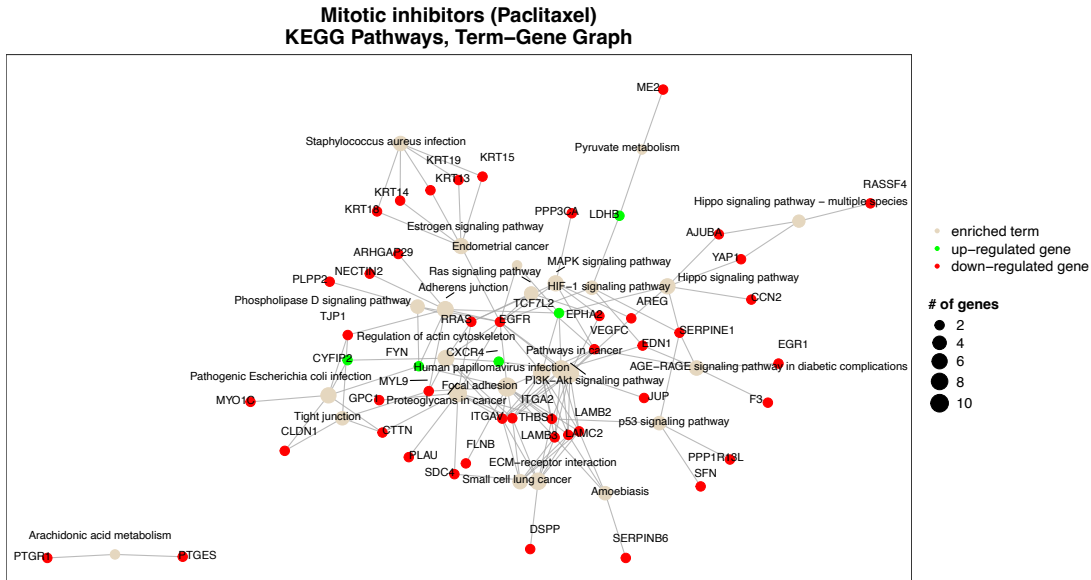


Figure 26. Term-gene graph illustrating the interactions between enriched KEGG terms and their associated genes for paclitaxel based on enrichment results from the best performing mitotic inhibitor model

“Chemical carcinogenesis - receptor activation” was inactive, in which EGFR, ADRB2, KLF5, and KLF4 were all downregulated (Tables 8-11 in Appendix 2). The downregulation of this term has been reported to contribute to drug resistance in cancer through reduced detoxification, altered drug metabolism, increased activity of efflux pumps, metabolic reprogramming, immune evasion, and enhanced survival (152).

“MicroRNAs in cancer” was only active with the upregulation of EZH2, in which EGFR, CYP24A1, PLAU, and THBS1 were downregulated (Tables 8 and 10 in Appendix 2). EZH2 is a histone methyltransferase enzyme that might increase cancer cell survival by silencing tumor suppressor genes and contribute to drug resistance by increasing gene expression of oncogenes. It also promotes EMT, evade chemotherapy, survive DNA damage, modulate TME, and evade immune response.

It is recognized to play a critical role in the development of drug resistance in cancer (153).

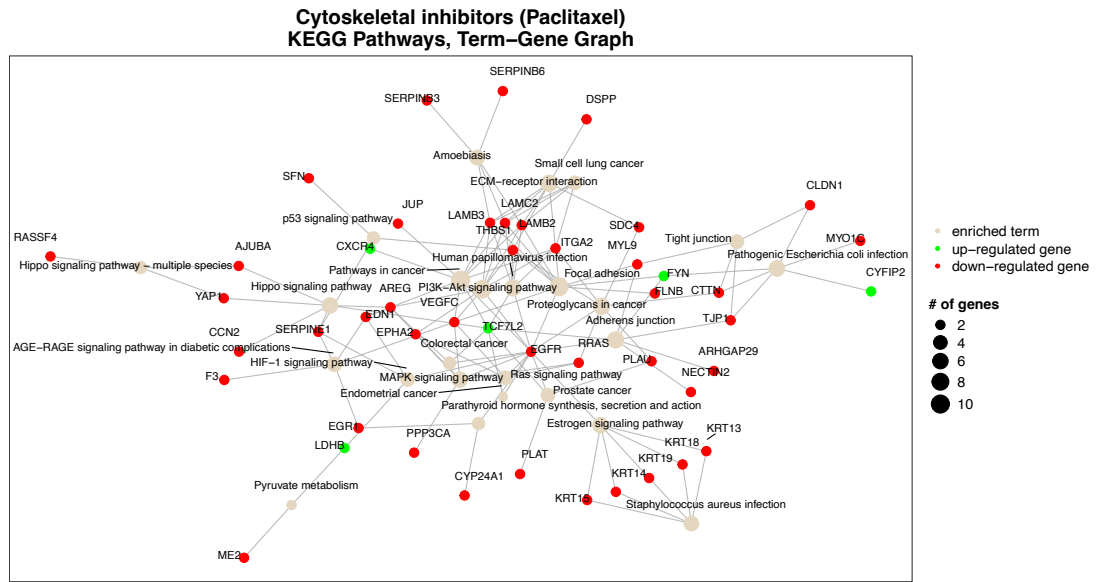


Figure 27. Term-gene graph illustrating the interactions between enriched KEGG terms and their associated genes for paclitaxel based on enrichment results from the best performing cytoskeletal inhibitor model

In addition to general resistance mechanisms, comparison of the first and second strategy pathway enrichment results showed that paclitaxel-specific resistance mechanisms and adaptations might exist. For this purpose, the focus was specifically placed on terms unique to the first strategy or terms containing unique genes identified under the first strategy. “Pyruvate metabolism” was a term that was unique in this way (Tables 8-11 in Appendix 2). In the term, only LDHB was upregulated while ME was downregulated, meaning the activity was LDHB-mediated (Figures 24-25). LDHB converts lactate to pyruvate for sustaining oxidative phosphorylation. In this way, cancer cells might adapt changing oxygen levels and nutrient-deprived environments. With this metabolic programming they can evade apoptosis. LDHB also plays a crucial role in the development of drug resistance by preventing the accumulation of ROS and modifying the TME (154).

Focusing only on genes that were unique under the first strategy revealed LDHB, FYN, FLNB, ODC1, CYP24A1, and MMP7 were the unique DEGs for paclitaxel in a total of 21 mitotic and cytoskeletal inhibitors. However, MYL9, EDN1, and ARHGAP29 were unique DEGs for paclitaxel among the 11 cytoskeletal inhibitors, but not in the mitotic inhibitors. Changes in the terms to which these genes belong may be potential paclitaxel-specific adaptations or resistance mechanisms. Thus, the enriched terms associated with each of these genes were examined in detail in the following paragraphs.

In addition to the “Pyruvate metabolism”, LDHB was the only upregulated gene in the “HIF-1 signaling pathway” and “Central carbon metabolism in cancer” (Tables 8-11 in Appendix 2). The HIF-1 signaling pathway plays a role in adaptation to hypoxic conditions in the TME through metabolic reprogramming, promote angiogenesis, increase heterogeneity, develop resistance to ROS-generating therapies, interact with drug efflux mechanisms, and evade immune system (155). In this term, EDN1 was downregulated, which seems cytoskeleton-specific adaptation against paclitaxel (Tables 9 and 11 in Appendix 2). On the other hand, “Central carbon metabolism in cancer” is linked to the development of resistance through mechanisms such as evading apoptosis, altering the TME via lactate accumulation, suppressing immune cell activity, promoting angiogenesis, supporting drug detoxification, and enhancing metabolic flexibility (156). The presence and upregulation of LDHB in the HIF-1 signaling pathway, central carbon metabolism, and pyruvate metabolism collectively suggest a paclitaxel-specific resistance mechanism.

FYN is a member of the Src family of kinases. Drug resistance association in cancer has been reported for FYN through enhancement of survival signaling, cytoskeletal remodeling, focal adhesion and ECM interaction, immune evasion, and drug efflux (157). FYN, which was upregulated individually in the most significant term “Focal adhesion”, was accompanied by downregulated FLNB and MYL9 (Figures 24-27). This suggests a potential paclitaxel-specific resistance mechanism

mediated by focal adhesion dynamics. Similarly, in the “Adherens junction”, FYN, MYL9, and ARHGAP29 were coexisting, but FYN was the only upregulated gene among them (Figures 24-27). This is likely another paclitaxel-specific resistance mechanism mediated by adherens junction dynamics. FYN was also the only upregulated gene in the “Phospholipase D signaling pathway”. Phospholipase D signaling produces phosphatidic acid to promote cell survival, inhibit apoptosis, promote metabolic adaptation, enhance drug efflux but decrease influx, improve cytoskeletal organization, enhance angiogenesis, and tolerate oxidative stress, leading to drug resistance in cancer (158).

ODC1 was the only upregulated gene driving activity in the “Glutathione metabolism” and “Efferocytosis” (Tables 8-11 in Appendix 2). Glutathione metabolism plays a crucial role in drug detoxification, maintaining redox balance, supporting cancer cell survival under chemotherapy-induced stress, and contributing to the MDR in cancer (159). Similarly, efferocytosis may cause drug resistance in cancer by means of immunosuppressive TME, pro-survival and proliferation signals, nutrient recycling and metabolic adaptation, promotion of angiogenesis, evasion of apoptosis, and MDR (160). Therefore, ODC1 seems to be linked to paclitaxel-specific resistance through its association with these two terms.

CYP24A1 was downregulated in the “MicroRNAs in cancer” (Tables 8-11 in Appendix 2). CYP24A1 is recognized for inactivation of active vitamin D metabolites and has been reported to contribute to the development of drug resistance through immune evasion and TME modulation (161).

MMP7 was downregulated in the “Wnt signaling pathway”, while TCF7L2-mediated activity maintained the pathway (Tables 8-11 in Appendix 2). Mostly, upregulation of MMP7 is reported to be associated with drug resistance since it might aid ECM remodeling and alter TME (162). Its downregulation likely to show complex reprogramming of the TME and signaling pathways.

In addition to the “Focal adhesion”, MYL9 was downregulated in the “Adherens junction”, “Tight junction”, and “Axon guidance” (Figures 24-25, Tables 8-11 in Appendix 2). All of these terms are critical in the development of drug resistance. In addition, it is reported in the previous studies that downregulation of MYL9 causes a weakening in cellular architecture and interactions, and promotes migration and metastasis (163). ARHGAP29 was also downregulated in the “Adherens junction” (Figures 24-25).

EDN1 was downregulated in the “HIF-1 signaling pathway”, “TNF signaling pathway”, “Renin secretion”, and “Melanogenesis” (Tables 9 and 11 in Appendix 2). Reduced TNF signaling can lead to resistance through reduced apoptosis, immune evasion, resistance to ROS-generating therapies, and microenvironment adaptation (164). Decreased renin is linked to impaired immune modulation and resistance to anti-angiogenic therapies (165). Melanogenesis has been reported to contribute to drug resistance through metabolic adaptations, evasion of immune responses, and modification of the TME (166). While TCF7L2 was upregulated, only EDN1 was downregulated in this term (Tables 9 and 11 in Appendix 2). EDN1, along with MYL9, was also downregulated in the “cAMP signaling pathway”, which was suppressed (Tables 9 and 11 in Appendix 2). The deficiency in cAMP signaling can trigger resistance through metabolic reprogramming, immune evasion, and apoptosis inhibition (167).

4.2.2. Resistance mechanisms of 5-FU

Pathway enrichment analysis for 5-FU showed a trend opposite to that of paclitaxel. Pathways for 5-FU were highly active although most enriched terms were common to both analyses (Tables 12 and 13 in Appendix 2). PI3K/Akt, MAPK, NF- κ B, JAK/STAT, Ras, Hippo, Wnt, and FoxO signaling pathways associated with cancer cell survival were all highly active, suggesting the maintenance of general pro-survival resistance mechanisms (Tables 12 and 13 in Appendix 2). The two most significant terms “ECM-receptor interaction” and “Focal adhesion” were also highly

active (Figure 28). ITGA4 was the only downregulated gene in the “ECM-receptor interaction”, while ITGA4, CCND2, RAC2, and VAV1 were downregulated in the “Focal adhesion” (Tables 12 and 13 in Appendix 2). More interestingly, TCF7L2 and CXCR4, which mainly maintained pathway activities in paclitaxel, were always downregulated although they were found in the same terms. Another interesting finding was that the epithelial markers CDH1 and TJP1 were upregulated in many terms, while the mesenchymal marker SNAI2 was upregulated in the “Adherens junction” and “Hippo signaling pathway” (Tables 12 and 13 in Appendix 2). The co-upregulation of both epithelial and mesenchymal markers has been associated with partial EMT and cellular plasticity (168).

A comparison of the first and second strategy results revealed unique terms under the first strategy including the “Neurotrophin signaling pathway”, “Adipocytokine signaling pathway”, and “Renal cell carcinoma” possibly indicated their roles in 5-FU-specific resistance mechanisms (Table 12 in Appendix 2). Neurotrophin signaling is critical in the development of drug resistance in cancer by promoting cancer cell survival and anti-apoptotic effects, increasing proliferation, facilitating stress adaptation, maintaining cancer stem cell phenotypes, interacting with drug efflux mechanisms, and promoting metastasis (169). Adipocytokine signaling similarly plays a role in resistance development through creating a tumor-supportive microenvironment with chronic inflammation, metabolic reprogramming, pro-survival signaling, EMT and metastasis, and angiogenesis (170).

Unique DEGs for 5-FU were identified among a total of 13 DNA replication inhibitors, with enrichment observed exclusively under the first strategy. MAGED1, PINK1, PREX1, LYN, TNFSF13B, and MYO1F emerged as unique DEGs (Table 12 in Appendix 2). The terms associated with these genes are linked to 5-FU-specific resistance mechanisms. Therefore, the following paragraphs elaborate on the roles and activities of these genes within the associated terms.

MAGED1 was uniquely upregulated in the “Neurotrophin signaling pathway” (Table 12 in Appendix 2). As described above paragraph, this pathway activates key pro-survival signaling cascades including PI3K/AKT and MAPK, that promote cell survival, proliferation, and resistance to apoptosis (169). The upregulation of MAGED1 may enhance these effects by stabilizing neurotrophin-receptor interactions or modulating downstream signaling. Additionally, the “Neurotrophin signaling pathway” is linked to cellular plasticity and the maintenance of cancer stem-like cells (CSCs), which exhibit increased resistance to chemotherapy due to their quiescence, enhanced DNA repair capacity, and increased drug efflux (169).

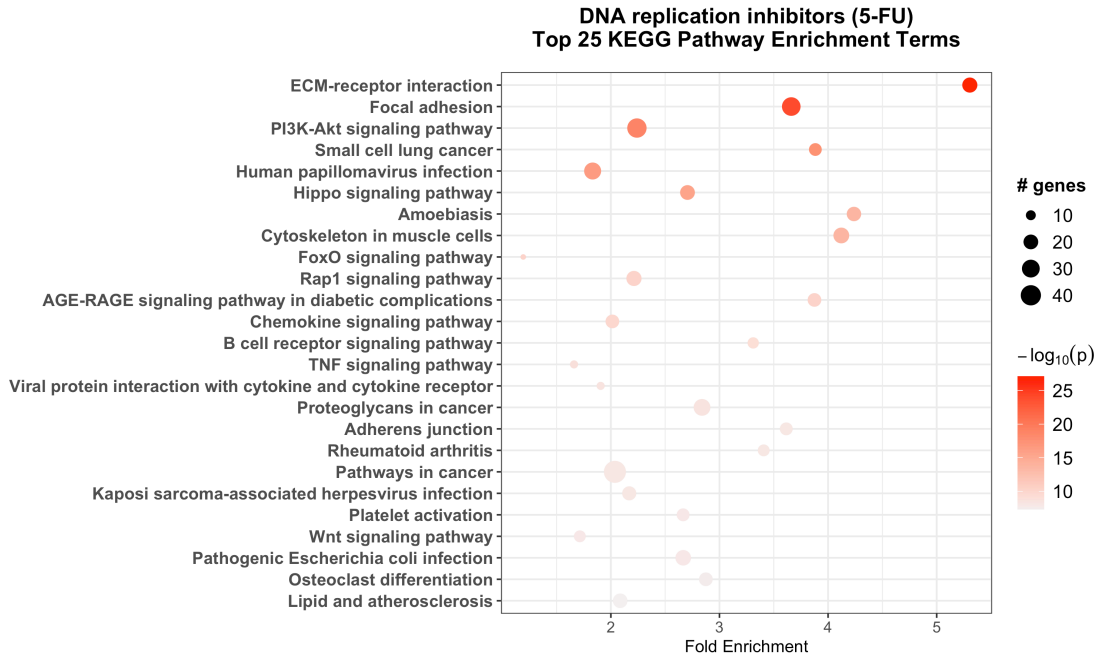


Figure 28. Fold enrichment plot showing the top 25 KEGG pathways for 5-FU based on the gene features of the best performing DNA replication inhibitor model

PINK1, along with RRAS, JUN, and MITF, was upregulated in the “Mitophagy” and no downregulated gene was present in the term (Table 12 in Appendix 2). Mitophagy is critical for maintaining cellular homeostasis by removing damaged mitochondria and preventing ROS accumulation. This process aids in metabolic flexibility and adaptation to hypoxic or nutrient-deprived conditions, enabling

resistance to chemotherapy-induced stress. In addition, it supports CSCs and aids in repairing cellular damage under stress (171).

PREX1 and LYN were together downregulated in the “Chemokine signaling pathway”, with LYN also downregulated in the “NF- κ B signaling pathway” and “B-cell receptor signaling pathway” (Table 12 in Appendix 2). TNFSF13B was also downregulated in the “NF- κ B signaling pathway” and additionally in the “Cytokine-cytokine receptor interaction” (Table 12 in Appendix 2). Chemokine signaling regulates immune cell recruitment and the TME, while NF- κ B and B-cell receptor signaling promote cell survival, proliferation, and immune evasion (172). Cytokine-cytokine receptor interaction is associated with drug resistance through critical activities in immune modulation, enhanced survival and proliferation, drug detoxification, EMT and metastasis, and interaction with the TME (173). However, these terms were not suppressed. 5-FU specific downregulation of PREX1, LYN, and TNFSF13B seems to be supported by other genes that maintain the activity of the pathways (Table 12 in Appendix 2).

MYO1F was downregulated in the “Motor proteins” although many genes were upregulated in the term (Table 12 in Appendix 2). Motor proteins are closely associated with drug resistance through altered microtubule dynamics, drug efflux, enhanced spindle checkpoint function, EMT, and autophagy (174). Up-regulation of genes associated with tubulin and actin dynamics in the term indicates likely cytoskeletal remodeling. However, downregulation of MYO1F, an actin-based motor protein, signifies its 5-FU specific role.

There were also terms unique under the second strategy, likely representing general resistance mechanisms. These mechanisms may have been overshadowed under the first strategy by the presence of 5-FU in the training dataset and the dominance of 5-FU-specific features, but became apparent when 5-FU was excluded. These unique terms were “TGF- β signaling pathway”, “Glutathione metabolism”, “Platinum drug resistance”, and “Signaling pathways regulating

pluripotency of stem cells”. These terms were highly active (Table 13 in Appendix 2). The “TGF- β signaling pathway” is associated with EMT, immune evasion, and plasticity, while “Glutathione metabolism” aids resistance by detoxifying ROS (175, 176). “Platinum drug resistance” highlights mechanisms such as DNA repair and drug efflux, whereas “Signaling pathways regulating pluripotency of stem cells” is linked with CSCs, cellular quiescence, and survival under stress (177, 178).

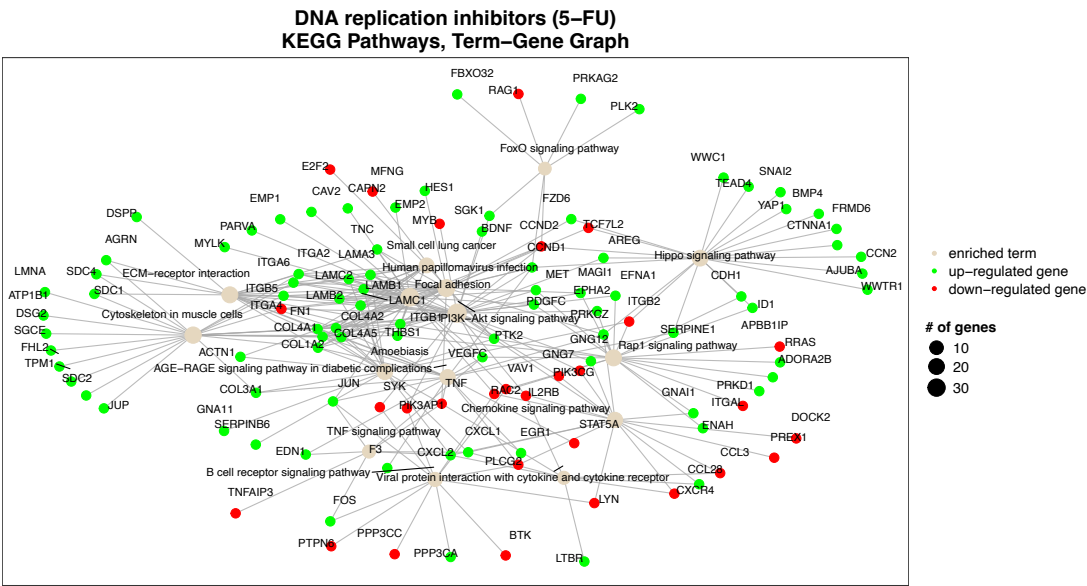
Figure 29 presents the term-gene graph illustrating the interactions among the top 15 terms identified in the enrichment analysis. The graph highlights that highly enriched and significant terms are notably active, with many genes playing roles across multiple terms.

4.2.3. Resistant mechanisms of gemcitabine

Initially pathway enrichment analysis for gemcitabine could not be performed due to insufficient DEGs. To perform the enrichment analysis, the $p\text{-value} < 0,05$ criterion was applied instead of adjusted $p\text{-value}$. In this way, various enriched terms were found. The most significant enriched terms were “Glutathione metabolism”, “TGF- β signaling pathway”, and “Signaling pathways regulating pluripotency of stem cells” (Figure 30). Most of the enriched terms were active, with resistant samples exhibiting higher gene expression levels (Figure 31). In addition, the enrichment analysis for gemcitabine revealed no unique DEGs between the first and second strategies, indicating that all identified enriched terms are associated with general resistance mechanisms rather than being specific to gemcitabine resistance (Tables 14 and 15 in Appendix 2).

“Glutathione metabolism” emerged as a key term, with the upregulation of MGST2, which associated with detoxification processes that may reduce intracellular gemcitabine concentrations (Figure 31) (176). In the “TGF- β signaling pathway”, the upregulation of INHBB and BMP4 was observed (Figure 31). “Signaling pathways regulating pluripotency of stem cells” showed the upregulation of FZD6, INHBB,

and BMP4. Both the Hippo and Wnt signaling pathways were active and FZD6 were upregulated in both terms (Figure 31). Additional terms “Drug metabolism - cytochrome P450” and “Chemical carcinogenesis - DNA adducts” as well as “Metabolism of xenobiotics by cytochrome P450”, also demonstrated MGST2 upregulation (Figure 31). Its upregulation is reported to be related with enhanced detoxification and resistance to gemcitabine (176).



contribute to resistance mechanisms. Hence, both partial EMT and cellular plasticity could allow cancer cells to adapt to and survive therapy (182).

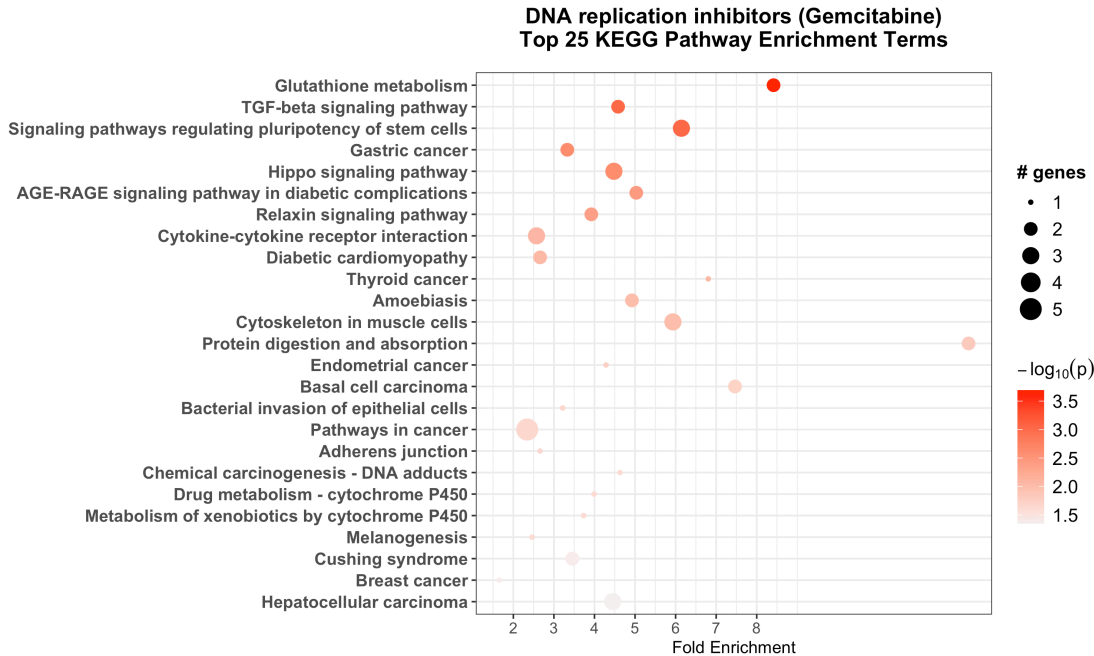


Figure 30. Fold enrichment plot showing the top 25 KEGG pathways for gemcitabine based on the gene features of the best performing DNA replication inhibitor model

4.2.4. Resistance mechanisms of cetuximab

pathfindR analysis identified 167 enriched terms for cetuximab and included many cancer-related signaling pathways (Figure 33, Tables 16 and 17 in Appendix 2). Similar to paclitaxel enrichment results, most genes within these terms were downregulated, while a few upregulated genes maintained activity. Enrichment analysis revealed that the active pathways associated with survival included PI3K/Akt, MAPK, NF- κ B, Ras, Wnt, and Hippo signaling pathways (Tables 16 and 17 in Appendix 2). In addition, immune signaling pathways such as B-cell and T-cell receptor signaling pathways were active. Additional active signaling pathways were Chemokine, Relaxin, Rap1, Apelin, Oxytocin, Calcium, cAMP, Prolactin, Phospholipase D signaling pathways, with only a few upregulated genes. In contrast, several pathways were significantly suppressed such as Estrogen, TNF, IL-17, p53,

FoxO, HIF-1, ErbB, VEGF, JAK-STAT, Neurothropin, Adipocytokine signaling pathways.

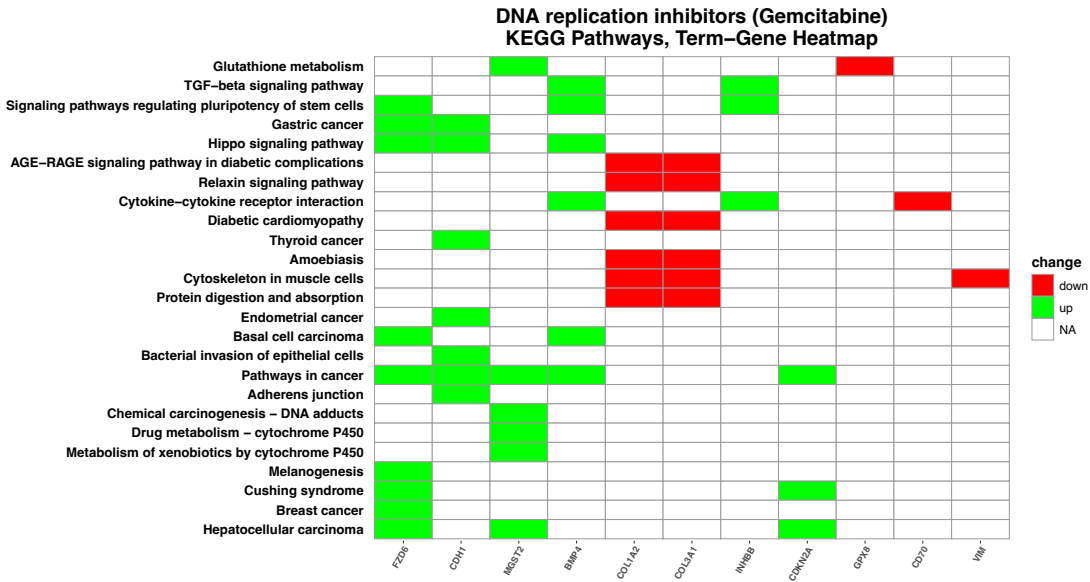


Figure 31. Term-gene heatmap plot for gemcitabine based on the gene features of the best performing DNA replication inhibitor model

Additional enriched terms related to cell-cell interaction and cellular adhesion, metabolism, cytoskeletal organization, immune response, TME, microRNA regulation, and drug resistance were identified (Tables 16 and 17 in Appendix 2). The top three enriched terms were “ECM-receptor interaction”, “Focal adhesion”, and “PI3K-Akt signaling pathway” (Figure 33).

“EGFR TKI resistance” was suppressed, indicating that cetuximab resistance mechanisms may bypass EGFR signaling and rely on alternative signaling pathways. In contrast, “Endocrine resistance” remained active only through the upregulation of E2F2 (Tables 16 and 17 in Appendix 2). It was also upregulated in multiple cancer-related terms, including pancreatic, gastric, and breast cancer, as well as in the “MicroRNAs in cancer” (Tables 16 and 17 in Appendix 2). This term also remained active with the upregulation of ZEB1 and VIM, both of which are mesenchymal markers (183). E2F2 was also exclusively upregulated in the “Cellular senescence”. Cellular senescence is highly associated with drug resistance in cancer through pro-

survival signaling, microenvironment remodeling, drug efflux, metabolic reprogramming, and immune evasion (184). The role of E2F2 in drug resistance has been reported through bypassing cell cycle checkpoints, enhancing DNA damage response and repair, evading apoptosis, inducing stemness and drug tolerance, and engaging in crosstalk with survival pathways (185).

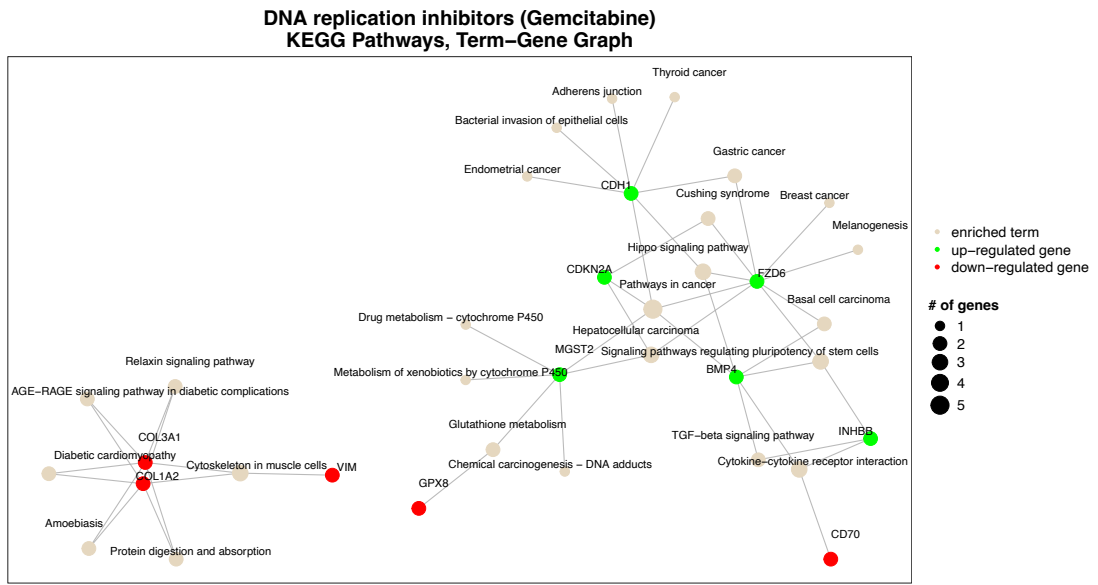


Figure 32. Term-gene graph illustrating the interactions between enriched KEGG terms and their associated genes for gemcitabine based on enrichment results from the best performing DNA replication inhibitor model

EMT markers CDH1 and TJP1, associated with epithelial characteristics, were consistently downregulated across multiple enriched terms (Tables 16 and 17 in Appendix 2). However, mesenchymal markers VIM and ZEB1 were upregulated. These expression changes highlight a potential role for EMT in cetuximab resistance. Interestingly, another mesenchymal marker SNAI2 was found to be downregulated, suggesting cellular plasticity and partial EMT (186).

The comparison of the first and second strategies showed that “Signaling pathways regulating pluripotency of stem cells” and “Viral carcinogenesis” were unique under the first strategy (Table 16 in Appendix 2). Both terms were however suppressed. This suggests that cetuximab resistance relies on alternative mechanisms

other than CSC plasticity and immune evasion. In addition, no cetuximab-specific DEG was present in the first strategy enrichment results, suggesting a reliance on general resistance mechanisms rather than cetuximab-specific molecular changes.

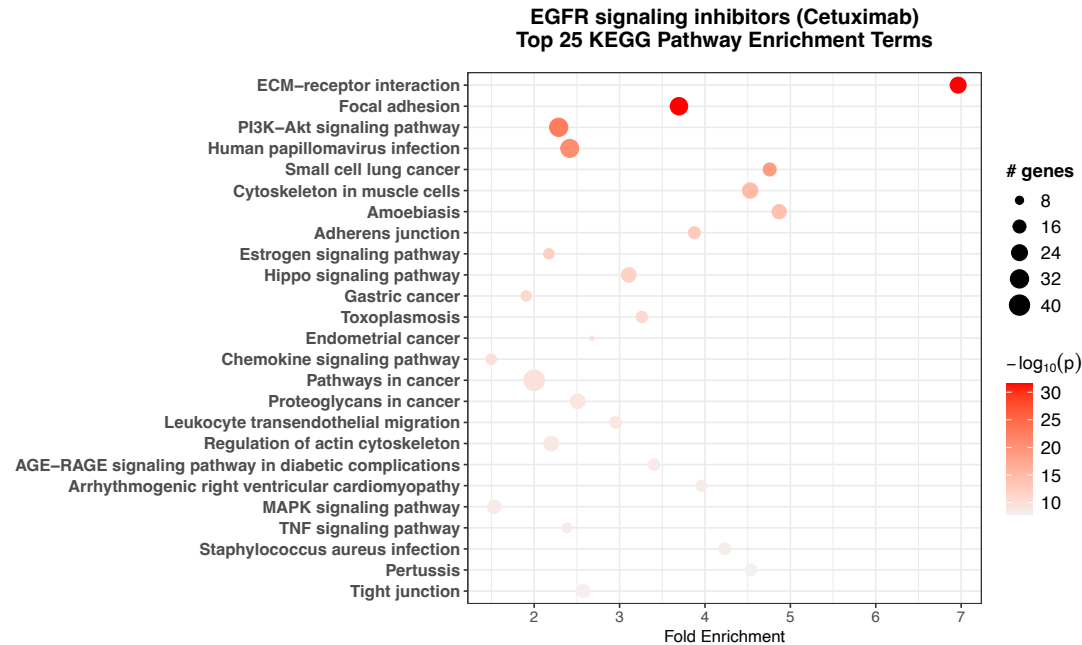


Figure 33. Fold enrichment plot showing the top 25 KEGG pathways for cetuximab based on the gene features of the best performing EGFR signaling inhibitor model

Figure 34, the term-gene graph, illustrates that the enriched terms for cetuximab predominantly consist of downregulated genes. However, a few upregulated genes, including E2F2, LEF1, TCF7L2, FYN, and CXCR4, sustain the activity of certain pathways. Notably, TCF7L2, FYN, and CXCR4 were also crucial in maintaining pathway activity in numerous similar terms associated with paclitaxel (Figures 24-25). Similarly, many pathways were suppressed in paclitaxel resistance. This suggests that some general resistance mechanisms are shared between paclitaxel and cetuximab resistance and that these three genes significantly contribute to the development of resistance in both drugs.

4.3. Association of Mutation and CNA Statuses with Drug Resistance

Using FET with the GDSC paclitaxel and 5-FU samples, the resistance associations of mutation and CNA statuses for unique DEGs identified as paclitaxel- and 5-FU-specific in the enrichment analysis revealed no significant mutation-related associations. However, CNA status showed significant associations for some 5-FU-specific genes, suggesting that resistance to 5-FU may be driven not only by transcriptional changes but also by CNAs.

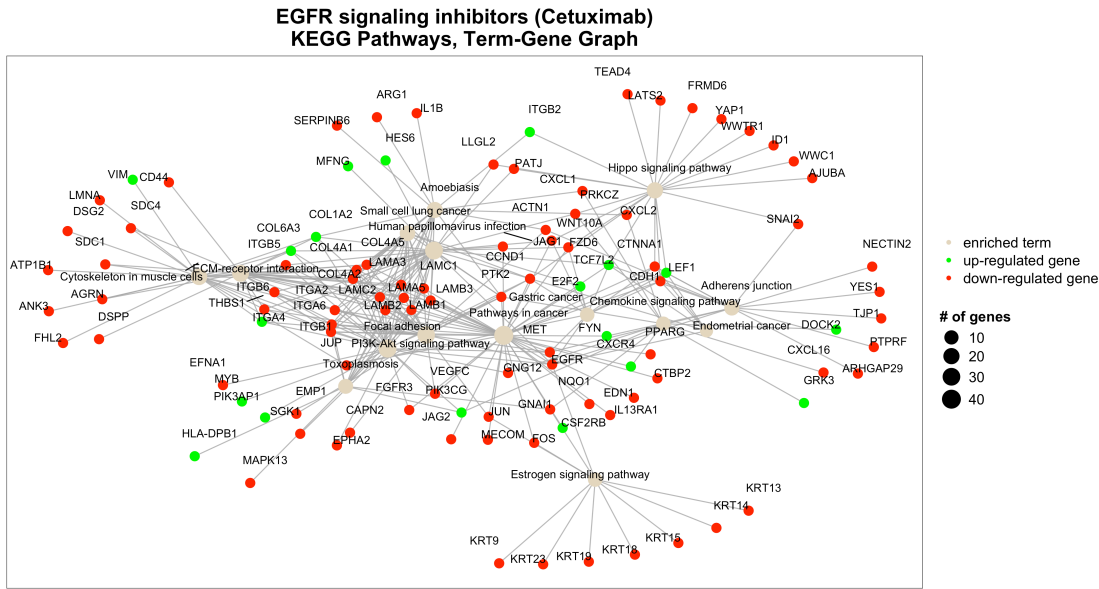


Figure 34. Term-gene graph illustrating the interactions between enriched KEGG terms and their associated genes for cetuximab based on enrichment results from the best performing EGFR signaling inhibitor model

A potentially significant association between paclitaxel resistance and CNA status was observed only for LDHB among paclitaxel-unique DEGs in the GDSC paclitaxel samples, showing a high OR and a borderline p-value (OR = 2,87, p-value = 0,05). FET was also performed for TCF7L2, CXCR4, EZH2, and CYFIP2, which are involved in maintaining pathway activities. Among these, only the CNA status of TCF7L2 showed a significant association (OR = 4,11, p-value = 0,02).

With GDSC 5-FU samples, FET showed significant associations between CNA status and resistance for PINK1 (OR = 2,34, p-value = 0,003), PREX1 (OR = 2,68, p-value = $5,70 \times 10^{-5}$), TNFSF13B (OR = 2,13, p-value = 0,002), and MYO1F (OR = 2,72, p-value = 0,0005).

Although not specific to 5-FU or gemcitabine, several genes from the feature list for these drugs exhibited significant associations between mutation status and drug resistance. For 5-FU, significant associations were observed with TJP1 (OR = 6,83, p-value = 0,02) and CNTRL (OR = 6,41, p-value = 0,03), while for gemcitabine, NUP210 (OR = ∞ , p-value = 0,04) and COL3A1 (OR = ∞ , p-value = 0,04) were significant. Similarly, for CNA status, several genes from the feature list of paclitaxel, 5-FU, and gemcitabine were significantly associated with resistance. For paclitaxel, several genes within enriched terms demonstrated significant associations, including PLAU (OR = 5,87, p-value = 0,007) and DKK1 (OR = 3,51, p-value = 0,04). Other genes from the paclitaxel feature list that showed significant associations are presented in Table 18 in Appendix 3. For 5-FU, around 500 genes exhibited significant associations (Table 19 in Appendix 3), while for gemcitabine, over 100 genes showed CNA-related resistance links (Table 20 in Appendix 3).

4.4. Predictive Analysis of Resistance Mechanisms and Genomic Interactions with Survival Outcomes

In the univariate penalized CoxPH regression models for FYN, FLNB, and MYL9, which are unique DEGs in the “Focal adhesion” that emerged most significant in the enrichment analysis for paclitaxel, the C-index was highest for FYN (C-index = 0,89). In the multivariate penalized CoxPH regression models, the combination of FYN and MYL9 yielded the highest C-index (0,90). Similarly, the C-indices of pathways derived from multivariate models containing resistance-related genes were also notably high. The highest was for the “Axon guidance”, which also included FYN and MYL9 (C-index = 0,94). The C-index for terms containing only FYN was almost identical, indicating that FYN alone was sufficient for strong

predictive performance. In contrast, MYL9 alone had a lower C-index (0,80) although higher values were observed for most terms containing only MYL9 except for the “Motor proteins”. The top-performing MYL9-only term was “Cytoskeleton in muscle cells” (C-index = 0,92) followed by the “Regulation of actin cytoskeleton” (C-index = 0,90). Notably, “Focal adhesion” containing all three genes had a lower C-index (0,87) than FYN alone. These results indicate a strong predictive association between CNA status of FYN and paclitaxel resistance with the “Axon guidance” emerging as the most predictive term, outperforming “Focal adhesion”, which was the most significant term in the enrichment analysis (Figure 35).

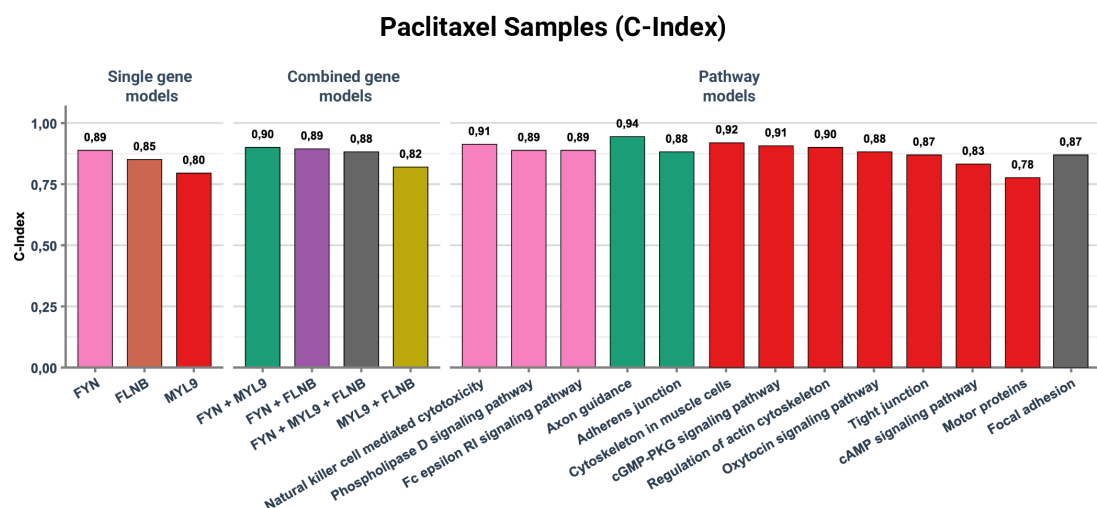


Figure 35. C-indices from penalized univariate and multivariate CoxPH regression models for paclitaxel-unique genes (FYN, FLNB, and MYL9) within the most enriched focal adhesion pathway

HRs of TCF7L2 and LDHB, whose CNA association with resistance to paclitaxel was significant with FET, was quite high for response and CNA x response interaction with univariate penalized CoxPH regression models. HRs were also high in the model in which two genes were used together with bivariate penalized CoxPH regression models. Although the C-indices were high individually, it was superior when combined, achieving a C-index = 0,90. This suggests that CNAs of FYN and LDHB are likely driving mechanisms of paclitaxel resistance (Figure 36).

When the LRT was applied separately to unique DEGs for paclitaxel and TCF7L2, the worst prognosis for TCF7L2, LDHB, MMP7, and FLNB was observed in the group with combined CNA and resistant samples. All of them showed a p-value < 0,05, suggesting that CNA alterations in these genes play a significant role in driving resistance to paclitaxel (Figure 37).

