

SpaRx: Elucidate single-cell spatial heterogeneity of drug responses for personalized treatment

Ziyang Tang¹, Xiang Liu², Zuotian Li^{2,3}, Tonglin Zhang⁴, Baijian Yang¹, Jing Su^{2,*}, and Qianqian Song^{5,6,*}

¹Department of Computer and Information Technology, Purdue University, Indiana, USA; ²Department of Biostatistics and Health Data Science, Indiana University School of Medicine, Indiana, USA; ³Department of Computer Graphics Technology, Purdue University, Indiana, USA; ⁴Department of Statistics, Purdue University, Indiana, USA; ⁵Center for Cancer Genomics and Precision Oncology, Atrium Health Wake Forest Baptist Comprehensive Cancer Center, North Carolina, USA, ⁶Department of Cancer Biology, Wake Forest University School of Medicine, North Carolina, USA

* Corresponding authors: Qianqian Song, PhD: qsong@wakehealth.edu; Jing Su, PhD: su1@iu.edu

Biographical Note

Ziyang Tang is the Ph.D. Candidate in the Department of Computer and Information Technology, Purdue University, Indiana, USA. His research focuses on developing novel artificial intelligence methods in interdisciplinary science.

Xiang Liu is the postdoctoral researcher in the Department of Biostatistics and Health Data Science, Indiana University School of Medicine, Indiana, USA. His research focuses on developing novel deep learning methods in biomedical informatics.

Zuotian Li is the Ph.D. Candidate in the Department of Computer Graphics Technology, Purdue University, Indiana, USA. Her research focuses on developing web tools for data visualization.

Tonglin Zhang is the Associate Professor in the Department of Statistics, Purdue University, Indiana, USA. His research focuses on developing novel statistical model for interdisciplinary research.

Baijian Yang is the Professor in the Department of Computer and Information Technology, Purdue University, Indiana, USA. His research focuses on developing novel statistical model for interdisciplinary research.

Jing Su is an Assistant Professor in the Department of Biostatistics and Health Data Science, Indiana University School of Medicine, Indiana, USA. His research focuses on graph artificial intelligence and machine learning in biomedical informatics and precision health.

Qianqian Song is an Assistant Professor in the Center for Cancer Genomics and Precision Oncology, Wake Forest University School of Medicine, North Carolina, USA. Her research focuses on machine learning and deep learning in bioinformatics and precision oncology.

Key Points

- We have developed a novel graph-based domain adaption model named SpaRx, to reveal the heterogeneity of spatial cellular response to different types of drugs, which bridges the gap between pharmacogenomics knowledgebase and single-cell spatial transcriptomics data.
- SpaRx is developed tailored for single-cell spatial transcriptomics data and is provided available as a ready-to-use open-source software, which demonstrates high accuracy and robust performance.
- SpaRx uncovers that tumor cells located in different areas within tumor lesion exhibit varying levels of sensitivity or resistance to drugs. Moreover, SpaRx reveals that tumor cells interact with themselves and the surrounding microenvironment to form an ecosystem capable of drug resistance.

ABSTRACT

Spatial cellular heterogeneity contributes to differential drug responses in a tumor lesion and potential therapeutic resistance. Recent emerging spatial technologies such as CosMx SMI, MERSCOPE, and Xenium delineate the spatial gene expression patterns at the single cell resolution. This provides unprecedented opportunities to identify spatially localized cellular resistance and to optimize the treatment for individual patients. In this work, we present a graph-based domain adaptation model, SpaRx, to reveal the heterogeneity of spatial cellular response to drugs. SpaRx transfers the knowledge from pharmacogenomics profiles to single-cell spatial transcriptomics data, through hybrid learning with dynamic adversarial adaptation. Comprehensive benchmarking demonstrates the superior and robust performance of SpaRx at different dropout rates, noise levels, and transcriptomics coverage. Further application of SpaRx to the state-of-art single-cell spatial transcriptomics data reveals that tumor cells in different locations of a tumor lesion present heterogeneous sensitivity or resistance to drugs. Moreover, resistant tumor cells interact with themselves or the surrounding constituents to form an ecosystem for drug resistance. Collectively, SpaRx characterizes the spatial therapeutic variability, unveils the molecular mechanisms underpinning drug resistance, and identifies personalized drug targets and effective drug combinations.

Keywords: Graph transformer, adversarial learning, spatial cellular drug response, single-cell spatial transcriptomics

INTRODUCTION

Understanding how different cells spatially localize and communicate in their microenvironment are critical to personalized treatment^{1,2}. For instance, tumor cells of one patient are spatially heterogeneous that present differential responses to treatment. For those tumor cells, their protective microenvironment and intercellular communications contribute to therapeutic failure and disease relapse^{3,4}. Therefore, in order to precisely treat a patient, it is crucial to identify that, inside its tumor lesion, which tumor cells are resistant to a candidate drug and what cell-cell communications are responsible for such tumor cell resistance. Unfortunately, these issues are not satisfactorily addressed, largely due to the lack of biotechnologies to accurately delineate the spatial heterogeneity of cells within tissues.

Recently, the emerging single-cell spatial transcriptomics (SCST) technologies, such as NanoString's CosMxTM SMI⁵ and Vizgen's MERSCOPE⁶, hold the promise to unravel the spatial tissue architectures at subcellular level and to further our understanding of the underlying functional mechanisms of tumor metastasis⁷ and drug resistance⁸. The SCST technologies provide the spatial locations of cells as well as their gene expression patterns, which offers a unique opportunity to investigate the therapeutic heterogeneity in the tumor microenvironment. Moreover, the existing pharmacogenomics databases, including the Cancer Cell Line Encyclopedia (CCLE)⁹ and the Genomics of Drugs Sensitivity in Cancer (GDSC)^{10,11}, provide valuable references relating gene expression patterns to drug response and treatment efficacy. The integration of these existing resources, along with SCST data, presents unprecedented opportunities to elucidate how individual cells within complex tissues will differentially respond to drugs, thus meeting the needs of personalized treatment.

Given the available data from drug screening cell lines, several studies have explored the connection between cell line profiles and single-cell RNA sequencing (scRNA-seq) data to investigate drug response at the single-cell level. For instance, Gambardella et.al.¹² proposed a method called DREEP, which utilizes scRNA-seq data to predict the drug sensitivity of individual cells. They discovered that cells exhibiting transcriptional heterogeneity displayed varying degrees of drug sensitivity. Chen et.al.¹³ developed a deep transfer learning framework, scDEAL, which integrates large-scale bulk cell-line data with scRNA-seq data to predict the response of single cells to cancer drugs. Similarly, Zheng et.al.¹⁴ introduced SCAD, an adversarial discriminative domain adaptation framework that leverages scRNA-seq data from the GDSC database to identify drug sensitivities. However, these methods are not directly applicable to SCST data, as they do not take into account the spatial cell locations. To address this limitation, graph-based domain adaptation models have emerged as promising solutions for uncovering the spatial cellular response to drugs. Such graph-based model aggregates information not only from individual cells but also from their spatial neighbors, enabling a more comprehensive understanding of drug

response within a spatial context.

Herein, we present SpaRx, a graph-based domain adaptation model, to reveal spatial therapeutic complexity with distinct drug responses, through leveraging the high-throughput pharmacological profiles and the SCST datasets. SpaRx is able to identify the cellular drug responses within a complex tissue, the spatial surrounding microenvironment of resistant cells, and the cell-cell communications involved in drug resistance. SpaRx can accurately transfer drug response predictors trained on a source domain (e.g., cell lines) to a target domain (e.g., spatial tumor cells). Our model explicitly considers the fundamental differences between source and target domain rather than modeling these differences as technical batch effects only. SpaRx will facilitate mechanistic studies for overcoming drug resistance and advance therapeutic research for precision medicine. It also holds promises to prioritize candidate drugs for individual patients, provide therapeutic guides for synergistic drug combinations, and repurpose anti-cancer drugs for other diseases such as Alzheimer's Disease.