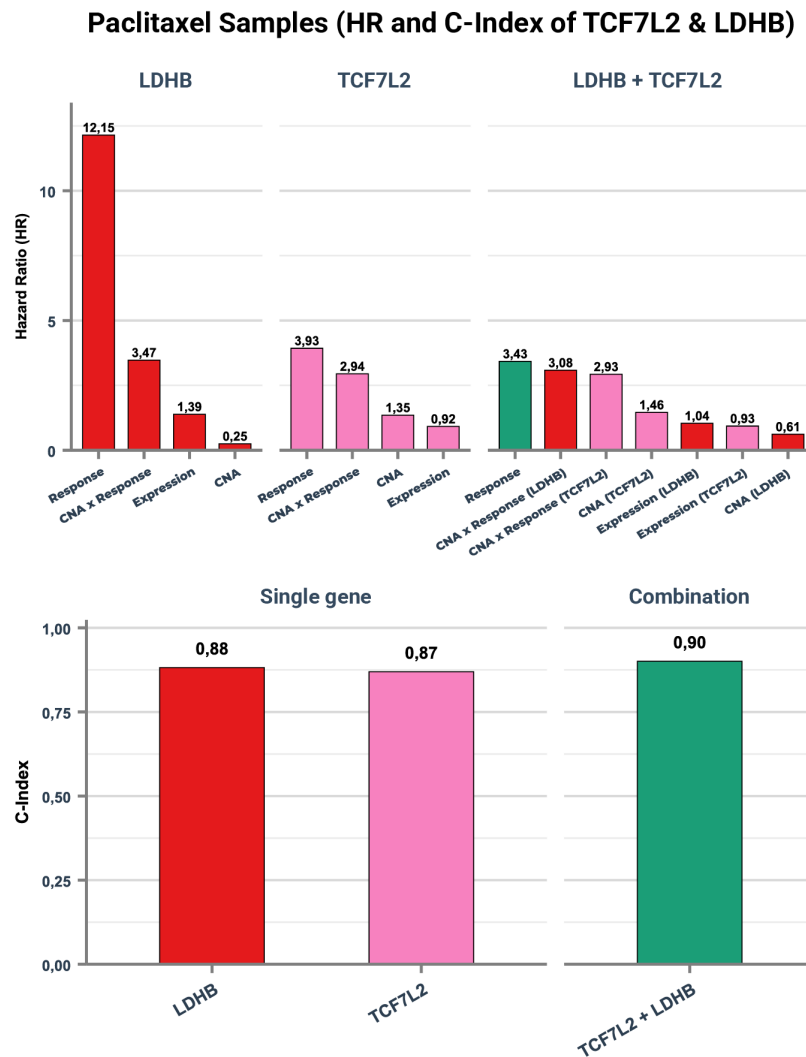


Figure 36. HRs and C-indices from penalized univariate CoxPH regression models for LDHB and TCF7L2, identified as significant in the pathway enrichment analysis for paclitaxel resistance, and from penalized bivariate CoxPH regression model when evaluated together

In the TCGA STAD cohort for 5-FU, the DEGs identified as 5-FU-specific showed the highest C-index score for LYN alone (C-index = 0,88) with the

univariate penalized CoxPH regression models. The next with the highest C-index was that for MYO1F with C-index = 0,81. There were increases in C-index when combinations were used with the multivariate penalized CoxPH regression models. The three different combinations involving LYN and PINK1 achieved the highest value with C-index = 0,93. Interestingly, when only LYN and PINK1 were used together, the C-index decreased to 0,80. However, the C-index increased when MYO1F or TNFSF13B was added to the model. The addition of PREX1 to the models did not result in any improvement. LYN, MYO1F, and PINK1 were enriched in different terms, while TNFSF13B was enriched only in the “NF-κB signaling pathway” with LYN. This demonstrated that combining genes from different terms improved the predictive relationship between CNA and resistance (Figure 38).

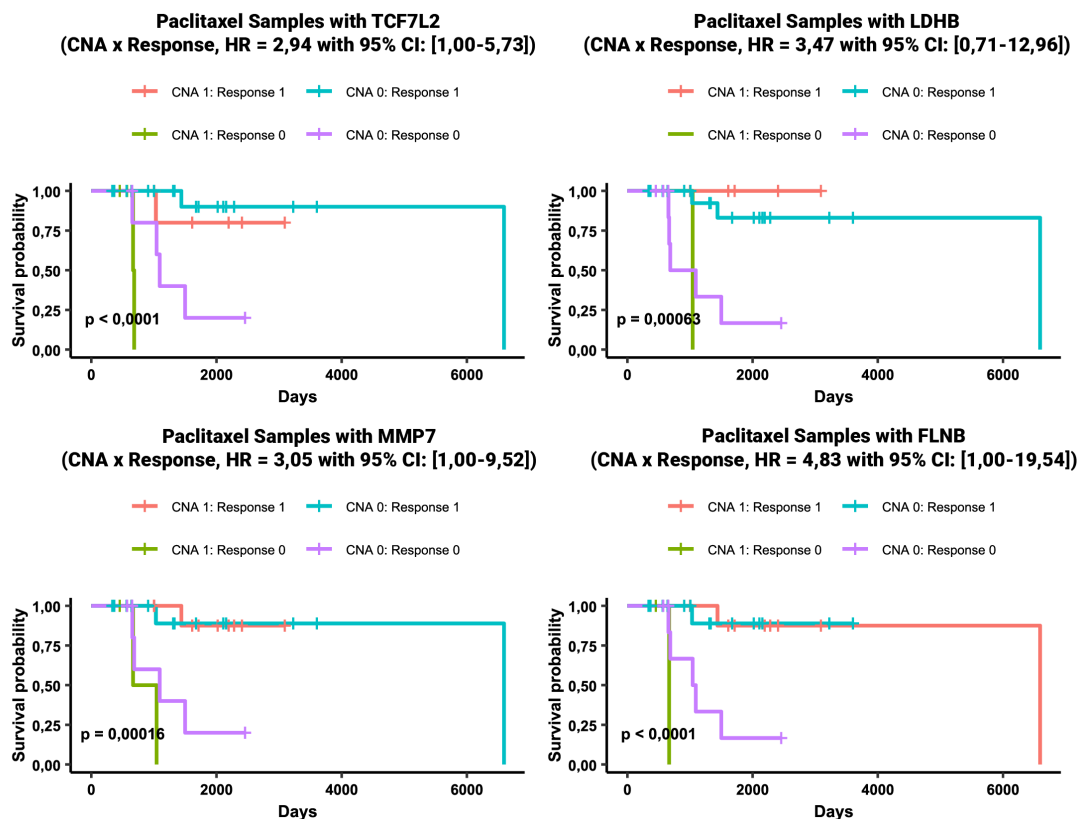


Figure 37. KM survival curves for paclitaxel-resistance-associated genes (TCF7L2, LDHB, MMP7, and FLNB) identified in the pathway enrichment analysis of paclitaxel

C-index comparison was conducted based on the terms to which LYN and TNFSF13B belong in the TCGA STAD cohort for 5-FU samples. This analysis was not conducted for MYO1F and PINK1 as they were exclusively enriched in the resistance-related terms “Motor proteins” and “Mitophagy” respectively. LYN had nearly identical C-indices for the “NF-κB signaling pathway” and “Fc gamma R-mediated phagocytosis” with the highest C-index = 0,89. However, the situation was different for TNFSF13B as the C-index increased to 0,90 for the “Cytokine-cytokine receptor interaction” and to 0,89 for the “NF-κB signaling pathway”. The C-indices of other terms also improved. These results indicate that LYN alone and TNFSF13B with the activity of the pathways it is associated with can effectively predict the relationship between CNA and resistance (Figure 39).

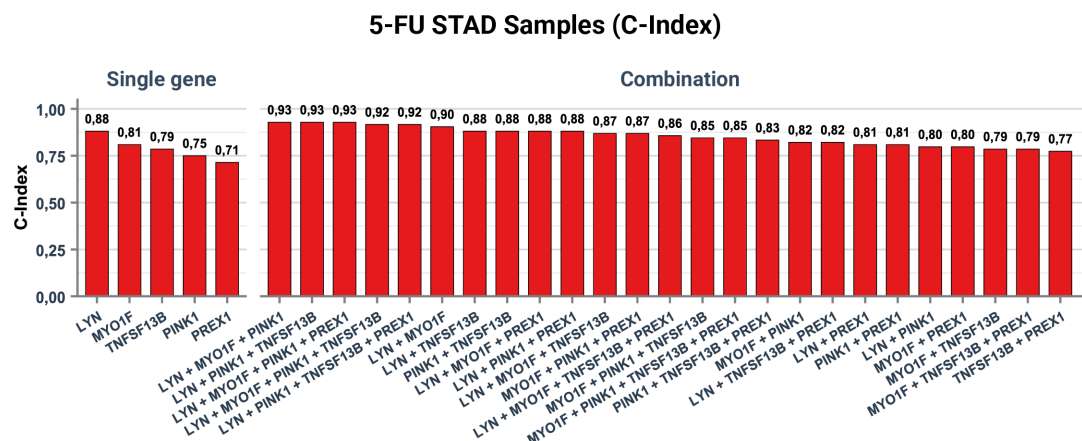


Figure 38. C-indices from penalized univariate and multivariate CoxPH regression models for 5-FU-unique genes (LYN, MYO1F, TNFSF13B, PINK1, and PREX1). Genes are evaluated individually and in combination

BAMBI, which was upregulated in the “Wnt signaling pathway” and “TGF-β signaling pathway” according to enrichment analysis for the 5-FU samples, was associated with poor prognosis based on mutation status (p-value = 0,00027) with the LRT. The univariate penalized CoxPH regression model yielded HR = 7,28 (95% CI: 1,00-13,02). This substantial effect size highlights BAMBI’s critical role in driving resistance and poor survival. Furthermore, CDH1, which was upregulated in several terms such as “Hippo signaling pathway” and “Adherens junction” according to

enrichment analysis for the 5-FU STAD samples and is also an epithelial marker, was associated with poor prognosis based on mutation status ($p\text{-value} = 0,044$) with the LRT. The univariate penalized CoxPH regression model yielded $HR = 1,18$ (95% CI: 1,00-1,22). This suggests that the mutation status of CDH1 contributes to worse outcomes in 5-FU resistance in the STAD cohort (Figure 40).

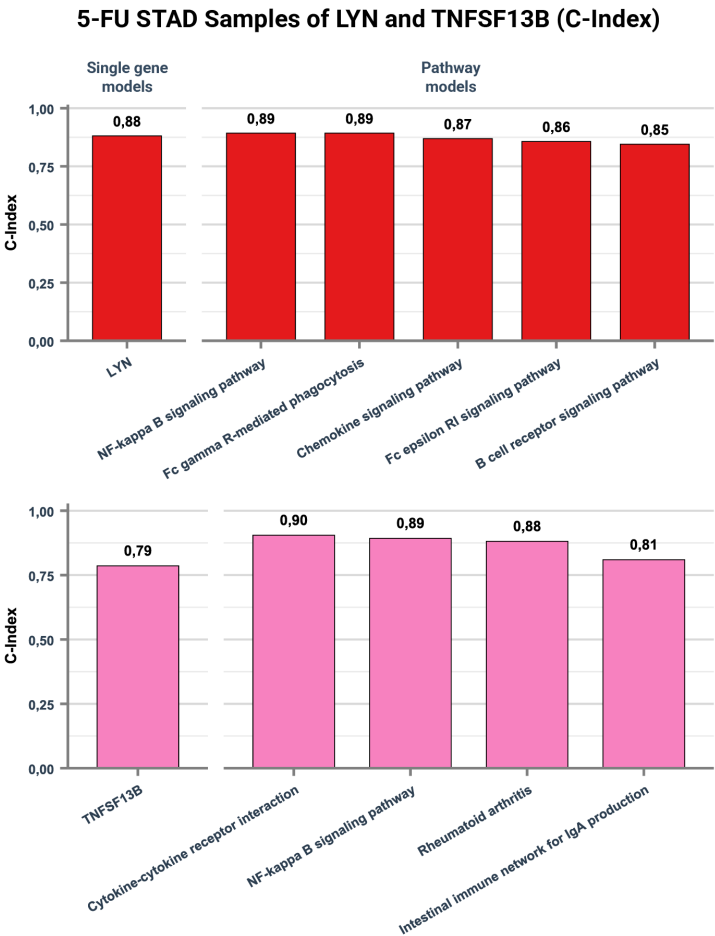


Figure 39. C-indices from penalized univariate CoxPH regression models for LYN and TNFSF13B, identified as significant in the pathway enrichment analysis for 5-FU resistance

When the univariate penalized CoxPH regression model for gemcitabine in the TCGA PAAD cohort was applied for GPX8, BMP4, MGST2, and CDH1, identified as key genes related to drug resistance in the enrichment analysis, the highest individual C-index (C-index = 0,64) was observed for GPX8. However, when all four genes were included in a multivariate penalized CoxPH regression model, the

C-index increased to 0,70, indicating an improved predictive performance. Additional improvements were observed with other combinations of the genes. Enrichment analysis revealed that GPX8 and MGST2 were enriched together in the “Glutathione metabolism”, while CDH1 and BMP4 were enriched together in the “Hippo signaling pathway”. The higher C-index in the multivariate model suggests that resistance to gemcitabine is multi-factorial, involving metabolic and signaling pathways that are captured more effectively when these genes are combined (Figure 41).

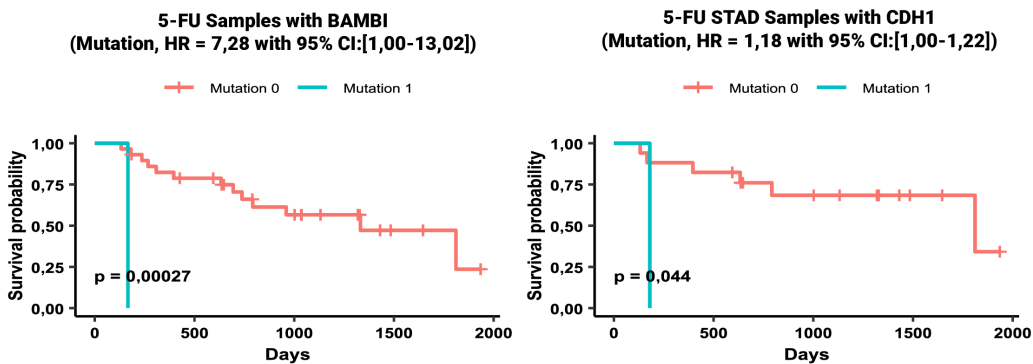


Figure 40. KM survival curves for BAMBI and CDH1, identified as resistance-associated in the pathway enrichment analysis for 5-FU, stratified by mutation status

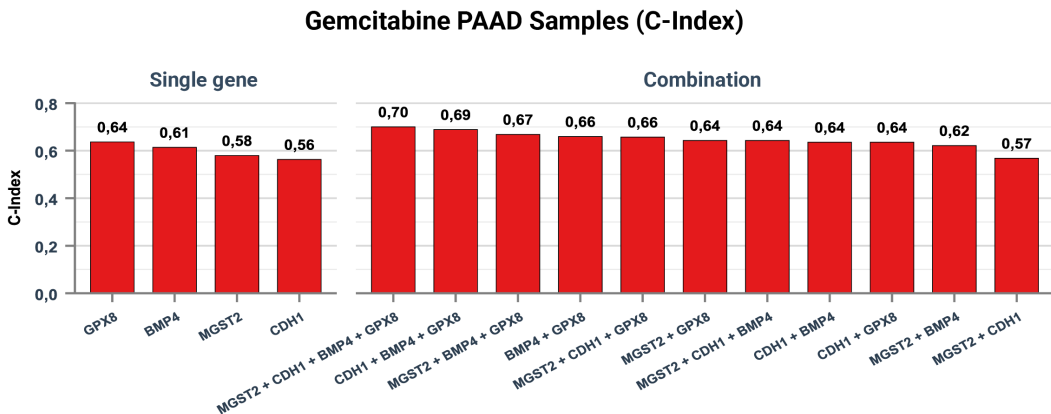


Figure 41. C-indices from penalized univariate and multivariate CoxPH regression models for gemcitabine, including genes from the top enriched pathways (GPX8, BMP4, MGST2, and CDH1)

When the LRT was applied to NQO1 and CLIC3, the features of the pan-drug models with the strongest prediction of response to gemcitabine based on mutation

status, the worst prognosis was observed in the group with mutations but no response. Both results were significant (p -value $< 0,0001$). The univariate penalized CoxPH regression models yielded $HR = 1,30$ (95% CI: 1,00-1,30) for the mutation $\times$ response interaction of NQO1, and $HR = 8,66$ (95% CI: 2,89-22,20) for the mutation $\times$ response interaction of CLIC3. These findings suggest mutations in NQO1 and CLIC3 are associated with resistance to gemcitabine and poor survival outcomes. The significant interaction terms further emphasize the presence of these mutations modifies the influence of these genes on the response to gemcitabine (Figure 42).

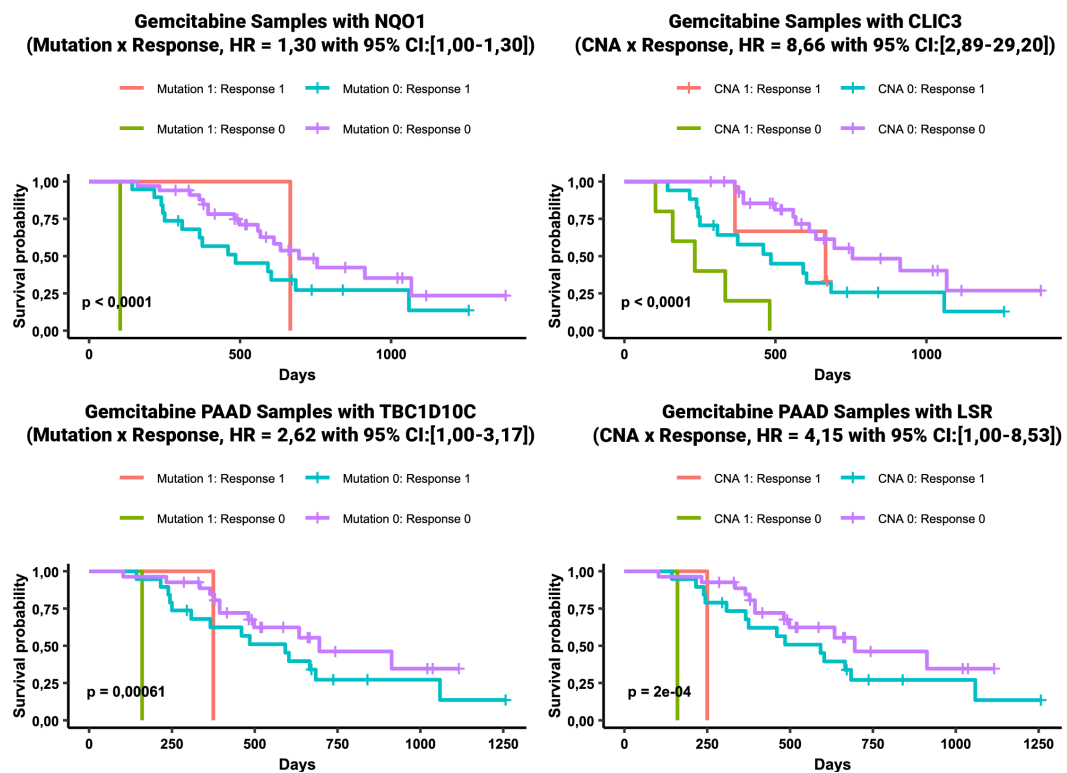


Figure 42. KM survival curves for gemcitabine-resistance-associated genes (NQO1, CLIC3, TBC1D10C, and LSR)

When the LRT was applied to TBC1D10C and LSR, which were also predictive features for gemcitabine in the TCGA PAAD cohort, the worst prognosis was observed in the groups with mutation and no response for TBC1D10C and CNA and no response for LSR, respectively. The p -values were highly significant for both. The univariate penalized CoxPH regression models yielded $HR = 2,62$ (95% CI: 1,00-3,20) for the mutation $\times$ response interaction for TBC1D10C, and $HR = 4,15$

(95% CI: 1,00-8,53) for the CNA $\times$ response interaction for LSR. These results indicated TBC1D10C mutations and LSR CNAs predict poor prognosis and resistance to gemcitabine in the PAAD cohort (Figure 42).

5. DISCUSSION

Our enrichment analyses identified that LDHB-driven “Pyruvate metabolism” was specifically activated in response to paclitaxel. Among all mitotic, cytoskeletal, DNA replication, and EGFR signaling inhibitors, LDHB was upregulated solely in paclitaxel-treated samples, and “Pyruvate metabolism” was uniquely enriched for paclitaxel. In addition, FET results for LDHB pointed to a potential link between paclitaxel resistance and its CNA status (OR = 2,87, p-value = 0,05). The univariate CoxPH regression model also showed CNA x resistance interaction was significant for LDHB. This suggests a multi-level role for LDHB in drug resistance through both transcriptional regulation and CNA-mediated effects. These findings emphasize LDHB’s dual role in resistance and its prognostic significance, particularly in paclitaxel-treated cases. These observations imply that both transcriptional upregulation and CNAs of LDHB may contribute to paclitaxel resistance. Consistent with this, earlier studies have shown that resistant cancer cells may enhance LDHB expression through metabolic reprogramming, converting lactate into pyruvate and channeling it into oxidative phosphorylation under hypoxic conditions (154, 187). Furthermore, LDHB upregulation was associated with paclitaxel resistance in oral squamous cell carcinoma (188). Additionally, silencing LDHB in Merkel cell carcinoma cell lines increased paclitaxel-induced apoptosis, which was reversed by LDHB re-expression (189).

Paclitaxel-resistant cell lines are likely to exist in a drug-tolerant persister (DTP) state, which is linked to the emergence of drug-resistant phenotypes (190). In this state, cancer cells enter reversible quiescence in response to drug-induced stress, prioritizing survival over proliferation with the Wnt signaling pathway frequently reported as active (191). During DTP, stress response and compensatory mechanisms are activated, accompanied by epigenetic and metabolic reprogramming, acquisition of stem-like features, and partial or complete suppression of pathway activity often with only a few genes remaining active (192). In our enrichment results for paclitaxel, most pathways were either suppressed or partially active, with limited

gene upregulation across multiple terms. Despite the suppression of numerous survival-associated pathways, compensatory Wnt and Hippo signaling driven by TCF7L2 appeared to support survival (192, 193). FYN-mediated “Focal adhesion” may also facilitate apoptotic resistance and promote cell survival (194). In addition to its role in survival, FYN upregulation was noteworthy as its activation has been characterized as a hallmark of the DTP phenotype (195). Furthermore, LDHB-mediated “Pyruvate metabolism” and ODC1-mediated “Glutathione metabolism” reflected metabolic reprogramming, while EZH2-mediated “MicroRNAs in cancer” suggested epigenetic adaptation (196).

In our enrichment results for paclitaxel, terms associated with cellular adhesion and cell-cell interaction such as “ECM-receptor interaction” and “Tight junction” were suppressed. In contrast, FYN-mediated “Focal adhesion” and TCF7L2- and FYN-mediated “Adherens junction” were partially active. Additionally, survival-related signaling pathways including p53 and PI3K/Akt were found to be suppressed. These results suggest the presence of anoikis resistance, in which FYN upregulation may represent a paclitaxel-specific survival mechanism. FYN was identified as a DEG exclusively for paclitaxel among 21 mitotic and cytoskeletal inhibitors. Our observation that FYN was the only upregulated gene within the most significantly enriched term, “Focal adhesion”, may indicate that paclitaxel-resistant cell lines engage FYN-driven adhesion signaling to evade apoptosis. Supporting this, Chou et al. (2021) demonstrated FYN upregulation in paclitaxel-resistant basal-like TNBC (197). Similarly, Egeland et al. (2024) reported FYN upregulation in resistant TNBC patients and in rapidly progressing metastatic lesions (198). However, FLNB and MYL9, two other paclitaxel-specific DEGs identified within the “Focal adhesion”, were downregulated. These genes are known to contribute to actin cytoskeleton remodeling and contraction, and their downregulation may support cell survival during mitotic arrest by reducing cytoskeletal tension (199, 200). In particular, the upregulation of FYN and downregulation of MYL9 were also observed in the “Adherens junction” and “Axon guidance”. In addition, FET revealed no association

of FYN with CNAs or mutations in the GDSC dataset, and univariate penalized CoxPH regression models indicated no significant interaction of CNA or mutation status with resistance. These findings suggest that FYN contributes to paclitaxel resistance primarily through its transcriptional activation and pathway-specific roles rather than genomic alterations.

Our enrichment results indicated that TCF7L2 upregulation may also contribute to paclitaxel resistance, potentially acting synergistically with FYN. TCF7L2 was found to be upregulated not only in the “Adherens junction” but also exclusively in the Wnt and Hippo signaling pathways, as well as in several cancer-related terms. Unlike FYN, however, FET suggested that CNAs in TCF7L2 may contribute to paclitaxel resistance alongside its transcriptional upregulation (OR = 4,11, p-value = 0,016). Mechanistically, FYN is capable of activating the Wnt signaling pathway either directly through phosphorylation of β -catenin, or indirectly by stabilizing β -catenin via inhibition of its degradation through interaction with GSK3 β within the β -catenin destruction complex (157, 201, 202). Therefore, TCF7L2 upregulation may be driven by FYN activation.

FYN may be a critical mediator of paclitaxel resistance not only by upregulating TCF7L2 but also by being activated through phosphorylation by CXCR4 (203). CXCR4, which was co-upregulated with FYN in the “Axon guidance”, was also co-upregulated with TCF7L2 in the “Pathways in cancer.” CXCR4 is known to be upregulated in hypoxic niches, where it facilitates chemotactic migration toward CXCL12-rich adaptive regions within the TME, thereby enabling cells to evade paclitaxel exposure (204). Moreover, CXCR4 was found to be upregulated in the “Regulation of actin cytoskeleton”. This suggests that actin cytoskeleton remodeling may enhance invasiveness, motility, metastasis, and EMT in paclitaxel-resistant cell lines, contributing to resistance through apoptotic evasion (205). Notably, CXCR4 exhibited the highest logFC among all genes upregulated by paclitaxel in this study. Overall, these findings suggest that CXCR4, FYN, and TCF7L2 may function in a

coordinated and interconnected regulatory mechanism contributing to paclitaxel resistance.

Similar to the enrichment results for paclitaxel, CXCR4, FYN, and TCF7L2 were predominantly upregulated in the same pathways in the enrichment results for cetuximab, suggesting that the two drugs may share common resistance mechanisms. Clinical trial findings demonstrating improved treatment responses with the combination of paclitaxel and cetuximab in metastatic or recurrent HNSC support the hypothesis that these genes may play a pivotal role in resistance and that co-targeting them could enhance therapeutic efficacy in combination strategies (206, 207). In our analysis with paclitaxel, FYN emerged as a unique DEG among mitotic and cytoskeletal inhibitors. While FYN was also upregulated for cetuximab, it did not meet the DEG threshold based on logFC. However, FYN was identified as a DEG for other EGFR-targeted therapies, including gefitinib, afatinib, and lapatinib, suggesting a broader role in resistance to EGFR inhibition. This finding aligns with previous studies reporting that FYN can be activated by EGFR pathway mutations and may contribute to therapeutic resistance and cancer progression (157, 208). Moreover, although the CNA status of TCF7L2 was significantly associated with paclitaxel resistance, no such association was observed for cetuximab, indicating that TCF7L2 may contribute to cetuximab resistance primarily through transcriptional upregulation. Among the genes upregulated in response to cetuximab, CXCR4 exhibited the highest logFC consistent with its expression pattern observed under paclitaxel treatment.

Our enrichment results for cetuximab suggest that cetuximab-resistant cell lines may exist in a DTP state similar to that observed in paclitaxel-resistant cell lines. In these results, most enriched terms were either partially active or suppressed, with a predominance of downregulated genes. Key survival-associated pathways including PI3K/Akt, MAPK, NF- κ B, Wnt, and Hippo were all partially active, reflecting limited activation of survival signaling. In addition, PIK3CG-mediated “Inositol phosphate metabolism”, together with ZEB1-, E2F2-, and VIM-mediated enrichment

in the “MicroRNAs in cancer”, indicated the presence of both metabolic and epigenetic reprogramming. Most importantly, FYN upregulation, which is a hallmark feature of the DTP phenotype, was also observed (195).

Our enrichment results for cetuximab also revealed critical insights into its resistance mechanisms. The suppression of the “EGFR tyrosine kinase inhibitor resistance” suggests that canonical EGFR-dependent mechanisms may not be the primary drivers of resistance and that resistant cells may survive by bypassing EGFR signaling (209). Supporting this, our data indicated that PIK3CG-mediated “Inositol phosphate metabolism” may activate the PI3K/Akt pathway through metabolic reprogramming in an EGFR-independent manner, thereby promoting cell survival and contributing to immune suppression (210, 211). Furthermore, the upregulation of VIM and ZEB1, along with the downregulation of CDH1 across multiple enriched terms, suggests that EMT may play a role in cetuximab resistance (212). Cetuximab may also contribute to resistance by inducing immune suppression through blockade of EGFR signaling (213). Given that EGFR inhibition can suppress the MAPK pathway, our findings are consistent with this mechanism as only partial activation of the MAP4K1-mediated “MAPK signaling pathway” was observed (214). Several immune-related terms such as “B-cell signaling pathways” and “T-cell signaling pathway” were also partially active, and key components of the AP-1 TF complex (JUN, FOS, and FOSL1), which are involved in immune signaling via MAPK, were downregulated within these terms (215). However, LCK, which is a Src family kinase, was upregulated in immune-related pathways either independently or in coordination with FYN, which is another Src kinase. Both FYN and LCK play key roles in T-cell activation (216). This pattern suggests that rather than triggering a full immune response cetuximab-resistant cell lines may exist in a partially immune-like signaling state by selectively activating Src kinases to support survival and cytoskeletal adaptation (217).

Our enrichment results for 5-FU revealed a contrasting pattern compared to paclitaxel and cetuximab with a large number of pathways being active and

containing numerous upregulated genes. Moreover, CSC markers such as ALDH1A1, ALDH1A3, SOX9, KLF5, and ITGB1 were upregulated across multiple enriched terms (218-222). These findings suggest that 5-FU-resistant cell lines may exhibit a stem-like phenotype, which is commonly associated with enhanced drug resistance (223, 224).

Enrichment analyses for 5-FU revealed that genes uniquely associated with 5-FU, identified from the union of gene features across 13 DNA replication inhibitors, were represented in the enriched terms. Among these genes, FET indicated a significant association between CNA status and drug resistance for PINK1, PREX1, TNFSF13B, and MYO1F, suggesting that both CNAs and transcriptional regulation of these genes may contribute to 5-FU-specific resistance. In our enrichment results, PINK1 was upregulated, whereas the other three genes were downregulated. The association between PINK1 upregulation and 5-FU resistance has been previously reported. Liu et al. (2024) demonstrated that reduced expression of the lncRNA ABALON led to PINK1 upregulation, promoting mitophagy and contributing to 5-FU resistance (225). Similarly, Yamashita et al. (2017) reported that PINK1 upregulation was linked to increased mitophagy, 5-FU resistance, and poor prognosis in esophageal squamous cell carcinoma (226). However, to our knowledge, no prior studies have shown that PINK1 CNAs may also contribute to 5-FU resistance. Therefore, this study is the first to suggest that both transcriptional upregulation and CNAs in PINK1 may be jointly associated with resistance to 5-FU.

Our enrichment results for gemcitabine based on the adjusted p-value were inconclusive as only SLFN11 met the threshold for differential expression under this criterion. Consequently, enrichment analysis was performed using genes identified based on the unadjusted p-value. The results suggest that gemcitabine likely exhibits cancer-type-specific activity and involves diverse resistance mechanisms. An alternative explanation is that epigenetic and post-transcriptional regulation, along with genomic alterations, may play a more prominent role in gemcitabine resistance than transcriptional changes alone (227-229).

The EM, EC and EMC pan-drug models exhibited variable performances in drug response according to the drugs across the four strategies used. For PDX, the EM models generally did not perform well. However, EC and EMC models were more robust. This demonstrated that CNAs are critical for building more powerful models for response prediction in the PDX dataset. It is usual for CNAs to be more important in experiments such as PDX, where controlled trials with less diversity in the lab environment are conducted (230). Immunosuppressed mice may not show intra-tumor heterogeneity compared to real patients. That is, they may all represent a single clonal population because tumors that are less dominant by selective pressure may not survive (231). This may have made the existing CNAs even more dominant. For the mutation, the opposite may have happened. They may have a lower mutation burden compared to actual patients. Because they are immunosuppressed, immune surveillance is weak and there are fewer environmental mutagenic factors (232). The small number of mutations in subclones may not contribute to drug response prediction because they do not show a dominant effect. Moreover, mutations may not be neutral in driver genes but in genes with no functional effect (233). In some cases, the EMC is the strongest model, indicating that mutations may also complement CNAs. However, their small number in large datasets with many drugs may indicate that the signals from them are diluted and cannot perform well in prediction with gene expression data alone. Since the most successful models in TL are those with the highest generalizability, it is not surprising to see poor performance of EM models in the PDX datasets.

In contrast to the PDX, there were drugs in which the mutation was highly predictive of response in TCGA response predictions. For example, the EM models for 5-FU and cetuximab were the strongest. In human patients, the number and even the impact of mutations may be higher because of the greater diversity and a truly functioning TME working together with the immune response (234). Moreover, in human patients, mutations that confer advantage among sub-clones may not be easily dominated by CNAs, which are dominant as in PDX (235). Driver genes are more

likely to have mutations. The development of resistance, especially in driver genes, may have been caused by mutations due to the application of drugs applied in the clinic for mutations in patients (233). Hypermutation is also high in some cohorts. CNAs are common with mutations in the TCGA datasets. Since sparsity is less than the PDX, response predictions are more realistic. The EMC models in which mutations complement CNAs are also seen in the TCGA, especially in cohort samples. This may be because the model can generalize cohort-specific mutations. Variable samples from different cohorts may dilute the complement signal between mutations and CNAs. Since the accumulation of mutations in cohorts may differ, cohort-specific model testing may contribute to more realistic assessments and higher performance (236).

There was a difference in the diversity of samples between the PDX and TCGA datasets. PDX consisted of one or two cohorts, while TCGA consisted of a larger number of cohorts. This is not unexpected across strategies in two datasets of different models when assessing the performance of all samples. What is interesting, however, is that different drug responses are found in the datasets even when samples from the same cohort are used. Paclitaxel was effective in a small number of samples in the PDX BRCA cohort, while the opposite was observed in the TCGA BRCA cohort. It is difficult for the trained model to predict successfully in both datasets in such a situation. However, generalizable, large and high diversity models can make strong predictions for both. Since such training datasets were constructed with the approach used in the study, strong predictions could be made. TL thus showed that despite class imbalance and low sample size, the datasets can be powerfully used in drug response prediction.

In this study, we integrated gene expression, somatic mutation, CNA, and drug response data from GDSC, PDX, and TCGA using a DNN-based TL approach. During the training phase, we combined drugs targeting mitosis, cytoskeleton, DNA replication, and EGFR signaling pathways from the GDSC dataset, and subsequently performed drug response prediction on independent PDX and TCGA test datasets.

Our models outperformed MOLI on both test datasets in terms of AUCPR for all six drugs assessed by MOLI. In particular, our models achieved superior performance for erlotinib and cetuximab, for which MOLI also applied a TL framework to predict responses in the PDX dataset. This underscores the effectiveness of training on pathway-specific drugs within a TL context. However, further improvements in predictive performance will likely require larger sample sizes (237). Moreover, the integration of additional omics layers such as metabolomics and proteomics may enhance prediction accuracy (238). Nevertheless, our findings indicate that incorporating all omics layers does not necessarily lead to improved model performance. The integration of multiple omics types increases model complexity, and sparse data across omics layers may impede effective learning. Thus, the inclusion of more data types does not inherently result in more accurate predictions (239). Moreover, this study did not incorporate drug-specific features such as structural similarity or target profiles. Incorporating such drug-based features in the training process could lead to the development of more robust and generalizable predictive models (240).

6. CONCLUSION

In this thesis, we demonstrated that MI plays a crucial role in achieving robust and high-performance drug response prediction. To this end, pan-drug training datasets were generated using TL-based DNNs with the GDSC data and robustly validated on in vivo PDX and ex vivo TCGA datasets.

Secondly, pathway enrichment analyses were used to uncover the resistance mechanisms of paclitaxel, 5-FU, gemcitabine, and cetuximab. By comparing enrichment results across different strategies, we identified general resistance mechanisms that were shared across strategies for the four drugs, and drug-specific resistance mechanisms that emerged only in the presence of drug response data for paclitaxel and 5-FU.

Thirdly, we investigated whether drug-specific resistance mechanisms were driven only by transcriptional changes or also influenced by genomic alterations. FET revealed that genomic changes played a significant role in some drug-specific resistance mechanisms, suggesting their importance beyond transcriptional regulation.

Fourth, we explored whether resistance mechanisms were attributable to the activity of individual genes or the collective effect of genes within pathways. In addition, we examined the interaction of resistance with mutations and CNAs. Using CoxPH regression models, we found that in some cases, individual genes better explained resistance, while in others, the combined effect of genes within pathways was more explanatory. Furthermore, it was determined that mutations and CNAs could have an interaction effect on resistance, suggesting that genomic alterations contribute significantly to the development of resistance.

Finally, we aimed to determine whether genes associated with resistance could also predict survival outcomes, which might be as both predictive and prognostic markers. Using the survival data from the TCGA dataset, LRT was performed and

KM survival curves were generated. Our findings revealed that genomic alterations and resistance were associated with poorer survival outcomes, indicating their dual role as predictive and prognostic markers.

In conclusion, this thesis provides a comprehensive approach to drug response prediction and elucidates drug resistance mechanisms. The identified resistance mechanisms, particularly for paclitaxel and 5-FU, hold potential for improving targeted therapy development, enabling personalized treatment opportunities, serving as biomarkers, and ultimately increasing treatment success.

7. REFERENCES

1. Sandhu APS, Tanvir ., Singh K, Singh S, Antaal H, Luthra S, et al. Decoding Cancer Risk: Understanding Gene-Environment Interactions in Cancer Development. *Cureus*. 2024;16(7).
2. Takabe K, Benesch MGK. Types of Cancer and Research Covered in World Journal of Oncology. *World J Oncol* [Internet]. 2022;13(6):325-8. Available from: <https://www.doi.org/10.14740/wjon1558>
3. Pinedo HM, Giaccone G. Chemotherapy. *Lancet*. 1997;349(SUPPL.2):7-9.
4. Duffy MJ, Crown J. A Personalized Approach to Cancer Treatment: How Biomarkers Can Help. *Clin Chem* [Internet]. 2008 Nov 1;54(11):1770-9. Available from: <https://doi.org/10.1373/clinchem.2008.110056>
5. Liu B, Zhou H, Tan L, Siu KTH, Guan XY. Exploring treatment options in cancer: Tumor treatment strategies. *Signal Transduct Target Ther* [Internet]. 2024;9(1). Available from: <http://dx.doi.org/10.1038/s41392-024-01856-7>
6. Liu J, Dang H, Wang XW. The significance of intertumor and intratumor heterogeneity in liver cancer. *Exp Mol Med* [Internet]. 2018;50(1):e416-8. Available from: <http://dx.doi.org/10.1038/emmm.2017.165>
7. Dagogo-Jack I, Shaw AT. Tumour heterogeneity and resistance to cancer therapies. *Nat Rev Clin Oncol*. 2018 Feb;15(2):81-94.
8. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, et al. Genomics of Drug Sensitivity in Cancer (GDSC): A resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res*. 2013;41(D1):955-61.
9. Gao H, Korn JM, Ferretti S, Monahan JE, Wang Y, Singh M, et al. High-throughput screening using patient-derived tumor xenografts to predict clinical trial drug response. *Nat Med*. 2015 Nov;21(11):1318-25.
10. McLendon R, Friedman A, Bigner D, Van Meir EG, Brat DJ, M. Mastrogiannis G, et al. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* [Internet]. 2008;455(7216):1061-8. Available from: <https://doi.org/10.1038/nature07385>
11. Ulgen E, Ozisik O, Sezerman OU. PathfindR: An R package for comprehensive identification of enriched pathways in omics data through active subnetworks. *Front Genet*. 2019;10(SEP):1-3
12. Brown JS, Amend SR, Austin RH, Gatenby RA, Hammarlund EU, Pienta KJ. Updating the Definition of Cancer. *Mol Cancer Res*. 2023 Nov;21(11):1142-7.
13. Somarelli JA. The Hallmarks of Cancer as Ecologically Driven Phenotypes. *Front Ecol Evol*. 2021 Apr;9.
14. Curtin SC, Tejada-Vera B, Bastian BA. Deaths: Leading Causes for 2020. *Natl vital Stat reports from Centers Dis Control Prev Natl Cent Heal Stat Natl Vital Stat Syst*. 2023 Dec;72(13):1-115.
15. Mattiuzzi C, Lippi G. Leading Causes of US Deaths in the 2022. *J Clin Med*. 2024 Nov;13(23).
16. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. *CA Cancer J Clin*. 2024;74(1):12-49.
17. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011 Mar;144(5):646-74.

18. Debela DT, Muzazu SG, Heraro KD, Ndalama MT, Mesele BW, Haile DC, et al. New approaches and procedures for cancer treatment: Current perspectives. *SAGE open Med.* 2021;9:20503121211034370.
19. Gambardella V, Tarazona N, Cejalvo JM, Lombardi P, Huerta M, Roselló S, et al. Personalized Medicine: Recent Progress in Cancer Therapy. *Cancers (Basel).* 2020 Apr;12(4).
20. Mansoori B, Mohammadi A, Davudian S, Shirjang S, Baradaran B. The Different Mechanisms of Cancer Drug Resistance: A Brief Review. *Adv Pharm Bull.* 2017 Sep;7(3):339-48.
21. Khan SU, Fatima K, Aisha S, Malik F. Unveiling the mechanisms and challenges of cancer drug resistance. *Cell Commun Signal [Internet].* 2024;22(1):109. Available from: <https://doi.org/10.1186/s12964-023-01302-1>
22. Burotto M, Wilkerson J, Stein WD, Bates SE, Fojo T. Adjuvant and neoadjuvant cancer therapies: A historical review and a rational approach to understand outcomes. *Semin Oncol.* 2019 Feb;46(1):83-99.
23. Hossain MB, Haldar Neer AH. Chemotherapy. *Cancer Treat Res.* 2023;185:49-58.
24. Amjad MT, Chidharla A, Kasi A. Cancer chemotherapy [Internet]. Treasure Island (FL): StatPearls Publishing; 2025 Jan- [updated 2023 Feb 27; cited 2025 Jul 4]. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK564367/>
25. Strobel H, Baisch T, Fitzel R, Schilberg K, Siegelin MD, Karpel-Massler G, et al. Temozolomide and Other Alkylating Agents in Glioblastoma Therapy. *Biomedicines.* 2019 Sep;7(3).
26. Al-Badr AA, Alodhaib MM. Dacarbazine. *Profiles Drug Subst Excip Relat Methodol.* 2016;41:323-77.
27. Anand U, Dey A, Chandel AKS, Sanyal R, Mishra A, Pandey DK, et al. Cancer chemotherapy and beyond: Current status, drug candidates, associated risks and progress in targeted therapeutics. *Genes Dis.* 2023 Jul;10(4):1367-401.
28. Minotti G, Menna P, Salvatorelli E, Cairo G, Gianni L. Anthracyclines: molecular advances and pharmacologic developments in antitumor activity and cardiotoxicity. *Pharmacol Rev.* 2004 Jun;56(2):185-229.
29. Balachandran RS, Kipreos ET. Addressing a weakness of anticancer therapy with mitotic inhibitors: Mitotic slippage. *Mol Cell Oncol.* 2017;4(2):e1277293.
30. Jang JY, Kim D, Kim ND. Recent Developments in Combination Chemotherapy for Colorectal and Breast Cancers with Topoisomerase Inhibitors. Vol. 24, *International Journal of Molecular Sciences.* 2023.
31. Zhang C, Xu C, Gao X, Yao Q. Platinum-based drugs for cancer therapy and anti-tumor strategies. *Theranostics.* 2022;12(5):2115-32.
32. Sahu M, Suryawanshi H. Immunotherapy: The future of cancer treatment. *J Oral Maxillofac Pathol.* 2021;25(2):371.
33. Cheever MA, Higano CS. PROVENGE (Sipuleucel-T) in prostate cancer: the first FDA-approved therapeutic cancer vaccine. *Clin cancer Res an Off J Am Assoc Cancer Res.* 2011 Jun;17(11):3520-6.
34. Du S, Yan J, Xue Y, Zhong Y, Dong Y. Adoptive cell therapy for cancer treatment. *Explor (Beijing, China).* 2023 Aug;3(4):20210058.

35. Sterner RC, Sterner RM. CAR-T-cell therapy: current limitations and potential strategies. *Blood Cancer J* [Internet]. 2021;11(4). Available from: <http://dx.doi.org/10.1038/s41408-021-00459-7>
36. Conlon KC, Miljkovic MD, Waldmann TA. Cytokines in the Treatment of Cancer. *J Interf cytokine Res Off J Int Soc Interf Cytokine Res*. 2019 Jan;39(1):6-21.
37. Gemmete JJ, Mukherji SK. Trastuzumab (herceptin). *AJNR Am J Neuroradiol*. 2011 Sep;32(8):1373-4.
38. Lin D, Shen Y, Liang T. Oncolytic virotherapy: basic principles, recent advances and future directions. *Signal Transduct Target Ther*. 2023;8(1).
39. Ferrucci PF, Pala L, Conforti F, Cocorocchio E. Talimogene Laherparepvec (T-VEC): An Intralesional Cancer Immunotherapy for Advanced Melanoma. *Cancers (Basel)*. 2021 Mar;13(6).
40. Smith CEP, Prasad V. Targeted Cancer Therapies. *Am Fam Physician*. 2021 Feb;103(3):155-63.
41. Yang Z, Hackshaw A, Feng Q, Fu X, Zhang Y, Mao C, et al. Comparison of gefitinib, erlotinib and afatinib in non-small cell lung cancer: A meta-analysis. *Int J cancer*. 2017 Jun;140(12):2805-19.
42. Sabnis AJ, Bivona TG. Principles of Resistance to Targeted Cancer Therapy: Lessons from Basic and Translational Cancer Biology. *Trends Mol Med*. 2019 Mar;25(3):185-97.
43. Jordan VC. The role of tamoxifen in the treatment and prevention of breast cancer. *Curr Probl Cancer*. 1992;16(3):129-76.
44. Choi YH, Yu AM. ABC transporters in multidrug resistance and pharmacokinetics, and strategies for drug development. *Curr Pharm Des*. 2014;20(5):793-807.
45. Girardi E, César-Razquin A, Lindinger S, Papakostas K, Konecka J, Hemmerich J, et al. A widespread role for SLC transmembrane transporters in resistance to cytotoxic drugs. *Nat Chem Biol*. 2020 Apr;16(4):469-78.
46. Liu X. ABC Family Transporters. *Adv Exp Med Biol*. 2019;1141:13-100.
47. Wang K, Zhu X, Yin Y. Maslinic Acid Enhances Docetaxel Response in Human Docetaxel-Resistant Triple Negative Breast Carcinoma MDA-MB-231 Cells via Regulating MELK-FoxM1-ABCB1 Signaling Cascade. *Front Pharmacol*. 2020;11:835.
48. Klejewski A, Świerczewska M, Zaorska K, Brązert M, Nowicki M, Zabel M, et al. New and Old Genes Associated with Topotecan Resistance Development in Ovarian Cancer Cell Lines. *Anticancer Res*. 2017 Apr;37(4):1625-36.
49. Besse A, Stolze SC, Rasche L, Weinhold N, Morgan GJ, Kraus M, et al. Carfilzomib resistance due to ABCB1/MDR1 overexpression is overcome by nelfinavir and lopinavir in multiple myeloma. *Leukemia*. 2018 Feb;32(2):391-401.
50. Puris E, Fricker G, Gynther M. The Role of Solute Carrier Transporters in Efficient Anticancer Drug Delivery and Therapy. *Pharmaceutics*. 2023 Jan;15(2).
51. Shuel SL. Targeted cancer therapies: Clinical pearls for primary care. *Can Fam Physician*. 2022 Jul;68(7):515-8.
52. Morales ML, García-Vicente R, Rodríguez-García A, Reyes-Palomares A, Vincelle-Nieto Á, Álvarez N, et al. Posttranslational splicing modifications as a key mechanism in cytarabine resistance in acute myeloid leukemia. *Leukemia* [Internet]. 2023;37(8):1649-59. Available from: <https://doi.org/10.1038/s41375-023-01963-4>

53. Roszkowska M. Multilevel Mechanisms of Cancer Drug Resistance. *Int J Mol Sci.* 2024 Nov;25(22).
54. Montoya S, Soong D, Nguyen N, Affer M, Munamarty SP, Taylor J. Targeted Therapies in Cancer: To Be or Not to Be, Selective. *Biomedicines.* 2021 Nov;9(11).
55. Araki T, Kanda S, Horinouchi H, Ohe Y. Current treatment strategies for EGFR-mutated non-small cell lung cancer: from first line to beyond osimertinib resistance. *Jpn J Clin Oncol.* 2023 Jun;53(7):547-61.
56. Marinello J, Delcuratolo M, Capranico G. Anthracyclines as Topoisomerase II Poisons: From Early Studies to New Perspectives. *Int J Mol Sci.* 2018 Nov;19(11).
57. Clusan L, Le Goff P, Flouriot G, Pakdel F. A Closer Look at Estrogen Receptor Mutations in Breast Cancer and Their Implications for Estrogen and Antiestrogen Responses. *Int J Mol Sci.* 2021 Jan;22(2).
58. Woods D, Turchi JJ. Chemotherapy induced DNA damage response: convergence of drugs and pathways. *Cancer Biol Ther.* 2013 May;14(5):379-89.
59. Duan M, Ulibarri J, Liu KJ, Mao P. Role of Nucleotide Excision Repair in Cisplatin Resistance. *Int J Mol Sci.* 2020 Dec;21(23).
60. Collins SP, Dritschilo A. The mismatch repair and base excision repair pathways: an opportunity for individualized (personalized) sensitization of cancer therapy. *Cancer Biol Ther.* 2009 Jun;8(12):1164-6.
61. Chernikova SB, Game JC, Brown JM. Inhibiting homologous recombination for cancer therapy. *Cancer Biol Ther.* 2012 Jan;13(2):61-8.
62. Huang R, Zhou PK. DNA damage repair: historical perspectives, mechanistic pathways and clinical translation for targeted cancer therapy. *Signal Transduct Target Ther [Internet].* 2021;6(1):254. Available from: <https://doi.org/10.1038/s41392-021-00648-7>
63. Li LY, Guan Y di, Chen XS, Yang JM, Cheng Y. DNA Repair Pathways in Cancer Therapy and Resistance. *Front Pharmacol.* 2020;11:629266.
64. Verbinen I, Boens S, Ferreira M, Székér K, Van Wijk L, Van Eynde A, et al. Enhanced DNA-repair capacity and resistance to chemically induced carcinogenesis upon deletion of the phosphatase regulator NIPP1. *Oncogenesis [Internet].* 2020;9(3):30. Available from: <https://doi.org/10.1038/s41389-020-0214-3>
65. Britten RA, Liu D, Tessier A, Hutchison MJ, Murray D. ERCC1 expression as a molecular marker of cisplatin resistance in human cervical tumor cells. *Int J cancer.* 2000 Sep;89(5):453-7.
66. Higuchi F, Nagashima H, Ning J, Koerner MVA, Wakimoto H, Cahill DP. Restoration of Temozolomide Sensitivity by PARP Inhibitors in Mismatch Repair Deficient Glioblastoma is Independent of Base Excision Repair. *Clin cancer Res an Off J Am Assoc Cancer Res.* 2020 Apr;26(7):1690-9.
67. Yamanishi K, Shen CS, Mizutani H. Apoptosis in the epidermis. *Methods Mol Biol.* 2005;289:171-4.
68. O'Leary B, Skinner H, Schoenfeld JD, Licitra L, Le Tourneau C, Esdar C, et al. Evasion of apoptosis and treatment resistance in squamous cell carcinoma of the head and neck. *Cancer Treat Rev.* 2024 Sep;129:102773.

69. O'Brien MA, Kirby R. Apoptosis: A review of pro-apoptotic and anti-apoptotic pathways and dysregulation in disease. Vol. 18, *Journal of Veterinary Emergency and Critical Care*. 2008. p. 572-85.
70. Vinay DS, Ryan EP, Pawelec G, Talib WH, Stagg J, Elkord E, et al. Immune evasion in cancer: Mechanistic basis and therapeutic strategies. *Semin Cancer Biol*. 2015 Dec;35 Suppl:S185-98.
71. Gong Y, Zhang Z. Alternative signaling pathways: when, where and why? *FEBS Lett*. 2005 Oct;579(24):5265-74.
72. Martin GS. Cell signaling and cancer. *Cancer Cell*. 2003 Sep;4(3):167-74.
73. Yang M, Mandal E, Liu FX, O'Hara RMJ, Leshner B, Sanborn RE. Non-small cell lung cancer with MET amplification: review of epidemiology, associated disease characteristics, testing procedures, burden, and treatments. *Front Oncol*. 2023;13:1241402.
74. Fang W, Huang Y, Gu W, Gan J, Wang W, Zhang S, et al. PI3K-AKT-mTOR pathway alterations in advanced NSCLC patients after progression on EGFR-TKI and clinical response to EGFR-TKI plus everolimus combination therapy. *Transl lung cancer Res*. 2020 Aug;9(4):1258-67.
75. Ruiz-Saenz A, Dreyer C, Campbell MR, Steri V, Gulizia N, Moasser MM. HER2 Amplification in Tumors Activates PI3K/Akt Signaling Independent of HER3. *Cancer Res*. 2018 Jul;78(13):3645-58.
76. Sweetlove M, Wrightson E, Kolekar S, Rewcastle GW, Baguley BC, Shepherd PR, et al. Inhibitors of pan-PI3K Signaling Synergize with BRAF or MEK Inhibitors to Prevent BRAF-Mutant Melanoma Cell Growth. *Front Oncol*. 2015;5:135.
77. Talwelkar SS, Mäyränpää MI, Schüller J, Linnavirta N, Hemmes A, Adinolfi S, et al. PI3K β inhibition enhances ALK-inhibitor sensitivity in ALK-rearranged lung cancer. *Mol Oncol*. 2023 May;17(5):747-64.
78. Jurj A, Ionescu C, Berindan-Neagoe I, Braicu C. The extracellular matrix alteration, implication in modulation of drug resistance mechanism: friends or foes? *J Exp Clin Cancer Res*. 2022 Sep;41(1):276.
79. Zhang Z, Wang R, Chen L. Drug Delivery System Targeting Cancer-Associated Fibroblast for Improving Immunotherapy. *Int J Nanomedicine*. 2025;20:483-503.
80. Chen Z, Han F, Du Y, Shi H, Zhou W. Hypoxic microenvironment in cancer: molecular mechanisms and therapeutic interventions. *Signal Transduct Target Ther*. 2023 Feb;8(1):70.
81. Tie Y, Tang F, Wei YQ, Wei XW. Immunosuppressive cells in cancer: mechanisms and potential therapeutic targets. *J Hematol Oncol*. 2022 May;15(1):61.
82. Prakash J, Shaked Y. The Interplay between Extracellular Matrix Remodeling and Cancer Therapeutics. *Cancer Discov*. 2024 Aug;14(8):1375-88.
83. Al Aboud NM, Tupper C, Jialal I. Genetics, Epigenetic Mechanism. In *Treasure Island (FL)*; 2025.
84. Adhikari S, Bhattacharya A, Adhikary S, Singh V, Gadad SS, Roy S, et al. The paradigm of drug resistance in cancer: an epigenetic perspective. *Biosci Rep*. 2022 Apr;42(4).
85. Esteller M. Epigenetics provides a new generation of oncogenes and tumour-suppressor genes. *Br J Cancer*. 2006 Jan;94(2):179-83.

86. Wang X, Wang Y, Xie M, Ma S, Zhang Y, Wang L, et al. Hypermethylation of CDKN2A CpG island drives resistance to PRC2 inhibitors in SWI/SNF loss-of-function tumors. *Cell Death Dis.* 2024 Nov;15(11):794.
87. Wang N, Ma T, Yu B. Targeting epigenetic regulators to overcome drug resistance in cancers. *Signal Transduct Target Ther.* 2023 Feb;8(1):69.
88. Jabbour AM, Heraud JE, Daunt CP, Kaufmann T, Sandow J, O'Reilly LA, et al. Puma indirectly activates Bax to cause apoptosis in the absence of Bid or Bim. *Cell Death Differ.* 2009 Apr;16(4):555-63.
89. Kang N, Eccleston M, Clermont PL, Latarani M, Male DK, Wang Y, et al. EZH2 inhibition: a promising strategy to prevent cancer immune editing. *Epigenomics.* 2020 Aug;12(16):1457-76.
90. Chawra HS, Agarwal M, Mishra A, Chandel SS, Singh RP, Dubey G, et al. MicroRNA-21's role in PTEN suppression and PI3K/AKT activation: Implications for cancer biology. *Pathol Res Pract.* 2024 Feb;254:155091.
91. Bhan A, Mandal SS. LncRNA HOTAIR: A master regulator of chromatin dynamics and cancer. *Biochim Biophys Acta.* 2015 Aug;1856(1):151-64.
92. Liberti M V, Locasale JW. The Warburg Effect: How Does it Benefit Cancer Cells? *Trends Biochem Sci.* 2016 Mar;41(3):211-8.
93. Li X, Yang Y, Zhang B, Lin X, Fu X, An Y, et al. Lactate metabolism in human health and disease. *Signal Transduct Target Ther.* 2022;7(1).
94. Matés JM, Campos-Sandoval JA, Santos-Jiménez J de los, Márquez J. Dysregulation of glutaminase and glutamine synthetase in cancer. *Cancer Lett [Internet].* 2019;467:29-39. Available from: <https://www.sciencedirect.com/science/article/pii/S0304383519304835>
95. Koundouros N, Poulogiannis G. Reprogramming of fatty acid metabolism in cancer. *Br J Cancer [Internet].* 2020;122(1):4-22. Available from: <http://dx.doi.org/10.1038/s41416-019-0650-z>
96. Jin P, Jiang J, Zhou L, Huang Z, Nice EC, Huang C, et al. Mitochondrial adaptation in cancer drug resistance: prevalence, mechanisms, and management. Vol. 15, *Journal of Hematology and Oncology.* 2022. 1-42 p.
97. Dart A. EMT in chemoresistance. *Nat Rev Cancer [Internet].* 2023;23(6):349. Available from: <https://doi.org/10.1038/s41568-023-00581-7>
98. Du B, Shim JS. Targeting Epithelial-Mesenchymal Transition (EMT) to Overcome Drug Resistance in Cancer. *Molecules.* 2016 Jul;21(7).
99. De Las Rivas J, Brozovic A, Izraely S, Casas-Pais A, Witz IP, Figueroa A. Cancer drug resistance induced by EMT: novel therapeutic strategies. *Arch Toxicol.* 2021 Jul;95(7):2279-97.
100. Lee AF, Chen MC, Chen CJ, Yang CJ, Huang MS, Liu YP. Reverse epithelial-mesenchymal transition contributes to the regain of drug sensitivity in tyrosine kinase inhibitor-resistant non-small cell lung cancer cells. *PLoS One.* 2017;12(7):e0180383.
101. Wu D ming, Zhang T, Liu Y bin, Deng S hua, Han R, Liu T, et al. The PAX6-ZEB2 axis promotes metastasis and cisplatin resistance in non-small cell lung cancer through PI3K/AKT signaling. *Cell Death Dis [Internet].* 2019;10(5):349. Available from: <https://doi.org/10.1038/s41419-019-1591-4>

102. Namba K, Shien K, Takahashi Y, Torigoe H, Sato H, Yoshioka T, et al. Activation of AXL as a Preclinical Acquired Resistance Mechanism Against Osimertinib Treatment in EGFR-Mutant Non-Small Cell Lung Cancer Cells. *Mol Cancer Res*. 2019 Feb;17(2):499-507.
103. Olmeda D, Moreno-Bueno G, Flores JM, Fabra A, Portillo F, Cano A. SNAIL is required for tumor growth and lymph node metastasis of human breast carcinoma MDA-MB-231 cells. *Cancer Res*. 2007 Dec;67(24):11721-31.
104. Kriegeskorte N, Golan T. Neural network models and deep learning. *Curr Biol*. 2019 Apr;29(7):R231-6.
105. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015 May;521(7553):436-44.
106. Almeida JS. Predictive non-linear modeling of complex data by artificial neural networks. *Curr Opin Biotechnol*. 2002 Feb;13(1):72-6.
107. Amrouche M, Stipanovic DM. A Formal Characterization of Activation Functions in Deep Neural Networks. *IEEE Trans neural networks Learn Syst*. 2024 Feb;35(2):2153-66.
108. Seo H, Bassenne M, Xing L. Closing the Gap Between Deep Neural Network Modeling and Biomedical Decision-Making Metrics in Segmentation via Adaptive Loss Functions. *IEEE Trans Med Imaging*. 2021 Feb;40(2):585-93.
109. Kalliola J, Kapočiūtė-Dzikienė J, Damaševičius R. Neural network hyperparameter optimization for prediction of real estate prices in Helsinki. *PeerJ Comput Sci*. 2021;7:e444.
110. Kang M, Ko E, Mersha TB. A roadmap for multi-omics data integration using deep learning. *Brief Bioinform*. 2022 Jan;23(1).
111. Theodoris C V, Xiao L, Chopra A, Chaffin MD, Al Sayed ZR, Hill MC, et al. Transfer learning enables predictions in network biology. *Nature*. 2023 Jun;618(7965):616-24.
112. Abu M, Zahri NAH, Amir A, Ismail MI, Yaakub A, Anwar SA, et al. A Comprehensive Performance Analysis of Transfer Learning Optimization in Visual Field Defect Classification. *Diagnostics (Basel, Switzerland)*. 2022 May;12(5).
113. Zhang X, Xue Y, Su X, Chen S, Liu K, Chen W, et al. A Transfer Learning Approach to Correct the Temporal Performance Drift of Clinical Prediction Models: Retrospective Cohort Study. *JMIR Med informatics*. 2022 Nov;10(11):e38053.
114. Nanni L, Interlenghi M, Brahnam S, Salvatore C, Papa S, Nemni R, et al. Comparison of Transfer Learning and Conventional Machine Learning Applied to Structural Brain MRI for the Early Diagnosis and Prognosis of Alzheimer's Disease. *Front Neurol*. 2020;11:576194.
115. Chen NC, Chyau CC, Lee YJ, Tseng HC, Chou FP. Promotion of mitotic catastrophe via activation of PTEN by paclitaxel with supplement of mulberry water extract in bladder cancer cells. *Sci Rep*. 2016;6(February):1-13.
116. Sharifi-Noghabi H, Zolotareva O, Collins CC, Ester M. MOLI: Multi-omics late integration with deep neural networks for drug response prediction. In: *Bioinformatics*. 2019.
117. Brazma A, Parkinson H, Sarkans U, Shojatalab M, Vilo J, Abeygunawardena N, et al. ArrayExpress--a public repository for microarray gene expression data at the EBI. *Nucleic Acids Res*. 2003 Jan;31(1):68-71.
118. Gautier L, Cope L, Bolstad BM, Irizarry RA. affy--analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics*. 2004 Feb;20(3):307-15.

119. Eschrich SA, Hoerter AM, Bloom GC, Fenstermacher DA. Tissue-specific RMA models to incrementally normalize Affymetrix GeneChip data. *Annu Int Conf IEEE Eng Med Biol Soc IEEE Eng Med Biol Soc Annu Int Conf*. 2008;2008:2419-22.
120. Sui Y, Zhao X, Speed TP, Wu Z. Background adjustment for DNA microarrays using a database of microarray experiments. *J Comput Biol a J Comput Mol cell Biol*. 2009 Nov;16(11):1501-15.
121. Dai M, Wang P, Boyd AD, Kostov G, Athey B, Jones EG, et al. Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic Acids Res [Internet]*. 2005 Nov 1;33(20):e175-e175. Available from: <https://doi.org/10.1093/nar/gni179>
122. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics [Internet]*. 2008;9(1):559. Available from: <https://doi.org/10.1186/1471-2105-9-559>
123. Broad Institute TCGA Genome Data Analysis Center. Firehose stddata analysis. Broad Institute of MIT and Harvard; [cited 2024 Dec 14]. Available from: <https://gdac.broadinstitute.org/>
124. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics*. 2012 Mar;28(6):882-3.
125. Chakravarty D, Gao J, Phillips SM, Kundra R, Zhang H, Wang J, et al. OncoKB: A Precision Oncology Knowledge Base. *JCO Precis Oncol*. 2017 Jul;2017.
126. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal*. 2013 Apr;6(269):p11.
127. Tate JG, Bamford S, Jubb HC, Sondka Z, Beare DM, Bindal N, et al. COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res*. 2019 Jan;47(D1):D941-7.
128. Bailey MH, Tokheim C, Porta-Pardo E, Sengupta S, Bertrand D, Weerasinghe A, et al. Comprehensive Characterization of Cancer Driver Genes and Mutations. *Cell*. 2018;173(2):371-385.e18.
129. Altshuler DM, Gibbs RA, Peltonen L, Altshuler DM, Gibbs RA, Peltonen L, et al. Integrating common and rare genetic variation in diverse human populations. *Nature*. 2010 Sep;467(7311):52-8.
130. Olshen AB, Venkatraman ES, Lucito R, Wigler M. Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics*. 2004 Oct;5(4):557-72.
131. The Genome Reference Consortium. Genome Reference Consortium Human Build 37 (GRCh37). Bethesda, MD: National Center for Biotechnology Information (NCBI); [cited 2024 Dec 14]. Available from: https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.13.
132. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47.
133. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, et al. PyTorch: An imperative style, high-performance deep learning library. *Adv Neural Inf Process Syst*. 2019;32(NeurIPS).

134. Kim M, Hwang KB. An empirical evaluation of sampling methods for the classification of imbalanced data. *PLoS One*. 2022;17(7):e0271260.
135. Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res*. 2012 Jan;40(Database issue):D109-14.
136. Jassal B, Matthews L, Viteri G, Gong C, Lorente P, Fabregat A, et al. The reactome pathway knowledgebase. *Nucleic Acids Res*. 2020 Jan;48(D1):D498-503.
137. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods*. 2020 Mar;17(3):261-72.
138. You M, Xie Z, Zhang N, Zhang Y, Xiao D, Liu S, et al. Signaling pathways in cancer metabolism: mechanisms and therapeutic targets. *Signal Transduct Target Ther*. 2023;8(1).
139. Zhan T, Rindtorff N, Boutros M. Wnt signaling in cancer. *Oncogene*. 2017;36(11):1461-73.
140. Eke I, Cordes N. Focal adhesion signaling and therapy resistance in cancer. *Semin Cancer Biol* [Internet]. 2015;31:65-75. Available from: <https://www.sciencedirect.com/science/article/pii/S1044579X14000984>
141. Kyuno D, Takasawa A, Kikuchi S, Takemasa I, Osanai M, Kojima T. Role of tight junctions in the epithelial-to-mesenchymal transition of cancer cells. *Biochim Biophys Acta - Biomembr* [Internet]. 2021;1863(3):183503. Available from: <https://www.sciencedirect.com/science/article/pii/S0005273620303461>
142. Steinmetz MO, Prota AE. Microtubule-Targeting Agents: Strategies To Hijack the Cytoskeleton. *Trends Cell Biol* [Internet]. 2018 Oct 1;28(10):776-92. Available from: <https://doi.org/10.1016/j.tcb.2018.05.001>
143. Peng Q, Ren B, Xin K, Liu W, Alam MS, Yang Y, et al. CYFIP2 serves as a prognostic biomarker and correlates with tumor immune microenvironment in human cancers. *Eur J Med Res*. 2023 Sep;28(1):364.
144. Chatterjee S, Behnam Azad B, Nimmagadda S. The intricate role of CXCR4 in cancer. *Adv Cancer Res*. 2014;124:31-82.
145. Chédotal A, Kerjan G, Moreau-Fauvarque C. The brain within the tumor: new roles for axon guidance molecules in cancers. *Cell Death Differ* [Internet]. 2005;12(8):1044-56. Available from: <https://doi.org/10.1038/sj.cdd.4401707>
146. Hanahan D, Monje M. Cancer hallmarks intersect with neuroscience in the tumor microenvironment. *Cancer Cell* [Internet]. 2023;41(3):573-80. Available from: <https://doi.org/10.1016/j.ccell.2023.02.012>
147. Wu L, Lian W, Zhao L. Calcium signaling in cancer progression and therapy. *FEBS J*. 2021;288(21):6187-205.
148. Nehme Z, Roehlen N, Dhawan P, Baumert TF. Tight Junction Protein Signaling and Cancer Biology. *Cells*. 2023 Jan;12(2).
149. Aasen T, Mesnil M, Naus CC, Lampe PD, Laird DW. Gap junctions and cancer: communicating for 50 years. *Nat Rev Cancer*. 2016 Dec;16(12):775-88.

150. Li M, Qi L, Xu JB, Zhong LY, Chan S, Chen SN, et al. Methylation of the Promoter Region of the Tight Junction Protein-1 by DNMT1 Induces EMT-like Features in Multiple Myeloma. *Mol Ther oncolytics*. 2020 Dec;19:197-207.
151. Ruffell B, Coussens LM. Macrophages and therapeutic resistance in cancer. *Cancer Cell*. 2015 Apr;27(4):462-72.
152. Guengerich FP. Metabolism of chemical carcinogens. *Carcinogenesis* [Internet]. 2000 Mar 1;21(3):345-51. Available from: <https://doi.org/10.1093/carcin/21.3.345>
153. Kaur P, Shankar E, Gupta S. EZH2-mediated development of therapeutic resistance in cancer. *Cancer Lett*. 2024 Apr;586:216706.
154. Mishra D, Banerjee D. Lactate Dehydrogenases as Metabolic Links between Tumor and Stroma in the Tumor Microenvironment. *Cancers (Basel)*. 2019 May;11(6).
155. Warfel NA, El-Deiry WS. HIF-1 signaling in drug resistance to chemotherapy. *Curr Med Chem*. 2014;21(26):3021-8.
156. Stine ZE, Schug ZT, Salvino JM, Dang C V. Targeting cancer metabolism in the era of precision oncology. *Nat Rev Drug Discov* [Internet]. 2022;21(2):141-62. Available from: <https://doi.org/10.1038/s41573-021-00339-6>
157. Peng S, Fu Y. FYN: emerging biological roles and potential therapeutic targets in cancer. *J Transl Med* [Internet]. 2023;21(1):84. Available from: <https://doi.org/10.1186/s12967-023-03930-0>
158. Bruntz RC, Lindsley CW, Brown HA. Phospholipase D signaling pathways and phosphatidic acid as therapeutic targets in cancer. *Pharmacol Rev*. 2014 Oct;66(4):1033-79.
159. Bansal A, Simon MC. Glutathione metabolism in cancer progression and treatment resistance. *J Cell Biol*. 2018 Jul;217(7):2291-8.
160. Zhou Y, Yao Y, Deng Y, Shao A. Regulation of efferocytosis as a novel cancer therapy. *Cell Commun Signal*. 2020 May;18(1):71.
161. Zeng R, Li H, Jia L, Lee SH, Jiang R, Zhang Y, et al. Association of CYP24A1 with survival and drug resistance in clinical cancer patients: a meta-analysis. *BMC Cancer*. 2022 Dec;22(1):1317.
162. Kumar P, Siripini S, Sreedhar AS. The matrix metalloproteinase 7 (MMP7) links Hsp90 chaperone with acquired drug resistance and tumor metastasis. *Cancer reports (Hoboken, NJ)*. 2022 Dec;5(12):e1261.
163. Feng M, Dong N, Zhou X, Ma L, Xiang R. Myosin light chain 9 promotes the proliferation, invasion, migration and angiogenesis of colorectal cancer cells by binding to Yes-associated protein 1 and regulating Hippo signaling. *Bioengineered*. 2022 Jan;13(1):96-106.
164. Sprowl JA, Reed K, Armstrong SR, Lanner C, Guo B, Kalatskaya I, et al. Alterations in tumor necrosis factor signaling pathways are associated with cytotoxicity and resistance to taxanes: a study in isogenic resistant tumor cells. *Breast Cancer Res*. 2012 Jan;14(1):R2.
165. Pinter M, Jain RK. Targeting the renin-angiotensin system to improve cancer treatment: Implications for immunotherapy. *Sci Transl Med*. 2017 Oct;9(410).
166. Chen KG, Valencia JC, Gillet JP, Hearing VJ, Gottesman MM. Involvement of ABC transporters in melanogenesis and the development of multidrug resistance of melanoma. *Pigment Cell Melanoma Res*. 2009 Dec;22(6):740-9.

167. Ahmed MB, Alghamdi AAA, Islam SU, Lee JS, Lee YS. cAMP Signaling in Cancer: A PKA-CREB and EPAC-Centric Approach. *Cells*. 2022 Jun;11(13).
168. Saitoh M. Involvement of partial EMT in cancer progression. *J Biochem*. 2018 Oct;164(4):257-64.
169. Chopin V, Lagadec C, Toillon RA, Le Bourhis X. Neurotrophin signaling in cancer stem cells. *Cell Mol Life Sci*. 2016 May;73(9):1859-70.
170. Li Y, Yu C, Deng W. Roles and mechanisms of adipokines in drug resistance of tumor cells. *Eur J Pharmacol* [Internet]. 2021;899:174019. Available from: <https://www.sciencedirect.com/science/article/pii/S0014299921001722>
171. Yan C, Li TS. Dual Role of Mitophagy in Cancer Drug Resistance. *Anticancer Res*. 2018 Feb;38(2):617-21.
172. Chow MT, Luster AD. Chemokines in cancer. *Cancer Immunol Res*. 2014 Dec;2(12):1125-31.
173. Yi M, Li T, Niu M, Zhang H, Wu Y, Wu K, et al. Targeting cytokine and chemokine signaling pathways for cancer therapy. *Signal Transduct Target Ther* [Internet]. 2024;9(1):176. Available from: <https://doi.org/10.1038/s41392-024-01868-3>
174. Hassan Ibrahim I, Balah A, Gomaa Abd Elfattah Hassan A, Gamal Abd El-Aziz H. Role of motor proteins in human cancers. *Saudi J Biol Sci* [Internet]. 2022;29(12):103436. Available from: <https://www.sciencedirect.com/science/article/pii/S1319562X22003527>
175. Zhang M, Zhang YY, Chen Y, Wang J, Wang Q, Lu H. TGF- β Signaling and Resistance to Cancer Therapy. *Front cell Dev Biol*. 2021;9:786728.
176. Townsend DM, Tew KD. The role of glutathione-S-transferase in anti-cancer drug resistance. *Oncogene*. 2003 Oct;22(47):7369-75.
177. Zhou J, Kang Y, Chen L, Wang H, Liu J, Zeng S, et al. The Drug-Resistance Mechanisms of Five Platinum-Based Antitumor Agents. *Front Pharmacol*. 2020;11:343.
178. Iglesias JM, Gumuzio J, Martin AG. Linking Pluripotency Reprogramming and Cancer. *Stem Cells Transl Med*. 2017 Feb;6(2):335-9.
179. Liyasova M, Mdmsci MS, Fortier J, Mba RV, Fiala M, Yang L, et al. The 65th ASH Annual Meeting Abstracts. *Blood* [Internet]. 2023;142(November):2292. Available from: <http://dx.doi.org/10.1182/blood-2023-186583>
180. Corda G, Sala A. Non-canonical WNT/PCP signalling in cancer: Fzd6 takes centre stage. *Oncogenesis* [Internet]. 2017;6(7):e364-e364. Available from: <https://doi.org/10.1038/oncsis.2017.69>
181. Bach DH, Luu TTT, Kim D, An YJ, Park S, Park HJ, et al. BMP4 Upregulation Is Associated with Acquired Drug Resistance and Fatty Acid Metabolism in EGFR-Mutant Non-Small-Cell Lung Cancer Cells. *Mol Ther Nucleic Acids*. 2018 Sep;12:817-28.
182. Shi ZD, Pang K, Wu ZX, Dong Y, Hao L, Qin JX, et al. Tumor cell plasticity in targeted therapy-induced resistance: mechanisms and new strategies. *Signal Transduct Target Ther* [Internet]. 2023;8(1):113. Available from: <https://doi.org/10.1038/s41392-023-01383-x>
183. Qin Y, Yu J, Zhang M, Qin F, Lan X. ZEB1 promotes tumorigenesis and metastasis in hepatocellular carcinoma by regulating the expression of vimentin. *Mol Med Rep*. 2019 Mar;19(3):2297-306.

184. Gordon RR, Nelson PS. Cellular senescence and cancer chemotherapy resistance. *Drug Resist Updat Rev Comment Antimicrob Anticancer Chemother*. 2012;15(1-2):123-31.
185. Gao Y, Qiao X, Liu Z, Zhang W. The role of E2F2 in cancer progression and its value as a therapeutic target. *Front Immunol*. 2024;15:1397303.
186. Fenouille N, Tichet M, Dufies M, Pottier A, Mogha A, Soo JK, et al. The epithelial-mesenchymal transition (EMT) regulatory factor SLUG (SNAI2) is a downstream target of SPARC and AKT in promoting melanoma cell invasion. *PLoS One*. 2012;7(7).
187. Brisson L, Bański P, Sboarina M, Dethier C, Danhier P, Fontenille MJ, et al. Lactate Dehydrogenase B Controls Lysosome Activity and Autophagy in Cancer. *Cancer Cell*. 2016 Sep;30(3):418-31.
188. Sun W, Zhang X, Ding X, Li H, Geng M, Xie Z, et al. Lactate dehydrogenase B is associated with the response to neoadjuvant chemotherapy in oral squamous cell carcinoma. *PLoS One*. 2015;10(5):e0125976.
189. Urbańska K, Orzechowski A. Unappreciated Role of LDHA and LDHB to Control Apoptosis and Autophagy in Tumor Cells. Vol. 20, *International Journal of Molecular Sciences*. 2019.
190. He J, Qiu Z, Fan J, Xie X, Sheng Q, Sui X. Drug tolerant persister cell plasticity in cancer: a revolutionary strategy for more effective anticancer therapies. *Signal Transduct Target Ther* [Internet]. 2024;9(1):209. Available from: <https://doi.org/10.1038/s41392-024-01891-4>
191. Mikubo M, Inoue Y, Liu G, Tsao MS. Mechanism of Drug Tolerant Persister Cancer Cells: The Landscape and Clinical Implication for Therapy. *J Thorac Oncol Off Publ Int Assoc Study Lung Cancer*. 2021 Nov;16(11):1798-809.
192. Zhang X, Li C, Wu Y, Cui P. The research progress of Wnt/ β -catenin signaling pathway in colorectal cancer. *Clin Res Hepatol Gastroenterol*. 2023 Mar;47(3):102086.
193. Fu M, Hu Y, Lan T, Guan KL, Luo T, Luo M. The Hippo signalling pathway and its implications in human health and diseases. *Signal Transduct Target Ther* [Internet]. 2022;7(1):376. Available from: <https://doi.org/10.1038/s41392-022-01191-9>
194. Kurenova E, Xu LH, Yang X, Baldwin ASJ, Craven RJ, Hanks SK, et al. Focal adhesion kinase suppresses apoptosis by binding to the death domain of receptor-interacting protein. *Mol Cell Biol*. 2004 May;24(10):4361-71.
195. Kim T, Park BS, Heo S, Jeon H, Kim J, Kim D, et al. Combinatorial CRISPR screen reveals FYN and KDM4 as targets for synergistic drug combination for treating triple negative breast cancer [Internet]. eLife Sciences Publications, Ltd; 2025. Available from: <http://dx.doi.org/10.7554/eLife.93921.2>
196. Gan L, Yang Y, Li Q, Feng Y, Liu T, Guo W. Epigenetic regulation of cancer progression by EZH2: from biological insights to therapeutic potential. *Biomark Res* [Internet]. 2018;6(1):10. Available from: <https://doi.org/10.1186/s40364-018-0122-2>
197. Chou CW, Huang YM, Chang YJ, Huang CY, Hung CS. Identified the novel resistant biomarkers for taxane-based therapy for triple-negative breast cancer. *Int J Med Sci*. 2021;18(12):2521-31.

198. Egeland EV, Seip K, Skourti E, Øy GF, Pettersen SJ, Pandya AD, et al. The SRC-family serves as a therapeutic target in triple negative breast cancer with acquired resistance to chemotherapy. *Br J Cancer* [Internet]. 2024;131(10):1656-67. Available from: <https://doi.org/10.1038/s41416-024-02875-5>
199. Aseervatham J. Cytoskeletal Remodeling in Cancer. *Biology (Basel)*. 2020 Nov;9(11).
200. Qin X, Zeng B, Sooranna SR, Li M. LAMB3 Promotes Myofibrogenesis and Cytoskeletal Reorganization in Endometrial Stromal Cells via the RhoA/ROCK1/MYL9 Pathway. *Cell Biochem Biophys*. 2024 Mar;82(1):127-37.
201. Li K, Zhang Y, Zhang Y, Jiang W, Shen J, Xu S, et al. Tyrosine kinase Fyn promotes osteoarthritis by activating the β -catenin pathway. *Ann Rheum Dis*. 2018 Jun;77(6):935-43.
202. McCubrey JA, Steelman LS, Bertrand FE, Davis NM, Abrams SL, Montalto G, et al. Multifaceted roles of GSK-3 and Wnt/ β -catenin in hematopoiesis and leukemogenesis: opportunities for therapeutic intervention. *Leukemia* [Internet]. 2014;28(1):15-33. Available from: <https://doi.org/10.1038/leu.2013.184>
203. Chernock RD, Cherla RP, Ganju RK. SHP2 and cbl participate in alpha-chemokine receptor CXCR4-mediated signaling pathways. *Blood*. 2001 Feb;97(3):608-15.
204. Guo Q, Gao BL, Zhang XJ, Liu GC, Xu F, Fan QY, et al. CXCL12-CXCR4 Axis Promotes Proliferation, Migration, Invasion, and Metastasis of Ovarian Cancer. *Oncol Res*. 2014;22(5-6):247-58.
205. Zi D, Li Q, Xu CX, Zhou ZW, Song GB, Hu CB, et al. CXCR4 knockdown enhances sensitivity of paclitaxel via the PI3K/Akt/mTOR pathway in ovarian carcinoma. *Aging (Albany NY)*. 2022 Jun;14(11):4673-98.
206. Chevalier T, Daste A, Saada-Bouazid E, Loundou A, Peyraud F, Lambert T, et al. Cetuximab combined with paclitaxel or paclitaxel alone for patients with recurrent or metastatic head and neck squamous cell carcinoma progressing after EXTREME. *Cancer Med*. 2021 Jun;10(12):3952-63.
207. Koyama T, Kiyota N, Boku S, Imamura Y, Shibata N, Satake H, et al. A phase II trial of paclitaxel plus biweekly cetuximab for patients with recurrent or metastatic head and neck cancer previously treated with both platinum-based chemotherapy and anti-PD-1 antibody. *ESMO Open* [Internet]. 2024 Jun 1;9(6). Available from: <https://doi.org/10.1016/j.esmoop.2024.103476>
208. Irwin ME, Johnson BP, Manshouri R, Amin HM, Chandra J. A NOX2/Egr-1/Fyn pathway delineates new targets for TKI-resistant malignancies. *Oncotarget*. 2015 Sep;6(27):23631-46.
209. Zhao B, Wang L, Qiu H, Zhang M, Sun L, Peng P, et al. Mechanisms of resistance to anti-EGFR therapy in colorectal cancer. *Oncotarget*. 2017 Jan;8(3):3980-4000.
210. Tan J, Yu CY, Wang ZH, Chen HY, Guan J, Chen YX, et al. Genetic variants in the inositol phosphate metabolism pathway and risk of different types of cancer. *Sci Rep* [Internet]. 2015;5(1):8473. Available from: <https://doi.org/10.1038/srep08473>
211. Di X, Pan Y, Yan J, Liu J, Wen D, Jiang H, et al. Therapeutic potential of anti-PIK3CG treatment for multiple myeloma via inhibiting c-Myc pathway. *Heliyon* [Internet]. 2024;10(1):e23165. Available from: <https://www.sciencedirect.com/science/article/pii/S2405844023103732>

212. Jung AR, Jung CH, Noh JK, Lee YC, Eun YG. Epithelial-mesenchymal transition gene signature is associated with prognosis and tumor microenvironment in head and neck squamous cell carcinoma. *Sci Rep*. 2020 Feb;10(1):3652.
213. Bardelli A, Jänne PA. The road to resistance: EGFR mutation and cetuximab. *Nat Med* [Internet]. 2012;18(2):199-200. Available from: <https://doi.org/10.1038/nm.2646>
214. Phuchareon J, McCormick F, Eisele DW, Tetsu O. EGFR inhibition evokes innate drug resistance in lung cancer cells by preventing Akt activity and thus inactivating Ets-1 function. *Proc Natl Acad Sci U S A*. 2015 Jul;112(29):E3855-63.
215. Atsaves V, Leventaki V, Rassidakis GZ, Claret FX. AP-1 Transcription Factors as Regulators of Immune Responses in Cancer. *Cancers (Basel)*. 2019 Jul;11(7).
216. Palacios EH, Weiss A. Function of the Src-family kinases, Lck and Fyn, in T-cell development and activation. *Oncogene* [Internet]. 2004;23(48):7990-8000. Available from: <https://doi.org/10.1038/sj.onc.1208074>
217. Sen B, Johnson FM. Regulation of SRC family kinases in human cancers. *J Signal Transduct*. 2011;2011:865819.
218. Wang W, He S, Zhang R, Peng J, Guo D, Zhang J, et al. ALDH1A1 maintains the cancer stem-like cells properties of esophageal squamous cell carcinoma by activating the AKT signal pathway and interacting with β -catenin. *Biomed Pharmacother* [Internet]. 2020;125:109940. Available from: <https://www.sciencedirect.com/science/article/pii/S075333222030130X>
219. Gan C, Pierscianek D, El Hindy N, Ahmadipour Y, Keyvani K, Sure U, et al. The predominant expression of cancer stem cell marker ALDH1A3 in tumor infiltrative area is associated with shorter overall survival of human glioblastoma. *BMC Cancer* [Internet]. 2020;20(1):672. Available from: <https://doi.org/10.1186/s12885-020-07153-0>
220. Kawai T, Yasuchika K, Ishii T, Miyauchi Y, Kojima H, Yamaoka R, et al. SOX9 is a novel cancer stem cell marker surrogated by osteopontin in human hepatocellular carcinoma. *Sci Rep* [Internet]. 2016;6(1):30489. Available from: <https://doi.org/10.1038/srep30489>
221. Morimoto Y, Mizushima T, Wu X, Okuzaki D, Yokoyama Y, Inoue A, et al. miR-4711-5p regulates cancer stemness and cell cycle progression via KLF5, MDM2 and TFDPI in colon cancer cells. *Br J Cancer*. 2020 Mar;122(7):1037-49.
222. Gou S, Wu A, Luo Z. Integrins in cancer stem cells. *Front cell Dev Biol*. 2024;12:1434378.
223. Phi LTH, Sari IN, Yang YG, Lee SH, Jun N, Kim KS, et al. Cancer Stem Cells (CSCs) in Drug Resistance and their Therapeutic Implications in Cancer Treatment. *Stem Cells Int*. 2018;2018:5416923.
224. Makena MR, Ranjan A, Thirumala V, Reddy AP. Cancer stem cells: Road to therapeutic resistance and strategies to overcome resistance. *Biochim Biophys Acta - Mol Basis Dis* [Internet]. 2020;1866(4):165339. Available from: <https://www.sciencedirect.com/science/article/pii/S0925443918304769>
225. Liu H, Huang Y, Zhao C, Wang G, Wang X. ABALON regulates mitophagy and 5-FU sensitivity in colorectal cancer via PINK1-Parkin pathway. *Transl Cancer Res*. 2024 Nov;13(11):6201-18.
226. Yamashita K, Miyata H, Makino T, Masuike Y, Furukawa H, Tanaka K, et al. High Expression of the Mitophagy-Related Protein Pink1 is Associated with a Poor Response to Chemotherapy and a

- Poor Prognosis for Patients Treated with Neoadjuvant Chemotherapy for Esophageal Squamous Cell Carcinoma. *Ann Surg Oncol*. 2017 Dec;24(13):4025-32.
227. Qin L, Dong Z, Zhang JT. Reversible epigenetic regulation of 14-3-3 σ expression in acquired gemcitabine resistance by uhrf1 and DNA methyltransferase 1. *Mol Pharmacol*. 2014 Nov;86(5):561-9.
 228. Cao P, Zhang W, Qiu J, Tang Z, Xue X, Feng T. Gemcitabine Inhibits the Progression of Pancreatic Cancer by Restraining the WTAP/MYC Chain in an m6A-Dependent Manner. *Cancer Res Treat*. 2024 Jan;56(1):259-71.
 229. Stukas D, Jasukaitiene A, Bartkeviciene A, Matthews J, Maimets T, Teino I, et al. Targeting AHR Increases Pancreatic Cancer Cell Sensitivity to Gemcitabine through the ELAVL1-DCK Pathway. *Int J Mol Sci*. 2023 Aug;24(17).
 230. Kelsey MD, Patrick-Lake B, Abdulai R, Broedl UC, Brown A, Cohn E, et al. Inclusion and diversity in clinical trials: Actionable steps to drive lasting change. *Contemp Clin Trials*. 2022 May;116:106740.
 231. Anderson ARA, Weaver AM, Cummings PT, Quaranta V. Tumor morphology and phenotypic evolution driven by selective pressure from the microenvironment. *Cell*. 2006 Dec;127(5):905-15.
 232. Thompson PA, Khatami M, Bagloli CJ, Sun J, Harris SA, Moon EY, et al. Environmental immune disruptors, inflammation and cancer risk. *Carcinogenesis* [Internet]. 2015 Jun 1;36(Suppl_1):S232-53. Available from: <https://doi.org/10.1093/carcin/bgv038>
 233. Ostroverkhova D, Przytycka TM, Panchenko AR. Cancer driver mutations: predictions and reality. *Trends Mol Med* [Internet]. 2023;29(7):554-66. Available from: <https://www.sciencedirect.com/science/article/pii/S1471491423000679>
 234. Cheong A, Nagel ZD. Human Variation in DNA Repair, Immune Function, and Cancer Risk. *Front Immunol*. 2022;13:899574.
 235. Manica M, Kim HR, Mathis R, Chouvarine P, Rutishauser D, De Vargas Roditi L, et al. Inferring clonal composition from multiple tumor biopsies. *NPJ Syst Biol Appl*. 2020 Aug;6(1):27.
 236. Smith JC, Sheltzer JM. Systematic identification of mutations and copy number alterations associated with cancer patient prognosis. *Elife*. 2018 Dec;7.
 237. van Zwet EW, Goodman SN. How large should the next study be? Predictive power and sample size requirements for replication studies. *Stat Med*. 2022 Jul;41(16):3090-101.
 238. Wu PY, Stich B, Weisweiler M, Shrestha A, Erban A, Westhoff P, et al. Improvement of prediction ability by integrating multi-omic datasets in barley. *BMC Genomics*. 2022 Mar;23(1):200.
 239. Li Y, Herold T, Mansmann U, Hornung R. Does combining numerous data types in multi-omics data improve or hinder performance in survival prediction? Insights from a large-scale benchmark study. *BMC Med Inform Decis Mak*. 2024 Sep;24(1):244.
 240. Ouyang X, Feng Y, Cui C, Li Y, Zhang L, Wang H. Improving generalizability of drug-target binding prediction by pre-trained multi-view molecular representations. *Bioinformatics*. 2024 Dec;41(1).