RESULTS

Overview of SpaRx

SpaRx employs a novel domain adaptation strategy to transfer the knowledge of drug responses from large-scale drug screening databases (source domain) to predict the drug-sensitivity of cells in SCST data (target domain). We hypothesize that there are domain-invariant relations between molecular profiles and drug responses. SpaRx is built to learn such transferable knowledge of drug responses across the source and the target domain through end-to-end adversarial training (**Fig. 1a**). As shown in (**Fig. 1b**), SpaRx consists of a feature extractor to project molecular profiles to a latent space, a drug response predictor to predict cellular sensitivity to a drug according to the latent representation of gene expression patterns, and three domain discriminators to distinguish the source domain from the target domain. A hybrid learning strategy is used, with an adversarial learning procedure (L_t) involving the feature extractor and the domain discriminators to learn domain-invariant molecular features, and a supervised procedure (L_y) involving the feature extractor and the predictor to learn molecular features that are responsible for drug responses. With this hybrid learning strategy, SpaRx can transfer the domain-invariant information from the source knowledgebase to predict the drug responses of individual cells in SCST data.

In the adversarial learning procedure, SpaRx includes dynamic adversarial adaptation learning to balance the learning of global and drug-specific domain-invariant gene expression patterns. To achieve this, three discriminators are used: a global discriminator that distinguishes the source domain from the target domain for all cells or cell lines, and two drug-specific discriminators to distinguish the source domain from the target domain under each category, drug-sensitive and drug-resistant. A dynamic learnable factor (ω) is used to balance the contribution of these discriminators. Trained through the predictor-based loss that captures the knowledge of drug response, as well as the domain adaptation loss regarding global and drug-specific distributions, SpaRx is able to predict cellular sensitivity to drugs in SCST data. In this way, SpaRx successfully translates preclinical knowledge into cell-level drug response in SCST data, which facilitates deep insights into the underlying mechanisms of drug resistance and advances therapeutic effectiveness.

SpaRx demonstrates accurate predictions of drug response

As SpaRx is the first method to predict drug response in SCST data, here we benchmark it against four deep learning (DL) models including SpaRx-GAT, SpaRx-GCN, scDEAL¹³, and SCAD¹⁴, as well as four machine learning (ML) methods including SVM, RF, LightGBM, and XGBoost (see **Materials and Methods**). Among them, SpaRx, SpaRx-GAT, and SpaRx-GCN share similar model architecture but use graph transformer¹⁵, GAT¹⁶, and GCN¹⁷ as the feature extractor. For benchmarking datasets, we randomly select a proportion (p) of cell lines from pharmacological database as the source domain. The remaining cell lines ($1-p$) are used to synthesize the single-cell gene expression data in target domain, with cellular complexity and drug response generated (see **Materials and Methods**). Benchmarking performance is evaluated by the accuracy of predicted drug response in target domain.

First, we randomly select 30% of cell lines ($p = 30\%$) for the source domain data, and the other 70% for

synthesizing the target domain data. The performance of SpaRx and other methods for predicting drug responses in the target domain are measured using the F1 score. SpaRx consistently demonstrates better performance across 80 drugs than the other methods including SpaRx-GAT (**Fig. 2a**), SpaRx-GCN (**Fig. 2b**), RF (**Fig. 2c**), and SVM (**Fig. 2d**), as well as SCAD, scDEAL, LightGBM, and XGBoost (**Supplementary Fig. 1a**). SpaRx achieves the highest accuracy (median F1 = 0.938, **Fig. 2e**), which is significantly higher than other DL models (median F1 of SpaRx-GAT: 0.787; SpaRx-GCN: 0.751; SCAD: 0.856, scDEAL: 0.669) and ML methods (RF: 0.628; SVM: 0.564; LightGBM, 0.576; XGBoost: 0.588). Meanwhile, SpaRx demonstrates particularly higher accuracy relative to SpaRx-GAT and SpaRx-GCN for certain drugs. For example, SpaRx shows noticeably higher F1 scores than SpaRx-GAT (F-1 scores, 0.923 vs 0.683) based on a hormone therapy drug tamoxifen, and higher than SpaRx-GCN (F-1 scores, 0.921 vs 0.445) for the other kinase inhibitor drug alisertib.