APPENDICES

APPENDIX 1: Common enriched terms for paclitaxel

Table 7. Common enriched terms for paclitaxel identified through enrichment analysis using gene features from pan-drug models trained with mitotic and cytoskeletal inhibitors

ID	Term Description
hsa04115	p53 signaling pathway
hsa04512	ECM-receptor interaction
hsa04510	Focal adhesion
hsa05222	Small cell lung cancer
hsa05165	Human papillomavirus infection
hsa04151	PI3K-Akt signaling pathway
hsa05205	Proteoglycans in cancer
hsa04390	Hippo signaling pathway
hsa05200	Pathways in cancer
hsa04392	Hippo signaling pathway - multiple species
hsa04933	AGE-RAGE signaling pathway in diabetic complications
hsa04530	Tight junction
hsa05130	Pathogenic Escherichia coli infection
hsa05213	Endometrial cancer
hsa05210	Colorectal cancer
hsa04520	Adherens junction
hsa05215	Prostate cancer
hsa04915	Estrogen signaling pathway
hsa05146	Amoebiasis
hsa04014	Ras signaling pathway
hsa04820	Cytoskeleton in muscle cells
hsa04010	MAPK signaling pathway

Table 7. Common enriched terms for paclitaxel identified through enrichment analysis using gene features from pan-drug models trained with mitotic and cytoskeletal inhibitors (continued)

ID	Term Description
hsa04934	Cushing syndrome
hsa05226	Gastric cancer
hsa05224	Breast cancer
hsa05225	Hepatocellular carcinoma
hsa05163	Human cytomegalovirus infection
hsa04610	Complement and coagulation cascades
hsa01521	EGFR TKI resistance
hsa05150	Staphylococcus aureus infection
hsa04916	Melanogenesis
hsa04066	HIF-1 signaling pathway
hsa05216	Thyroid cancer
hsa04072	Phospholipase D signaling pathway
hsa05412	Arrhythmogenic right ventricular cardiomyopathy
hsa04810	Regulation of actin cytoskeleton
hsa04926	Relaxin signaling pathway
hsa04668	TNF signaling pathway
hsa05120	Epithelial cell signaling in Helicobacter pylori infection
hsa04540	Gap junction
hsa04015	Rap1 signaling pathway
hsa04912	GnRH signaling pathway
hsa04371	Apelin signaling pathway
hsa04928	Parathyroid hormone synthesis, secretion and action
hsa05207	Chemical carcinogenesis - receptor activation
hsa05145	Toxoplasmosis
hsa05217	Basal cell carcinoma
hsa05418	Fluid shear stress and atherosclerosis

Table 7. Common enriched terms for paclitaxel identified through enrichment analysis using gene features from pan-drug models trained with mitotic and cytoskeletal inhibitors (continued)

ID	Term Description
hsa04921	Oxytocin signaling pathway
hsa05219	Bladder cancer
hsa05223	NSCLC
hsa05214	Glioma
hsa04012	ErbB signaling pathway
hsa05231	Choline metabolism in cancer
hsa01522	Endocrine resistance
hsa05131	Shigellosis
hsa05110	Vibrio cholerae infection
hsa00565	Ether lipid metabolism
hsa05132	Salmonella infection
hsa04936	Alcoholic liver disease
hsa04020	Calcium signaling pathway
hsa05160	Hepatitis C
hsa05230	Central carbon metabolism in cancer
hsa05206	MicroRNAs in cancer
hsa05212	Pancreatic cancer
hsa05218	Melanoma
hsa04310	Wnt signaling pathway
hsa04068	FoxO signaling pathway
hsa00564	Glycerophospholipid metabolism
hsa04666	Fc gamma R-mediated phagocytosis
hsa05167	Kaposi sarcoma-associated herpesvirus infection
hsa05235	PD-L1 expression and PD-1 checkpoint pathway in cancer
hsa04924	Renin secretion
hsa04720	Long-term potentiation

Table 7. Common enriched terms for paclitaxel identified through enrichment analysis using gene features from pan-drug models trained with mitotic and cytoskeletal inhibitors (continued)

ID	Term Description
hsa05221	Acute myeloid leukemia
hsa04625	C-type lectin receptor signaling pathway
hsa05031	Amphetamine addiction
hsa04630	JAK-STAT signaling pathway

APPENDIX 2: Pathway enrichment results

Table 8. pathfindR pathway enrichment analysis for paclitaxel utilizing the gene features of the pan-drug model trained with mitotic inhibitors (Strategy 1)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04510	Focal adhesion	4,6553661	2,3e-10	FYN	ITGA2, ITGAV, FLNB, MYL9, LAMB2, LAMB3, LAMC2, THBS1, EGFR, VEGFC
hsa04512	ECM-receptor interaction	7,7096534	1,8e-09		LAMB2, LAMB3, LAMC2, THBS1, ITGAV, SDC4, ITGA2, DSPP
hsa04115	p53 signaling pathway	4,5707231	4,5e-08		SFN, THBS1, SERPINE1, PPP1R13L
hsa05222	Small cell lung cancer	4,9375095	7,5e-08		ITGA2, ITGAV, LAMB2, LAMB3, LAMC2
hsa05165	Human papillomavirus infection	2,0708778	9,0e-08	TCF7L2	EGFR, ITGA2, ITGAV, LAMB2, LAMB3, LAMC2, THBS1
hsa04915	Estrogen signaling pathway	3,9338191	9,6e-08		EGFR, KRT13, KRT14, KRT15, KRT18, KRT19
hsa05205	Proteoglycans in cancer	3,7494213	1,2e-07		RRAS, CTTN, PLAU, THBS1, ITGA2, ITGAV, SDC4, GPC1, EGFR
hsa04151	PI3K-Akt signaling pathway	2,3735209	1,2e-07		EGFR, EPHA2, AREG, VEGFC, ITGA2, ITGAV, LAMB2, LAMB3, LAMC2, THBS1
hsa05150	Staphylococcus aureus infection	5,4045712	6,2e-07		KRT13, KRT14, KRT15, KRT18, KRT19

Table 8. pathfindR pathway enrichment analysis for paclitaxel utilizing the gene features of the pan-drug model trained with mitotic inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04390	Hippo signaling pathway	3,0962963	2,6e-06	TCF7L2	YAP1, CCN2, AREG, AJUBA, SERPINE1
hsa04933	AGE-RAGE signaling pathway in diabetic complications	4,3471551	9,9e-06		EGR1, EDN1, F3, VEGFC, SERPINE1
hsa04520	Adherens junction	6,4357883	1,1e-05	TCF7L2, FYN	EGFR, TJP1, NECTIN2, MYL9, ARHGAP29
hsa04392	Hippo signaling pathway - multiple species	8,5701058	2,6e-05		YAP1, RASSF4, AJUBA
hsa04810	Regulation of actin cytoskeleton	2,5566830	2,9e-05	CYFIP2, CXCR4	MYL9, ITGA2, ITGAV, RRAS, EGFR
hsa05200	Pathways in cancer	1,7811016	7,1e-05	TCF7L2, CXCR4	LAMB2, LAMB3, LAMC2, ITGA2, ITGAV, EGFR, JUP, VEGFC, EDN1
hsa04530	Tight junction	2,0122680	8,7e-05		TJP1, CLDN1, CTTN, MYL9
hsa05130	Pathogenic Escherichia coli infection	2,7741383	8,7e-05	FYN, CYFIP2	CTTN, CLDN1, TJP1, MYO1C
hsa00620	Pyruvate metabolism	3,6358025	1,0e-04	LDHB	ME2
hsa00590	Arachidonic acid metabolism	2,9086420	2,0e-04		PTGES, PTGR1
hsa04072	Phospholipase D signaling pathway	2,3018030	2,6e-04	FYN	RRAS, EGFR, PLPP2
hsa04066	HIF-1 signaling pathway	3,1995062	2,8e-04	LDHB	EGFR, SERPINE1, EDN1
hsa05146	Amoebiasis	3,4037300	3,6e-04		LAMB2, LAMB3, LAMC2, SERPINB6

Table 8. pathfindR pathway enrichment analysis for paclitaxel utilizing the gene features of the pan-drug model trained with mitotic inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05213	Endometrial cancer	2,9625057	3,8e-04	TCF7L2	EGFR
hsa04014	Ras signaling pathway	1,4543210	4,6e-04		RRAS, VEGFC, EGFR, EPHA2
hsa04010	MAPK signaling pathway	1,7974754	5,0e-04		RRAS, EGFR, EPHA2, PPP3CA, AREG, VEGFC
hsa05210	Colorectal cancer	3,1164021	7,5e-04	TCF7L2	EGFR, AREG
hsa04820	Cytoskeleton in muscle cells	4,7855862	7,7e-04		ITGA2, ITGAV, DSG2, JUP, THBS1, SDC4, MYL9
hsa04928	Parathyroid hormone synthesis, secretion and action	2,2853616	7,8e-04		EGFR, CYP24A1, EGR1
hsa05215	Prostate cancer	3,6775933	8,2e-04	TCF7L2	EGFR, PLAUI, PLAT
hsa05163	Human cytomegalovirus infection	1,5917941	1,5e-03	CXCR4	EGFR, ITGAV, PPP3CA
hsa04610	Complement and coagulation cascades	4,5707231	1,7e-03		F3, SERPINE1, PLAUI, PLAT
hsa01521	EGFR TKI resistance	2,2853616	1,7e-03		EGFR, AXL
hsa04934	Cushing syndrome	1,1938456	2,0e-03	TCF7L2	EGFR
hsa05226	Gastric cancer	1,7263523	2,0e-03	TCF7L2	JUP, EGFR
hsa05224	Breast cancer	1,1426808	2,0e-03	TCF7L2	EGFR
hsa05225	Hepatocellular carcinoma	1,0254827	2,4e-03	TCF7L2	EGFR
hsa04015	Rap1 signaling pathway	2,4116881	3,3e-03		VEGFC, EGFR, EPHA2, ADORA2B, RRAS, THBS1
hsa04912	GnRH signaling pathway	1,9044680	3,7e-03		EGFR, EGR1
hsa05206	MicroRNAs in cancer	1,5264820	3,7e-03	EZH2	EGFR, CYP24A1, PLAUI, THBS1

Table 8. pathfindR pathway enrichment analysis for paclitaxel utilizing the gene features of the pan-drug model trained with mitotic inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05412	Arrhythmogenic right ventricular cardiomyopathy	5,4786065	4,0e-03	TCF7L2	DSG2, JUP, ITGA2, ITGAV
hsa05120	Epithelial cell signaling in Helicobacter pylori infection	2,3876912	4,2e-03		TJP1, EGFR
hsa04540	Gap junction	2,1049383	4,2e-03		TJP1, EGFR
hsa04611	Platelet activation	1,3557230	5,0e-03	FYN	ITGA2
hsa04371	Apelin signaling pathway	3,1741133	5,0e-03		SERPINE1, RRAS, CCN2, EGR1, PLAT
hsa05410	Hypertrophic cardiomyopathy	2,7268519	5,1e-03		ITGA2, ITGAV, EDN1
hsa05221	Acute myeloid leukemia	2,6662551	5,2e-03	TCF7L2	JUP
hsa05414	Dilated cardiomyopathy	1,7201646	5,4e-03		ITGA2, ITGAV
hsa04360	Axon guidance	3,4141072	5,6e-03	CXCR4, FYN	PPP3CA, EPHA2, RHOD, RRAS, MYL9
hsa05216	Thyroid cancer	2,3525781	5,6e-03	TCF7L2	
hsa04814	Motor proteins	1,5683854	6,4e-03		MYO1C, MYL9
hsa04926	Relaxin signaling pathway	2,0335844	6,9e-03		EDN1, VEGFC, EGFR
hsa05207	Chemical carcinogenesis - receptor activation	1,7676830	7,6e-03		EGFR, ADRB2, KLF5, KLF4
hsa04145	Phagosome	1,9996914	9,5e-03		ITGAV, ITGA2, THBS1
hsa04024	cAMP signaling pathway	1,5607347	1,1e-02		ADRB2, RRAS, MYL9, EDN1
hsa05145	Toxoplasmosis	2,4996142	1,1e-02		LAMB2, LAMB3, LAMC2

Table 8. pathfindR pathway enrichment analysis for paclitaxel utilizing the gene features of the pan-drug model trained with mitotic inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05418	Fluid shear stress and atherosclerosis	3,0070547	1,2e-02		PLAT, EDN1, ITGAV, GPC1, SDC4
hsa04064	NF-κB signaling pathway	1,8387967	1,3e-02		PLAU, PPP1R13L
hsa04668	TNF signaling pathway	1,5839139	1,3e-02		EDN1, VEGFC
hsa05202	Transcriptional misregulation in cancer	3,0184020	1,3e-02		NFKBIZ, PLAU, JUP, PLAT
hsa04921	Oxytocin signaling pathway	1,8458689	1,4e-02		MYL9, PPP3CA, EGFR
hsa04020	Calcium signaling pathway	2,1716105	1,4e-02	CXCR4	PPP3CA, EGFR, ADORA2B, ADRB2, VEGFC
hsa05217	Basal cell carcinoma	1,2901235	1,6e-02	TCF7L2	
hsa04080	Neuroactive ligand-receptor interaction	1,5891587	1,8e-02		ADORA2B, F2RL1, ADRB2, EDN1, ADM, LYPD6B
hsa05219	Bladder cancer	4,3236570	1,8e-02		THBS1, EGFR
hsa04310	Wnt signaling pathway	2,0122680	2,0e-02	TCF7L2	PPP3CA, MMP7, DKK1
hsa05223	NSCLC	1,2498071	2,2e-02		EGFR
hsa05214	Glioma	1,1426808	2,2e-02		EGFR
hsa04012	ErbB signaling pathway	2,1618285	2,2e-02		EGFR, AREG
hsa05231	Choline metabolism in cancer	1,7974754	2,3e-02		EGFR, PLPP2
hsa01522	Endocrine resistance	0,8987377	2,3e-02		EGFR
hsa04148	Efferocytosis	2,9407226	2,5e-02	ODC1	ITGAV, AXL, THBS1, ADAM9
hsa05110	Vibrio cholerae infection	2,1049383	2,5e-02		TJP1

Table 8. pathfindR pathway enrichment analysis for paclitaxel utilizing the gene features of the pan-drug model trained with mitotic inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa00565	Ether lipid metabolism	1,7388621	2,6e-02		PLPP2
hsa05218	Melanoma	1,1426808	2,8e-02		EGFR
hsa05230	Central carbon metabolism in cancer	2,4996142	2,8e-02	LDHB	EGFR
hsa05212	Pancreatic cancer	1,1426808	2,8e-02		EGFR
hsa04664	Fc epsilon RI signaling pathway	1,2901235	2,8e-02	FYN	
hsa04924	Renin secretion	4,0671689	2,9e-02		ADRB2, PPP3CA, EDN1
hsa05235	PD-L1 expression and PD-1 checkpoint pathway in cancer	1,9509184	3,0e-02		PPP3CA, EGFR
hsa04725	Cholinergic synapse	0,9410312	3,0e-02	FYN	
hsa04916	Melanogenesis	1,7018650	3,1e-02	TCF7L2	EDN1
hsa04625	C-type lectin receptor signaling pathway	1,6839506	3,2e-02		RRAS, PPP3CA
hsa04068	FoxO signaling pathway	1,3006123	3,7e-02		EGFR, PLK2
hsa00480	Glutathione metabolism	2,9086420	3,7e-02	ODC1	GPX8
hsa04071	Sphingolipid signaling pathway	0,7406264	3,8e-02	FYN	
hsa04919	Thyroid hormone signaling pathway	0,7406264	3,8e-02		ITGAV
hsa04650	Natural killer cell mediated cytotoxicity	1,4543210	3,8e-02	FYN	PPP3CA
hsa05131	Shigellosis	1,1762890	3,9e-02		CTTN, EGFR, MYL9
hsa00564	Glycerophospholipid metabolism	0,9089506	3,9e-02		PLPP2

Table 8. pathfindR pathway enrichment analysis for paclitaxel utilizing the gene features of the pan-drug model trained with mitotic inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04666	Fc gamma R-mediated phagocytosis	0,8987377	3,9e-02		PLPP2
hsa04660	T-cell receptor signaling pathway	1,3790975	4,0e-02	FYN	PPP3CA
hsa04380	Osteoclast differentiation	1,2901235	4,1e-02	FYN	PPP3CA
hsa05416	Viral myocarditis	1,5683854	4,2e-02	FYN	
hsa04923	Regulation of lipolysis in adipocytes	1,5092010	4,3e-02		ADRB2
hsa04720	Long-term potentiation	1,3112730	4,3e-02		PPP3CA
hsa05143	African trypanosomiasis	2,9625057	4,4e-02		F2RL1
hsa05160	Hepatitis C	1,0957213	4,5e-02		CLDN1, EGFR
hsa05132	Salmonella infection	1,8096754	4,5e-02	TCF7L2, CYFIP2	AHNAK2, RRAS, MYL9
hsa05171	Coronavirus disease - COVID-19	0,5881445	4,7e-02		EGFR
hsa05031	Amphetamine addiction	1,2901235	4,7e-02		PPP3CA
hsa04218	Cellular senescence	1,6435819	4,9e-02		RRAS, PPP3CA, SERPINE1
hsa04630	JAK-STAT signaling pathway	0,4999228	5,0e-02		EGFR

Table 9. pathfindR pathway enrichment analysis for paclitaxel utilizing the gene features of the pan-drug model trained with mitotic inhibitors (Strategy 2)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04512	ECM-receptor interaction	9,1835578	1,8e-09		LAMB2, LAMB3, LAMC2, THBS1, ITGAV, SDC4, ITGA2, DSPP
hsa04510	Focal adhesion	4,5371148	3,4e-09		ITGA2, ITGAV, MYL9, LAMB2, LAMB3, LAMC2, THBS1, EGFR, VEGFC
hsa04115	p53 signaling pathway	5,4445378	4,5e-08		SFN, THBS1, SERPINE1, PPP1R13L
hsa05222	Small cell lung cancer	5,8814452	7,5e-08		ITGA2, ITGAV, LAMB2, LAMB3, LAMC2
hsa05165	Human papillomavirus infection	2,4667809	9,0e-08	TCF7L2	EGFR, ITGA2, ITGAV, LAMB2, LAMB3, LAMC2, THBS1
hsa05205	Proteoglycans in cancer	4,4662224	1,2e-07		RRAS, CTTN, PLAUI, THBS1, ITGA2, ITGAV, SDC4, GPC1, EGFR
hsa04151	PI3K-Akt signaling pathway	2,8272822	1,2e-07		EGFR, EPHA2, AREG, VEGFC, ITGA2, ITGAV, LAMB2, LAMB3, LAMC2, THBS1
hsa04390	Hippo signaling pathway	3,6882353	2,6e-06	TCF7L2	YAP1, CCN2, AREG, AJUBA, SERPINE1
hsa04933	AGE-RAGE signaling pathway in diabetic complications	5,1782289	3,8e-06		EGR1, EDN1, F3, VEGFC, SERPINE1
hsa04392	Hippo signaling pathway - multiple species	10,2085084	2,6e-05		YAP1, RASSF4, AJUBA

Table 9. pathfindR pathway enrichment analysis for paclitaxel utilizing the gene features of the pan-drug model trained with mitotic inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05200	Pathways in cancer	1,9287330	7,1e-05	TCF7L2, CXCR4	LAMB2, LAMB3, LAMC2, ITGA2, ITGAV, EGFR, VEGFC, EDN1
hsa04530	Tight junction	2,3969663	8,7e-05		TJP1, CLDN1, CTTN, MYL9
hsa05130	Pathogenic Escherichia coli infection	2,7537402	8,7e-05	CYFIP2	CTTN, CLDN1, TJP1, MYO1C
hsa05213	Endometrial cancer	3,5288671	1,9e-04	TCF7L2	EGFR
hsa00590	Arachidonic acid metabolism	3,4647059	2,0e-04		PTGES, PTGR1
hsa05210	Colorectal cancer	3,7121849	2,1e-04	TCF7L2	EGFR, AREG
hsa04520	Adherens junction	6,5709939	2,1e-04	TCF7L2	EGFR, TJP1, NECTIN2, MYL9, ARHGAP29
hsa05215	Prostate cancer	4,3806626	2,1e-04	TCF7L2	EGFR, PLAU, PLAT
hsa04915	Estrogen signaling pathway	3,1239151	2,1e-04		EGFR, KRT15, KRT18, KRT19
hsa05146	Amoebiasis	4,0544431	3,6e-04		LAMB2, LAMB3, LAMC2, SERPINB6
hsa04014	Ras signaling pathway	1,7323529	4,6e-04		RRAS, VEGFC, EGFR, EPHA2
hsa04010	MAPK signaling pathway	2,1411104	5,0e-04		RRAS, EGFR, EPHA2, PPP3CA, AREG, VEGFC
hsa04934	Cushing syndrome	1,4220808	5,0e-04	TCF7L2	EGFR
hsa05226	Gastric cancer	1,3709268	5,0e-04	TCF7L2	EGFR
hsa05224	Breast cancer	1,3611345	5,0e-04	TCF7L2	EGFR
hsa05225	Hepatocellular carcinoma	1,2215309	6,0e-04	TCF7L2	EGFR
hsa04820	Cytoskeleton in muscle cells	4,8861237	7,7e-04		ITGA2, ITGAV, DSG2, THBS1, SDC4, MYL9

Table 9. pathfindR pathway enrichment analysis for paclitaxel utilizing the gene features of the pan-drug model trained with mitotic inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05163	Human cytomegalovirus infection	1,8961077	1,5e-03	CXCR4	EGFR, ITGAV, PPP3CA
hsa04610	Complement and coagulation	5,4445378	1,7e-03		F3, SERPINE1, PLAUI, PLAT
hsa01521	EGFR TKI	2,7222689	1,7e-03		EGFR, AXL
hsa04015	Rap1 signaling pathway	2,8727461	2,2e-03		VEGFC, EGFR, EPHA2, ADORA2B, RRAS, THBS1
hsa05150	Staphylococcus aureus infection	3,8626789	2,5e-03		KRT15, KRT18, KRT19
hsa04916	Melanogenesis	2,0272215	2,5e-03	TCF7L2	EDN1
hsa04066	HIF-1 signaling pathway	2,8583824	2,5e-03		EGFR, SERPINE1, EDN1
hsa04912	GnRH signaling pathway	2,2685574	2,7e-03		EGFR, EGR1
hsa05216	Thyroid cancer	2,8023356	3,1e-03	TCF7L2	
hsa04072	Phospholipase D signaling pathway	2,0563902	3,4e-03		RRAS, EGFR, PLPP2
hsa04928	Parathyroid hormone synthesis, secretion and action	1,8148459	3,5e-03		EGFR, EGR1
hsa05412	Arrhythmogenic right ventricular cardiomyopathy	5,2207897	4,0e-03	TCF7L2	DSG2, ITGA2, ITGAV
hsa04810	Regulation of actin cytoskeleton	3,0454607	4,0e-03	CYFIP2, CXCR4	MYL9, ITGA2, ITGAV, RRAS, EGFR
hsa05120	Epithelial cell signaling in Helicobacter pylori infection	2,8441615	4,2e-03		TJP1, EGFR
hsa04540	Gap junction	2,5073529	4,2e-03		TJP1, EGFR
hsa04926	Relaxin signaling pathway	2,4223579	4,4e-03		EDN1, VEGFC, EGFR

Table 9. pathfindR pathway enrichment analysis for paclitaxel utilizing the gene features of the pan-drug model trained with mitotic inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04371	Apelin signaling pathway	3,7809290	4,9e-03		SERPINE1, RRAS, CCN2, EGR1, PLAT
hsa05410	Hypertrophic cardiomyopathy	3,2481618	5,1e-03		ITGA2, ITGAV, EDN1
hsa05414	Dilated cardiomyopathy	2,0490196	5,4e-03		ITGA2, ITGAV
hsa04814	Motor proteins	1,8682238	6,4e-03		MYO1C, MYL9
hsa05207	Chemical carcinogenesis - receptor activation	1,5792168	7,6e-03		EGFR, KLF5, KLF4
hsa05217	Basal cell	1,5367647	8,6e-03	TCF7L2	
hsa04064	NF-κB signaling pathway	2,1903313	8,7e-03		PLAU, PPP1R13L
hsa04668	TNF signaling pathway	1,8867210	8,7e-03		EDN1, VEGFC
hsa04145	Phagosome	2,3819853	9,5e-03		ITGAV, ITGA2, THBS1
hsa05145	Toxoplasmosis	2,9774816	1,1e-02		LAMB2, LAMB3, LAMC2
hsa05418	Fluid shear stress and atherosclerosis	3,5819328	1,2e-02		PLAT, EDN1, ITGAV, GPC1, SDC4
hsa05219	Bladder cancer	5,1502385	1,3e-02		THBS1, EGFR
hsa04921	Oxytocin signaling pathway	2,1987557	1,4e-02		MYL9, PPP3CA, EGFR
hsa05202	Transcriptional misregulation in cancer	2,6965871	2,0e-02		NFKBIZ, PLAU, PLAT
hsa05223	NSCLC	1,4887408	2,2e-02		EGFR
hsa05214	Glioma	1,3611345	2,2e-02		EGFR
hsa04012	ErbB signaling pathway	2,5751192	2,2e-02		EGFR, AREG
hsa05231	Choline metabolism in	2,1411104	2,3e-02		EGFR, PLPP2
hsa01522	Endocrine	1,0705552	2,3e-02		EGFR

Table 9. pathfindR pathway enrichment analysis for paclitaxel utilizing the gene features of the pan-drug model trained with mitotic inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04148	Efferocytosis	2,8023356	2,5e-02		ITGAV, AXL, THBS1, ADAM9
hsa05110	Vibrio cholerae infection	2,5073529	2,5e-02		TJP1
hsa00565	Ether lipid metabolism	2,0712916	2,6e-02		PLPP2
hsa05206	MicroRNAs in cancer	1,4546475	2,6e-02	EZH2	EGFR, PLAU, THBS1
hsa05230	Central carbon metabolism in	1,4887408	2,7e-02		EGFR
hsa05212	Pancreatic cancer	1,3611345	2,7e-02		EGFR
hsa05218	Melanoma	1,3611345	2,7e-02		EGFR
hsa04936	Alcoholic liver disease	0,7005839	3,0e-02	TCF7L2	
hsa05235	PD-L1 expression and PD-1 checkpoint pathway	2,3238881	3,1e-02		PPP3CA, EGFR
hsa04020	Calcium signaling pathway	2,1556428	3,3e-02	CXCR4	PPP3CA, EGFR, ADORA2B,
hsa05160	Hepatitis C	1,3051974	3,3e-02		CLDN1, EGFR
hsa04310	Wnt signaling pathway	1,7977248	3,5e-02	TCF7L2	PPP3CA, DKK1
hsa04625	C-type lectin receptor signaling	2,0058824	3,5e-02		RRAS, PPP3CA
hsa04068	FoxO signaling pathway	1,5492587	3,7e-02		EGFR, PLK2
hsa00480	Glutathione metabolism	1,7323529	3,7e-02		GPX8
hsa05131	Shigellosis	1,4011678	3,9e-02		CTTN, EGFR, MYL9
hsa00564	Glycerophospholipid metabolism	1,0827206	3,9e-02		PLPP2
hsa04666	Fc gamma R-mediated phagocytosis	1,0705552	3,9e-02		PLPP2

Table 9. pathfindR pathway enrichment analysis for paclitaxel utilizing the gene features of the pan-drug model trained with mitotic inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05167	Kaposi sarcoma-associated herpesvirus	1,0951657	3,9e-02	TCF7L2	PPP3CA
hsa04924	Renin secretion	3,2298106	4,3e-02		PPP3CA, EDN1
hsa04720	Long-term potentiation	1,5619576	4,3e-02		PPP3CA
hsa05143	African trypanosomiasis	3,5288671	4,4e-02		F2RL1
hsa05221	Acute myeloid leukemia	1,5879902	4,4e-02	TCF7L2	
hsa05132	Salmonella infection	2,1556428	4,5e-02	TCF7L2, CYFIP2	AHNAK2, RRAS, MYL9
hsa05031	Amphetamine addiction	1,5367647	4,7e-02		PPP3CA
hsa04630	JAK-STAT signaling pathway	0,5954963	5,0e-02		EGFR

Table 10. pathfindR pathway enrichment analysis for paclitaxel utilizing the gene features of the pan-drug model trained with mitotic inhibitors (Strategy 2)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04510	Focal adhesion	4,5105820	6,8e-09	FYN	ITGA2, FLNB, MYL9, LAMB2, LAMB3, LAMC2, THBS1, EGFR, VEGFC
hsa04915	Estrogen signaling pathway	4,1926230	9,6e-08		EGFR, KRT13, KRT14, KRT15, KRT18, KRT19
hsa04512	ECM-receptor interaction	7,1897590	1,3e-07		LAMB2, LAMB3, LAMC2, THBS1, SDC4, ITGA2, DSPP
hsa05150	Staphylococcus aureus infection	5,7601351	6,2e-07		KRT13, KRT14, KRT15, KRT18, KRT19
hsa04390	Hippo signaling pathway	3,3000000	2,6e-06	TCF7L2	YAP1, CCN2, AREG, AJUBA, SERPINE1
hsa04933	AGE-RAGE signaling pathway in diabetic complications	4,6331522	3,8e-06		EGR1, EDN1, F3, VEGFC, SERPINE1
hsa04115	p53 signaling pathway	3,6535714	4,4e-06		SFN, THBS1, SERPINE1
hsa05165	Human papillomavirus infection	1,9312298	5,7e-06	TCF7L2	EGFR, ITGA2, LAMB2, LAMB3, LAMC2, THBS1
hsa05222	Small cell lung cancer	4,2098765	5,7e-06		ITGA2, LAMB2, LAMB3, LAMC2
hsa04151	PI3K-Akt signaling pathway	2,2767062	7,0e-06		EGFR, EPHA2, AREG, VEGFC, ITGA2, LAMB2, LAMB3, LAMC2, THBS1
hsa05205	Proteoglycans in cancer	3,1080729	7,0e-06		RRAS, CTTN, PLAUI, THBS1, ITGA2, SDC4, EGFR

Table 10. pathfindR pathway enrichment analysis for paclitaxel utilizing the gene features of the pan-drug model trained with mitotic inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04520	Adherens junction	6,8591954	1,1e-05	TCF7L2, FYN	EGFR, TJP1, NECTIN2, MYL9, ARHGAP29
hsa04392	Hippo signaling pathway - multiple species	9,1339286	2,6e-05		YAP1, RASSF4, AJUBA
hsa04530	Tight junction	2,1446541	8,7e-05		TJP1, CLDN1, CTTN, MYL9
hsa05130	Pathogenic Escherichia coli infection	2,9566474	8,7e-05	FYN, CYFIP2	CTTN, CLDN1, TJP1, MYO1C
hsa00620	Pyruvate metabolism	3,8750000	1,0e-04	LDHB	ME2
hsa04066	HIF-1 signaling pathway	3,4100000	2,8e-04	LDHB	EGFR, SERPINE1, EDN1
hsa05213	Endometrial cancer	3,1574074	3,8e-04	TCF7L2	EGFR
hsa05146	Amoebiasis	4,5345745	4,1e-04		LAMB2, LAMB3, LAMC2, SERPINB6, SERPINB3
hsa04014	Ras signaling pathway	1,5500000	4,6e-04		RRAS, VEGFC, EGFR, EPHA2
hsa04010	MAPK signaling pathway	1,9157303	5,0e-04		RRAS, EGFR, EPHA2, PPP3CA, AREG, VEGFC
hsa05210	Colorectal cancer	3,3214286	7,5e-04	TCF7L2	EGFR, AREG
hsa04928	Parathyroid hormone synthesis, secretion and action	2,4357143	7,8e-04		EGFR, CYP24A1, EGR1
hsa05215	Prostate cancer	3,9195402	8,2e-04	TCF7L2	EGFR, PLAUI, PLAT

Table 10. pathfindR pathway enrichment analysis for paclitaxel utilizing the gene features of the pan-drug model trained with mitotic inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05200	Pathways in cancer	1,7257085	1,4e-03	TCF7L2, CXCR4	LAMB2, LAMB3, LAMC2, ITGA2, EGFR, JUP, VEGFC, EDN1
hsa04610	Complement and coagulation cascades	4,8714286	1,7e-03		F3, SERPINE1, PLAUI, PLAT
hsa01521	EGFR TKI resistance	2,4357143	1,7e-03		EGFR, AXL
hsa04934	Cushing syndrome	1,2723881	2,0e-03	TCF7L2	EGFR
hsa05226	Gastric cancer	1,8399281	2,0e-03	TCF7L2	JUP, EGFR
hsa05224	Breast cancer	1,2178571	2,0e-03	TCF7L2	EGFR
hsa05225	Hepatocellular carcinoma	1,0929487	2,4e-03	TCF7L2	EGFR
hsa04015	Rap1 signaling pathway	2,5703518	3,3e-03		VEGFC, EGFR, EPHA2, ADORA2B, RRAS, THBS1
hsa04072	Phospholipase D signaling pathway	2,4532374	3,4e-03	FYN	RRAS, EGFR, PLPP2
hsa05120	Epithelial cell signaling in Helicobacter pylori infection	2,5447761	3,6e-03		TJP1, EGFR
hsa04912	GnRH signaling pathway	2,0297619	3,7e-03		EGFR, EGR1
hsa05206	MicroRNAs in cancer	1,3015267	3,7e-03		EGFR, CYP24A1, PLAUI, THBS1
hsa04540	Gap junction	2,2434211	3,7e-03		TJP1, EGFR
hsa04371	Apelin signaling pathway	3,3829365	5,0e-03		SERPINE1, RRAS, CCN2, EGR1, PLAT
hsa05221	Acute myeloid leukemia	2,8416667	5,2e-03	TCF7L2	JUP

Table 10. pathfindR pathway enrichment analysis for paclitaxel utilizing the gene features of the pan-drug model trained with mitotic inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04360	Axon guidance	3,6387195	5,6e-03	CXCR4, FYN	PPP3CA, EPHA2, RHOD, RRAS, MYL9
hsa05216	Thyroid cancer	2,5073529	5,6e-03	TCF7L2	
hsa04926	Relaxin signaling pathway	2,1673729	6,9e-03		EDN1, VEGFC, EGFR
hsa05207	Chemical carcinogenesis - receptor activation	1,8839779	7,6e-03		EGFR, ADRB2, KLF5, KLF4
hsa04668	TNF signaling pathway	1,6881188	1,0e-02		EDN1, VEGFC
hsa05202	Transcriptional misregulation in cancer	3,2169811	1,0e-02		NFKBIZ, PLAUI, JUP, PLAT
hsa04921	Oxytocin signaling pathway	1,9673077	1,1e-02		MYL9, PPP3CA, EGFR
hsa04024	cAMP signaling pathway	1,6634146	1,1e-02		ADRB2, RRAS, MYL9, EDN1
hsa04814	Motor proteins	1,6715686	1,1e-02		MYO1C, MYL9
hsa04820	Cytoskeleton in muscle cells	4,3717949	1,3e-02		ITGA2, DSG2, JUP, THBS1, SDC4, MYL9
hsa05163	Human cytomegalovirus infection	1,2723881	1,3e-02	CXCR4	EGFR, PPP3CA
hsa04020	Calcium signaling pathway	2,3144796	1,3e-02	CXCR4	PPP3CA, EGFR, ADORA2B, ADRB2, VEGFC
hsa05145	Toxoplasmosis	2,6640625	1,4e-02		LAMB2, LAMB3, LAMC2
hsa05217	Basal cell carcinoma	1,3750000	1,6e-02	TCF7L2	
hsa04080	Neuroactive ligand-receptor interaction	1,4114238	1,8e-02		ADORA2B, ADRB2, EDN1, ADM, LYPD6B
hsa05219	Bladder cancer	4,6081081	1,8e-02		THBS1, EGFR

Table 10. pathfindR pathway enrichment analysis for paclitaxel utilizing the gene features of the pan-drug model trained with mitotic inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05412	Arrhythmogenic right ventricular cardiomyopathy	4,6712329	2,0e-02	TCF7L2	DSG2, JUP, ITGA2
hsa04310	Wnt signaling pathway	2,1446541	2,0e-02	TCF7L2	PPP3CA, MMP7, DKK1
hsa05223	NSCLC	1,3320312	2,2e-02		EGFR
hsa05214	Glioma	1,2178571	2,2e-02		EGFR
hsa04012	ErbB signaling pathway	2,3040541	2,2e-02		EGFR, AREG
hsa05231	Choline metabolism in cancer	1,9157303	2,3e-02		EGFR, PLPP2
hsa01522	Endocrine resistance	0,9578652	2,3e-02		EGFR
hsa00565	Ether lipid metabolism	1,8532609	2,6e-02		PLPP2
hsa04924	Renin secretion	4,3347458	2,6e-02		ADRB2, PPP3CA, EDN1
hsa04720	Long-term potentiation	1,3975410	2,6e-02		PPP3CA
hsa04810	Regulation of actin cytoskeleton	2,3356164	2,8e-02	CYFIP2, CXCR4	MYL9, ITGA2, RRAS, EGFR
hsa04916	Melanogenesis	1,8138298	3,1e-02	TCF7L2	EDN1
hsa05031	Amphetamine addiction	1,3750000	3,2e-02		PPP3CA
hsa05235	PD-L1 expression and PD-1 checkpoint pathway in cancer	2,0792683	3,5e-02		PPP3CA, EGFR
hsa05131	Shigellosis	1,2536765	3,5e-02		CTTN, EGFR, MYL9
hsa04370	VEGF signaling pathway	1,6394231	3,5e-02		PPP3CA
hsa04662	B-cell receptor signaling pathway	1,1217105	3,5e-02		PPP3CA
hsa04724	Glutamatergic synapse	0,9472222	3,6e-02		PPP3CA

Table 10. pathfindR pathway enrichment analysis for paclitaxel utilizing the gene features of the pan-drug model trained with mitotic inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05110	Vibrio cholerae infection	2,2434211	3,6e-02		TJP1
hsa04068	FoxO signaling pathway	1,3861789	3,7e-02		EGFR, PLK2
hsa05230	Central carbon metabolism in cancer	2,6640625	3,7e-02	LDHB	EGFR
hsa05212	Pancreatic cancer	1,2178571	3,7e-02		EGFR
hsa05218	Melanoma	1,2178571	3,7e-02		EGFR
hsa00564	Glycerophospholipid metabolism	0,9687500	3,9e-02		PLPP2
hsa04666	Fc gamma R-mediated phagocytosis	0,9578652	3,9e-02		PLPP2
hsa05132	Salmonella infection	1,9287330	4,0e-02	TCF7L2, CYFIP2	AHNAK2, RRAS, MYL9
hsa04922	Glucagon signaling pathway	1,6553398	4,1e-02	LDHB	PPP3CA
hsa04923	Regulation of lipolysis in adipocytes	1,6084906	4,3e-02		ADRB2
hsa04625	C-type lectin receptor signaling pathway	1,7947368	4,4e-02		RRAS, PPP3CA
hsa05160	Hepatitis C	1,1678082	4,5e-02		CLDN1, EGFR
hsa04728	Dopaminergic synapse	0,7349138	4,5e-02		PPP3CA
hsa05416	Viral myocarditis	1,6715686	4,7e-02	FYN	
hsa04664	Fc epsilon RI signaling pathway	1,3750000	4,7e-02	FYN	
hsa04114	Oocyte meiosis	0,6557692	4,9e-02		PPP3CA
hsa05418	Fluid shear stress and atherosclerosis	1,9229323	4,9e-02		PLAT, EDN1, SDC4
hsa04650	Natural killer cell mediated cytotoxicity	2,3250000	4,9e-02	FYN	PPP3CA, ULBP2

Table 10. pathfindR pathway enrichment analysis for paclitaxel utilizing the gene features of the pan-drug model trained with mitotic inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04630	JAK-STAT signaling pathway	0,5328125	5,0e-02		EGFR
hsa04218	Cellular senescence	1,7517123	5,0e-02		RRAS, PPP3CA, SERPINE1
hsa04022	cGMP-PKG signaling pathway	1,7164430	5,0e-02		MYL9, PPP3CA, ADRB2

Table 11. pathfindR pathway enrichment analysis for paclitaxel utilizing the gene features of the pan-drug model trained with cytoskeletal inhibitors (Strategy 2)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04115	p53 signaling pathway	5,6958242	4,5e-08		SFN, SERPINB5, THBS1, SERPINE1
hsa04915	Estrogen signaling pathway	4,9021438	9,6e-08		EGFR, KRT13, KRT14, KRT15, KRT18, KRT19
hsa04512	ECM-receptor interaction	8,4064875	9,7e-08		LAMB2, LAMB3, LAMC2, THBS1, SDC4, ITGA2, DSPP
hsa04510	Focal adhesion	3,6917379	9,7e-08		ITGA2, LAMB2, LAMB3, LAMC2, THBS1, EGFR, VEGFC
hsa05150	Staphylococcus aureus infection	6,7349272	6,2e-07		KRT13, KRT14, KRT15, KRT18, KRT19
hsa05165	Human papillomavirus infection	2,2580533	2,0e-06	TCF7L2	EGFR, ITGA2, LAMB2, LAMB3, LAMC2, THBS1
hsa04151	PI3K-Akt signaling pathway	2,6619950	2,3e-06		EGFR, EPHA2, AREG, VEGFC, ITGA2, LAMB2, LAMB3, LAMC2, THBS1
hsa05222	Small cell lung cancer	4,9223172	2,3e-06		ITGA2, LAMB2, LAMB3, LAMC2
hsa04390	Hippo signaling pathway	3,8584615	2,6e-06	TCF7L2	YAP1, CCN2, AREG, AJUBA, SERPINE1
hsa05205	Proteoglycans in cancer	3,6340545	2,9e-06		RRAS, CTTN, PLAU, THBS1, ITGA2, SDC4, EGFR
hsa04392	Hippo signaling pathway - multiple species	10,6796703	2,6e-05		YAP1, RASSF4, AJUBA

Table 11. pathfindR pathway enrichment analysis for paclitaxel utilizing the gene features of the pan-drug model trained with cytoskeletal inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04530	Tight junction	1,8806967	8,7e-05		TJP1, CLDN1, CTTN
hsa05130	Pathogenic Escherichia coli infection	2,8808359	8,7e-05	CYFIP2	CTTN, CLDN1, TJP1, MYO1C
hsa05213	Endometrial cancer	3,6917379	1,9e-04	TCF7L2	EGFR
hsa05210	Colorectal cancer	3,8835165	2,1e-04	TCF7L2	EGFR, AREG
hsa04520	Adherens junction	4,5828470	2,1e-04	TCF7L2	EGFR, TJP1, NECTIN2
hsa05215	Prostate cancer	4,5828470	2,1e-04	TCF7L2	EGFR, PLAU, PLAT
hsa04810	Regulation of actin cytoskeleton	2,2757288	2,2e-04	CYFIP2, CXCR4	ITGA2, RRAS, EGFR
hsa05146	Amoebiasis	5,3019640	2,5e-04		LAMB2, LAMB3, LAMC2, SERPINB6, SERPINB3
hsa04933	AGE-RAGE signaling pathway in diabetic complications	4,3337793	4,6e-04		EGR1, F3, VEGFC, SERPINE1
hsa04014	Ras signaling pathway	1,8123077	4,6e-04		RRAS, VEGFC, EGFR, EPHA2
hsa04010	MAPK signaling pathway	2,2399309	5,0e-04		RRAS, EGFR, EPHA2, PPP3CA, AREG, VEGFC
hsa04934	Cushing syndrome	1,4877153	5,0e-04	TCF7L2	EGFR
hsa05226	Gastric cancer	2,1513005	5,0e-04	TCF7L2	JUP, EGFR
hsa05224	Breast cancer	1,4239560	5,0e-04	TCF7L2	EGFR
hsa05200	Pathways in cancer	1,8159763	5,8e-04	TCF7L2, CXCR4	LAMB2, LAMB3, LAMC2, ITGA2, EGFR, JUP, VEGFC

Table 11. pathfindR pathway enrichment analysis for paclitaxel utilizing the gene features of the pan-drug model trained with cytoskeletal inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05225	Hepatocellular carcinoma	1,2779093	6,0e-04	TCF7L2	EGFR
hsa04610	Complement and coagulation cascades	5,6958242	1,7e-03		F3, SERPINE1, PLAUI, PLAT
hsa01521	EGFR TKI resistance	2,8479121	1,7e-03		EGFR, AXL
hsa04015	Rap1 signaling pathway	3,0053344	2,2e-03		VEGFC, EGFR, EPHA2, ADORA2B, RRAS, THBS1
hsa05221	Acute myeloid leukemia	3,3225641	2,6e-03	TCF7L2	JUP
hsa04912	GnRH signaling pathway	2,3732601	2,7e-03		EGFR, EGR1
hsa04657	IL-17 signaling pathway	2,3453394	3,0e-03		MMP13, LCN2
hsa05216	Thyroid cancer	2,9316742	3,1e-03	TCF7L2	
hsa04072	Phospholipase D signaling pathway	2,1513005	3,4e-03		RRAS, EGFR, PLPP2
hsa04928	Parathyroid hormone synthesis, secretion and action	2,8479121	3,5e-03		EGFR, MMP13, EGR1
hsa05120	Epithelial cell signaling in Helicobacter pylori infection	2,9754305	3,6e-03		TJP1, EGFR
hsa04540	Gap junction	2,6230769	3,7e-03		TJP1, EGFR
hsa04926	Relaxin signaling pathway	2,5341591	4,4e-03		VEGFC, MMP13, EGFR
hsa04371	Apelin signaling pathway	3,9554335	4,9e-03		SERPINE1, RRAS, CCN2, EGR1, PLAT
hsa05207	Chemical carcinogenesis - receptor activation	2,2028049	7,6e-03		EGFR, ADRB2, KLF5, KLF4

Table 11. pathfindR pathway enrichment analysis for paclitaxel utilizing the gene features of the pan-drug model trained with cytoskeletal inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05217	Basal cell carcinoma	1,6076923	8,6e-03	TCF7L2	
hsa05145	Toxoplasmosis	3,1149038	9,7e-03		LAMB2, LAMB3, LAMC2
hsa05412	Arrhythmogenic right ventricular cardiomyopathy	5,4617492	1,1e-02	TCF7L2	DSG2, JUP, ITGA2
hsa04921	Oxytocin signaling pathway	1,5334911	1,1e-02		PPP3CA, EGFR
hsa05219	Bladder cancer	5,3879418	1,3e-02		THBS1, EGFR
hsa05163	Human cytomegalovirus infection	1,4877153	1,3e-02	CXCR4	EGFR, PPP3CA
hsa04020	Calcium signaling pathway	2,7061608	1,3e-02	CXCR4	PPP3CA, EGFR, ADORA2B, ADRB2, VEGFC
hsa04916	Melanogenesis	1,0603928	1,6e-02	TCF7L2	
hsa04820	Cytoskeleton in muscle cells	4,2596976	1,6e-02		ITGA2, DSG2, JUP, THBS1, SDC4
hsa05230	Central carbon metabolism in cancer	1,5574519	1,6e-02		EGFR
hsa05223	NSCLC	1,5574519	1,6e-02		EGFR
hsa05212	Pancreatic cancer	1,4239560	1,6e-02		EGFR
hsa05214	Glioma	1,4239560	1,6e-02		EGFR
hsa05218	Melanoma	1,4239560	1,6e-02		EGFR
hsa04012	ErbB signaling pathway	2,6939709	1,7e-02		EGFR, AREG
hsa05235	PD-L1 expression and PD-1 checkpoint pathway in cancer	2,4311445	1,7e-02		PPP3CA, EGFR
hsa05231	Choline metabolism in cancer	2,2399309	1,8e-02		EGFR, PLPP2

Table 11. pathfindR pathway enrichment analysis for paclitaxel utilizing the gene features of the pan-drug model trained with cytoskeletal inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa01522	Endocrine resistance	1,1199654	1,8e-02		EGFR
hsa04625	C-type lectin receptor signaling pathway	2,0984615	1,9e-02		RRAS, PPP3CA
hsa04066	HIF-1 signaling pathway	1,9935385	2,0e-02		EGFR, SERPINE1
hsa04611	Platelet activation	0,8447197	2,5e-02		ITGA2
hsa04068	FoxO signaling pathway	1,6207630	2,5e-02		EGFR, PLK2
hsa00565	Ether lipid metabolism	2,1668896	2,6e-02		PLPP2
hsa04924	Renin secretion	3,3788787	2,6e-02		ADRB2, PPP3CA
hsa04720	Long-term potentiation	1,6340479	2,6e-02		PPP3CA
hsa05206	MicroRNAs in cancer	1,5217851	2,6e-02		EGFR, SERPINB5, PLAUI, THBS1
hsa05171	Coronavirus disease - COVID-19	0,7329186	3,0e-02		EGFR
hsa04936	Alcoholic liver disease	0,7329186	3,0e-02	TCF7L2	
hsa04218	Cellular senescence	2,0481560	3,0e-02		RRAS, PPP3CA, SERPINE1
hsa05160	Hepatitis C	1,3654373	3,0e-02		CLDN1, EGFR
hsa05031	Amphetamine addiction	1,6076923	3,2e-02		PPP3CA
hsa04140	Autophagy - animal	0,6601121	3,2e-02		RRAS
hsa04630	JAK-STAT signaling pathway	0,6229808	3,4e-02		EGFR
hsa04360	Axon guidance	3,0389306	3,4e-02	CXCR4	PPP3CA, EPHA2, RHOD, RRAS
hsa05131	Shigellosis	0,9772247	3,5e-02		CTTN, EGFR

Table 11. pathfindR pathway enrichment analysis for paclitaxel utilizing the gene features of the pan-drug model trained with cytoskeletal inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04310	Wnt signaling pathway	1,8806967	3,5e-02	TCF7L2	PPP3CA, DKK1
hsa04370	VEGF signaling pathway	1,9168639	3,5e-02		PPP3CA
hsa04662	B-cell receptor signaling pathway	1,3115385	3,5e-02		PPP3CA
hsa04724	Glutamatergic synapse	1,1075214	3,6e-02		PPP3CA
hsa04668	TNF signaling pathway	0,9869002	3,7e-02		VEGFC
hsa00564	Glycerophospholipid metabolism	1,1326923	3,9e-02		PLPP2
hsa04666	Fc gamma R-mediated phagocytosis	1,1199654	3,9e-02		PLPP2
hsa05167	Kaposi sarcoma-associated herpesvirus infection	1,1457118	3,9e-02	TCF7L2	PPP3CA
hsa04922	Glucagon signaling pathway	0,9677371	4,1e-02		PPP3CA
hsa04923	Regulation of lipolysis in adipocytes	1,8806967	4,3e-02		ADRB2
hsa05208	Chemical carcinogenesis - ROS	0,5246154	4,4e-02		EGFR
hsa05110	Vibrio cholerae infection	2,6230769	4,4e-02		TJP1
hsa04728	Dopaminergic synapse	0,8592838	4,5e-02		PPP3CA
hsa04024	cAMP signaling pathway	0,9724578	4,7e-02		ADRB2, RRAS
hsa04114	Oocyte meiosis	0,7667456	4,9e-02		PPP3CA
hsa05418	Fluid shear stress and atherosclerosis	1,4989011	4,9e-02		PLAT, SDC4
hsa04650	Natural killer cell mediated cytotoxicity	1,8123077	4,9e-02		PPP3CA, ULBP2

Table 11. pathfindR pathway enrichment analysis for paclitaxel utilizing the gene features of the pan-drug model trained with cytoskeletal inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04022	cGMP-PKG signaling pathway	1,3379453	5,0e-02		PPP3CA, ADRB2
hsa05132	Salmonella infection	1,8041072	5,0e-02	TCF7L2, CYFIP2	AHNAK2, RRAS

Table 12. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 1)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04512	ECM-receptor interaction	5,3050649	7,8e-28	LAMA3, LAMB1, LAMB2, LAMC1, LAMC2, COL1A2, COL4A1, COL4A2, COL4A5, TNC, FN1, THBS1, ITGB1, ITGA6, SDC1, SDC4, ITGB5, ITGA2, AGRN, DSPP	ITGA4
hsa04510	Focal adhesion	3,6610161	1,2e-24	ITGB1, ITGB5, ITGA6, ITGA2, MYLK, PARVA, ACTN1, PTK2, CAPN2, COL1A2, COL4A1, COL4A2, COL4A5, FN1, TNC, LAMA3, LAMB1, LAMB2, LAMC1, LAMC2, THBS1, CAV2, CCND1, MET, PDGFC, VEGFC, JUN, EMP2, EMP1	ITGA4, CCND2, RAC2, VAV1
hsa04151	PI3K-Akt signaling pathway	2,2398663	2,1e-19	EPHA2, MET, CCND1, EFNA1, AREG, PDGFC, BDNF, VEGFC, GNG12, ITGA6, ITGA2, ITGB1, ITGB5, COL1A2, COL4A1, COL4A2, COL4A5, FN1, TNC, LAMA3, LAMB1, LAMB2, LAMC1, LAMC2, THBS1, PTK2, SGK1, MAGI1	PIK3CG, SYK, GNG7, ITGA4, IL2RB, MYB, PIK3AP1, CCND2

Table 12. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05222	Small cell lung cancer	3,8828958	2,3e-18	CCND1, ITGB1, ITGA6, ITGA2, PTK2, COL4A1, COL4A2, COL4A5, FN1, LAMA3, LAMB1, LAMB2, LAMC1, LAMC2	E2F2
hsa05165	Human papillomavirus infection	1,8321237	4,1e-17	CCND1, ITGA6, ITGA2, ITGB1, ITGB5, COL1A2, COL4A1, COL4A2, COL4A5, FN1, TNC, LAMA3, LAMB1, LAMB2, LAMC1, LAMC2, THBS1, PTK2, MAGI1, PRKCZ, HES1, FZD6	ITGA4, TNF, CCND2, MFNG, TCF7L2
hsa04390	Hippo signaling pathway	2,7055016	3,6e-16	WWC1, FRMD6, YAP1, WWTR1, CTNNA1, CCN2, AREG, CDH1, TEAD4, AJUBA, PRKCZ, SERPINE1, FZD6, CCND1, SNAI2, BMP4, ID1	ITGB2, TCF7L2, CCND2
hsa05146	Amoebiasis	4,2381395	2,2e-14	FN1, LAMA3, LAMB1, LAMB2, LAMC1, LAMC2, CXCL1, CXCL2, GNA11, COL1A2, COL4A1, COL4A2, COL4A5, ACTN1, COL3A1, PTK2, SERPINB6	TNF, ITGB2

Table 12. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04820	Cytoskeleton in muscle cells	4,1218433	3,5e-14	SGCE, ITGA6, ITGA2, ITGB1, ITGB5, DSG2, JUP, ATP1B1, LMNA, COL1A2, COL3A1, FN1, THBS1, COL4A1, COL4A2, COL4A5, AGRN, SDC1, SDC2, SDC4, FHL2, TPM1	ITGA4
hsa04068	FoxO signaling pathway	1,1932802	4,2e-11	PRKAG2, CCND1, PLK2, SGK1, FBXO32	CCND2, RAG1
hsa04015	Rap1 signaling pathway	2,2126653	7,0e-11	EFNA1, PDGFC, VEGFC, EPHA2, MET, ENAH, CDH1, ITGB1, ADORA2B, MAGI1, PRKD1, GNAI1, RRAS, ID1, THBS1,	APBB1IP, RAC2, ITGAL, ITGB2, VAV1
hsa04933	AGE-RAGE signaling pathway in diabetic complications	3,8744548	7,0e-11	EGR1, EDN1, F3, JUN, CCND1, COL1A2, COL3A1, COL4A1, COL4A2, COL4A5, VEGFC, SERPINE1, PRKCZ, FN1	TNF, STAT5A, PLCG2
hsa04062	Chemokine signaling pathway	2,0138409	1,3e-10	CXCL1, CXCL2, CCL28, PTK2, GNAI1, GNG12, PRKCZ	CXCR4, CCL3, LYN, PLCG2, GNG7, VAV1, DOCK2, RAC2, PREX1, PIK3CG

Table 12. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04662	B-cell receptor signaling pathway	3,3106796	6,5e-10	FOS, JUN, PPP3CA	VAV1, PPP3CC, RAC2, PLCG2, BTK, LYN, SYK, PTPN6, PIK3AP1
hsa04668	TNF signaling pathway	1,6608030	1,7e-09	FOS, JUN, CXCL1, CXCL2, EDN1, VEGFC	TNF, TNFAIP3
hsa04061	Viral protein interaction with cytokine and cytokine receptor	1,9061489	3,7e-09	CCL28, CXCL1, CXCL2, LTBR	CCL3, CXCR4, IL2RB, TNF
hsa05205	Proteoglycans in cancer	2,8393676	4,4e-09	RRAS, CTTN, CCND1, PLAU, PTK2, MET, TIMP3, THBS1, LUM, SDC1, COL1A2, ITGA2, ITGB1, ITGB5, FZD6, SDC2, SDC4, FN1, CTSL, CAV2	ITPR1, HCLS1, VAV1, TNF, PLCG2, PTPN6
hsa04520	Adherens junction	3,6151099	6,9e-09	CDH1, CTNNA1, YES1, SNAI2, MET, PTPRF, PTPRM, TJP1, ACTN1, NECTIN3, NECTIN2, ARHGAP29	TCF7L2, PTPN6, RAC2
hsa05323	Rheumatoid arthritis	3,4072411	8,0e-09	CTSL, CXCL1, CXCL2, FOS, JUN	TNF, HLA-DMA, HLA-DPB1, HLA-DQB1, TNFSF13B, ITGAL, ITGB2, CCL3

Table 12. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05200	Pathways in cancer	2,0373413	9,1e-09	CDH1, CTNNA1, ITGB1, COL4A1, COL4A2, COL4A5, FN1, LAMA3, LAMB1, LAMB2, LAMC1, LAMC2, ITGA6, ITGA2, PTK2, FZD6, MET, CCND1, JUN, FOS, MECOM, CTBP2, PPARG, JUP, MTF, BMP4, VEGFC, GNG12, GNAI1, GNAI1, HES1, NFE2L2, NQO1, GSTA4, MGST1, IL13RA1, EDN1	STAT5A, PLCG2, RAC2, TCF7L2, CCND2, E2F2, GNG7, CXCR4, PIM2, CSF2RB, IL2RB
hsa05167	Kaposi sarcoma-associated herpesvirus infection	2,1690660	9,1e-09	C3, GNG12, CXCL1, CXCL2, PPP3CA, FOS, JUN, CCND1	E2F2, GNG7, SYK, PPP3CC, ITPR1, PLCG2, LYN, PIK3CG, PREX1, TCF7L2
hsa04611	Platelet activation	2,6653777	1,3e-08	ITGA2, ITGB1, MYLK, GNAI1, COL1A2, COL3A1, PRKCZ	ITPR1, PLCG2, APBB1IP, PIK3CG, FERMT3, LYN, BTK, SYK
hsa04310	Wnt signaling pathway	1,7143351	1,3e-08	PPP3CA, FZD6, CCND1, FOSL1, JUN, TLE1, CTBP2, DKK1, BAMBI	PPP3CC, RAC2, CCND2, TCF7L2

Table 12. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05130	Pathogenic Escherichia coli infection	2,6664048	1,4e-08	CTTN, CLDN4, CLDN1, ITGB1, TUBB3, TUBB2A, FOS, JUN, NCKAP1, CYFIP1, TJP1, MYO1B, MYO1C, MYO6	HCLS1, WIPF1, CASP1, TNF, PTPN6, CYFIP2, MYO1F, MYO1G
hsa04380	Osteoclast differentiation	2,8745955	3,5e-08	TNFRSF11B, FOS, JUN, FOSL1, PPP3CA, MITF, FHL2, PPARG	BTK, SYK, PLCG2, PPP3CC, TNF, CYBA, NCF2, NCF4, ITPR1
hsa05417	Lipid and atherosclerosis	2,0857859	4,8e-08	FOS, JUN, NFE2L2, CXCL1, CXCL2, PPP3CA, HSPA2, PTK2, PPARG	CASP1, CYBA, NCF2, NCF4, CCL3, TNF, PPP3CC, LYN, VAV1, ITPR1
hsa04658	Th1 and Th2 cell differentiation	2,7134590	5,9e-08	PPP3CA, JUN, FOS	PPP3CC, CD3D, IL2RB, STAT5A, HLA-DMA, HLA-DPB1, HLA-DQB1, RUNX3
hsa05140	Leishmaniasis	4,4476807	6,5e-08	ITGB1, C3, FOS, JUN	ITGA4, ITGB2, HLA-DMA, HLA-DPB1, HLA-DQB1, TNF, PTPN6, CYBA, NCF2, NCF4
hsa04810	Regulation of actin cytoskeleton	2,2978233	7,1e-08	NCKAP1, CYFIP1, CFL2, GSN, ACTN1, MYLK, ENAH, GNG12, PTK2, ITGA6, ITGA2, ITGB1, ITGB5, FN1, RRAS, PDGFC	CYFIP2, RAC2, IQGAP2, VAV1, ITGA4, ITGAL, ITGB2, CXCR4

Table 12. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05226	Gastric cancer	1,2067705	9,8e-08	CCND1, FZD6, MET, CDH1, CTNNA1, JUP	TCF7L2, E2F2
hsa04664	Fc epsilon RI signaling pathway	2,3673139	1,0e-07		TNF, BTK, SYK, LYN, PLCG2, RAC2, VAV1
hsa04659	Th17 cell differentiation	2,2836041	1,0e-07	PPP3CA, JUN, FOS, AHR	PPP3CC, IL2RB, STAT5A, HLA-DMA, HLA-DPB1, HLA-DQB1, CD3D
hsa05145	Toxoplasmosis	2,8393676	1,1e-07	GNAI1, ITGA6, ITGB1, LAMA3, LAMB1, LAMB2, LAMC1, LAMC2	TNF, HLA-DMA, HLA-DPB1, HLA-DQB1, PIK3CG
hsa04670	Leukocyte transendothelial migration	3,3627343	1,1e-07	CLDN4, CLDN1, ITGB1, PTK2, CTNNA1, GNAI1, ACTN1	ITGB2, ITGAL, ITGA4, CYBA, NCF2, NCF4, VAV1, RAC2, CXCR4, PLCG2
hsa04360	Axon guidance	2,4291775	1,2e-07	GNAI1, PPP3CA, PTK2, ITGB1, EPHA2, EFNA1, MET, SEMA3A, SEMA3C, RHOD, CFL2, RRAS, PRKCZ, ENAH	CXCR4, RAC2, PPP3CC, FES, PLCG2
hsa05166	Human T-cell leukemia virus 1 infection	2,0857859	1,5e-07	PPP3CA, LTBR, JUN, CCND1, FOS, EGR1, FOSL1	ITGAL, ITGB2, PPP3CC, IL2RB, STAT5A, CCND2, CD3D, E2F2, HLA-DMA, HLA-DPB1, HLA-DQB1, TNF

Table 12. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05412	Arrhythmogenic right ventricular cardiomyopathy	3,1595070	2,1e-07	LMNA, DSG2, JUP, ITGA6, ITGA2, ITGB1, ITGB5, CTNNA1, SGCE	ITGA4, TCF7L2
hsa04974	Protein digestion and absorption	3,5944521	2,7e-07	ATP1B1, COL1A2, COL3A1, COL4A1, COL4A2, COL4A5	
hsa05170	Human immunodeficiency virus 1 infection	1,9478133	3,3e-07	GNG12, PPP3CA, GNAI1, GNAI1, PTK2, FOS, JUN, CFL2, AP1M2	GNG7, PPP3CC, TNF, CXCR4, ITPR1, PLCG2, RAC2, CD3D
hsa04915	Estrogen signaling pathway	1,2030612	3,8e-07	FOS, JUN, GNAI1, KRT15, KRT18, KRT19	ITPR1
hsa05418	Fluid shear stress and atherosclerosis	3,1530282	4,1e-07	FOS, JUN, NFE2L2, NQO1, BMP4, PRKCZ, GSTA4, MGST1, PLAT, CTSL, EDN1, PTK2, CAV2, SDC1, SDC2, SDC4	CYBA, NCF2, RAC2, TNF
hsa04371	Apelin signaling pathway	2,3297375	4,6e-07	PRKAG2, SERPINE1, GNAI1, GNG12, RRAS, CCN2, MYLK, CCND1, EGR1, PLAT, CDH1	GNG7, PIK3CG, ITPR1
hsa04657	IL-17 signaling pathway	1,7267466	4,6e-07	FOS, JUN, FOSL1, CXCL1, CXCL2	TNFAIP3, TNF
hsa05100	Bacterial invasion of epithelial cells	2,6209547	7,5e-07	CTNNA1, CDH1, CTTN, ITGB1, PTK2, MET, FN1, CAV2	HCLS1

Table 12. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05163	Human cytomegalovirus infection	1,4604325	8,7e-07	GNG12, PPP3CA, GNAI1, GNAI1, CCND1, PTK2	TNF, GNG7, PPP3CC, ITPR1, CCL3, RAC2, CXCR4, E2F2
hsa04612	Antigen processing and presentation	3,5538369	9,2e-07	CTSB, CTSB, LGMN, HSPA2	CD74, HLA-DMA, HLA-DPB1, HLA-DQB1, CTSS, TNF
hsa04630	JAK-STAT signaling pathway	0,9173341	9,8e-07	IL13RA1, CCND1	STAT5A, IL2RB, CCND2, PTPN6, CSF2RB
hsa05213	Endometrial cancer	1,5531583	1,3e-06	CTNNA1, CDH1, CCND1	TCF7L2
hsa04530	Tight junction	1,8462071	1,6e-06	TJP1, PRKCZ, MAGI1, ACTN1, MICALL2, RAB13, CLDN4, CLDN1, PRKAG2, JUN, CCND1, CTTN, ITGB1	HCLS1
hsa05162	Measles	1,3419288	1,6e-06	CCND1, FOS, JUN	IL2RB, CD3D, STAT5A, TNFAIP3, CCND2
hsa04926	Relaxin signaling pathway	2,3099940	1,6e-06	GNG12, PRKCZ, EDN1, GNAI1, VEGFC, COL1A2, COL3A1, COL4A1, COL4A2, COL4A5, FOS, JUN	GNG7
hsa05133	Pertussis	3,6465457	2,3e-06	ITGB1, GNAI1, CFL2, C1R, C1S, C3, FOS, JUN	ITGB2, TNF, IRF8, CASP1

Table 12. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04650	Natural killer cell mediated cytotoxicity	2,4779935	2,3e-06	PPP3CA	PPP3CC, PLCG2, TNF, LAT2, SYK, PTPN6, CD48, ICAM2, RAC2, VAV1, ITGB2, ITGAL
hsa04917	Prolactin signaling pathway	1,2707659	2,7e-06	CCND1, FOS	STAT5A, CCND2
hsa04014	Ras signaling pathway	1,1436893	4,9e-06	RRAS, EFNA1, PDGFC, BDNF, VEGFC, EPHA2, MET, GNG12	RAC2, RASAL3, GNG7, PLCG2
hsa04060	Cytokine-cytokine receptor interaction	1,3201846	5,1e-06	INHBB, LTBR, IL13RA1, TNFRSF12A, TNFRSF11B, CXCL1, CXCL2, CCL28, GDF15, BMP4	IL2RB, CSF2RB, TNFSF13B, CD70, CXCR4, TNF, CCL3
hsa04921	Oxytocin signaling pathway	1,6128952	5,2e-06	MYLK, PRKAG2, PPP3CA, JUN, FOS, CCND1, GNAI1	ITPR1, PPP3CC, PIK3CG
hsa04010	MAPK signaling pathway	1,4920791	6,0e-06	RRAS, GNG12, EPHA2, MET, MECOM, FOS, JUN, PPP3CA, EFNA1, AREG, PDGFC, BDNF, VEGFC	DUSP2, PTPN7, TNF, RAC2, PPP3CC, MAP4K1
hsa04713	Circadian entrainment	1,2785145	6,2e-06	GNG12, GNAI1, FOS	GNG7, ITPR1
hsa04725	Cholinergic synapse	1,9734247	6,2e-06	GNAI1, GNAI1, GNG12, FOS	ITPR1, CHAT, GNG7, PIK3CG

Table 12. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04613	Neutrophil extracellular trap formation	1,9595923	1,1e-05	C3	SYK, PLCG2, CASP1, ITGB2, CYBA, NCF2, RAC2, NCF4, ITGAL
hsa04064	NF-κB signaling pathway	2,6510806	1,1e-05	LTBR, PLAU, CXCL1, CXCL2	TNF, TNFSF13B, TNFAIP3, PLCG2, SYK, LYN, BTK
hsa05216	Thyroid cancer	2,4667809	1,2e-05	PPARG, CDH1, CCND1	TCF7L2
hsa05221	Acute myeloid leukemia	1,7473031	1,2e-05	JUP, CCND1	STAT5A, TCF7L2, PIM2
hsa05410	Hypertrophic cardiomyopathy	2,6209547	1,2e-05	ITGA6, ITGA2, ITGB1, ITGB5, LMNA, TPM1, PRKAG2, EDN1, SGCE	ITGA4, TNF
hsa04020	Calcium signaling pathway	1,1385143	1,3e-05	PPP3CA, MET, GNA11, MYLK, ADORA2B, ASPH, PDGFC, VEGFC	PPP3CC, PLCG2, ITPR1, CXCR4
hsa04012	ErbB signaling pathway	1,4167323	1,3e-05	PTK2, JUN, AREG	STAT5A, PLCG2
hsa04625	C-type lectin receptor signaling pathway	1,9864078	1,3e-05	RRAS, PPP3CA, JUN	SYK, CASP1, PLCG2, PPP3CC,
hsa04922	Glucagon signaling pathway	1,0178465	1,3e-05	PYGL, PPP3CA, PRKAG2	PPP3CC, ITPR1
hsa05224	Breast cancer	1,0483819	1,3e-05	FOS, JUN, CCND1, FZD6, HES1	TCF7L2, E2F2
hsa05210	Colorectal cancer	1,6338419	1,4e-05	CCND1, JUN, FOS, AREG	RAC2, TCF7L2
hsa04666	Fc gamma R-mediated phagocytosis	2,8270972	1,4e-05	CFL2, ASAP2, MARCKS, GSN, PLPP2, PLPP3	LYN, SYK, PTPRC, PLCG2, VAV1, RAC2

Table 12. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05414	Dilated cardiomyopathy	2,0291262	1,4e-05	ITGA6, ITGA2, ITGB1, ITGB5, LMNA, TPM1, SGCE	ITGA4, TNF
hsa04066	HIF-1 signaling pathway	1,0483819	1,5e-05	HK1, SERPINE1, EDN1, LTBR	PLCG2
hsa04660	T-cell receptor signaling pathway	1,6267995	1,6e-05	FOS, JUN, PPP3CA	TNF, PPP3CC, VAV1, CD3D, PTPRC, PTPN6
hsa00564	Glycerophospholipid metabolism	1,1913430	1,8e-05	PLPP2, PLPP3, MBOAT2, CDS1	CHAT
hsa05142	Chagas disease	1,9454509	1,9e-05	C3, FOS, JUN, SERPINE1, GNAI1, GNA11	TNF, CCL3, CD3D
hsa04971	Gastric acid secretion	1,4460440	2,0e-05	ATP1B1, GNAI1, MYLK	ITPR1
hsa04728	Dopaminergic synapse	1,2652885	2,0e-05	GNAI1, GNG12, FOS, PPP3CA	ITPR1, GNG7, PPP3CC
hsa04514	Cell adhesion molecules	3,3019902	2,0e-05	ITGB1, ITGA6, NECTIN2, CDH1, NECTIN3, PTPRM, SDC1, SDC2, SDC4, PTPRF, CLDN4, CLDN1	PTPRC, ITGB2, ITGAL, ITGA4, ICAM2, HLA-DMA, HLA-DPB1, HLA-DQB1
hsa05135	Yersinia infection	1,8305080	2,3e-05	ITGB1, FN1, PTK2, FOS, JUN	WIPF1, CASP1, ITGA4, VAV1, TNF, RAC2
hsa04814	Motor proteins	1,8500857	3,0e-05	DYNLL1, TUBB3, TUBB2A, MYO1B, MYO1C, MYO6, TPM1	MYO1F, MYO1G

Table 12. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05169	Epstein-Barr virus infection	1,9913399	4,0e-05	JUN, HES1, CCND1	HLA-DMA, HLA-DPB1, HLA-DQB1, ITGAL, TNFAIP3, LYN, SYK, PLCG2, RUNX3, E2F2, CCND2, BTK, TNF, CD3D
hsa04924	Renin secretion	2,1323021	4,4e-05	GNAI1, CTSB, PPP3CA, EDN1	ITPR1, PPP3CC
hsa04022	cGMP-PKG signaling pathway	1,1257792	4,4e-05	GNAI1, PPP3CA, ATP1B1, MYLK, GNAI1	ITPR1, PPP3CC, PIK3CG
hsa04720	Long-term potentiation	1,0311953	4,4e-05	PPP3CA	PPP3CC, ITPR1
hsa04724	Glutamatergic synapse	1,3978425	4,9e-05	GNAI1, GNG12, PPP3CA	ITPR1, GNG7, PPP3CC
hsa04726	Serotonergic	1,2942986	6,8e-05	GNAI1, GNG12,	ITPR1, GNG7
hsa05415	Diabetic cardiomyopathy	0,8435256	8,1e-05	COL1A2, COL3A1, PRKCZ	CYBA, NCF2, RAC2, NCF4
hsa04210	Apoptosis	1,9220334	8,4e-05	CAPN2, LMNA, JUN, FOS, CTSB, CTSB, CTSV	CSF2RB, TNF, CTSS, ITPR1
hsa04072	Phospholipase D signaling pathway	1,0559242	8,8e-05	RRAS, PLPP2, PLPP3, PDGFC	PLCG2, PIK3CG, SYK
hsa05152	Tuberculosis	1,4063659	9,2e-05	PPP3CA, C3	PPP3CC, ITGB2, TNF, CD74, HLA-DMA, HLA-DPB1, HLA-DQB1, CTSS, SYK
hsa05207	Chemical carcinogenesis - receptor activation	1,1584330	9,9e-05	FOS, JUN, AHR, GSTA4, MGST1, CCND1, KLF5, KLF4, GNAI1	STAT5A

Table 12. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05416	Viral myocarditis	3,2890412	1,0e-04	CCND1, SGCE	HLA-DMA, HLA-DPB1, HLA-DQB1, ITGAL, ITGB2, RAC2
hsa04370	VEGF signaling pathway	2,0161190	1,1e-04	PPP3CA, PTK2	PLCG2, PPP3CC, RAC2
hsa04621	NOD-like receptor signaling pathway	1,7595220	1,4e-04	JUN, CTSB, CXCL1, CXCL2	CASP1, CARD8, AIM2, ITPR1, CYBA, IFI16, TNFAIP3, TNF
hsa05225	Hepatocellular carcinoma	1,3440793	1,5e-04	CCND1, FZD6, MET, NFE2L2, NQO1, GSTA4, MGST1	TCF7L2, E2F2, PLCG2
hsa05120	Epithelial cell signaling in Helicobacter pylori infection	2,5035985	1,5e-04	JUN, CXCL1, CXCL2, MET, TJP1	PLCG2, CSK, LYN
hsa05321	Inflammatory bowel disease	1,8075550	1,6e-04	JUN	TNF, HLA-DMA, HLA-DPB1, HLA-DQB1
hsa05020	Prion disease	1,1851273	2,0e-04	LAMC1, EGR1, TUBB3, TUBB2A, PPP3CA, CAV2	ITPR1, PPP3CC, TNF, CYBA, NCF2, RAC2, NCF4
hsa01522	Endocrine resistance	1,1779572	2,0e-04	CCND1, JUN, FOS, PTK2	E2F2
hsa04723	Retrograde endocannabinoid signaling	0,9890395	2,0e-04	GNAI1, GNG12, MGLL	ITPR1, GNG7
hsa04935	Growth hormone synthesis, secretion and action	1,3717146	2,0e-04	GNAI1, FOS, PTK2, GNAI1	ITPR1, PLCG2, STAT5A
hsa04142	Lysosome	1,6553398	2,5e-04	CTSB, CTSL, CTSV, LGMN,	CTSS

Table 12. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04024	cAMP signaling pathway	1,0228116	2,8e-04	ATP1B1, GNAI1, SOX9, RRAS, BDNF, FOS, JUN, EDN1	VAV1, RAC2
hsa04620	Toll-like receptor signaling pathway	0,8646448	2,9e-04	FOS, JUN	TNF, CCL3
hsa05134	Legionellosis	2,6767197	2,9e-04	C3, CXCL1, CXCL2	CASP1, ITGB2, TNF
hsa04115	p53 signaling pathway	1,4976884	3,0e-04	SFN, THBS1, SERPINE1, CCND1	CCND2
hsa00983	Drug metabolism - other enzymes	0,8986130	3,0e-04	UPP1, GSTA4, MGST1	
hsa01521	EGFR TKI resistance	1,4976884	3,2e-04	MET, PDGFC, GAS6, AXL	PLCG2
hsa05310	Asthma	3,8122977	3,3e-04		HLA-DMA, HLA-DPB1, HLA-DQB1, TNF
hsa05144	Malaria	3,5798406	3,9e-04	SDC1, SDC2, MET, THBS1	TNF, ITGAL, ITGB2
hsa04727	GABAergic synapse	0,9828580	4,4e-04	GNAI1, GNG12	GNG7
hsa04916	Melanogenesis	1,1152999	4,9e-04	MITF, GNAI1, FZD6, EDN1	TCF7L2
hsa05132	Salmonella infection	1,8026476	5,1e-04	FOS, JUN, NCKAP1, CYFIP1, MYO6, SNX9, AHNAK2, AHNAK, RRAS, ANXA2, DYNLL1, TUBB3, TUBB2A	TNF, CASP1, TCF7L2, CYFIP2, PIK3CG, PTPRC
hsa04392	Hippo signaling pathway - multiple species	5,9907536	5,3e-04	WWC1, FRMD6, YAP1, WWTR1, TEAD4, RASSF4, AJUBA	RASSF2
hsa04114	Oocyte meiosis	0,6451581	5,3e-04	PPP3CA	FBXO43, ITPR1, PPP3CC

Table 12. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05330	Allograft rejection	3,1063167	5,4e-04		HLA-DMA, HLA-DPB1, HLA-DQB1, TNF
hsa05332	Graft-versus-host disease	2,9953768	5,6e-04		HLA-DMA, HLA-DPB1, HLA-DQB1, TNF
hsa04940	Type I diabetes mellitus	2,8920879	5,9e-04		HLA-DMA, HLA-DPB1, HLA-DQB1, TNF
hsa05032	Morphine addiction	1,1489116	7,0e-04	GNAI1, GNG12	GNG7, PDE7A
hsa05203	Viral carcinogenesis	1,5417381	7,2e-04	JUN, CCND1, GSN, ACTN1, C3, LTBR	STAT5A, LYN, SYK, CCND2
hsa04936	Alcoholic liver disease	1,0792166	7,7e-04	PRKAG2, CCND1, C3, CXCL1, CXCL2	TNF, TCF7L2
hsa04218	Cellular senescence	1,2925256	7,8e-04	RRAS, CCND1, CAPN2, PPP3CA, SERPINE1	CCND2, E2F2, PPP3CC, ITPR1
hsa04928	Parathyroid hormone synthesis, secretion and action	1,1981507	9,5e-04	GNA11, GNAI1, FOS, EGR1, SGK1	ITPR1
hsa04071	Sphingolipid signaling pathway	0,7765792	1,1e-03	GNAI1, PRKCZ	TNF, RAC2
hsa04672	Intestinal immune network for IgA production	4,5335433	1,1e-03	CCL28, LTBR	TNFSF13B, HLA-DMA, HLA-DPB1, HLA-DQB1, ITGA4, CXCR4
hsa05208	Chemical carcinogenesis - ROS	1,3242718	1,4e-03	MET, NQO1, PRKD1, PTK2, FOS, JUN, NFE2L2, AHR, GSTA4, MGST1	CYBA, NCF2

Table 12. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05031	Amphetamine addiction	1,3527508	1,5e-03	PPP3CA, FOS, JUN	PPP3CC
hsa05131	Shigellosis	1,6445206	1,5e-03	JUN, C3, WIP1, CTTN, ITGB1, PTK2, ACTN1, CAPN2, HK1	TNF, TIFA, CASP1, FNBP1, HCLS1, ITPR1, PLCG2
hsa05320	Autoimmune thyroid disease	1,4976884	1,5e-03		HLA-DMA, HLA-DPB1, HLA-DQB1
hsa04934	Cushing syndrome	1,4082742	1,7e-03	GNA11, FZD6, GNAI1, CCND1, AHR, MEN1	ITPR1, TCF7L2, E2F2
hsa04145	Phagosome	2,9704153	1,7e-03	C1R, ITGB1, ITGB5, ITGA2, TUBB3, TUBB2A, C3, THBS1, CTSL	ITGB2, CYBA, NCF4, NCF2, HLA-DMA, HLA-DPB1, HLA-DQB1, CTSS
hsa05223	NSCLC	1,6380967	2,2e-03	CCND1, MET	E2F2, PLCG2, STAT5A
hsa05034	Alcoholism	0,9707240	2,3e-03	BDNF, GNAI1, GNG12,	GNG7
hsa00565	Ether lipid	0,9116364	2,4e-03	PLPP2, PLPP3	
hsa04960	Aldosterone-regulated sodium	2,0967638	2,4e-03	ATP1B1, SGK1, SFN	
hsa04148	Efferocytosis	1,8500857	2,4e-03	GULP1, ITGB5, AXL, PTK2, GAS6, THBS1, PPARG, ADAM9, SGK1	DUSP2, PTPN6, CASP1
hsa05171	Coronavirus disease - COVID-19	1,3875642	2,4e-03	FOS, JUN, C1R, C1S, C3	TNF, SYK, PLCG2, CASP1
hsa05322	Systemic lupus erythematosus	3,3548220	2,5e-03	C3, C1S, C1R, ACTN1	TNF, HLA-DMA, HLA-DPB1, HLA-
hsa05220	Chronic myeloid leukemia	1,4560859	2,5e-03	MECOM, CTBP2, CCND1	STAT5A, E2F2

Table 12. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05161	Hepatitis B	0,6942926	2,5e-03	FOS, JUN	TNF, STAT5A, E2F2
hsa05164	Influenza A	0,7488442	2,8e-03		CASP1, HLA-DMA, HLA-DPB1, HLA-DQB1, TNF
hsa05235	PD-L1 expression and PD-1 checkpoint pathway	2,0456232	2,8e-03	FOS, JUN, PPP3CA	PPP3CC, CD3D, PTPN6, BATE, BATE3
hsa00860	Porphyrin metabolism	1,7473031	3,1e-03	BLVRA, BLVRB, FTH1	
hsa04750	Inflammatory mediator regulation of TRP channels	0,4820147	3,1e-03		PLCG2, ITPR1
hsa05204	Chemical carcinogenesis - DNA adducts	0,8387055	3,1e-03	GSTA4, MGST1	
hsa00830	Retinol metabolism	0,8064476	3,1e-03	ALDH1A1, ALDH1A3	
hsa05030	Cocaine addiction	1,3383598	3,2e-03	GNAI1, JUN,	
hsa00982	Drug metabolism - cytochrome P450	0,7230220	3,2e-03	GSTA4, MGST1	
hsa00980	Metabolism of xenobiotics by cytochrome P450	0,6763754	3,2e-03	GSTA4, MGST1	
hsa04070	Phosphatidylinositol signaling system	0,6837273	3,4e-03	CDS1	PLCG2, ITPR1
hsa05150	Staphylococcus aureus infection	3,1168110	3,4e-03	C3, C1S, C1R, KRT15, KRT18, KRT19	HLA-DMA, HLA-DPB1, HLA-DQB1, ITGB2, ITGAL
hsa05217	Basal cell carcinoma	1,0145631	3,5e-03	BMP4, FZD6	TCF7L2
hsa00561	Glycerolipid metabolism	1,8392665	3,7e-03	PLPP2, PLPP3, MBOAT2, MGLL	LPL
hsa04930	Type II diabetes mellitus	1,6553398	4,3e-03	PRKCZ, HK1	TNF

Table 12. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05231	Choline metabolism in cancer	1,4135486	4,3e-03	PDGFC, PLPP2, PLPP3, FOS, JUN	RAC2
hsa05218	Melanoma	1,7972261	4,4e-03	CDH1, CCND1, MITF, MET, PDGFC	E2F2
hsa04640	Hematopoietic cell lineage	1,9969179	5,1e-03	ITGA6, ITGA2	ITGA4, CD3D, HLA-DMA, HLA-DPB1, HLA-DQB1, TNF
hsa04931	Insulin resistance	1,0697774	5,2e-03	PTPRF, PRKCZ, PYGL, PRKAG2	TNF
hsa05202	Transcriptional misregulation in cancer	2,1758869	5,3e-03	PPARG, MET, PLAU, JUP, PTK2, MEN1, PLAT, MITF	IL2RB, CCND2, LYL1
hsa00562	Inositol phosphate metabolism	0,5990754	5,4e-03		PLCG2, PIK3CG
hsa04270	Vascular smooth muscle contraction	1,2438429	6,2e-03	MYLK, GNA11, CALD1, ADORA2B, EDN1, ADM	ITPR1
hsa05215	Prostate cancer	1,4460440	6,3e-03	CCND1, PDGFC, PLAU, PLAT	TCF7L2, E2F2
hsa04918	Thyroid hormone synthesis	1,3312786	6,5e-03	ATP1B1, GPX2, GPX8	ITPR1
hsa04110	Cell cycle	0,8960529	6,6e-03	SFN, CCND1, STAG1	CCND2, E2F2
hsa05010	Alzheimer disease	0,8544805	7,5e-03	BACE2, APP, CAPN2, PPP3CA, WIPI1, TUBB3, TUBB2A, FZD6, DKK1	TNF, ITPR1, PPP3CC, LPL
hsa04932	Non-alcoholic fatty liver disease	0,7596970	7,7e-03	PRKAG2, FOS, JUN, PPARG	TNF
hsa04970	Salivary secretion	0,6166952	7,7e-03	ATP1B1	ITPR1

Table 12. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04910	Insulin signaling pathway	0,8064476	9,1e-03	PTPRF, PRKAG2, PYGL, HK1, PRKCZ	
hsa04540	Gap junction	1,9312298	1,0e-02	GNAI1, TJP1, TUBB3, TUBB2A, PDGFC, GNA11	ITPR1
hsa04610	Complement and coagulation cascades	2,6958391	1,1e-02	F3, SERPINE1, SERPINE2, PLAUI, PLAT, C3, C1R, C1S	ITGB2
hsa05212	Pancreatic cancer	0,8986130	1,1e-02	CCND1	RAC2, E2F2
hsa04925	Aldosterone synthesis and	1,0483819	1,1e-02	GNA11, PRKD1, ATP1B1	ITPR1
hsa04912	GnRH signaling pathway	0,9984589	1,2e-02	GNA11, JUN, EGR1	ITPR1
hsa05219	Bladder cancer	2,2667716	1,3e-02	THBS1, CDH1, CCND1	E2F2
hsa01523	Antifolate resistance	0,9984589	1,5e-02		TNF
hsa05022	Pathways of neurodegeneration - multiple diseases	0,8202499	1,7e-02	TUBB3, TUBB2A, PPP3CA, PINK1, WIP1, BDNF, GPX2, GPX8, CAPN2, FZD6, DKK1, APP	ITPR1, PPP3CC, RAB39B, TNF
hsa04213	Longevity regulating pathway - multiple species	0,3956158	2,2e-02	PRKAG2	
hsa04340	Hedgehog signaling pathway	0,7624595	2,3e-02	CCND1	CCND2
hsa05110	Vibrio cholerae infection	1,1035599	2,4e-02	TJP1	PLCG2
hsa04929	GnRH secretion	1,0483819	2,6e-02	GNA11	ITPR1
hsa05211	Renal cell	0,6552387	2,9e-02	MET, JUN	
hsa04920	Adipocytokine signaling pathway	0,9530744	3,0e-02	PRKAG2	ACSL5, TNF

Table 12. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05014	Amyotrophic lateral sclerosis	0,8356667	3,1e-02	GPX2, GPX8, PPP3CA, PINK1, TUBB3, TUBB2A, WIP1	TNF, CASP1, PPP3CC, RAB39B
hsa05214	Glioma	0,8986130	3,4e-02	CCND1	E2F2, PLCG2
hsa05206	MicroRNAs in cancer	0,9603498	3,4e-02	TPM1, MET, CCND1, TIMP3, MARCKS, TNC, PLAU, EFNA1, THBS1	CCND2, E2F2, PLCG2
hsa05160	Hepatitis C	0,8616837	3,4e-02	CLDN4, CLDN1, IFIT1, CCND1	TNF, E2F2
hsa04962	Vasopressin-regulated water	1,0228116	3,4e-02	DYNLL1	ARHGDIB
hsa04261	Adrenergic signaling in	0,6212633	3,4e-02	TPM1, ATP1B1, GNAI1	PIK3CG
hsa04730	Long-term	1,6445206	3,6e-02	GNA11, GNAI1	LYN, ITPR1
hsa04137	Mitophagy - animal	0,9116364	3,7e-02	PINK1, RRAS, JUN, MITF	
hsa04211	Longevity regulating pathway	0,5114058	3,8e-02	PRKAG2, PPARG	
hsa04927	Cortisol synthesis and secretion	0,8064476	3,9e-02	GNA11	ITPR1
hsa04919	Thyroid hormone signaling pathway	0,7765792	3,9e-02	ATP1B1, BMP4, CCND1	PLCG2
hsa04975	Fat digestion and absorption	1,3104773	4,2e-02	PLPP2, PLPP3	
hsa04152	AMPK signaling pathway	0,5566629	4,7e-02	PRKAG2, PPARG, CCND1	
hsa04961	Endocrine and other factor-regulated calcium	0,8387055	4,8e-02	ATP1B1	KLK1
hsa04923	Regulation of lipolysis in	0,7912316	4,9e-02	MGLL, GNAI1	
hsa04722	Neurotrophin signaling pathway	0,9530744	5,0e-02	BDNF, JUN, MAGED1	PLCG2, ARHGDIB