Moreover, we evaluate the performance of SpaRx in the settings of different p (50%, 30%, 10%) based on the F1 score. Across these different settings, SpaRx (median F1 = 0.934, **Fig. 2b**, **Supplementary Fig. 1b**) is consistently superior to other DL models (median F1; SpaRx-GAT: 0.784; SpaRx-GCN: 0.765; SCAD: 0.873; scDEAL: 0.669), and ML methods (median F1; RF: 0.638; SVM: 0.581; LightGBM: 0.585, XGBoost: 0.604). For example, for a commonly used liver cancer drug mitoxantrone, SpaRx shows superior performance (median F1 = 0.957) than other DL models (SpaRx-GAT: 0.844; SpaRx-GCN: 0.657; SCAD: 0.678; scDEAL: 0.662). In addition to F1 score, metrics including AUROC, AUPR, precision, and recall (**Supplementary Figs 2-5, Supplementary File 1**), further demonstrate that SpaRx not only presents superior performance with different sizes of source data, but also achieves accurate response predictions for different types of drugs.

SpaRx accurately predicts drug response in different scenarios

We further evaluate the performance of SpaRx in the scenarios of different noise levels, dropout rates, and numbers of genes in the source and the target data (see **Materials and Methods**). Benchmarking methods including four DL models (SpaRx-GAT, SpaRx-GCN, SCAD, scDEAL) and four ML models (RF, SVM, LightGBM, XGBoost).

Fig. 3a shows the F1 scores achieved by different methods for each drug at the noise level of 1. SpaRx achieves higher F1 scores than SpaRx-GAT and SpaRx-GCN (median F1; 0.938, 0.787, 0.751), and also performs significantly better than SCAD and scDEAL (median F1: 0.856, 0.669, **Supplementary Fig. 1b**). Moreover, when the noise level increases (noise level = 1, 1.5, 2), SpaRx maintains accurate predictions with median F1 as 0.938, 0.921, and 0.893, respectively (**Fig. 3b**). In contrast, the other methods, such as SpaRx-GAT and SpaRx-GCN, are affected by the increased noise in source and target data, indicating these methods are more likely to be undermined by data noise. The other metrics including AUROC, AUPR, precision, and recall (**Supplementary Figs 2-5, Supplementary File 1**) supports that SpaRx is robust to data noise in real applications.

In addition, we evaluate the performance of SpaRx with different dropout rates. When the dropout rate is 70%, SpaRx remains more accurate than SpaRx-GAT and SpaRx-GCN (median F1: 0.916, 0.738, 0.751; **Fig. 3c**), as well as SCAD and scDEAL (median F1: 0.846, 0.650; **Supplementary Fig. 1b**). When the dropout rate increases (**Fig. 3d, Supplementary Fig. 1b**), SpaRx is still superior to the other DL methods (median F1; SpaRx: 0.882; SpaRx-GAT: 0.709; SpaRx-GCN: 0.707; SCAD: 0.833; scDEAL: 0.620) and ML models (median F1; RF: 0.675, SVM: 0.709, LightGBM: 0.641, XGBoost: 0.615). Other metrics including AUROC, AUPR, precision, and recall (**Supplementary Figs 2-5, Supplementary File 1**) demonstrate that SpaRx provides accurate predictions at different dropout levels.