Table 13. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 2)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04512	ECM-receptor interaction	5,4642169	2,7e-28	LAMA3, LAMB1, LAMB2, LAMC1, LAMC2, COL1A2, COL4A1, COL4A2, COL4A5, TNC, FN1, THBS1, ITGB1, ITGA6, SDC1, SDC4, ITGB5, ITGA2, AGRN, DSPP	ITGA4
hsa04510	Focal adhesion	3,7708466	3,3e-25	ITGB1, ITGB5, ITGA6, ITGA2, MYLK, PARVA, ACTN1, PTK2, CAPN2, COL1A2, COL4A1, COL4A2, COL4A5, FN1, TNC, LAMA3, LAMB1, LAMB2, LAMC1, LAMC2, THBS1, CAV2, CCND1, MET, PDGFC, VEGFC, JUN, EMP2, EMP1	ITGA4, CCND2, RAC2, VAV1

Table 13. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04151	PI3K-Akt signaling pathway	2,3070623	6,0e-20	EPHA2, MET, CCND1, EFNA1, AREG, PDGFC, BDNF, VEGFC, GNG12, ITGA6, ITGA2, ITGB1, ITGB5, COL1A2, COL4A1, COL4A2, COL4A5, FN1, TNC, LAMA3, LAMB1, LAMB2, LAMC1, LAMC2, THBS1, PTK2, SGK1, MAGI1	PIK3CG, SYK, GNG7, ITGA4, IL2RB, MYB, PIK3API, CCND2
hsa05222	Small cell lung cancer	3,9993827	1,1e-18	CCND1, ITGB1, ITGA6, ITGA2, PTK2, COL4A1, COL4A2, COL4A5, FN1, LAMA3, LAMB1, LAMB2, LAMC1, LAMC2	E2F2

Table 13. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05165	Human papillomavirus infection	1,8870874	1,5e-17	CCND1, ITGA6, ITGA2, ITGB1, ITGB5, COL1A2, COL4A1, COL4A2, COL4A5, FN1, TNC, LAMA3, LAMB1, LAMB2, LAMC1, LAMC2, THBS1, PTK2, MAGI1, PRKCZ, HES1, FZD6	ITGA4, TNF, CCND2, MFNG, TCF7L2
hsa04390	Hippo signaling pathway	2,7866667	3,6e-16	WWC1, FRMD6, YAP1, WWTR1, CTNNA1, CCN2, AREG, CDH1, TEAD4, AJUBA, PRKCZ, SERPINE1, FZD6, CCND1, SNAI2, BMP4, ID1	ITGB2, TCF7L2, CCND2

Table 13. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04820	Cytoskeleton in muscle cells	4,2454986	1,9e-14	SGCE, ITGA6, ITGA2, ITGB1, ITGB5, DSG2, JUP, ATP1B1, LMNA, COL1A2, COL3A1, FN1, THBS1, COL4A1, COL4A2, COL4A5, AGRN, SDC1, SDC2, SDC4, FHL2, TPM1	ITGA4
hsa05146	Amoebiasis	4,3652837	2,2e-14	FN1, LAMA3, LAMB1, LAMB2, LAMC1, LAMC2, CXCL1, CXCL2, GNA11, COL1A2, COL4A1, COL4A2, COL4A5, ACTN1, COL3A1, PTK2, SERPINB6	TNF, ITGB2
hsa04068	FoxO signaling pathway	1,2290786	4,2e-11	PRKAG2, CCND1, PLK2, SGK1, FBXO32	CCND2, RAG1

Table 13. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04015	Rap1 signaling pathway	2,2790452	7,0e-11	EFNA1, PDGFC, VEGFC, EPHA2, MET, ENAH, CDH1, ITGB1, ADORA2B, MAGI1, PRKD1, GNAI1, RRAS, ID1, THBS1, PRKCZ	APBB1IP, RAC2, ITGAL, ITGB2, VAV1
hsa04933	AGE-RAGE signaling pathway in diabetic complications	3,9906884	7,0e-11	EGR1, EDN1, F3, JUN, CCND1, COL1A2, COL3A1, COL4A1, COL4A2, COL4A5, VEGFC, SERPINE1, PRKCZ, FN1	TNF, STAT5A, PLCG2
hsa04062	Chemokine signaling pathway	1,8302260	1,3e-10	CXCL1, CXCL2, CCL28, PTK2, GNAI1, GNG12, PRKCZ	CXCR4, CCL3, PLCG2, GNG7, VAV1, DOCK2, RAC2, PIK3CG
hsa04668	TNF signaling pathway	1,7106271	1,7e-09	FOS, JUN, CXCL1, CXCL2, EDN1, VEGFC	TNF, TNFAIP3

Table 13. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05205	Proteoglycans in cancer	2,9245486	2,7e-09	RRAS, CTTN, CCND1, PLAUI, PTK2, MET, TIMP3, THBS1, LUM, SDC1, COL1A2, ITGA2, ITGB1, ITGB5, FZD6, SDC2, SDC4, FN1, CTSL, CAV2	ITPR1, HCLS1, VAV1, TNF, PLCG2, PTPN6
hsa04061	Viral protein interaction with cytokine and cytokine receptor	1,9633333	3,7e-09	CCL28, CXCL1, CXCL2, LTBR	CCL3, CXCR4, IL2RB, TNF

Table 13. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05200	Pathways in cancer	2,0984615	4,5e-09	CDH1, CTNNA1, ITGB1, COL4A1, COL4A2, COL4A5, FN1, LAMA3, LAMB1, LAMB2, LAMC1, LAMC2, ITGA6, ITGA2, PTK2, FZD6, MET, CCND1, JUN, FOS, MECOM, CTBP2, PPARG, JUP, MITF, BMP4, VEGFC, GNG12, GNA11, GNAI1, HES1, NFE2L2, NQO1, GSTA4, MGST1, IL13RA1, EDN1	STAT5A, PLCG2, RAC2, TCF7L2, CCND2, E2F2, GNG7, CXCR4, PIM2, CSF2RB, IL2RB
hsa04520	Adherens junction	3,7235632	6,9e-09	CDH1, CTNNA1, YES1, SNAI2, MET, PTPRF, PTPRM, TJP1, ACTN1, NECTIN3, NECTIN2, ARHGAP29	TCF7L2, PTPN6, RAC2

Table 13. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05323	Rheumatoid arthritis	3,2395000	8,0e-09	CTSL, CXCL1, CXCL2, FOS, JUN	TNF, HLA-DMA, HLA-DPB1, HLA-DQB1, ITGAL, ITGB2, CCL3
hsa05167	Kaposi sarcoma-associated herpesvirus infection	1,9859004	9,1e-09	C3, GNG12, CXCL1, CXCL2, PPP3CA, FOS, JUN, CCND1	E2F2, GNG7, SYK, PPP3CC, ITPR1, PLCG2, PIK3CG, TCF7L2
hsa04310	Wnt signaling pathway	1,7657652	1,3e-08	PPP3CA, FZD6, CCND1, FOSL1, JUN, TLE1, CTBP2, DKK1, BAMBI	PPP3CC, RAC2, CCND2, TCF7L2
hsa04662	B-cell receptor signaling pathway	3,1258333	1,6e-08	FOS, JUN, PPP3CA	VAV1, PPP3CC, RAC2, PLCG2, BTK, SYK, PTPN6, PIK3AP1
hsa04380	Osteoclast differentiation	2,9608333	3,5e-08	TNFRSF11B, FOS, JUN, FOSL1, PPP3CA, MIF, FHL2, PPARG	BTK, SYK, PLCG2, PPP3CC, TNF, CYBA, NCF2, NCF4, ITPR1

Table 13. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04810	Regulation of actin cytoskeleton	2,3667580	4,0e-08	NCKAP1, CYFIP1, CFL2, GSN, ACTN1, MYLK, ENAH, GNG12, PTK2, ITGA6, ITGA2, ITGB1, ITGB5, FN1, RRAS, PDGFC	CYFIP2, RAC2, IQGAP2, VAV1, ITGA4, ITGAL, ITGB2, CXCR4
hsa05417	Lipid and atherosclerosis	1,9222164	4,8e-08	FOS, JUN, NFE2L2, CXCL1, CXCL2, PPP3CA, HSPA2, PTK2, PPARG	CYBA, NCF2, NCF4, CCL3, TNF, PPP3CC, VAV1, ITPR1
hsa04658	Th1 and Th2 cell differentiation	2,7948627	5,9e-08	PPP3CA, JUN, FOS	PPP3CC, CD3D, IL2RB, STAT5A, HLA-DMA, HLA-DPB1, HLA-DQB1, RUNX3
hsa05140	Leishmaniasis	4,5811111	6,5e-08	ITGB1, C3, FOS, JUN	ITGA4, ITGB2, HLA-DMA, HLA-DPB1, HLA-DQB1, TNF, PTPN6, CYBA, NCF2, NCF4
hsa05145	Toxoplasmosis	2,9245486	7,7e-08	GNAI1, ITGA6, ITGB1, LAMA3, LAMB1, LAMB2, LAMC1, LAMC2	TNF, HLA-DMA, HLA-DPB1, HLA-DQB1, PIK3CG

Table 13. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04210	Apoptosis	1,9796944	8,7e-08	CAPN2, LMNA, JUN, FOS, CTSB, CTSL, CTSV	CSF2RB, TNF, CTSS, ITPR1
hsa04670	Leukocyte transendothelial migration	3,4636164	9,1e-08	CLDN4, CLDN1, ITGB1, PTK2, CTNNA1, GNAI1, ACTN1	ITGB2, ITGAL, ITGA4, CYBA, NCF2, NCF4, VAV1, RAC2, CXCR4, PLCG2
hsa04360	Axon guidance	2,5020528	9,1e-08	GNAI1, PPP3CA, PTK2, ITGB1, EPHA2, EFNA1, MET, SEMA3A, SEMA3C, RHOD, CFL2, RRAS, PRKCZ, ENAH	CXCR4, RAC2, PPP3CC, FES, PLCG2
hsa05226	Gastric cancer	1,2429736	9,8e-08	CCND1, FZD6, MET, CDH1, CTNNA1, JUP	TCF7L2, E2F2
hsa04659	Th17 cell differentiation	2,3521122	1,0e-07	PPP3CA, JUN, FOS, AHR	PPP3CC, IL2RB, STAT5A, HLA-DMA, HLA-DPB1, HLA-DQB1, CD3D
hsa04611	Platelet activation	2,5623164	1,1e-07	ITGA2, ITGB1, MYLK, GNAI1, COL1A2, COL3A1, PRKCZ	ITPR1, PLCG2, APBB1IP, PIK3CG, FERMT3, BTK, SYK

Table 13. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05130	Pathogenic Escherichia coli infection	2,4967245	1,3e-07	CTTN, CLDN4, CLDN1, ITGB1, TUBB3, TUBB2A, FOS, JUN, NCKAP1, CYFIP1, TJP1, MYO1B, MYO1C, MYO6	HCLS1, WIPF1, TNF, PTPN6, CYFIP2, MYO1G
hsa05412	Arrhythmogenic right ventricular cardiomyopathy	3,2542922	1,4e-07	LMNA, DSG2, JUP, ITGA6, ITGA2, ITGB1, ITGB5, CTNNA1, SGCE	ITGA4, TCF7L2
hsa05166	Human T-cell leukemia virus 1 infection	2,1483595	1,5e-07	PPP3CA, LTBR, JUN, CCND1, FOS, EGR1, FOSL1	ITGAL, ITGB2, PPP3CC, IL2RB, STAT5A, CCND2, CD3D, E2F2, HLA-DMA, HLA-DPB1, HLA-DQB1, TNF
hsa00480	Glutathione metabolism	1,5706667	2,7e-07	GPX2, GPX8, GSTA4, MGST1	
hsa04974	Protein digestion and absorption	3,7022857	2,7e-07	ATP1B1, COL1A2, COL3A1, COL4A1, COL4A2, COL4A5	
hsa00561	Glycerolipid metabolism	1,8944444	3,2e-07	PLPP2, PLPP3, MBOAT2, MGLL	LPL

Table 13. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05170	Human immunodeficiency virus 1 infection	2,0062477	3,3e-07	GNG12, PPP3CA, GNAI1, GNAI1, PTK2, FOS, JUN, CFL2, AP1M2	GNG7, PPP3CC, TNF, CXCR4, ITPR1, PLCG2, RAC2, CD3D
hsa04915	Estrogen signaling pathway	1,2391530	3,8e-07	FOS, JUN, GNAI1, KRT15, KRT18, KRT19	ITPR1
hsa04530	Tight junction	1,9015933	4,0e-07	TJP1, PRKCZ, MAGI1, ACTN1, MICALL2, RAB13, CLDN4, CLDN1, PRKAG2, JUN, CCND1, CTTN, ITGB1	HCLS1
hsa05418	Fluid shear stress and atherosclerosis	3,2476190	4,1e-07	FOS, JUN, NFE2L2, NQO1, BMP4, PRKCZ, GSTA4, MGST1, PLAT, CTSL, EDN1, PTK2, CAV2, SDC1, SDC2, SDC4	CYBA, NCF2, RAC2, TNF

Table 13. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04371	Apelin signaling pathway	2,3996296	4,6e-07	PRKAG2, SERPINE1, GNAI1, GNG12, RRAS, CCN2, MYLK, CCND1, EGR1, PLAT, CDH1	GNG7, PIK3CG, ITPR1
hsa04657	IL-17 signaling pathway	1,7785490	4,6e-07	FOS, JUN, FOSL1, CXCL1, CXCL2	TNFAIP3, TNF
hsa05100	Bacterial invasion of epithelial cells	2,6995833	7,5e-07	CTNNA1, CDH1, CTTN, ITGB1, PTK2, MET, FN1, CAV2	HCLS1
hsa05163	Human cytomegalovirus infection	1,5042454	8,7e-07	GNG12, PPP3CA, GNAI1, GNAI1, CCND1, PTK2	TNF, GNG7, PPP3CC, ITPR1, CCL3, RAC2, CXCR4, E2F2
hsa04612	Antigen processing and presentation	3,6604520	9,2e-07	CTSB, CTSL, LGMN, HSPA2	CD74, HLA-DMA, HLA-DPB1, HLA-DQB1, CTSS, TNF
hsa04630	JAK-STAT signaling pathway	0,9448542	9,8e-07	IL13RA1, CCND1	STAT5A, IL2RB, CCND2, PTPN6, CSF2RB
hsa04664	Fc epsilon RI signaling pathway	2,0900000	1,1e-06		TNF, BTK, SYK, PLCG2, RAC2, VAV1

Table 13. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04650	Natural killer cell mediated cytotoxicity	2,5523333	1,1e-06	PPP3CA	PPP3CC, PLCG2, TNF, LAT2, SYK, PTPN6, CD48, ICAM2, RAC2, VAV1, ITGB2, ITGAL
hsa05213	Endometrial cancer	1,5997531	1,3e-06	CTNNA1, CDH1, CCND1	TCF7L2
hsa05162	Measles	1,3821867	1,6e-06	CCND1, FOS, JUN	IL2RB, CD3D, STAT5A, TNFAIP3, CCND2
hsa04926	Relaxin signaling pathway	2,3792938	1,6e-06	GNG12, PRKCZ, EDN1, GNAI1, VEGFC, COL1A2, COL3A1, COL4A1, COL4A2, COL4A5, FOS, JUN	GNG7
hsa05133	Pertussis	3,4429469	2,3e-06	ITGB1, GNAI1, CFL2, C1R, C1S, C3, FOS, JUN	ITGB2, TNF, IRF8
hsa04917	Prolactin signaling	1,3088889	2,7e-06	CCND1, FOS	STAT5A, CCND2
hsa04014	Ras signaling pathway	1,1780000	4,9e-06	RRAS, EFNA1, PDGFC, BDNF, VEGFC, EPHA2, MET, GNG12	RAC2, RASAL3, GNG7, PLCG2

Table 13. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04060	Cytokine-cytokine receptor interaction	1,2798025	5,1e-06	INHBB, LTBR, IL13RA1, TNFRSF12A, TNFRSF11B, CXCL1, CXCL2, CCL28, GDF15, BMP4	IL2RB, CSF2RB, CD70, CXCR4, TNF, CCL3
hsa04921	Oxytocin signaling pathway	1,6612821	5,2e-06	MYLK, PRKAG2, PPP3CA, JUN, FOS, CCND1, GNAI1	ITPR1, PPP3CC, PIK3CG
hsa04010	MAPK signaling pathway	1,5368414	6,0e-06	RRAS, GNG12, EPHA2, MET, MECOM, FOS, JUN, PPP3CA, EFNA1, AREG, PDGFC, BDNF, VEGFC	DUSP2, PTPN7, TNF, RAC2, PPP3CC, MAP4K1
hsa04713	Circadian entrainment	1,3168699	6,2e-06	GNG12, GNAI1, FOS	GNG7, ITPR1
hsa04725	Cholinergic synapse	2,0326275	6,2e-06	GNA11, GNAI1, GNG12, FOS	ITPR1, CHAT, GNG7, PIK3CG
hsa05410	Hypertrophic cardiomyopathy	2,6995833	8,8e-06	ITGA6, ITGA2, ITGB1, ITGB5, LMNA, TPM1, PRKAG2, EDN1, SGCE	ITGA4, TNF

Table 13. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05414	Dilated cardiomyopathy	2,0900000	1,0e-05	ITGA6, ITGA2, ITGB1, ITGB5, LMNA, TPM1, SGCE	ITGA4, TNF
hsa04613	Neutrophil extracellular trap formation	1,8165421	1,1e-05	C3	SYK, PLCG2, ITGB2, CYBA, NCF2, RAC2, NCF4, ITGAL
hsa05216	Thyroid cancer	2,5407843	1,2e-05	PPARG, CDH1, CCND1	TCF7L2
hsa05221	Acute myeloid leukemia	1,7997222	1,2e-05	JUP, CCND1	STAT5A, TCF7L2, PIM2
hsa04020	Calcium signaling pathway	1,1726697	1,3e-05	PPP3CA, MET, GNA11, MYLK, ADORA2B, ASPH, PDGFC, VEGFC	PPP3CC, PLCG2, ITPR1, CXCR4
hsa04012	ErbB signaling pathway	1,4592342	1,3e-05	PTK2, JUN, AREG	STAT5A, PLCG2
hsa04625	C-type lectin receptor signaling	1,8186667	1,3e-05	RRAS, PPP3CA, JUN	SYK, PLCG2, PPP3CC, ITPR1, TNF
hsa04922	Glucagon signaling pathway	1,0483819	1,3e-05	PYGL, PPP3CA, PRKAG2	PPP3CC, ITPR1
hsa05224	Breast cancer	1,0798333	1,3e-05	FOS, JUN, CCND1, FZD6, HES1	TCF7L2, E2F2
hsa05210	Colorectal cancer	1,6828571	1,4e-05	CCND1, JUN, FOS, AREG	RAC2, TCF7L2
hsa04066	HIF-1 signaling pathway	1,0798333	1,5e-05	HK1, SERPINE1, EDN1, LTBR	PLCG2

Table 13. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04064	NF-κB signaling pathway	2,2341379	1,6e-05	LTBR, PLAU, CXCL1, CXCL2	TNF, TNFAIP3, PLCG2, SYK, BTK
hsa04660	T-cell receptor signaling pathway	1,6756034	1,6e-05	FOS, JUN, PPP3CA	TNF, PPP3CC, VAV1, CD3D, PTPRC, PTPN6
hsa00564	Glycerophospholipid metabolism	1,2270833	1,8e-05	PLPP2, PLPP3, MBOAT2, CDS1	CHAT
hsa04971	Gastric acid secretion	1,4894253	2,0e-05	ATP1B1, GNAI1, MYLK	ITPR1
hsa04728	Dopaminergic synapse	1,3032471	2,0e-05	GNAI1, GNG12, FOS, PPP3CA	ITPR1, GNG7, PPP3CC
hsa04514	Cell adhesion molecules	3,4010499	2,0e-05	ITGB1, ITGA6, NECTIN2, CDH1, NECTIN3, PTPRM, SDC1, SDC2, SDC4, PTPRF, CLDN4, CLDN1	PTPRC, ITGB2, ITGAL, ITGA4, ICAM2, HLA-DMA, HLA-DPB1, HLA-DQB1
hsa05142	Chagas disease	2,0038144	2,2e-05	C3, FOS, JUN, SERPINE1, GNAI1, GNA11	TNF, CCL3, CD3D
hsa05135	Yersinia infection	1,7140212	2,3e-05	ITGB1, FN1, PTK2, FOS, JUN	WIPF1, ITGA4, VAV1, TNF, RAC2

Table 13. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05169	Epstein-Barr virus infection	1,9304283	4,0e-05	JUN, HES1, CCND1	HLA-DMA, HLA-DPB1, HLA-DQB1, ITGAL, TNFAIP3, SYK, PLCG2, RUNX3, E2F2, CCND2, BTK, TNF, CD3D
hsa04924	Renin secretion	2,1962712	4,4e-05	GNAI1, CTSB, PPP3CA, EDN1	ITPR1, PPP3CC
hsa04022	cGMP-PKG signaling pathway	1,1595526	4,4e-05	GNAI1, PPP3CA, ATP1B1, MYLK, GNAI1	ITPR1, PPP3CC, PIK3CG
hsa04720	Long-term potentiation	1,0621311	4,4e-05	PPP3CA	PPP3CC, ITPR1
hsa04724	Glutamatergic synapse	1,4397778	4,9e-05	GNAI1, GNG12, PPP3CA	ITPR1, GNG7, PPP3CC
hsa04726	Serotonergic synapse	1,3331276	6,8e-05	GNAI1, GNG12, APP	ITPR1, GNG7
hsa05208	Chemical carcinogenesis - ROS	1,3640000	7,3e-05	MET, NQO1, PRKD1, PTK2, FOS, JUN, NFE2L2, AHR, GSTA4, MGST1	CYBA, NCF2
hsa05415	Diabetic cardiomyopathy	0,8688314	8,1e-05	COL1A2, COL3A1, PRKCZ	CYBA, NCF2, RAC2, NCF4
hsa04072	Phospholipase D signaling pathway	1,0876019	8,8e-05	RRAS, PLPP2, PLPP3, PDGFC	PLCG2, PIK3CG, SYK

Table 13. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05152	Tuberculosis	1,4485569	9,2e-05	PPP3CA, C3	PPP3CC, ITGB2, TNF, CD74, HLA-DMA, HLA-DPB1, HLA-DQB1, CTSS, SYK
hsa05207	Chemical carcinogenesis - receptor activation	1,1931860	9,9e-05	FOS, JUN, AHR, GSTA4, MGST1, CCND1, KLF5, KLF4, GNAI1	STAT5A
hsa05416	Viral myocarditis	3,3877124	1,0e-04	CCND1, SGCE	HLA-DMA, HLA-DPB1, HLA-DQB1, ITGAL, ITGB2, RAC2
hsa04142	Lysosome	1,7050000	1,1e-04	CTSB, CTSL, CTSV, LGMN, AP1M2	CTSS
hsa04370	VEGF signaling pathway	2,0766026	1,1e-04	PPP3CA, PTK2	PLCG2, PPP3CC, RAC2
hsa04620	Toll-like receptor signaling	0,8905842	1,1e-04	FOS, JUN	TNF, CCL3
hsa04621	NOD-like receptor signaling pathway	1,6612821	1,4e-04	JUN, CTSB, CXCL1, CXCL2	CARD8, AIM2, ITPR1, CYBA, IFI16, TNFAIP3, TNF
hsa05225	Hepatocellular carcinoma	1,3844017	1,5e-04	CCND1, FZD6, MET, NFE2L2, NQO1, GSTA4, MGST1	TCF7L2, E2F2, PLCG2
hsa05120	Epithelial cell signaling in Helicobacter pylori infection	2,2563682	1,5e-04	JUN, CXCL1, CXCL2, MET, TJP1	PLCG2, CSK

Table 13. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05321	Inflammatory bowel disease	1,8617816	1,6e-04	JUN	TNF, HLA-DMA, HLA-DPB1, HLA-DQB1
hsa04666	Fc gamma R-mediated phagocytosis	2,6692509	1,6e-04	CFL2, ASAP2, MARCKS, GSN, PLPP2, PLPP3	SYK, PTPRC, PLCG2, VAV1, RAC2
hsa05020	Prion disease	1,2206812	2,0e-04	LAMC1, EGR1, TUBB3, TUBB2A, PPP3CA, CAV2	ITPR1, PPP3CC, TNF, CYBA, NCF2, RAC2, NCF4
hsa01522	Endocrine resistance	1,2132959	2,0e-04	CCND1, JUN, FOS, PTK2	E2F2
hsa04723	Retrograde endocannabinoid signaling	1,0187107	2,0e-04	GNAI1, GNG12, MGLL	ITPR1, GNG7
hsa04935	Growth hormone synthesis, secretion and action	1,4128660	2,0e-04	GNAI1, FOS, PTK2, GNAI1	ITPR1, PLCG2, STAT5A
hsa04024	cAMP signaling pathway	1,0534959	2,8e-04	ATP1B1, GNAI1, SOX9, RRAS, BDNF, FOS, JUN, EDN1	VAV1, RAC2
hsa04218	Cellular senescence	1,3313014	2,9e-04	RRAS, CCND1, CAPN2, PPP3CA, SERPINE1	CCND2, E2F2, PPP3CC, ITPR1
hsa05134	Legionellosis	2,2975177	2,9e-04	C3, CXCL1, CXCL2	ITGB2, TNF
hsa04115	p53 signaling pathway	1,5426190	3,0e-04	SFN, THBS1, SERPINE1, CCND1	CCND2

Table 13. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa00983	Drug metabolism - other enzymes	0,9255714	3,0e-04	UPP1, GSTA4, MGST1	
hsa04814	Motor proteins	1,6938562	3,1e-04	DYNLL1, TUBB3, TUBB2A, MYO1B, MYO1C, MYO6, TPM1	MYO1G
hsa01521	EGFR TKI resistance	1,5426190	3,2e-04	MET, PDGFC, GAS6, AXL	PLCG2
hsa05310	Asthma	3,9266667	3,3e-04		HLA-DMA, HLA-DPB1, HLA-DQB1, TNF
hsa04350	TGF- β signaling pathway	1,9244554	3,3e-04	THBS1, ID1, ID4, INHBB, BMP4, BAMBI, TGIF1	TNF, NRROS
hsa05171	Coronavirus disease - COVID-19	1,2703922	3,6e-04	FOS, JUN, C1R, C1S, C3	TNF, SYK, PLCG2
hsa05144	Malaria	3,6872358	3,9e-04	SDC1, SDC2, MET, THBS1	TNF, ITGAL, ITGB2
hsa04727	GABAergic synapse	1,0123438	4,4e-04	GNAI1, GNG12	GNG7
hsa05161	Hepatitis B	0,7151214	4,7e-04	FOS, JUN	TNF, STAT5A, E2F2
hsa04916	Melanogenesis	1,1487589	4,9e-04	MITE, GNAI1, FZD6, EDN1	TCF7L2
hsa04392	Hippo signaling pathway - multiple species	6,1704762	5,3e-04	WWC1, FRMD6, YAP1, WWTR1, TEAD4, RASSF4, AJUBA	RASSF2

Table 13. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04114	Oocyte meiosis	0,6645128	5,3e-04	PPP3CA	FBXO43, ITPR1, PPP3CC
hsa05330	Allograft rejection	3,1995062	5,4e-04		HLA-DMA, HLA-DPB1, HLA-DQB1, TNF
hsa04340	Hedgehog signaling	0,7853333	5,4e-04	CCND1	CCND2
hsa05332	Graft-versus-host disease	3,0852381	5,6e-04		HLA-DMA, HLA-DPB1, HLA-DQB1, TNF
hsa04940	Type I diabetes mellitus	2,9788506	5,9e-04		HLA-DMA, HLA-DPB1, HLA-DQB1, TNF
hsa05031	Amphetamine addiction	1,3933333	6,5e-04	PPP3CA, FOS, JUN	PPP3CC
hsa05032	Morphine addiction	1,1833790	7,0e-04	GNAI1, GNG12	GNG7, PDE7A
hsa05203	Viral carcinogenesis	1,4291912	7,2e-04	JUN, CCND1, GSN, ACTN1, C3, LTBR	STAT5A, SYK, CCND2
hsa04936	Alcoholic liver disease	1,1115931	7,7e-04	PRKAG2, CCND1, C3, CXCL1, CXCL2	TNF, TCF7L2
hsa04928	Parathyroid hormone synthesis, secretion and	1,2340952	9,5e-04	GNAI1, GNAI1, FOS, EGR1, SGK1	ITPR1
hsa05204	Chemical carcinogenesis - DNA adducts	0,8638667	1,0e-03	GSTA4, MGST1	
hsa00982	Drug metabolism - cytochrome	0,7447126	1,0e-03	GSTA4, MGST1	
hsa01524	Platinum drug resistance	0,7447126	1,0e-03	GSTA4, MGST1	
hsa04071	Sphingolipid signaling	0,7998765	1,1e-03	GNAI1, PRKCZ	TNF, RAC2

Table 13. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04672	Intestinal immune network for IgA production	4,0858559	1,1e-03	CCL28, LTBR	HLA-DMA, HLA-DPB1, HLA-DQB1, ITGA4, CXCR4
hsa00980	Metabolism of xenobiotics by cytochrome P450	0,6966667	1,1e-03	GSTA4, MGST1	
hsa05235	PD-L1 expression and PD-1 checkpoint pathway in	2,1069919	1,3e-03	FOS, JUN, PPP3CA	PPP3CC, CD3D, PTPN6, BATF, BATF3
hsa05132	Salmonella infection	1,7590045	1,4e-03	FOS, JUN, NCKAP1, CYFIP1, MYO6, SNX9, AHNAK2, AHNAK, RRAS, ANXA2, DYNLL1, TUBB3, TUBB2A	TNF, TCF7L2, CYFIP2, PIK3CG, PTPRC
hsa05131	Shigellosis	1,5879902	1,5e-03	JUN, C3, WIP1, CTTN, ITGB1, PTK2, ACTN1, CAPN2, HK1	TNF, TIFA, FNBP1, HCLS1, ITPR1, PLCG2
hsa05320	Autoimmune thyroid disease	1,5426190	1,5e-03		HLA-DMA, HLA-DPB1, HLA-DQB1
hsa04934	Cushing syndrome	1,4505224	1,7e-03	GNAI1, FZD6, GNAI1, CCND1, AHR, MEN1	ITPR1, TCF7L2, E2F2

Table 13. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04145	Phagosome	3,0595278	1,7e-03	C1R, ITGB1, ITGB5, ITGA2, TUBB3, TUBB2A, C3, THBS1, CTSL	ITGB2, CYBA, NCF4, NCF2, HLA-DMA, HLA-DPB1, HLA-DQB1, CTSS
hsa05223	NSCLC	1,6872396	2,2e-03	CCND1, MET	E2F2, PLCG2, STAT5A
hsa05034	Alcoholism	0,9998457	2,3e-03	BDNF, GNAI1, GNG12, ADORA2B	GNG7
hsa00565	Ether lipid metabolism	0,9389855	2,4e-03	PLPP2, PLPP3	
hsa04960	Aldosterone-regulated sodium reabsorption	2,1596667	2,4e-03	ATP1B1, SGK1, SFN	
hsa04148	Efferocytosis	1,7467892	2,4e-03	GULP1, ITGB5, AXL, PTK2, GAS6, THBS1, PPARG, ADAM9, SGK1	DUSP2, PTPN6
hsa05322	Systemic lupus erythematosus	3,4554667	2,5e-03	C3, C1S, C1R, ACTN1	TNF, HLA-DMA, HLA-DPB1, HLA-DQB1
hsa05220	Chronic myeloid leukemia	1,4997685	2,5e-03	MECOM, CTBP2, CCND1	STAT5A, E2F2
hsa05164	Influenza A	0,6170476	2,8e-03		HLA-DMA, HLA-DPB1, HLA-DQB1, TNF
hsa00860	Porphyrin metabolism	1,7997222	3,1e-03	BLVRA, BLVRB, FTH1	
hsa04110	Cell cycle	0,9229345	3,1e-03	SFN, CCND1, STAG1	CCND2, E2F2

Table 13. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04750	Inflammatory mediator regulation of TRP channels	0,4964751	3,1e-03		PLCG2, ITPR1
hsa00830	Retinol metabolism	0,8306410	3,1e-03	ALDH1A1, ALDH1A3	
hsa05030	Cocaine addiction	1,3785106	3,2e-03	GNAI1, JUN, BDNF	
hsa04070	Phosphatidylinositol signaling	0,7042391	3,4e-03	CDS1	PLCG2, ITPR1
hsa05150	Staphylococcus aureus infection	3,2103153	3,4e-03	C3, C1S, C1R, KRT15, KRT18, KRT19	HLA-DMA, HLA-DPB1, HLA-DQB1, ITGB2,
hsa05217	Basal cell carcinoma	1,0450000	3,5e-03	BMP4, FZD6	TCF7L2
hsa04930	Type II diabetes mellitus	1,7050000	4,3e-03	PRKCZ, HK1	TNF
hsa05231	Choline metabolism in cancer	1,4559551	4,3e-03	PDGFC, PLPP2, PLPP3, FOS, JUN	RAC2
hsa05218	Melanoma	1,8511429	4,4e-03	CDH1, CCND1, MITF, MET, PDGFC	E2F2
hsa04640	Hematopoietic cell lineage	2,0568254	5,1e-03	ITGA6, ITGA2	ITGA4, CD3D, HLA-DMA, HLA-DPB1, HLA-DQB1, TNF
hsa04932	Non-alcoholic fatty liver disease	0,7824879	5,2e-03	PRKAG2, FOS, JUN, PPARG	TNF
hsa05202	Transcriptional misregulation in cancer	2,2411635	5,3e-03	PPARG, MET, PLAUI, JUP, PTK2, MEN1, PLAT, MITF	IL2RB, CCND2, LYL1
hsa00562	Inositol phosphate	0,6170476	5,4e-03		PLCG2, PIK3CG

Table 13. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04270	Vascular smooth muscle contraction	1,2811582	6,2e-03	MYLK, GNA11, CALD1, ADORA2B, EDN1, ADM	ITPR1
hsa05215	Prostate cancer	1,4894253	6,3e-03	CCND1, PDGFC, PLAUI, PLAT	TCF7L2, E2F2
hsa04918	Thyroid hormone synthesis	1,3712169	6,5e-03	ATP1B1, GPX2, GPX8	ITPR1
hsa05010	Alzheimer disease	0,8801149	7,5e-03	BACE2, APP, CAPN2, PPP3CA, WIP1, TUBB3, TUBB2A, FZD6, DKK1	TNF, ITPR1, PPP3CC, LPL
hsa04970	Salivary secretion	0,6351961	7,7e-03	ATP1B1	ITPR1
hsa04610	Complement and coagulation cascades	2,7767143	8,3e-03	F3, SERPINE1, SERPINE2, PLAUI, PLAT, C3, C1R, C1S	ITGB2
hsa04540	Gap junction	1,9891667	1,0e-02	GNAI1, TJP1, TUBB3, TUBB2A, PDGFC, GNA11	ITPR1
hsa05212	Pancreatic cancer	0,9255714	1,1e-02	CCND1	RAC2, E2F2
hsa04925	Aldosterone synthesis and secretion	1,0798333	1,1e-02	GNA11, PRKD1, ATP1B1	ITPR1
hsa04550	Signaling pathways regulating pluripotency of stem cells	1,1467257	1,2e-02	FZD6, KLF4, INHBB, BMP4, ID1, ID4	

Table 13. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04912	GnRH signaling pathway	1,0284127	1,2e-02	GNA11, JUN, EGR1	ITPR1
hsa05219	Bladder cancer	2,3347748	1,3e-02	THBS1, CDH1, CCND1	E2F2
hsa01523	Antifolate resistance	1,0284127	1,5e-02		TNF
hsa04140	Autophagy - animal	0,8581457	1,6e-02	RRAS, WIPI1, CTSL, CTSB	ITPR1, RAB39B
hsa05022	Pathways of neurodegeneration - multiple diseases	0,7920538	1,7e-02	TUBB3, TUBB2A, PPP3CA, WIPI1, BDNF, GPX2, GPX8, CAPN2, FZD6, DKK1, APP	ITPR1, PPP3CC, RAB39B, TNF
hsa05110	Vibrio cholerae infection	1,1366667	2,4e-02	TJP1	PLCG2
hsa04929	GnRH secretion	1,0798333	2,6e-02	GNA11	ITPR1
hsa05160	Hepatitis C	0,8875342	2,6e-02	CLDN4, CLDN1, IFIT1, CCND1	TNF, E2F2
hsa04931	Insulin resistance	1,1018707	3,1e-02	PTPRF, PRKCZ, PYGL, PRKAG2	TNF
hsa04910	Insulin signaling pathway	0,8306410	3,1e-02	PTPRF, PRKAG2, PYGL, HK1, PRKCZ	
hsa05014	Amyotrophic lateral sclerosis	0,7042391	3,1e-02	GPX2, GPX8, PPP3CA, TUBB3, TUBB2A, WIPI1	TNF, PPP3CC, RAB39B

Table 13. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04919	Thyroid hormone signaling pathway	0,7998765	3,1e-02	ATP1B1, BMP4, CCND1	PLCG2
hsa05214	Glioma	0,9255714	3,4e-02	CCND1	E2F2, PLCG2
hsa05206	MicroRNAs in cancer	0,9891603	3,4e-02	TPM1, MET, CCND1, TIMP3, MARCKS, TNC, PLAUI, EFNA1, THBS1	CCND2, E2F2, PLCG2
hsa04962	Vasopressin-regulated water reabsorption	1,0534959	3,4e-02	DYNLL1	ARHGDIB
hsa04261	Adrenergic signaling in cardiomyocytes	0,6399012	3,4e-02	TPM1, ATP1B1, GNAI1	PIK3CG
hsa04975	Fat digestion and absorption	1,3497917	3,5e-02	PLPP2, PLPP3	
hsa04730	Long-term depression	1,2703922	3,6e-02	GNA11, GNAI1	ITPR1
hsa04137	Mitophagy - animal	0,7042391	3,7e-02	RRAS, JUN, MITF	
hsa04927	Cortisol synthesis and	0,8306410	3,9e-02	GNA11	ITPR1
hsa04152	AMPK signaling pathway	0,5733628	4,7e-02	PRKAG2, PPARG, CCND1	
hsa04961	Endocrine and other factor-regulated calcium	0,4319333	4,8e-02	ATP1B1	
hsa04973	Carbohydrate digestion and absorption	1,0044961	4,8e-02	ATP1B1, HK1	

Table 13. pathfindR pathway enrichment analysis for 5-FU utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04923	Regulation of lipolysis in adipocytes	0,8149686	4,9e-02	MGLL, GNAI1	
hsa04213	Longevity regulating pathway -	0,4074843	4,9e-02	PRKAG2	

Table 14. pathfindR pathway enrichment analysis for gemcitabine utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 1)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa00480	Glutathione metabolism	8,414286	0,00020	MGST2	GPX8
hsa04350	TGF- β signaling pathway	4,582037	0,00091	INHBB, BMP4	
hsa04550	Signaling pathways regulating pluripotency of stem cells	6,143173	0,00091	FZD6, INHBB, BMP4	
hsa05226	Gastric cancer	3,329394	0,00237	FZD6, CDH1	
hsa04390	Hippo signaling pathway	4,478571	0,00237	CDH1, FZD6, BMP4	
hsa04933	AGE-RAGE signaling pathway in diabetic	5,030279	0,00387		COL1A2, COL3A1
hsa04926	Relaxin signaling pathway	3,921913	0,00410		COL1A2, COL3A1
hsa04060	Cytokine-cytokine receptor interaction	2,571032	0,00843	INHBB, BMP4	CD70
hsa05415	Diabetic cardiomyopathy	2,659688	0,00879		COL1A2, COL3A1
hsa05216	Thyroid cancer	6,805672	0,00941	CDH1	
hsa05146	Amoebiasis	4,923252	0,00999		COL1A2, COL3A1
hsa04820	Cytoskeleton in muscle cells	5,933150	0,00999		VIM, COL1A2, COL3A1
hsa04974	Protein digestion and absorption	13,222449	0,01491		COL1A2, COL3A1
hsa05213	Endometrial cancer	4,285053	0,01796	CDH1	
hsa05217	Basal cell carcinoma	7,464286	0,01897	BMP4, FZD6	
hsa05100	Bacterial invasion of epithelial cells	3,213790	0,02134	CDH1	
hsa05200	Pathways in cancer	2,342033	0,02223	CDH1, FZD6, CDKN2A, BMP4,	
hsa04520	Adherens junction	2,659688	0,02339	CDH1	

Table 14. pathfindR pathway enrichment analysis for gemcitabine utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05204	Chemical carcinogenesis - DNA adducts	4,627857	0,02371	MGST2	
hsa00982	Drug metabolism - cytochrome P450	3,989532	0,02371	MGST2	
hsa00980	Metabolism of xenobiotics by cytochrome P450	3,732143	0,02371	MGST2	
hsa04916	Melanogenesis	2,461626	0,02428	FZD6	
hsa04934	Cushing syndrome	3,453625	0,04032	FZD6, CDKN2A	
hsa05224	Breast cancer	1,652806	0,04034	FZD6	
hsa05225	Hepatocellular carcinoma	4,449863	0,04454	FZD6, MGST2, CDKN2A	
hsa04310	Wnt signaling pathway	1,455301	0,04454	FZD6	

Table 15. pathfindR pathway enrichment analysis for gemcitabine utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 2)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa00480	Glutathione metabolism	7,600000	0,00020	MGST2	GPX8
hsa04350	TGF- β signaling pathway	4,138614	0,00091	INHBB, BMP4	
hsa04550	Signaling pathways regulating pluripotency of stem cells	5,548673	0,00091	FZD6, INHBB, BMP4	
hsa05226	Gastric cancer	3,007194	0,00237	FZD6, CDH1	
hsa04390	Hippo signaling pathway	4,045161	0,00237	CDH1, FZD6, BMP4	
hsa04933	AGE-RAGE signaling pathway in diabetic complications	4,543478	0,00387		COL1A2, COL3A1
hsa04926	Relaxin signaling pathway	3,542373	0,00410		COL1A2, COL3A1
hsa04060	Cytokine-cytokine receptor interaction	2,322222	0,00843	INHBB, BMP4	CD70
hsa05415	Diabetic cardiomyopathy	2,402299	0,00879		COL1A2, COL3A1
hsa05216	Thyroid cancer	6,147059	0,00941	CDH1	
hsa05146	Amoebiasis	4,446809	0,00999		COL1A2, COL3A1
hsa04820	Cytoskeleton in muscle cells	5,358974	0,00999		VIM, COL1A2, COL3A1
hsa04974	Protein digestion and absorption	11,942857	0,01491		COL1A2, COL3A1
hsa04510	Focal adhesion	2,211640	0,01770		COL1A2, CAV1
hsa05205	Proteoglycans in cancer	3,265625	0,01770	FZD6	COL1A2, CAV1
hsa05213	Endometrial cancer	3,870370	0,01796	CDH1	
hsa05217	Basal cell carcinoma	6,741936	0,01897	BMP4, FZD6	
hsa05100	Bacterial invasion of epithelial cells	5,805556	0,02134	CDH1	CAV1

Table 15. pathfindR pathway enrichment analysis for gemcitabine utilizing the gene features of the pan-drug model trained with DNA replication inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05200	Pathways in cancer	2,115385	0,02223	CDH1, FZD6, CDKN2A, BMP4, MGST2	
hsa04520	Adherens junction	2,402299	0,02339	CDH1	
hsa05204	Chemical carcinogenesis - DNA adducts	4,180000	0,02371	MGST2	
hsa00982	Drug metabolism - cytochrome P450	3,603448	0,02371	MGST2	
hsa00980	Metabolism of xenobiotics by cytochrome P450	3,370968	0,02371	MGST2	
hsa04916	Melanogenesis	2,223404	0,02428	FZD6	
hsa04934	Cushing syndrome	3,119403	0,04032	FZD6, CDKN2A	
hsa05224	Breast cancer	1,492857	0,04034	FZD6	
hsa05225	Hepatocellular carcinoma	4,019231	0,04454	FZD6, MGST2, CDKN2A	
hsa04310	Wnt signaling pathway	1,314465	0,04454	FZD6	

Table 16. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 1)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04512	ECM-receptor interaction	6,7346090	5,8e-30	COL1A2, COL6A3, ITGA4	LAMA3, LAMA5, LAMB1, LAMB2, LAMB3, LAMC1, LAMC2, COL4A1, COL4A2, THBS1, ITGB1, ITGA6, SDC1, SDC4, ITGB6, ITGB5, CD44, ITGA2, DSPP
hsa04510	Focal adhesion	3,7641249	5,8e-30	ITGA4, FYN, COL1A2, COL6A3	ITGB1, ITGB5, ITGB6, ITGA6, ITGA2, ACTN1, PTK2, CAPN2, COL4A1, COL4A2, LAMA3, LAMA5, LAMB1, LAMB2, LAMB3, LAMC1, LAMC2, THBS1, CCND1, EGFR, MET, VEGFC, JUN, EMP1
hsa04151	PI3K-Akt signaling pathway	2,3372200	2,4e-22	PIK3CG, ITGA4, COL1A2, COL6A3, MYB, PIK3AP1	EGFR, EPHA2, FGFR3, MET, CCND1, EFNA1, VEGFC, GNG12, ITGA6, ITGA2, ITGB1, ITGB5, ITGB6, COL4A1, COL4A2, LAMA3, LAMA5, LAMB1, LAMB2, LAMB3, LAMC1, LAMC2, THBS1, PTK2, SGK1

Table 16. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05165	Human papillomavirus infection	2,3023288	7,8e-20	ITGA4, COL1A2, COL6A3, HES6, TCF7L2	EGFR, CCND1, ITGA6, ITGA2, ITGB1, ITGB5, ITGB6, COL4A1, COL4A2, LAMA3, LAMA5, LAMB1, LAMB2, LAMB3, LAMC1, LAMC2, THBS1, PTK2, LLGL2, PATJ, JAG1, FZD6, WNT10A
hsa05222	Small cell lung cancer	4,7051561	6,2e-18	E2F2	CCND1, ITGB1, ITGA6, ITGA2, PTK2, COL4A1, COL4A2, LAMA3, LAMA5, LAMB1, LAMB2, LAMB3, LAMC1, LAMC2
hsa04820	Cytoskeleton in muscle cells	4,3432210	3,5e-14	ITGA4, VIM, COL1A2, COL6A3	ITGA6, ITGA2, ITGB1, ITGB5, ITGB6, DSG2, JUP, ATP1B1, LMNA, THBS1, COL4A1, COL4A2, SDC1, SDC4, FHL2, ANK3
hsa04520	Adherens junction	4,0886184	1,8e-13	LEF1, TCF7L2, FYN	CDH1, CTNNA1, YES1, SNAI2, EGFR, MET, PTPRF, TJP1, ACTN1, NECTIN2, ARHGAP29

Table 16. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05146	Amoebiasis	4,8653317	2,4e-13	ITGB2, COL1A2	IL1B, LAMA3, LAMA5, LAMB1, LAMB2, LAMB3, LAMC1, LAMC2, CXCL1, CXCL2, ARG1, COL4A1, COL4A2, ACTN1, PTK2, SERPINB6
hsa04915	Estrogen signaling pathway	2,2908711	3,7e-13		EGFR, FOS, JUN, GNAI1, KRT23, KRT9, KRT13, KRT14, KRT15, KRT18, KRT19
hsa04390	Hippo signaling pathway	3,1145098	7,2e-13	ITGB2, LEF1, TCF7L2	WWC1, FRMD6, YAP1, WWTR1, LATS2, CTNNA1, PATJ, CDH1, LLGL2, TEAD4, AJUBA, WNT10A, FZD6, CCND1, SNAI2, ID1
hsa05226	Gastric cancer	2,0106926	7,0e-12	LEF1, TCF7L2, E2F2	CCND1, FZD6, WNT10A, MET, CDH1, CTNNA1, JUP, EGFR
hsa05145	Toxoplasmosis	3,4406454	8,0e-12	HLA-DPB1, PIK3CG	MAPK13, GNAI1, ITGA6, ITGB1, LAMA3, LAMA5, LAMB1, LAMB2, LAMB3, LAMC1, LAMC2
hsa05213	Endometrial cancer	2,8230937	7,1e-11	LEF1, TCF7L2	CTNNA1, CDH1, EGFR, CCND1

Table 16. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05205	Proteoglycans in cancer	2,6466503	1,7e-10	HCLS1, COL1A2	CD44, ANK3, RRAS, CTTN, CCND1, PLAU, PTK2, MET, THBS1, SDC1, ITGA2, ITGB1, ITGB5, FZD6, SDC4, MAPK13, EGFR, WNT10A
hsa04670	Leukocyte transendothelial migration	2,8763596	7,2e-10	ITGB2, ITGAL, ITGA4, CXCR4	CLDN4, F11R, ITGB1, PTK2, MAPK13, CTNNA1, GNAI1, ACTN1
hsa05200	Pathways in cancer	1,9544495	1,9e-09	LEF1, TCF7L2, E2F2, CXCR4	CDH1, WNT10A, CTNNA1, ITGB1, COL4A1, COL4A2, LAMA3, LAMA5, LAMB1, LAMB2, LAMB3, LAMC1, LAMC2, ITGA6, ITGA2, PTK2, FZD6, EGFR, MET, FGFR3, CCND1, JUN, FOS, MECOM, CTBP2, PPARG, JUP, VEGFC, GNG12, GNAI1, JAG1, JAG2, NQO1, EDN1
hsa04933	AGE-RAGE signaling pathway in diabetic complications	3,0378943	2,4e-09	COL1A2	EGR1, MAPK13, EDN1, F3, IL1B, JUN, CCND1, COL4A1, COL4A2, VEGFC
hsa05412	Arrhythmogenic right ventricular cardiomyopathy	4,1766317	2,5e-09	ITGA4, LEF1, TCF7L2	LMNA, DSG2, JUP, ITGA6, ITGA2, ITGB1, ITGB5, ITGB6, CTNNA1

Table 16. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04010	MAPK signaling pathway	1,6177278	2,7e-09	MAP4K1	DUSP1, DUSP5, RRAS, GNG12, EGFR, EPHA2, FGFR3, MET, MAPK13, IL1B, MECOM, FOS, JUN, PPP3CA, EFNA1, VEGFC
hsa04668	TNF signaling pathway	2,5156280	2,8e-09		MAPK13, FOS, JUN, CXCL1, CXCL2, JAG1, IL1B, MMP14, EDN1, VEGFC
hsa04062	Chemokine signaling pathway	1,2919242	3,6e-09	CXCR4, DOCK2, PIK3CG	CXCL1, CXCL2, CXCL16, PTK2, GNAI1, GNG12
hsa05150	Staphylococcus aureus infection	4,4635400	5,4e-09	HLA-DPB1, ITGB2, ITGAL	C3, C1S, C1R, KRT23, KRT9, KRT13, KRT14, KRT15, KRT18, KRT19
hsa04530	Tight junction	2,5567641	1,3e-08	HCLS1	PATJ, TJP1, F11R, ACTN1, MICALL2, RAB13, LLGL2, MARVELD3, CLDN4, PRKAG2, TUBA4A, JUN, CCND1, CTTN, ITGB1
hsa04810	Regulation of actin cytoskeleton	2,2043334	1,6e-08	CYFIP2, ITGA4, ITGAL, ITGB2, CXCR4	NCKAP1, SSH3, GSN, ACTN1, GNG12, PTK2, ITGA6, ITGA2, ITGB1, ITGB5, ITGB6, RRAS, EGFR, FGFR3

Table 16. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04657	IL-17 signaling pathway	2,3913264	2,0e-08		FOS, JUN, FOSL1, MAPK13, CXCL1, CXCL2, S100A7, IL1B
hsa05133	Pertussis	4,7869849	3,4e-08	ITGB2, IRF8	ITGB1, IL1B, GNAI1, C1R, C1S, C3, MAPK13, FOS, JUN, CASP1, PYCARD
hsa05130	Pathogenic Escherichia coli infection	3,0841891	3,9e-08	HCLS1, FYN, WIPF1, CYFIP2, MYO1G	TUBA4A, CTTN, CLDN4, ITGB1, TUBB6, PYCARD, CASP1, IL1B, MAPK13, FOS, JUN, NCKAP1, TJP1, MYO1B, MYO1C, MYO6
hsa04310	Wnt signaling pathway	1,9175731	5,7e-08	LEF1, TCF7L2	PPP3CA, FZD6, MMP7, CCND1, FOSL1, JUN, TLE1, CTBP2, DKK1, WNT10A
hsa01522	Endocrine resistance	2,5693325	7,3e-08	E2F2	EGFR, MAPK13, JAG1, JAG2, CCND1, JUN, FOS, PTK2
hsa05225	Hepatocellular carcinoma	1,4658371	7,8e-08	LEF1, TCF7L2, E2F2	EGFR, CCND1, FZD6, WNT10A, MET, NQO1
hsa05216	Thyroid cancer	3,7364475	7,8e-08	LEF1, TCF7L2	PPARG, CDH1, CCND1
hsa04380	Osteoclast differentiation	2,0490196	8,7e-08	FYN, LCK	MAPK13, IL1B, FOS, JUN, FOSL1, PPP3CA, FHL2, PPARG

Table 16. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05410	Hypertrophic cardiomyopathy	2,5985294	2,2e-07	ITGA4	ITGA6, ITGA2, ITGB1, ITGB5, ITGB6, LMNA, PRKAG2, EDN1
hsa04625	C-type lectin receptor signaling pathway	1,8721569	2,4e-07		RRAS, CASP1, PYCARD, IL1B, PPP3CA, MAPK13, JUN
hsa05414	Dilated cardiomyopathy	1,9124183	2,5e-07	ITGA4	ITGA6, ITGA2, ITGB1, ITGB5, ITGB6, LMNA
hsa05132	Salmonella infection	2,2993523	2,9e-07	LEF1, TCF7L2, CYFIP2, PIK3CG, PTPRC	MAPK13, IL1B, FOS, JUN, PYCARD, CASP1, NCKAP1, MYO6, SNX9, AHNAK2, AHNAK, RRAS, KLC3, TUBA4A, TUBB6
hsa05210	Colorectal cancer	1,9798319	3,6e-07	LEF1, TCF7L2	CCND1, JUN, FOS, EGFR
hsa04926	Relaxin signaling pathway	2,3685278	4,4e-07	COL1A2	GNG12, EDN1, GNAI1, VEGFC, COL4A1, COL4A2, FOS, JUN, MAPK13, EGFR
hsa05135	Yersinia infection	2,2181450	4,5e-07	WIPF1, ITGA4, LCK	ITGB1, PTK2, PYCARD, CASP1, IL1B, FOS, JUN, MAPK13
hsa04934	Cushing syndrome	1,7064969	1,3e-06	LEF1, TCF7L2, E2F2	FZD6, EGFR, GNAI1, CCND1, AHR, WNT10A
hsa05224	Breast cancer	1,9963305	1,5e-06	LEF1, TCF7L2, E2F2	EGFR, JAG1, JAG2, FOS, JUN, CCND1, FZD6, WNT10A

Table 16. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04015	Rap1 signaling pathway	2,1705193	1,9e-06	APBB1IP, ITGAL, ITGB2	EFNA1, VEGFC, EGFR, EPHA2, FGFR3, MET, MAPK13, CDH1, ITGB1, ADORA2B, GNAI1, RRAS, ID1, THBS1
hsa04928	Parathyroid hormone synthesis, secretion and	1,6938562	2,8e-06		GNAI1, MMP14, EGFR, CYP24A1, FOS, EGR1, SGK1
hsa05417	Lipid and atherosclerosis	1,5963043	3,3e-06		FOS, JUN, IL1B, MAPK13, PYCARD, CASP1, CXCL1, CXCL2, PPP3CA, HSPA2, PTK2, PPARG
hsa05131	Shigellosis	2,1173203	3,7e-06	TIFA, FNBP1, HCLS1	IL1B, MAPK13, JUN, PYCARD, CASP1, C3, CTTN, ITGB1, CD44, PTK2, ACTN1, CAPN2, EGFR, HK1
hsa05221	Acute myeloid leukemia	1,6938562	4,5e-06	LEF1, TCF7L2	JUP, CCND1
hsa05215	Prostate cancer	2,0443092	5,4e-06	LEF1, TCF7L2, E2F2, ZEB1	CCND1, EGFR, PLAU
hsa04371	Apelin signaling pathway	1,8148459	6,8e-06	PIK3CG	PRKAG2, GNAI1, GNG12, RRAS, CCND1, EGR1, JAG1, CDH1
hsa05418	Fluid shear stress and atherosclerosis	2,1014006	7,9e-06		MAPK13, FOS, JUN, NQO1, ASS1, IL1B, EDN1, PTK2, SDC1, SDC4, DUSP1

Table 16. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05163	Human cytomegalovirus infection	1,2640718	7,9e-06	CXCR4, E2F2	EGFR, GNG12, PPP3CA, GNAI1, CCND1, IL1B, PTK2, MAPK13
hsa04360	Axon guidance	2,1689622	8,1e-06	CXCR4, FYN	GNAI1, PPP3CA, PTK2, ITGB1, EPHA2, EFNA1, MET, SEMA3C, RHOD, RRAS, SSH3, ABLIM1
hsa05167	Kaposi sarcoma-associated herpesvirus infection	1,8982871	8,6e-06	E2F2, PIK3CG, LEF1, TCF7L2	C3, GNG12, CXCL1, CXCL2, PPP3CA, FOS, JUN, CCND1, MAPK13
hsa05235	PD-L1 expression and PD-1 checkpoint	1,8591105	1,2e-05	LCK	FOS, JUN, PPP3CA, EGFR, MAPK13
hsa05323	Rheumatoid arthritis	2,5407843	1,3e-05	HLA-DPB1, ITGAL, ITGB2	IL1B, CXCL1, CXCL2, FOS, JUN
hsa04611	Platelet activation	1,9378863	1,5e-05	APBB1IP, PIK3CG, COL1A2, FERMT3, FYN	ITGA2, ITGB1, MAPK13, GNAI1
hsa05171	Coronavirus disease - COVID-19	1,6814014	1,6e-05		IL1B, FOS, JUN, EGFR, MAPK13, CASP1, C1R, C1S, C3
hsa04064	NF-κB signaling pathway	2,0443092	1,8e-05	LCK	IL1B, LTBR, PLAU, CXCL1, CXCL2, PPP1R13L
hsa04621	NOD-like receptor signaling pathway	1,2437406	1,8e-05		PYCARD, CASP1, IL1B, MAPK13, JUN, CXCL1, CXCL2
hsa05142	Chagas disease	1,5716192	2,3e-05		C3, MAPK13, FOS, JUN, IL1B, GNAI1
hsa05140	Leishmaniasis	3,4647059	3,5e-05	ITGA4, ITGB2, HLA-DPB1	ITGB1, C3, IL1B, FOS, MAPK13, JUN

Table 16. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05100	Bacterial invasion of epithelial cells	2,4702070	4,4e-05	HCLS1	CTNNA1, CDH1, CTTN, ITGB1, PTK2, MET
hsa04660	T-cell receptor signaling pathway	1,5332319	4,7e-05	FYN, LCK, PTPRC	FOS, JUN, PPP3CA, MAPK13
hsa05230	Central carbon metabolism in	1,5879902	5,2e-05		EGFR, FGFR3, MET, HK1
hsa05207	Chemical carcinogenesis - receptor activation	1,2633734	5,4e-05		EGFR, FOS, JUN, AHR, CCND1, KLF5, KLF4, JAG1, GNAI1
hsa05170	Human immunodeficiency virus 1 infection	1,1107254	5,4e-05	CXCR4	GNG12, PPP3CA, GNAI1, PTK2, MAPK13, FOS, JUN
hsa05169	Epstein-Barr virus infection	1,1355461	5,8e-05	HLA-DPB1, ITGAL, VIM, E2F2	CD44, MAPK13, JUN, CCND1
hsa05217	Basal cell	1,6392157	7,0e-05	LEF1, TCF7L2	WNT10A, FZD6
hsa05120	Epithelial cell signaling in Helicobacter pylori infection	3,4129939	7,5e-05	CSK	MAPK13, JUN, CXCL1, CXCL2, MET, F11R, TJP1, EGFR
hsa04921	Oxytocin signaling pathway	1,5635596	8,7e-05	PIK3CG	PRKAG2, PPP3CA, JUN, FOS, CCND1, GNAI1, EGFR
hsa05218	Melanoma	1,8148459	9,8e-05	E2F2	CDH1, CCND1, EGFR, MET
hsa04936	Alcoholic liver disease	1,8682238	1,1e-04	LEF1, TCF7L2	IL1B, PRKAG2, MAPK13, CCND1, C3, CXCL1, CXCL2, ALDH3A2
hsa04014	Ras signaling pathway	1,0394118	1,1e-04	RASAL3	RRAS, EFNA1, VEGFC, EGFR, EPHA2, FGFR3, MET, GNG12

Table 16. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04020	Calcium signaling pathway	0,9197409	1,1e-04	CXCR4	PPP3CA, EGFR, FGFR3, MET, ADORA2B, ASPH, VEGFC
hsa04912	GnRH signaling pathway	1,8148459	1,4e-04	CGA	MAPK13, MMP14, EGFR, JUN, EGR1
hsa04658	Th1 and Th2 cell differentiation	2,3913264	1,4e-04	LCK, HLA-DPB1	PPP3CA, MAPK13, JUN, FOS, JAG1, JAG2
hsa04620	Toll-like receptor signaling pathway	1,0477461	1,4e-04		MAPK13, IL1B, FOS, JUN
hsa04659	Th17 cell differentiation	2,0125024	1,5e-04	LCK, HLA-DPB1	PPP3CA, MAPK13, JUN, FOS, IL1B, AHR
hsa04061	Viral protein interaction with cytokine and cytokine receptor	1,4436275	1,9e-04	CXCR4	CXCL1, CXCL2, IL20RB, LTBR
hsa04814	Motor proteins	1,7436755	2,0e-04	MYO1G	TUBA4A, TUBB6, KLC3, MYO1B, MYO1C, MYO6
hsa04514	Cell adhesion molecules	3,0009264	2,2e-04	PTPRC, ITGB2, ITGAL, ITGA4, ICAM2, HLA-DPB1	ITGB1, ITGA6, NECTIN2, CDH1, SDC1, SDC4, PTPRF, F11R, CLDN4
hsa04916	Melanogenesis	1,6217772	2,4e-04	LEF1, TCF7L2	GNAI1, FZD6, WNT10A, EDN1
hsa04115	p53 signaling pathway	1,4518768	2,7e-04		SFN, THBS1, CCND1, PPP1R13L
hsa04060	Cytokine-cytokine receptor interaction	0,8469281	2,8e-04	CXCR4	IL20RB, IL1B, LTBR, CXCL16, TNFRSF12A, CXCL1, CXCL2, GDF15

Table 16. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05134	Legionellosis	3,7841469	3,0e-04	ITGB2	PYCARD, CASP1, IL1B, C3, CXCL1, CXCL2
hsa00220	Arginine biosynthesis	2,3098039	3,3e-04		ASS1, ARG1
hsa04974	Protein digestion and absorption	3,6296919	3,4e-04	COL1A2, COL6A3	ATP1B1, COL4A1, COL4A2
hsa04935	Growth hormone synthesis, secretion and action	0,9498259	3,5e-04		FOS, PTK2, MAPK13, GNAI1
hsa05208	Chemical carcinogenesis - ROS	1,0698039	4,3e-04		EGFR, MET, NQO1, PTK2, FOS, JUN, MAPK13, AHR
hsa04392	Hippo signaling pathway - multiple species	8,1668067	4,4e-04	RASSF2	WWC1, FRMD6, YAP1, WWTR1, LATS2, TEAD4, RASSF4, AJUBA
hsa04932	Non-alcoholic fatty liver disease	1,1046888	4,5e-04		PRKAG2, FOS, JUN, IL1B, MAPK13, PPARG
hsa04725	Cholinergic synapse	1,7934948	4,8e-04	CHAT, PIK3CG, FYN	GNAI1, GNG12, FOS
hsa04724	Glutamatergic synapse	0,8469281	4,8e-04		GNAI1, GNG12, PPP3CA
hsa05219	Bladder cancer	4,1201908	4,8e-04	E2F2	THBS1, EGFR, CDH1, FGFR3,
hsa05166	Human T-cell leukemia virus 1 infection	1,7293296	5,3e-04	ITGAL, ITGB2, E2F2, HLA-DPB1, LCK	PPP3CA, LTBR, JUN, CCND1, FOS, EGR1, FOSL1, MMP7
hsa05144	Malaria	3,7182209	5,4e-04	ITGAL, ITGB2	SDC1, MET, THBS1, IL1B
hsa04068	FoxO signaling pathway	1,2394070	5,5e-04		PRKAG2, EGFR, CCND1, PLK2, SGK1, MAPK13

Table 16. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04024	cAMP signaling pathway	0,8675849	5,5e-04	CGA	ATP1B1, GNAI1, RRAS, FOS, JUN, EDN1
hsa04066	HIF-1 signaling pathway	1,0163137	6,6e-04		EGFR, HK1, EDN1, LTBR
hsa04917	Prolactin signaling pathway	1,5398693	6,6e-04	CGA	MAPK13, CCND1, FOS
hsa05162	Measles	0,8130510	7,6e-04		IL1B, CCND1, FOS, JUN
hsa01521	EGFR TKI resistance	1,4518768	7,9e-04		EGFR, MET, FGFR3, GAS6
hsa04148	Efferocytosis	2,4286909	7,9e-04		GULP1, ITGB5, PTK2, GAS6, P2RY2, THBS1, PPARG, ARG1, ADAM9, DUSP5, MAPK13, SGK1, CASP1
hsa04623	Cytosolic DNA-sensing pathway	1,3611345	8,7e-04		IL1B, PYCARD, CASP1
hsa04728	Dopaminergic synapse	1,0951657	9,2e-04		GNAI1, GNG12, MAPK13, FOS, PPP3CA
hsa04210	Apoptosis	1,2703922	9,3e-04		CAPN2, LMNA, JUN, FOS, CTSV, TUBA4A
hsa04012	ErbB signaling pathway	1,0300477	9,5e-04		EGFR, PTK2, JUN
hsa04370	VEGF signaling pathway	1,4658371	1,2e-03		PPP3CA, PTK2, MAPK13
hsa04960	Aldosterone-regulated sodium reabsorption	2,5407843	1,2e-03		ATP1B1, SGK1, SFN
hsa04726	Serotonergic synapse	1,2547083	1,3e-03		GNAI1, GNG12, DUSP1, APP
hsa04713	Circadian entrainment	0,9295552	1,3e-03		GNG12, GNAI1, FOS

Table 16. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05223	NSCLC	1,5879902	1,6e-03	E2F2	CCND1, EGFR, MET
hsa05231	Choline metabolism in	1,4274069	1,7e-03		EGFR, SLC44A3, PLPP3, FOS, JUN
hsa04022	cGMP-PKG signaling pathway	0,6820897	1,8e-03	PIK3CG	PPP3CA, ATP1B1, GNAI1
hsa05161	Hepatitis B	0,6730554	1,9e-03	E2F2	FOS, JUN, MAPK13
hsa04145	Phagosome	2,1173203	2,1e-03	ITGB2, HLA-DPB1	C1R, ITGB1, ITGB5, ITGA2, TUBA4A, TUBB6, C3, THBS1
hsa04924	Renin secretion	1,2919242	2,4e-03		GNAI1, PPP3CA, EDN1
hsa04723	Retrograde endocannabinoid signaling	0,7190899	3,1e-03		GNAI1, GNG12, MAPK13
hsa04650	Natural killer cell mediated cytotoxicity	1,8478431	3,3e-03	FYN, LAT2, LCK, CD48, ICAM2, ITGB2, ITGAL	PPP3CA
hsa04540	Gap junction	1,6715686	4,0e-03		GNAI1, TJP1, TUBB6, TUBA4A, EGFR
hsa00562	Inositol phosphate metabolism	0,7259384	4,2e-03	PIK3CG	INPP1
hsa04613	Neutrophil extracellular trap formation	1,4247389	4,5e-03	HMGB1, ITGB2, ITGAL	CASP1, C3, MAPK13
hsa04722	Neurotrophin signaling pathway	0,4619608	4,7e-03		JUN, MAPK13
hsa04640	Hematopoietic cell lineage	1,8148459	5,6e-03	ITGA4, HLA-DPB1	ITGA6, ITGA2, CD44, IL1B
hsa05168	Herpes simplex virus 1 infection	0,4252359	6,2e-03	HLA-DPB1, ZNF100, ZNF92, ZFP82, ZNF107	NECTIN2, C3, IL1B

Table 16. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04261	Adrenergic signaling in	0,7528250	6,4e-03	PIK3CG	ATP1B1, GNAI1, MAPK13
hsa04072	Phospholipase D signaling pathway	0,9139512	6,7e-03	PIK3CG, FYN	RRAS, EGFR, PLPP3
hsa05164	Influenza A	0,7259384	6,7e-03	HLA-DPB1	PYCARD, CASP1, IL1B
hsa04217	Necroptosis	1,4004323	6,7e-03	HMGB1	PYGL, IL1B, PYCARD, CASP1, CAPN2, CHMP4C
hsa04218	Cellular senescence	1,0441579	7,3e-03	E2F2	RRAS, CCND1, MAPK13, CAPN2, PPP3CA
hsa05321	Inflammatory bowel disease	1,3141988	8,6e-03	HLA-DPB1	IL1B, JUN
hsa05212	Pancreatic cancer	1,0889076	9,0e-03	E2F2	EGFR, CCND1
hsa05214	Glioma	1,0889076	9,0e-03	E2F2	EGFR, CCND1
hsa05206	MicroRNAs in cancer	1,1637180	9,3e-03	ZEB1, E2F2, VIM	MET, CCND1, EGFR, CYP24A1, CD44, PLAU, EFNA1, THBS1, FGFR3
hsa04662	B-cell receptor signaling pathway	1,6715686	1,1e-02	CD72, PIK3AP1	FOS, JUN, PPP3CA
hsa04664	Fc epsilon RI signaling pathway	0,8196078	1,2e-02	FYN	MAPK13
hsa04213	Longevity regulating pathway - multiple species	0,4793933	1,6e-02		PRKAG2
hsa05416	Viral myocarditis	2,4909650	1,8e-02	HLA-DPB1, ITGAL, ITGB2, FYN	CCND1
hsa01523	Antifolate resistance	1,2098973	1,9e-02		IL1B
hsa05160	Hepatitis C	0,8701316	2,1e-02	E2F2	CLDN4, IFIT1, EGFR, CCND1
hsa05211	Renal cell carcinoma	0,7939951	2,2e-02		MET, JUN

Table 16. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04920	Adipocytokine signaling pathway	0,3849673	2,2e-02		PRKAG2
hsa04071	Sphingolipid signaling pathway	0,7057734	2,2e-02	FYN	GNAI1, MAPK13
hsa04914	Progesterone-mediated oocyte maturation	0,5031256	2,2e-02		GNAI1, MAPK13
hsa04211	Longevity regulating pathway	0,6197035	2,3e-02		PRKAG2, PPARG
hsa04630	JAK-STAT signaling pathway	0,4763971	2,5e-02		IL20RB, EGFR, CCND1
hsa04919	Thyroid hormone signaling pathway	0,4705156	2,6e-02		ATP1B1, CCND1
hsa05034	Alcoholism	0,7057734	2,6e-02		GNAI1, GNG12, ADORA2B
hsa05031	Amphetamine addiction	1,2294118	2,8e-02		PPP3CA, FOS, JUN
hsa04152	AMPK signaling pathway	0,6745445	3,0e-02		PRKAG2, PPARG, CCND1
hsa04730	Long-term depression	0,4981930	3,0e-02		GNAI1
hsa05020	Prion disease	0,9942199	3,2e-02	FYN	LAMC1, EGR1, IL1B, KLC3, TUBA4A, TUBB6, PPP3CA, MAPK13
hsa04137	Mitophagy - animal	0,8285166	3,5e-02		OPTN, RRAS, JUN
hsa05202	Transcriptional misregulation in cancer	1,6778764	3,5e-02	ZEB1	PPARG, MET, NFKBIZ, PLAU, JUP, PTK2
hsa04672	Intestinal immune network for IgA production	2,7467939	3,6e-02	HLA-DPB1, ITGA4, CXCR4	LTBR
hsa04971	Gastric acid secretion	0,8761325	3,7e-02		ATP1B1, GNAI1
hsa04973	Carbohydrate digestion and absorption	1,1817601	3,9e-02		ATP1B1, HK1

Table 16. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 1) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04931	Insulin resistance	0,7777911	3,9e-02		PTPRF, PYGL, PRKAG2
hsa04720	Long-term potentiation	0,4165220	4,0e-02		PPP3CA
hsa04922	Glucagon signaling pathway	0,7400343	4,1e-02		PYGL, PPP3CA, PRKAG2
hsa04727	GABAergic synapse	0,7939951	4,2e-02		GNAI1, GNG12
hsa05415	Diabetic cardiomyopathy	0,2920442	4,3e-02	COL1A2	MAPK13
hsa05030	Cocaine addiction	1,0811848	4,4e-02		GNAI1, JUN
hsa05203	Viral carcinogenesis	1,1209343	4,4e-02		JUN, CCND1, GSN, ACTN1, C3, LTBR
hsa04550	Signaling pathways regulating pluripotency of	1,3490890	4,6e-02		WNT10A, FZD6, KLF4, FGFR3, MAPK13, ID1
hsa04750	Inflammatory mediator regulation of TRP	1,1681767	4,8e-02		P2RY2, IL1B, MAPK13, PRKCH
hsa05032	Morphine addiction	0,6961053	4,8e-02		GNAI1, GNG12

Table 17. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 2)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04512	ECM-receptor interaction	6,9904693	2,7e-32	COL1A2, COL6A3, ITGA4	LAMA3, LAMA5, LAMB1, LAMB2, LAMB3, LAMC1, LAMC2, COL4A1, COL4A2, COL4A5, THBS1, ITGB1, ITGA6, SDC1, SDC4, ITGB6, ITGB5, CD44, ITGA2, AGRN, DSPP
hsa04510	Focal adhesion	3,7094488	2,7e-32	ITGA4, FYN, COL1A2, COL6A3	ITGB1, ITGB5, ITGB6, ITGA6, ITGA2, ACTN1, PTK2, CAPN2, COL4A1, COL4A2, COL4A5, LAMA3, LAMA5, LAMB1, LAMB2, LAMB3, LAMC1, LAMC2, THBS1, CCND1, EGFR, MET, VEGFC, JUN, EMP1
hsa04151	PI3K-Akt signaling pathway	2,2955844	1,5e-23	PIK3CG, ITGA4, COL1A2, COL6A3, MYB, PIK3AP1	EGFR, EPHA2, FGFR3, MET, CCND1, EFNA1, VEGFC, GNG12, ITGA6, ITGA2, ITGB1, ITGB5, ITGB6, COL4A1, COL4A2, COL4A5, LAMA3, LAMA5, LAMB1, LAMB2, LAMB3, LAMC1, LAMC2, THBS1, PTK2, SGK1

Table 17. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05165	Human papillomavirus infection	2,4253611	1,5e-21	ITGA4, COL1A2, COL6A3, HES6, MFNG, TCF7L2	EGFR, CCND1, ITGA6, ITGA2, ITGB1, ITGB5, ITGB6, COL4A1, COL4A2, COL4A5, LAMA3, LAMA5, LAMB1, LAMB2, LAMB3, LAMC1, LAMC2, THBS1, PTK2, PRKCZ, LLGL2, PATJ, JAG1, FZD6, WNT10A
hsa05222	Small cell lung cancer	4,7753823	6,9e-20	E2F2	CCND1, ITGB1, ITGA6, ITGA2, PTK2, COL4A1, COL4A2, COL4A5, LAMA3, LAMA5, LAMB1, LAMB2, LAMB3, LAMC1, LAMC2
hsa04820	Cytoskeleton in muscle cells	4,5457967	1,2e-15	ITGA4, VIM, COL1A2, COL6A3	ITGA6, ITGA2, ITGB1, ITGB5, ITGB6, DSG2, JUP, ATP1B1, LMNA, THBS1, COL4A1, COL4A2, COL4A5, AGRN, SDC1, SDC4, FHL2, ANK3
hsa05146	Amoebiasis	4,8865116	5,3e-15	ITGB2, COL1A2	IL1B, LAMA3, LAMA5, LAMB1, LAMB2, LAMB3, LAMC1, LAMC2, CXCL1, CXCL2, ARG1, COL4A1, COL4A2, COL4A5, ACTN1, PTK2, SERPINB6

Table 17. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04520	Adherens junction	3,8902899	1,8e-13	LEF1, TCF7L2, FYN	CDH1, CTNNA1, YES1, SNAI2, EGFR, MET, PTPRF, TJP1, ACTN1, NECTIN2, ARHGAP29
hsa04915	Estrogen signaling pathway	2,1797468	3,7e-13		EGFR, FOS, JUN, GNAI1, KRT23, KRT9, KRT13, KRT14, KRT15, KRT18, KRT19
hsa04390	Hippo signaling pathway	3,1194030	7,2e-13	ITGB2, LEF1, TCF7L2	WWC1, FRMD6, YAP1, WWTR1, LATS2, CTNNA1, PATJ, CDH1, LLGL2, TEAD4, AJUBA, PRKCZ, WNT10A, FZD6, CCND1, SNAI2, ID1
hsa05226	Gastric cancer	1,9131590	7,0e-12	LEF1, TCF7L2, E2F2	CCND1, FZD6, WNT10A, MET, CDH1, CTNNA1, JUP, EGFR
hsa05145	Toxoplasmosis	3,2737484	1,2e-11	HLA-DPB1, PIK3CG	MAPK13, GNAI1, ITGA6, ITGB1, LAMA3, LAMA5, LAMB1, LAMB2, LAMB3, LAMC1, LAMC2
hsa05213	Endometrial cancer	2,6861526	7,1e-11	LEF1, TCF7L2	CTNNA1, CDH1, EGFR, CCND1
hsa04062	Chemokine signaling pathway	1,5024243	9,4e-11	CXCR4, GRK3, DOCK2, PIK3CG	CXCL1, CXCL2, CXCL16, PTK2, GNAI1, GNG12, PRKCZ

Table 17. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05200	Pathways in cancer	2,0064581	1,6e-10	LEF1, TCF7L2, E2F2, CXCR4, CSF2RB	CDH1, WNT10A, CTNNA1, ITGB1, COL4A1, COL4A2, COL4A5, LAMA3, LAMA5, LAMB1, LAMB2, LAMB3, LAMC1, LAMC2, ITGA6, ITGA2, PTK2, FZD6, EGFR, MET, FGFR3, CCND1, JUN, FOS, MECOM, CTBP2, PPARG, JUP, VEGFC, GNG12, GNAI1, JAG1, JAG2, NQO1, IL13RA1, EDN1
hsa05205	Proteoglycans in cancer	2,5182680	6,2e-10	HCLS1, COL1A2	CD44, ANK3, RRAS, CTTN, CCND1, PLAU, PTK2, MET, THBS1, SDC1, ITGA2, ITGB1, ITGB5, FZD6, SDC4, MAPK13, EGFR, WNT10A
hsa04670	Leukocyte transendothelial migration	2,9649043	7,2e-10	ITGB2, ITGAL, ITGA4, NCF4, CXCR4	CLDN4, F11R, ITGB1, PTK2, MAPK13, CTNNA1, GNAI1, ACTN1
hsa04810	Regulation of actin cytoskeleton	2,2077966	1,1e-09	CYFIP2, ITGA4, ITGAL, ITGB2, CXCR4	NCKAP1, CYFIP1, SSH3, GSN, ACTN1, GNG12, PTK2, ITGA6, ITGA2, ITGB1, ITGB5, ITGB6, RRAS, EGFR, FGFR3

Table 17. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04933	AGE-RAGE signaling pathway in diabetic complications	3,4160853	2,4e-09	COL1A2	EGR1, MAPK13, EDN1, F3, IL1B, JUN, CCND1, COL4A1, COL4A2, COL4A5, VEGFC, PRKCZ
hsa05412	Arrhythmogenic right ventricular cardiomyopathy	3,9740339	2,5e-09	ITGA4, LEF1, TCF7L2	LMNA, DSG2, JUP, ITGA6, ITGA2, ITGB1, ITGB5, ITGB6, CTNNA1
hsa04010	MAPK signaling pathway	1,5392560	2,7e-09	MAP4K1	DUSP1, DUSP5, RRAS, GNG12, EGFR, EPHA2, FGFR3, MET, MAPK13, IL1B, MECOM, FOS, JUN, PPP3CA, EFNA1, VEGFC
hsa04668	TNF signaling pathway	2,3936013	2,8e-09		MAPK13, FOS, JUN, CXCL1, CXCL2, JAG1, IL1B, MMP14, EDN1, VEGFC
hsa05150	Staphylococcus aureus infection	4,2470250	5,4e-09	HLA-DPB1, ITGB2, ITGAL	C3, C1S, C1R, KRT23, KRT9, KRT13, KRT14, KRT15, KRT18, KRT19
hsa05133	Pertussis	4,5547804	1,1e-08	ITGB2, IRF8	ITGB1, IL1B, GNAI1, C1R, C1S, C3, MAPK13, FOS, JUN, CASP1, PYCARD

Table 17. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04530	Tight junction	2,5847883	1,3e-08	HCLS1	PATJ, TJP1, PRKCZ, F11R, ACTN1, MICALL2, RAB13, LLGL2, MARVELD3, CLDN4, PRKAG2, TUBA4A, JUN, CCND1, CTTN, ITGB1
hsa04657	IL-17 signaling pathway	2,2753292	2,0e-08		FOS, JUN, FOSL1, MAPK13, CXCL1, CXCL2, S100A7, IL1B
hsa05130	Pathogenic Escherichia coli infection	3,0743249	3,9e-08	HCLS1, FYN, WIPF1, CYFIP2, MYO1G	TUBA4A, CTTN, CLDN4, ITGB1, TUBB6, PYCARD, CASP1, IL1B, MAPK13, FOS, JUN, NCKAP1, CYFIP1, TJP1, MYO1B, MYO1C, MYO6
hsa04310	Wnt signaling pathway	1,8245565	5,7e-08	LEF1, TCF7L2	PPP3CA, FZD6, MMP7, CCND1, FOSL1, JUN, TLE1, CTBP2, DKK1, WNT10A
hsa04015	Rap1 signaling pathway	2,1867172	6,4e-08	APBB1IP, ITGAL, ITGB2	EFNA1, VEGFC, EGFR, EPHA2, FGFR3, MET, MAPK13, CDH1, ITGB1, ADORA2B, GNAI1, RRAS, ID1, THBS1, PRKCZ
hsa05225	Hepatocellular carcinoma	1,3947331	7,8e-08	LEF1, TCF7L2, E2F2	EGFR, CCND1, FZD6, WNT10A, MET, NQO1
hsa05216	Thyroid cancer	3,5552019	7,8e-08	LEF1, TCF7L2	PPARG, CDH1, CCND1

Table 17. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04380	Osteoclast differentiation	2,3395522	8,7e-08	FYN, LCK, BTK, NCF4	MAPK13, IL1B, FOS, JUN, FOSL1, PPP3CA, FHL2, PPARG
hsa04926	Relaxin signaling pathway	2,6633886	9,2e-08	COL1A2	GNG12, PRKCZ, EDN1, GNAI1, VEGFC, COL4A1, COL4A2, COL4A5, FOS, JUN, MAPK13, EGFR
hsa01522	Endocrine resistance	2,7163341	1,1e-07	E2F2	EGFR, MAPK13, JAG1, JAG2, CCND1, JUN, FOS, BIK, PTK2
hsa05410	Hypertrophic cardiomyopathy	2,4724813	2,2e-07	ITGA4	ITGA6, ITGA2, ITGB1, ITGB5, ITGB6, LMNA, PRKAG2, EDN1
hsa05135	Yersinia infection	2,1105484	2,3e-07	WIPF1, ITGA4, LCK	ITGB1, PTK2, PYCARD, CASP1, IL1B, FOS, JUN, MAPK13
hsa05414	Dilated cardiomyopathy	1,8196517	2,5e-07	ITGA4	ITGA6, ITGA2, ITGB1, ITGB5, ITGB6, LMNA
hsa04330	Notch signaling pathway	2,5009007	3,5e-07	DTX1, MFNG	TLE1, CTBP2, JAG1, JAG2
hsa05210	Colorectal cancer	1,8837953	3,6e-07	LEF1, TCF7L2	CCND1, JUN, FOS, EGFR
hsa04625	C-type lectin receptor signaling pathway	1,7813433	6,0e-07		RRAS, CASP1, PYCARD, IL1B, PPP3CA, MAPK13, JUN

Table 17. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05132	Salmonella infection	2,2972074	9,5e-07	LEF1, TCF7L2, CYFIP2, PIK3CG, PTPRC	MAPK13, IL1B, FOS, JUN, PYCARD, CASP1, NCKAP1, CYFIP1, MYO6, SNX9, AHNAK2, AHNAK, RRAS, KLC3, TUBA4A, TUBB6
hsa04360	Axon guidance	2,2111622	1,0e-06	CXCR4, FYN	GNAI1, PPP3CA, PTK2, ITGB1, EPHA2, EFNA1, MET, SEMA3C, RHOD, RRAS, PRKCZ, SSH3, ABLIM1
hsa04934	Cushing syndrome	1,6237191	1,3e-06	LEF1, TCF7L2, E2F2	FZD6, EGFR, GNAI1, CCND1, AHR, WNT10A
hsa05417	Lipid and atherosclerosis	1,6454442	1,4e-06	NCF4	FOS, JUN, IL1B, MAPK13, PYCARD, CASP1, CXCL1, CXCL2, PPP3CA, HSPA2, PTK2, PPARG
hsa05418	Fluid shear stress and atherosclerosis	2,1812367	1,5e-06		MAPK13, FOS, JUN, NQO1, PRKCZ, ASS1, IL1B, EDN1, PTK2, SDC1, SDC4, DUSP1
hsa05224	Breast cancer	1,8994936	1,5e-06	LEF1, TCF7L2, E2F2	EGFR, JAG1, JAG2, FOS, JUN, CCND1, FZD6, WNT10A
hsa05131	Shigellosis	2,0146144	1,7e-06	TIFA, FNBP1, HCLS1	IL1B, MAPK13, JUN, PYCARD, CASP1, C3, CTTN, ITGB1, CD44, PTK2, ACTN1, CAPN2, EGFR, HK1

Table 17. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04611	Platelet activation	2,2536365	2,1e-06	APBB1IP, PIK3CG, COL1A2, FERMT3, FYN, BTK	ITGA2, ITGB1, MAPK13, GNAI1, PRKCZ
hsa05221	Acute myeloid leukemia	1,6116915	4,5e-06	LEF1, TCF7L2	JUP, CCND1
hsa04928	Parathyroid hormone synthesis, secretion and action	1,6116915	4,6e-06		GNAI1, MMP14, EGFR, CYP24A1, FOS, EGR1, SGK1
hsa05215	Prostate cancer	1,9451450	5,4e-06	LEF1, TCF7L2, E2F2, ZEB1	CCND1, EGFR, PLAU
hsa05140	Leishmaniasis	3,6629353	6,0e-06	ITGA4, ITGB2, HLA-DPB1, NCF4	ITGB1, C3, IL1B, FOS, MAPK13, JUN
hsa05142	Chagas disease	1,4953839	6,5e-06		C3, MAPK13, FOS, JUN, IL1B, GNAI1
hsa05120	Epithelial cell signaling in Helicobacter pylori infection	3,2474382	8,0e-06	CSK	MAPK13, JUN, CXCL1, CXCL2, MET, F11R, TJP1, EGFR
hsa05167	Kaposi sarcoma-associated herpesvirus infection	1,8062060	8,6e-06	E2F2, PIK3CG, LEF1, TCF7L2	C3, GNG12, CXCL1, CXCL2, PPP3CA, FOS, JUN, CCND1, MAPK13
hsa05171	Coronavirus disease - COVID-19	1,5998409	8,9e-06		IL1B, FOS, JUN, EGFR, MAPK13, CASP1, C1R, C1S, C3
hsa05100	Bacterial invasion of epithelial cells	2,3503835	9,6e-06	HCLS1	CTNNA1, CDH1, CTTN, ITGB1, PTK2, MET

Table 17. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04621	NOD-like receptor signaling pathway	1,1834099	1,0e-05		PYCARD, CASP1, IL1B, MAPK13, JUN, CXCL1, CXCL2
hsa04974	Protein digestion and absorption	4,1443497	1,0e-05	COL1A2, COL6A3	ATP1B1, COL4A1, COL4A2, COL4A5
hsa04371	Apelin signaling pathway	1,7268124	1,1e-05	PIK3CG	PRKAG2, GNAI1, GNG12, RRAS, CCND1, EGR1, JAG1, CDH1
hsa05323	Rheumatoid arthritis	2,4175373	1,2e-05	HLA-DPB1, ITGAL, ITGB2	IL1B, CXCL1, CXCL2, FOS, JUN
hsa05235	PD-L1 expression and PD-1 checkpoint pathway in cancer	1,7689297	1,2e-05	LCK	FOS, JUN, PPP3CA, EGFR, MAPK13
hsa05163	Human cytomegalovirus infection	1,2027549	1,5e-05	CXCR4, E2F2	EGFR, GNG12, PPP3CA, GNAI1, CCND1, IL1B, PTK2, MAPK13
hsa04064	NF-κB signaling pathway	2,2230228	1,6e-05	BTK, LCK	IL1B, LTBR, PLAU, CXCL1, CXCL2, PPP1R13L
hsa05169	Epstein-Barr virus infection	1,2155216	4,6e-05	HLA-DPB1, ITGAL, VIM, E2F2, BTK	CD44, MAPK13, JUN, CCND1
hsa04660	T-cell receptor signaling pathway	1,4588587	4,7e-05	FYN, LCK, PTPRC	FOS, JUN, PPP3CA, MAPK13
hsa05230	Central carbon metabolism in cancer	1,5109608	5,2e-05		EGFR, FGFR3, MET, HK1
hsa05217	Basal cell carcinoma	1,5597015	7,0e-05	LEF1, TCF7L2	WNT10A, FZD6

Table 17. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04936	Alcoholic liver disease	1,7776010	8,1e-05	LEF1, TCF7L2	IL1B, PRKAG2, MAPK13, CCND1, C3, CXCL1, CXCL2, ALDH3A2
hsa04620	Toll-like receptor signaling pathway	0,9969226	8,6e-05		MAPK13, IL1B, FOS, JUN
hsa04659	Th17 cell differentiation	1,9148810	8,7e-05	LCK, HLA-DPB1	PPP3CA, MAPK13, JUN, FOS, IL1B, AHR
hsa05207	Chemical carcinogenesis - receptor activation	1,2020904	9,4e-05		EGFR, FOS, JUN, AHR, CCND1, KLF5, KLF4, JAG1, GNAI1
hsa05170	Human immunodeficiency virus 1 infection	1,1889528	9,4e-05	CXCR4, AP1S2	GNG12, PPP3CA, GNAI1, PTK2, MAPK13, FOS, JUN
hsa04514	Cell adhesion molecules	2,8553590	9,4e-05	PTPRC, ITGB2, ITGAL, ITGA4, ICAM2, HLA-DPB1	ITGB1, ITGA6, NECTIN2, CDH1, SDC1, SDC4, PTPRF, F11R, CLDN4
hsa05218	Melanoma	1,7268124	9,8e-05	E2F2	CDH1, CCND1, EGFR, MET
hsa04921	Oxytocin signaling pathway	1,4877153	1,0e-04	PIK3CG	PRKAG2, PPP3CA, JUN, FOS, CCND1, GNAI1, EGFR
hsa04014	Ras signaling pathway	0,9889925	1,1e-04	RASAL3	RRAS, EFNA1, VEGFC, EGFR, EPHA2, FGFR3, MET, GNG12
hsa04020	Calcium signaling pathway	0,8751266	1,1e-04	CXCR4	PPP3CA, EGFR, FGFR3, MET, ADORA2B, ASPH, VEGFC
hsa04630	JAK-STAT signaling pathway	0,7554804	1,5e-04	CSF2RB	IL13RA1, IL20RB, EGFR, CCND1

Table 17. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa01521	EGFR TKI resistance	1,3814499	1,5e-04		EGFR, MET, FGFR3, GAS6
hsa04658	Th1 and Th2 cell differentiation	2,2753292	1,6e-04	LCK, HLA-DPB1	PPP3CA, MAPK13, JUN, FOS, JAG1, JAG2
hsa04912	GnRH signaling pathway	1,7268124	1,9e-04	CGA	MAPK13, MMP14, EGFR, JUN, EGR1
hsa04061	Viral protein interaction with cytokine and cytokine receptor	1,3736007	1,9e-04	CXCR4	CXCL1, CXCL2, IL20RB, LTBR
hsa04814	Motor proteins	1,6590942	2,0e-04	MYO1G	TUBA4A, TUBB6, KLC3, MYO1B, MYO1C, MYO6
hsa04916	Melanogenesis	1,5431089	2,4e-04	LEF1, TCF7L2	GNAI1, FZD6, WNT10A, EDN1
hsa05134	Legionellosis	3,6005875	2,5e-04	ITGB2	PYCARD, CASP1, IL1B, C3, CXCL1, CXCL2
hsa04932	Non-alcoholic fatty liver disease	1,0511032	2,5e-04		PRKAG2, FOS, JUN, IL1B, MAPK13, PPARG
hsa04060	Cytokine-cytokine receptor interaction	0,9849226	2,8e-04	CSF2RB, CXCR4	IL20RB, IL1B, LTBR, IL13RA1, CXCL16, TNFRSF12A, CXCL1, CXCL2, GDF15
hsa00220	Arginine biosynthesis	2,1977612	3,3e-04		ASS1, ARG1
hsa04115	p53 signaling pathway	1,3814499	4,4e-04		SFN, THBS1, CCND1, PPP1R13L
hsa04392	Hippo signaling pathway - multiple species	7,7706557	4,4e-04	RASSF2	WWC1, FRMD6, YAP1, WWTR1, LATS2, TEAD4, RASSF4, AJUBA

Table 17. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04725	Cholinergic synapse	1,7064969	4,8e-04	CHAT, PIK3CG, FYN	GNAI1, GNG12, FOS
hsa04724	Glutamatergic synapse	1,0744610	4,8e-04	GRK3	GNAI1, GNG12, PPP3CA
hsa05219	Bladder cancer	3,9203308	4,8e-04	E2F2	THBS1, EGFR, CDH1, FGFR3, CCND1
hsa04935	Growth hormone synthesis, secretion and action	0,9037523	4,9e-04		FOS, PTK2, MAPK13, GNAI1
hsa04623	Cytosolic DNA-sensing pathway	1,2951093	5,9e-04		IL1B, PYCARD, CASP1
hsa05144	Malaria	3,5378595	5,9e-04	ITGAL, ITGB2	SDC1, MET, THBS1, IL1B
hsa05162	Measles	0,7736119	6,0e-04		IL1B, CCND1, FOS, JUN
hsa04066	HIF-1 signaling pathway	0,9670149	6,6e-04		EGFR, HK1, EDN1, LTBR
hsa05208	Chemical carcinogenesis - ROS	1,0179104	6,6e-04		EGFR, MET, NQO1, PTK2, FOS, JUN, MAPK13, AHR
hsa04917	Prolactin signaling pathway	1,4651741	7,3e-04	CGA	MAPK13, CCND1, FOS
hsa05166	Human T-cell leukemia virus 1 infection	1,6454442	8,4e-04	ITGAL, ITGB2, E2F2, HLA-DPB1, LCK	PPP3CA, LTBR, JUN, CCND1, FOS, EGR1, FOSL1, MMP7
hsa04068	FoxO signaling pathway	1,1792865	8,5e-04		PRKAG2, EGFR, CCND1, PLK2, SGK1, MAPK13
hsa04024	cAMP signaling pathway	0,8255005	9,0e-04	CGA	ATP1B1, GNAI1, RRAS, FOS, JUN, EDN1
hsa04728	Dopaminergic synapse	1,0420419	9,2e-04		GNAI1, GNG12, MAPK13, FOS, PPP3CA

Table 17. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04148	Efferocytosis	2,3108813	1,2e-03		GULP1, ITGB5, PTK2, GAS6, P2RY2, THBS1, PPARG, ARG1, ADAM9, DUSP5, MAPK13, SGK1, CASP1
hsa04370	VEGF signaling pathway	1,3947331	1,3e-03		PPP3CA, PTK2, MAPK13
hsa04012	ErbB signaling pathway	0,9800827	1,4e-03		EGFR, PTK2, JUN
hsa04210	Apoptosis	1,4102301	1,4e-03	CSF2RB	CAPN2, LMNA, JUN, FOS, CTSV, TUBA4A
hsa04960	Aldosterone-regulated sodium reabsorption	2,4175373	1,5e-03		ATP1B1, SGK1, SFN
hsa05223	NSCLC	1,5109608	1,6e-03	E2F2	CCND1, EGFR, MET
hsa04540	Gap junction	1,5904851	1,7e-03		GNAI1, TJP1, TUBB6, TUBA4A, EGFR
hsa04726	Serotonergic synapse	1,1938456	1,8e-03		GNAI1, GNG12, DUSP1, APP
hsa04022	cGMP-PKG signaling pathway	0,6490033	1,8e-03	PIK3CG	PPP3CA, ATP1B1, GNAI1
hsa04713	Circadian entrainment	0,8844649	1,8e-03		GNG12, GNAI1, FOS
hsa04924	Renin secretion	1,2292563	2,4e-03		GNAI1, PPP3CA, EDN1
hsa05231	Choline metabolism in cancer	1,3581670	2,4e-03		EGFR, SLC44A3, PLPP3, FOS, JUN
hsa05161	Hepatitis B	0,6404072	2,7e-03	E2F2	FOS, JUN, MAPK13

Table 17. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04145	Phagosome	2,2160759	3,3e-03	ITGB2, NCF4, HLA-DPB1	C1R, ITGB1, ITGB5, ITGA2, TUBA4A, TUBB6, C3, THBS1
hsa05212	Pancreatic cancer	1,0360874	3,8e-03	E2F2	EGFR, CCND1
hsa04071	Sphingolipid signaling pathway	0,8953842	4,1e-03	FYN	GNAI1, PRKCZ, MAPK13
hsa04723	Retrograde endocannabinoid signaling	0,6842087	4,1e-03		GNAI1, GNG12, MAPK13
hsa00562	Inositol phosphate metabolism	0,6907249	4,2e-03	PIK3CG	INPP1
hsa04650	Natural killer cell mediated cytotoxicity	1,7582090	4,2e-03	FYN, LAT2, LCK, CD48, ICAM2, ITGB2, ITGAL	PPP3CA
hsa04217	Necroptosis	1,3325009	4,5e-03	HMGB1	PYGL, IL1B, PYCARD, CASP1, CAPN2, CHMP4C
hsa04613	Neutrophil extracellular trap formation	1,5815665	4,6e-03	HMGB1, ITGB2, NCF4, ITGAL	CASP1, C3, MAPK13
hsa04931	Insulin resistance	0,9867499	4,7e-03		PTPRF, PRKCZ, PYGL, PRKAG2
hsa04722	Neurotrophin signaling pathway	0,4395522	5,4e-03		JUN, MAPK13
hsa05164	Influenza A	0,6907249	5,8e-03	HLA-DPB1	PYCARD, CASP1, IL1B
hsa05168	Herpes simplex virus 1 infection	0,4046088	6,2e-03	HLA-DPB1, ZNF100, ZNF92, ZFP82, ZNF107	NECTIN2, C3, IL1B
hsa05032	Morphine addiction	0,9935085	6,5e-03	GRK3	GNAI1, GNG12
hsa05321	Inflammatory bowel disease	1,2504503	7,7e-03	HLA-DPB1	IL1B, JUN

Table 17. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04910	Insulin signaling pathway	0,9298220	7,7e-03		PTPRF, PRKAG2, PYGL, HK1, PRKCZ
hsa04261	Adrenergic signaling in cardiomyocytes	0,7163074	8,0e-03	PIK3CG	ATP1B1, GNAI1, MAPK13
hsa05020	Prion disease	1,0511032	8,2e-03	FYN, NCF4	LAMC1, EGR1, IL1B, KLC3, TUBA4A, TUBB6, PPP3CA, MAPK13
hsa04640	Hematopoietic cell lineage	1,7268124	8,2e-03	ITGA4, HLA-DPB1	ITGA6, ITGA2, CD44, IL1B
hsa04072	Phospholipase D signaling pathway	0,8696177	8,3e-03	PIK3CG, FYN	RRAS, EGFR, PLPP3
hsa04218	Cellular senescence	0,9935085	9,2e-03	E2F2	RRAS, CCND1, MAPK13, CAPN2, PPP3CA
hsa04664	Fc epsilon RI signaling pathway	1,1697761	1,0e-02	FYN, BTK	MAPK13
hsa05214	Glioma	1,0360874	1,1e-02	E2F2	EGFR, CCND1
hsa05160	Hepatitis C	0,8279237	1,2e-02	E2F2	CLDN4, IFIT1, EGFR, CCND1
hsa05415	Diabetic cardiomyopathy	0,5557557	1,2e-02	NCF4, COL1A2	MAPK13, PRKCZ
hsa05206	MicroRNAs in cancer	1,1072690	1,3e-02	ZEB1, E2F2, VIM	MET, CCND1, EGFR, CYP24A1, CD44, PLAU, EFNA1, THBS1, FGFR3
hsa04662	B-cell receptor signaling pathway	1,9085821	1,3e-02	BTK, CD72, PIK3AP1	FOS, JUN, PPP3CA
hsa05030	Cocaine addiction	1,0287393	1,4e-02		GNAI1, JUN
hsa01523	Antifolate resistance	1,1512082	1,7e-02		IL1B

Table 17. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa04213	Longevity regulating pathway - multiple species	0,4561391	1,9e-02		PRKAG2
hsa05416	Viral myocarditis	2,3701346	2,1e-02	HLA-DPB1, ITGAL, ITGB2, FYN	CCND1
hsa05031	Amphetamine addiction	1,1697761	2,2e-02		PPP3CA, FOS, JUN
hsa04211	Longevity regulating pathway	0,5896432	2,3e-02		PRKAG2, PPARG
hsa04622	RIG-I-like receptor signaling pathway	0,3608265	2,4e-02		MAPK13
hsa05211	Renal cell carcinoma	0,7554804	2,4e-02		MET, JUN
hsa04920	Adipocytokine signaling pathway	0,3662935	2,4e-02		PRKAG2
hsa04914	Progesterone-mediated oocyte maturation	0,4787203	2,7e-02		GNAI1, MAPK13
hsa05340	Primary immunodeficiency	3,1194030	3,0e-02	BTK, PTPRC, LCK, RFXAP	
hsa04144	Endocytosis	1,4128465	3,0e-02	CXCR4, GRK3, WIPF1	CHMP4C, PRKCZ, EGFR, FGFR3, RAB11FIP5, ASAP2
hsa04152	AMPK signaling pathway	0,6418241	3,0e-02		PRKAG2, PPARG, CCND1
hsa04730	Long-term depression	0,4740269	3,0e-02		GNAI1
hsa04919	Thyroid hormone signaling pathway	0,4476921	3,1e-02		ATP1B1, CCND1
hsa05034	Alcoholism	0,6715381	3,1e-02		GNAI1, GNG12, ADORA2B

Table 17. pathfindR pathway enrichment analysis for cetuximab utilizing the gene features of the pan-drug model trained with EGFR signaling inhibitors (Strategy 2) (continued)

ID	Term Description	Fold Enrichment	lowest_p	Up regulated	Down regulated
hsa05202	Transcriptional misregulation in cancer	1,5964869	3,3e-02	ZEB1	PPARG, MET, NFKBIZ, PLAUI, JUP, PTK2
hsa04930	Type II diabetes mellitus	1,2723881	3,5e-02		PRKCZ, HK1
hsa04672	Intestinal immune network for IgA production	2,6135539	3,6e-02	HLA-DPB1, ITGA4, CXCR4	LTBR
hsa04971	Gastric acid secretion	0,8336336	3,7e-02		ATP1B1, GNAI1
hsa04750	Inflammatory mediator regulation of TRP channels	1,1115114	3,9e-02		P2RY2, IL1B, MAPK13, PRKCH
hsa04137	Mitophagy - animal	0,7883274	4,0e-02		OPTN, RRAS, JUN
hsa04720	Long-term potentiation	0,3963176	4,0e-02		PPP3CA
hsa04727	GABAergic synapse	0,7554804	4,2e-02		GNAI1, GNG12
hsa04973	Carbohydrate digestion and absorption	1,1244360	4,5e-02		ATP1B1, HK1
hsa04922	Glucagon signaling pathway	0,7041371	4,7e-02		PYGL, PPP3CA, PRKAG2

APPENDIX 3: Associations between CNA status and drug resistance

Table 18. Table of associations between CNA status and paclitaxel resistance

Gene Symbol	OR	p-value
PAPSS2	6,43	0,003983105192262734
BAG3	6,19	0,004113611749071865
IFIT1	6,03	0,004362061506488537
COL17A1	6,03	0,004362061506488537
PLAU	5,87	0,007133122105907046
HHEX	5,87	0,007133122105907046
PCBD1	5,72	0,00723329062860378
MYOF	5,64	0,0073624606414624
FAM241B	5,49	0,012506661146092456
TSPAN15	5,49	0,012506661146092456
SRGN	5,27	0,012419704730508036
HTRA1	4,27	0,010125420865656772
PLEKHA1	4,27	0,010125420865656772
TACC2	4,21	0,016227594338153954
ACSL5	4,11	0,016209847740137985
TCF7L2	4,11	0,016209847740137985
DUSP5	3,85	0,026985326648223586
DPCD	3,80	0,026919055526623874
TAF5	3,70	0,027209763799450986
PIK3AP1	3,60	0,028067798630273144
PDLIM1	3,56	0,04375719500729744
ANK3	3,56	0,04375719500729744
DKK1	3,51	0,04350248786099126
AVPI1	3,46	0,043448673556278114

Table 19. Table of associations between CNA status and 5-FU resistance

Gene Symbol	OR	p-value
CFL2	4,69	5,081433839242822e-08
PRKD1	4,66	8,56091581914088e-08
AJUBA	4,58	2,4707215031305434e-07
MAP4K1	4,55	6,44468236955747e-07
BMP4	4,50	1,6180355671897593e-07
AXL	4,49	6,901133074580304e-07
FERMT2	4,47	1,661251641090647e-07
FRMD6	4,39	2,9560694722027296e-07
GMFG	4,26	8,715208145880064e-07
RRAS	4,22	3,3587310481866544e-06
C19orf33	4,15	1,4835304187574832e-06
SPINT2	4,15	1,4835304187574832e-06
PYGL	4,12	6,293393801903408e-07
ATF5	4,09	5,676801028705717e-06
LIMCH1	4,08	2,4512552258918705e-06
BLVRB	4,05	2,512965597840349e-06
BCAR3	3,96	9,510909892105268e-06
LRRC8C	3,93	9,61085197388555e-06
ARHGAP29	3,93	9,61085197388555e-06
F3	3,91	9,844237943507272e-06
CNN3	3,91	9,844237943507272e-06
BATF	3,88	8,116994959611325e-07
PLPP2	3,85	3,181742070821569e-06
LSR	3,85	3,181742070821569e-06
FOS	3,81	1,3404317746720272e-06
LGALS3	3,81	1,3404317746720272e-06
ARHGAP45	3,75	5,294783798212752e-06

Table 19. Table of associations between CNA status and 5-FU resistance (continued)

Gene Symbol	OR	p-value
ARID3A	3,75	5,294783798212752e-06
MOB3A	3,68	8,573468812891452e-06
LGMN	3,68	1,5741937126662551e-06
IFI27	3,66	2,574981175317691e-06
CD53	3,64	1,9540424252448987e-05
SIX4	3,64	2,5774625780843076e-06
KLK1	3,62	1,988728796619056e-05
FAM114A1	3,57	1,4411183726243418e-05
NECTIN2	3,57	1,4411183726243418e-05
PNMA8A	3,55	1,4788944394444845e-05
IGFLR1	3,54	1,0119988819226802e-05
SLC16A4	3,53	3,2108395919985387e-05
ID1	3,51	5,163221610177107e-07
LYL1	3,46	5,126182757309181e-05
CARD8	3,46	2,397011002176386e-05
CD70	3,42	3,82266488122667e-05
VAV1	3,42	3,82266488122667e-05
PTPN22	3,40	3,8177924913033976e-05
C3	3,40	3,8177924913033976e-05
HSPA2	3,38	5,469477943552048e-06
BCL11B	3,36	8,553513521852822e-06
BDNF	3,33	2,8213454308595425e-05
MOCOS	3,31	3,482723474317755e-06
CD59	3,31	2,9052132969316526e-05
EMB	3,30	2,011662628146173e-05
RAB17	3,27	0,00017161360369053816
RBM47	3,27	4,48616756031473e-05

Table 19. Table of associations between CNA status and 5-FU resistance (continued)

Gene Symbol	OR	p-value
CHI3L2	3,25	6,431523059007829e-05
DENND2D	3,25	6,431523059007829e-05
CDS1	3,25	2,1881002974407114e-05
ACAP1	3,24	9,577464269147405e-06
GPX2	3,22	1,5005448029448967e-05
PTGFRN	3,21	4,785606697545499e-05
CLDN7	3,21	1,0324355060549617e-05
PPP3CA	3,21	2,2744406513530162e-05
MGST2	3,18	2,4293831284739977e-05
LAMB2	3,16	1,0513938129910519e-05
PIMREG	3,13	1,6915260123358668e-05
SH3PXD2B	3,13	7,526868406796156e-05
PPARG	3,13	1,6915260123358668e-05
IQGAP2	3,13	0,00015388826494810716
RHPN2	3,11	5,497368721441369e-05
GNA11	3,11	5,497368721441369e-05
RASAL3	3,11	0,0001544177761992473
MDK	3,11	0,0001544177761992473
GNG7	3,09	8,290279458207295e-05
CKB	3,08	2,6925874883877594e-05
GIPC1	3,07	0,0001603687867180443
AHNAK2	3,06	2,756410430021426e-05
RFTN1	3,06	2,756410430021426e-05
ANXA2	3,05	8,429474442507451e-05
DHRS3	3,03	0,00012099390445192873
MAP1B	3,03	0,00023961819507490174
MITF	3,03	1,8791424271917532e-05

Table 19. Table of associations between CNA status and 5-FU resistance (continued)

Gene Symbol	OR	p-value
SVIL	3,03	2,8640959030444736e-05
BAMBI	3,03	2,8640959030444736e-05
DSPP	3,03	4,192305715906043e-05
CD44	3,03	8,634153554329955e-05
OSBPL1A	3,02	2,8966446675692328e-05
TAF4B	3,02	1,9145235932137343e-05
APBB1IP	3,02	2,8966446675692328e-05
CCL3	3,01	0,00024163540212807967
LMO2	2,99	0,0001314351618934765
GATA6	2,98	6,556770015841987e-05
MFAP4	2,97	3,070817644056558e-05
ADGRL2	2,97	0,00037106919401136613
NUP210	2,97	4,577537760942657e-05
FJX1	2,96	0,00013468335698758416
VEGFC	2,96	4,7329099629449175e-05
GPX8	2,94	0,00013835716195963876
SNX7	2,94	0,0002795232199245255
TLCD1	2,94	0,0002795232199245255
STAP2	2,94	0,00013835716195963876
PLK2	2,94	0,00013835716195963876
RAB34	2,94	0,0002795232199245255
MLPH	2,93	0,0004914542862495707
YES1	2,93	3,237962460378655e-05
LAMA3	2,91	4,923906591110208e-05
ANXA3	2,89	0,00010563871033372174
CD151	2,89	0,00010563871033372174
ITGA2	2,89	0,00014930028543893065

Table 19. Table of associations between CNA status and 5-FU resistance (continued)

Gene Symbol	OR	p-value
TSPAN4	2,89	0,00010563871033372174
TGIF1	2,89	3,377118562434714e-05
TMC4	2,88	0,0002085658298800406
CDH2	2,87	3,51597926845458e-05
CXCL2	2,87	0,00010786492287210282
TRIM22	2,87	0,00015165395912139287
SATB1	2,86	7,74913519862489e-05
PDGFC	2,86	0,00011461871543020976
ITPR1	2,86	5,2979077825686e-05
NMU	2,86	0,0001111804189013088
TMEM98	2,85	0,00042786441931499475
PLEK2	2,84	7,868785233654939e-05
CDC42BPB	2,84	7,868785233654939e-05
DSG2	2,84	3,6139575464875995e-05
ADA	2,82	2,1430481703997206e-05
UACA	2,82	0,0002324628627009189
PLCL2	2,81	8,352382767371508e-05
PROSER2	2,81	8,255562651046297e-05
TMEM51	2,81	0,0003204932707963296
CCND1	2,81	8,352382767371508e-05
THAP10	2,80	0,000234654009439914
NTN4	2,80	0,00139131708652148
CDCP1	2,80	8,369321410305579e-05
HACD1	2,80	8,369321410305579e-05
LARP6	2,80	0,000234654009439914
ACTN1	2,79	0,00012389038600485325
TNNT1	2,79	0,0003252354075652026

Table 19. Table of associations between CNA status and 5-FU resistance (continued)

Gene Symbol	OR	p-value
LEF1	2,79	0,00012124628040503114
LOX	2,78	0,0003329882148432277
EVI2B	2,78	0,0003329882148432277
TMEM131L	2,78	0,00018119989330188943
NFIB	2,77	3,080317938804086e-05
NAV2	2,77	0,00024573743981889354
PTPRM	2,76	6,046147919744074e-05
PPIC	2,76	0,0004882829989075886
MND1	2,76	0,00018166447585902273
RBM11	2,75	0,00025820533370554016
RHOD	2,75	0,00025820533370554016
SLFN11	2,75	0,0006493260371330247
JUP	2,74	0,0004884982492393793
NRIP1	2,74	0,0002601660770519664
ADORA2B	2,74	8,984459572277992e-05
IFI30	2,74	0,0006534777148429196
TJP1	2,73	0,00036022267248425664
PARVA	2,72	0,0002645358452090272
TNFAIP8	2,72	0,0004930761058351523
KRT15	2,72	0,0004930761058351523
MYO1F	2,72	0,0004930761058351523
KRT19	2,72	0,0004930761058351523
SAMSN1	2,72	0,0002645358452090272
STARD10	2,71	0,00019648256819218516
GDF15	2,70	0,0006789868483020865
PMP22	2,69	0,0001402981664638441
ATP8B1	2,69	5,567050989165651e-05

Table 19. Table of associations between CNA status and 5-FU resistance (continued)

Gene Symbol	OR	p-value
PALLD	2,69	0,00019848048894500938
PREX1	2,68	5,69281047939222e-05
SHROOM3	2,68	0,0002846480255786225
SCARB2	2,68	0,0002846480255786225
ANXA5	2,68	0,0002846480255786225
P3H4	2,67	0,0007382623224058333
ADGRA3	2,67	0,0003992702923407126
FKBP10	2,67	0,0007382623224058333
CXCL1	2,66	0,0002888310931523048
CRIP2	2,65	0,00021172635297913572
SERPINH1	2,65	0,00021172635297913572
AKR1C1	2,65	0,00015011640961202404
HTATIP2	2,65	0,0005524287121288075
DOCK2	2,65	0,0005524287121288075
AREG	2,65	0,0002956096210966914
FOSL1	2,63	0,0005612141155172828
CPE	2,63	0,0003076748273108565
AKR1C2	2,63	0,00022200023151845494
TIFA	2,63	0,0003076748273108565
IFTAP	2,63	0,0005612141155172828
ARMH1	2,61	0,0018639799694397663
TSPAN8	2,60	0,001872074303162927
P2RX5	2,60	0,00022674582753247317
LUM	2,60	0,001872074303162927
SDC4	2,60	9,603937560401715e-05
GRAMD2B	2,59	0,0011060945990207514
ITGB1	2,58	0,00023217695492087301

Table 19. Table of associations between CNA status and 5-FU resistance (continued)

Gene Symbol	OR	p-value
VIM	2,58	0,00023217695492087301
JUN	2,58	0,0014572713485740628
ITGA6	2,58	0,003460456582810915
TRIM2	2,57	0,0004728475675668677
CTTN	2,57	0,00033226499197019184
MTCL1	2,57	0,0002362053063331618
CELF2	2,57	0,0002362053063331618
WWC1	2,57	0,0008385253207771411
SULF2	2,56	9,93986197760777e-05
BEND5	2,56	0,0014776569862311436
ATP9A	2,56	9,93986197760777e-05
CENPV	2,55	0,0002413520160264947
BST2	2,54	0,001641475437536807
TACSTD2	2,54	0,0021397411712895152
RAB31	2,54	0,0002486393426456968
THBS1	2,54	0,0009160268681112723
SERPINE2	2,52	0,0027278357942437876
TBC1D10C	2,52	0,0006746382175669528
MYO1C	2,52	0,0003643164843184079
MAGI1	2,52	0,0002488119621352871
PTPRF	2,51	0,0021336744362755156
CEP15	2,51	0,0002531607995896679
WWC2	2,49	0,0003717533850893728
STAT5A	2,49	0,0012598368084143392
AP1M2	2,49	0,0012598368084143392
CTDSPL	2,49	0,0005366906824957704
FN1	2,49	0,0027937522071696764

Table 19. Table of associations between CNA status and 5-FU resistance (continued)

Gene Symbol	OR	p-value
ADAMTS1	2,48	0,0013682689792152368
TUBB6	2,48	0,00038016026081189025
CD276	2,48	0,0010008886298872193
LOXL1	2,46	0,001007024123170836
RBM38	2,46	0,00024219169581614719
FAT1	2,46	0,00039174596256122364
SSBP2	2,46	0,0018231557083772847
SQOR	2,46	0,001363177240110283
RASSF2	2,46	0,000270105735017489
TM6SF1	2,46	0,0007394731842433309
TPM1	2,45	0,0013683481979701677
HYI	2,44	0,003097603432206119
GNG12	2,43	0,0031164829641750144
MEF2C	2,42	0,0014094826088402734
RAG1	2,42	0,0014094826088402734
IGFBP2	2,42	0,004007813791098916
GALNT3	2,42	0,004007813791098916
EFEMP1	2,41	0,003157336417597761
MEN1	2,41	0,0018735852379626852
PLPP3	2,41	0,0024508868565868845
TMBIM1	2,40	0,005671114428022635
CYFIP2	2,38	0,002661276481025297
PPM1H	2,36	0,004470249487873121
APP	2,36	0,0021913288740393728
PAM	2,36	0,00205013822295858
FERMT3	2,36	0,002675029775134667
SRXN1	2,35	0,0006281700511698794

Table 19. Table of associations between CNA status and 5-FU resistance (continued)

Gene Symbol	OR	p-value
VAMP8	2,35	0,0057035715770571566
WIPF1	2,35	0,0057035715770571566
PINK1	2,34	0,0029329362460409998
ADM	2,34	0,0021860100533366775
CKAP4	2,34	0,005789147404234545
SEPSECS	2,33	0,0016395756205296506
CTSZ	2,33	0,0004194095099726527
EPSTI1	2,32	0,0009271009495277065
BACE2	2,31	0,0022205244322998714
SLC17A9	2,31	0,0006487263696265651
TGFBI	2,30	0,004981267960774359
PXK	2,29	0,0009563761524047857
PAPSS2	2,29	0,0023622921585425747
ST8SIA4	2,28	0,0030055396468474907
CCL28	2,27	0,0009720273512174912
P4HA2	2,27	0,003936389241143331
MYO6	2,25	0,003201053923625333
N4BP2L1	2,25	0,0014270358739646755
PLAU	2,25	0,0034238066418099036
FAM241B	2,23	0,003418211455462491
KRTCAP3	2,23	0,008962028851912763
CSK	2,23	0,003418211455462491
ASAP2	2,23	0,005514573616678707
DOCK10	2,22	0,007086944026082452
IFIT1	2,22	0,0034341128177793412
NFE2L2	2,22	0,008990725476680244
MGST1	2,21	0,001989084310501768

Table 19. Table of associations between CNA status and 5-FU resistance (continued)

Gene Symbol	OR	p-value
FGFBP1	2,21	0,0026177364447594995
LPXN	2,20	0,005989903458392842
TSC22D1	2,20	0,0015082100661720776
HK1	2,20	0,00347152362893637
EAF2	2,20	0,003646977159508743
MBOAT2	2,20	0,007790275381265522
HVCN1	2,19	0,012581487237403715
CDKN2A	2,19	0,0010902800998742806
HCLS1	2,19	0,0036553178679107883
TOM1L1	2,19	0,002750929181550202
MYLK	2,19	0,0036553178679107883
ITGB5	2,19	0,0036553178679107883
SLAIN1	2,18	0,0015256222021316243
GULP1	2,18	0,009916371786522809
PHLDA1	2,18	0,009916371786522809
TFPI	2,18	0,009916371786522809
PRSS23	2,17	0,0029268152945716303
RAB11FIP5	2,17	0,01562755029290513
RAI14	2,17	0,0015197289913663004
ID2	2,17	0,00777004850769641
AFAP1	2,17	0,003863080915336622
HOXD13	2,16	0,01252862441062002
ELF1	2,16	0,0021609674825203
ANKRD1	2,16	0,004904777409015839
EPHA2	2,16	0,006479753675991196
MYOF	2,16	0,004904777409015839
UGT3A2	2,15	0,0022118779928772146

Table 19. Table of associations between CNA status and 5-FU resistance (continued)

Gene Symbol	OR	p-value
COL3A1	2,15	0,009991980851725608
PCBD1	2,15	0,0052260476902616145
ITGA4	2,15	0,009991980851725608
AHR	2,14	0,0014317250263740397
SPARC	2,13	0,01086287673743729
TNFSF13B	2,13	0,002221562951606004
AHNAK	2,13	0,008440883196855532
ANXA6	2,13	0,01086287673743729
CTNNA1	2,13	0,01086287673743729
DAB2	2,12	0,0022453393024268785
SRGN	2,12	0,006955376360665105
MAP7	2,11	0,0041081334861974
WLS	2,11	0,010851036501455906
AMIGO2	2,11	0,010851036501455906
NR2F2	2,11	0,0041081334861974
CLIP4	2,11	0,01733203908279707
MYO1B	2,10	0,013780716337657466
CCDC80	2,10	0,00417030196028584
GCSAM	2,09	0,00579636679694474
COL4A2	2,09	0,002314152443097373
CHST11	2,09	0,01730473395667995
ST14	2,08	0,003267670256777085
CD74	2,08	0,01100154425708224
RASSF4	2,08	0,00738260312908024
ZNF22	2,08	0,00738260312908024
EPS8	2,07	0,004398644223119914
FARP1	2,07	0,003286261034048092

Table 19. Table of associations between CNA status and 5-FU resistance (continued)

Gene Symbol	OR	p-value
PRXL2A	2,07	0,007391037471650555
COL4A1	2,07	0,003286261034048092
POGLUT3	2,06	0,004523858094839605
RHOBTB3	2,06	0,009190847730218034
EGR1	2,06	0,015016038614443512
EMP1	2,06	0,005826809345997201
ERRFI1	2,06	0,007439004727525098
MYB	2,06	0,005826809345997201
ARHGDIB	2,05	0,005889707030768514
NECTIN3	2,05	0,005889707030768514
KLF5	2,05	0,0033786252822527087
GPRC5A	2,05	0,005889707030768514
TANC1	2,04	0,018958434275629793
JAKMIP2	2,04	0,011866662580818463
BAG3	2,04	0,007526588783929068
SYTL2	2,04	0,006261658896657276
GAS2L3	2,03	0,023685849860354054
SPATS2L	2,03	0,018984552800121977
SDC1	2,03	0,015139367625047258
FLI1	2,03	0,004758181972858891
NCKAP1	2,03	0,015139367625047258
DYNLL1	2,03	0,018984552800121977
DKK 1,00	2,03	0,010339401521924505
POU3F2	2,02	0,007899944285254215
LIMA1	2,02	0,019102607916775073
PXDN	2,02	0,01622806086935365
ANXA4	2,02	0,019102607916775073

Table 19. Table of associations between CNA status and 5-FU resistance (continued)

Gene Symbol	OR	p-value
PLBD1	2,02	0,008173036240594383
TESC	2,02	0,019102607916775073
HSPB8	2,02	0,019102607916775073
UPP1	2,01	0,0034076648466529304
IRAG2	2,01	0,008524513847640497
TNFRSF12A	2,01	0,012704102451824528
NLRC3	2,01	0,012704102451824528
SPR	2,01	0,023858483475386254
RUBCNL	2,01	0,00482067168505877
SGK1	2,01	0,008188302497886615
CORO7	2,00	0,012827395050027224
AGR2	2,00	0,004529340131544032
TSPAN13	2,00	0,004529340131544032
PIK3AP1	1,99	0,010439612331441517
YAP1	1,99	0,006614733269865049
COL18A1	1,98	0,01057616994344832
RNASEH2B	1,98	0,006816356263136086
FHL2	1,97	0,02582165458779814
PDLIM1	1,97	0,014274951721507136
FNBP1	1,97	0,010956226074380623
MGLL	1,96	0,011073134844638843
VASN	1,96	0,01740489299387192
BLVRA	1,96	0,004901769534322364
BLM	1,96	0,011397004174574826
HTRA1	1,96	0,011073134844638843
MICALL2	1,96	0,004901769534322364
PLEKHA1	1,96	0,011073134844638843

Table 19. Table of associations between CNA status and 5-FU resistance (continued)

Gene Symbol	OR	p-value
FSTL1	1,95	0,015038869436505791
TWF1	1,95	0,022124361355346376
PCDH9	1,95	0,006905953979975017
RAB32	1,95	0,011866902167273943
COL8A1	1,95	0,015038869436505791
MYO1G	1,94	0,006837574819928797
DUSP2	1,94	0,02613225609066125
PHLDB2	1,93	0,01188389448611836
TPST1	1,93	0,007079455073344889
SFN	1,93	0,014491712636860263
TPBG	1,93	0,014491712636860263
CASP1	1,92	0,009414839349458913
TACC2	1,92	0,015057363444908052
POU2AF1	1,92	0,009559965006742402
NIBAN2	1,92	0,015057363444908052
BIN2	1,92	0,034800369106540344
KRT7	1,92	0,034800369106540344
AVPI1	1,92	0,01954876325230869
AEBP1	1,91	0,007110686740985474
NELL2	1,91	0,023628430825604497
WDFY4	1,91	0,015168046324709278
TNFAIP3	1,91	0,012083312414725463
KRT80	1,90	0,034857556819136055
TNC	1,90	0,01571890137577925
TCF7L2	1,90	0,01534655460223955
PAWR	1,90	0,029893616280680064
KRT81	1,90	0,034857556819136055

Table 19. Table of associations between CNA status and 5-FU resistance (continued)

Gene Symbol	OR	p-value
RILPL2	1,89	0,028380837893656044
NPDC1	1,89	0,020494710083148484
CHAT	1,89	0,020494710083148484
E2F2	1,88	0,01968543515061638
GYPC	1,88	0,037437451725445006
TIMP2	1,88	0,016384856146423987
RUNX3	1,88	0,01968543515061638
SLC43A1	1,88	0,025040832652029738
EPCAM	1,88	0,037437451725445006
SLC43A3	1,88	0,025040832652029738
BCL9L	1,88	0,013239373832527542
LGALS3BP	1,88	0,016384856146423987
SHTN1	1,87	0,0205007557706098
AKNA	1,87	0,016168767292418337
TMEM45A	1,86	0,021393118290481634
IGFBP6	1,86	0,04646320132078671
KRT8	1,86	0,04646320132078671
ASPHD1	1,86	0,030233410652096494
NT5E	1,86	0,020592046779958996
KRT18	1,85	0,046455389926261775
C16orf54	1,85	0,04000341390257748
TNFRSF21	1,85	0,025496944174856644
CLDN4	1,85	0,013488271113965528
FMNL2	1,84	0,03784764049351708
LY75	1,84	0,03784764049351708
ITPRIPL2	1,84	0,03992123596685215
TUSC3	1,84	0,010172624628348675

Table 19. Table of associations between CNA status and 5-FU resistance (continued)

Gene Symbol	OR	p-value
GAS6	1,84	0,013740570962944596
FES	1,84	0,022231096755560995
CLEC2D	1,84	0,022231096755560995
INHBB	1,83	0,049784954851602455
C1R	1,83	0,022239766294180213
CCDC28B	1,83	0,0320155822847873
C1S	1,83	0,022239766294180213
DCBLD2	1,82	0,0218066147160142
CCN2	1,82	0,022336417630240656
ABCC3	1,82	0,022336417630240656
NUPR1	1,81	0,04020476612386403
PLOD2	1,80	0,023664176790172693
PTPN6	1,80	0,022794348061579296
GPRC5C	1,80	0,022794348061579296
GLCCI1	1,80	0,013348853575986676
PRKCB	1,80	0,042327294675582994
CD3D	1,79	0,01878704843069236
FAM117A	1,79	0,02989079528931792
WIP1	1,78	0,023952583805986943
ICA1	1,78	0,018231743265803825
METRNL	1,78	0,03554383691739045
ICAM2	1,77	0,024232331976866202
ACSL5	1,77	0,02947851065147232
UGCG	1,77	0,03010969683928561
ALDH1A3	1,77	0,030881801691504458
EMP2	1,77	0,04457743123706435
IFI6	1,76	0,03703344623387968

Table 19. Table of associations between CNA status and 5-FU resistance (continued)

Gene Symbol	OR	p-value
GSN	1,76	0,038381380578870125
FTH1	1,76	0,04464630620164789
ST6GALNAC2	1,75	0,030991016538945843
GNAI1	1,75	0,019629231759650395
CNTRL	1,75	0,03843956252781229
ME1	1,75	0,037287625338066364
PTGR1	1,74	0,03977113854370233
GALNT11	1,74	0,025848000896277187
FUCA2	1,74	0,031210179489273696
GTF2IRD1	1,74	0,025580948333354163
LAT2	1,74	0,025580948333354163
SERINC2	1,74	0,04524470039940246
ARHGEF10	1,73	0,02518618374018598
SEMA3C	1,72	0,026132361823828842
ITGB2	1,72	0,04877235273639598
LRP11	1,72	0,04099630170877379
SOX9	1,72	0,0328530616293784
DUSP4	1,71	0,026400953007884337
MARCKS	1,71	0,04104891542687775
SUSD3	1,71	0,040137188813792024
CD9	1,70	0,041234829444845064
PRKAG2	1,69	0,03489195358726929
LTBR	1,68	0,053870394477331926
TM4SF1	1,64	0,057610030452848224
PPP3CC	1,64	0,04617947288032642
LAMB1	1,63	0,04666768246239733
TRIP6	1,63	0,04642602505986762

Table 19. Table of associations between CNA status and 5-FU resistance (continued)

Gene Symbol	OR	p-value
NRROS	1,63	0,04651960289872276
BAIAP2L1	1,63	0,046771858541076086
CTSB	1,63	0,046371091224976846
CAV2	1,61	0,047490787626765084

Table 20. Table of associations between CNA status and gemcitabine resistance

Gene Symbol	OR	p-value
XYLT1	4,03	0,0011302289032104238
ITPRIPL2	3,98	0,001151061603131553
PRKCB	3,44	0,0023144039793556613
HS3ST4	3,37	0,0036302230346763036
ASPHD1	3,29	0,0036320389089941844
C16orf54	3,26	0,0036684918772198006
NUPR1	3,2	0,005594227983982539
ITGAL	3,16	0,0055867602139632305
TGFB1I1	3,14	0,005614248474602529
UACA	3,08	0,0021112114127222375
CORO1A	3,07	0,008592529853150713
GIPC1	3,05	0,004340055453314377
EMP2	3,05	0,004340055453314377
POU3F2	2,86	0,0036585871421277024
POPDC3	2,86	0,0036585871421277024
JUP	2,85	0,005082239714507561
KRT19	2,83	0,005136860872907472
KRT15	2,83	0,005136860872907472
AP1M2	2,79	0,007733443032636823
FKBP10	2,78	0,007701913586761892
P3H4	2,78	0,007701913586761892
NLRC3	2,71	0,007825046388555907
TNFRSF12A	2,71	0,007825046388555907
ANXA2	2,67	0,00588745369153997
LARP6	2,66	0,005986742905315529
THAP10	2,66	0,005986742905315529
VASN	2,64	0,011590114060600221

Table 20. Table of associations between CNA status and gemcitabine resistance (continued)

Gene Symbol	OR	p-value
LYL1	2,62	0,011603988360062944
MARCKS	2,61	0,00629788860765517
CCL3	2,59	0,011735535494924444
SQOR	2,56	0,008911479401617914
TPM1	2,55	0,009039218439726802
MAP7	2,50	0,0069301311665923414
PLCG2	2,50	0,0069301311665923414
CD70	2,49	0,01306630952365235
VAV1	2,49	0,01306630952365235
C3	2,47	0,013153745850058803
TMEM51	2,46	0,013277885397994377
PTPN22	2,46	0,013277885397994377
RAI14	2,45	0,004994096287648542
STAT5A	2,44	0,013438869599841453
MYB	2,44	0,010326262825334087
CNKS3	2,43	0,010347380247070805
PTPRK	2,43	0,010347380247070805
UGT3A2	2,42	0,005073020152826431
CD276	2,42	0,014287853810887347
JUN	2,41	0,025052504027628267
GNA11	2,41	0,01429982497551881
METRNL	2,41	0,01917741748551264
LOXL1	2,41	0,01429982497551881
MYO1F	2,41	0,01917741748551264
SGK1	2,39	0,01058529448496421
CSK	2,38	0,014438458753640833

Table 20. Table of associations between CNA status and gemcitabine resistance (continued)

Gene Symbol	OR	p-value
HOOK1	2,38	0,02509298275330364
TACSTD2	2,37	0,025207270990306736
WLS	2,37	0,025207270990306736
BEND5	2,37	0,025207270990306736
LRP11	2,36	0,015460260848477961
PTPRF	2,35	0,02538440653855817
SNX9	2,34	0,015428024713865264
MYOF	2,33	0,02106366731763873
CCN2	2,32	0,01548226651656542
CORO7	2,32	0,02791204574536754
TMEM98	2,32	0,02791204574536754
THBS1	2,31	0,02099982031870227
AKAP12	2,30	0,015568787150844662
TM6SF1	2,30	0,015568787150844662
HYI	2,29	0,035939993899973974
STAP2	2,28	0,02114066695001493
GNG12	2,28	0,035986504469043966
CYBA	2,27	0,016405182316252376
RASAL3	2,26	0,027958086377254013
SLFN11	2,26	0,027958086377254013
CNN3	2,25	0,028132421759223165
PTGFRN	2,24	0,02173716412621282
MICALL2	2,24	0,00820429885282492
FTH1	2,23	0,028372146091817443
LAPTM5	2,23	0,028372146091817443
SERPINH1	2,19	0,01742314740200212

Table 20. Table of associations between CNA status and gemcitabine resistance (continued)

Gene Symbol	OR	p-value
RUNX3	2,18	0,030456050226507055
IRF8	2,18	0,02402281317646026
GDF15	2,18	0,03970437271435687
SERINC2	2,18	0,03970437271435687
E2F2	2,18	0,030456050226507055
EVI2B	2,17	0,030621733865391597
DAB2	2,17	0,01266108042478933
CAMK2N1	2,16	0,030854690548277816
PLPP2	2,16	0,03263310412015483
TNFAIP3	2,16	0,02412577043706035
PINK1	2,14	0,031155119419814123
WWTR1	2,13	0,02520249918279198
AFAP1	2,13	0,02444472859081484
CCDC28B	2,12	0,04066658560855228
PLPP3	2,12	0,04066658560855228
ARID3A	2,11	0,032830356089165624
FUCA2	2,11	0,024980791359767857
ARHGAP45	2,11	0,032830356089165624
DHRS3	2,10	0,043085920790253515
NQO1	2,08	0,03334121480710836
MOB3A	2,08	0,03334121480710836
CBFA2T3	2,08	0,025794883888811283
CD53	2,06	0,04362475081589075
AJUBA	2,06	0,03474245044254622
ERRFI1	2,06	0,04612899313885674
DENND2D	2,05	0,04397658079481634

Table 20. Table of associations between CNA status and gemcitabine resistance (continued)

Gene Symbol	OR	p-value
CHI3L2	2,05	0,04397658079481634
TJP1	2,04	0,04611013998666569
RBM47	2,03	0,04617835660522318
IFTAP	2,02	0,04633375044751878
RHOD	2,02	0,03556702148773297
RAG1	2,02	0,04633375044751878
TBC1D10C	2,01	0,04881199253410558
FAM114A1	2,01	0,046576465572228734

8. CURRICULUM VITAE