Finally, we evaluate the performance of SpaRx with reduced number of genes in source and target data. Based on only 2,000 genes, SpaRx remains superior to SpaRx-GAT and SpaRx-GCN (median F1; 0.960, 0.823, 0.883; **Fig. 3e**) as well as SCAD and scDEAL (median F1: 0.853, 0.669; **Supplementary Fig. 1b**). With the number of genes decreasing to around 1k genes and 500 genes captured by NanoString CosMx and Vizgen MERSCOPE respectively, SpaRx maintains much more reliable performance than the other methods (median F1; SpaRx: 0.930; SpaRx-GAT: 0.753; SpaRx-GCN: 0.866; SCAD: 0.754; scDEAL: 0.595; **Fig. 3f and Supplementary Fig. 1b**). The other metrics including AUROC, AUPR, precision, and recall (**Supplementary Figs 2-5, Supplementary File 1**) shows that SpaRx outperforms benchmarking methods.

Collectively, these results demonstrate that SpaRx achieves superior predictions in different scenarios, even when the target data has extra noises, high dropout rates, and limited number of genes. These evaluations demonstrate the effectiveness of SpaRx in transferring drug-related intrinsic information across different biological domains, which enable to predict cellular drug response in single-cell spatial transcriptomics data.

SpaRx reveals the spatial cellular heterogeneity of drug response in lung cancer

To reveal spatial cell variability in drug response, we first apply SpaRx to the NanoString CosMx lung cancer SCST data with different cell types on eight Filed Of View (FOV)⁵ (**Fig. 4a**). The zoomed-in image on the right shows that tumor cells (colored in light blue) are infiltrated with immune cells such as macrophage (colored in orange) and B cells (colored in green). Based on this tissue slice, we apply SpaRx to predict the tumor cells' response to a typical lung cancer drug, cisplatin, for which the mechanism of action is to cause DNA damage in cancer cells, blocking cell division and leading to apoptotic cell death. As in **Fig. 4b**, SpaRx uncovers tumor cells' response to cisplatin, which exhibits strong heterogeneity of sensitivity and resistance. Interestingly, in contrast to the agminated resistant cells, the zoom-in FOV presents a scattered pattern of sensitive cells.

Given that those tumor cells respond differentially to cisplatin, we further interrogate if the surrounding microenvironment of resistant tumor cells is different from that of sensitive ones. As shown in **Fig. 4c**, for each FOV, the spatial distributions (i.e., proportions) of cell types adjacent to resistant and sensitive cells are different. Such differences also appear to be distinct across different FOVs. Noteworthy, across FOV 1-3 (**Fig. 4c**), CD8 memory T cells, B cells, and natural killer (NK) cells are consistently reduced in the surroundings of resistant cells. Heatmap in **Fig. 4d** further quantitatively delineates the distinctive microenvironment between sensitive and resistant cells, where CD4 and CD8 memory T cells are less infiltrated in the surrounding of resistant cells. Moreover, fewer B cells, dendritic cells (DC), and NK cells are present in the microenvironment of resistant cells. Further averaging of the surrounding cell type proportions across the eight FOVs (**Fig. 4e**) shows that most cell types except macrophages are more abundant in the microenvironment of sensitive cells. Similar patterns are also observed in the TCGA lung cancer patients receiving cisplatin treatment (**Fig. 4f**). Specifically, after bulk RNA-seq decomposition by CIBERSORT¹⁸, CD4 and CD8 memory T cells, B cells, and NK cells are shown to be more prevalent in responders than non-responders to cisplatin. These results indicate that the surrounding microenvironment may be relevant to or modulate the tumor cells' responses to cisplatin.

Spatial cellular crosstalk mediates drug resistance

Given the distinctive microenvironment surrounding sensitive and resistant tumor cells, further characterization of cell-cell communications can explain how neighboring cells modulate the differential responses of tumor cells to cisplatin. Using spaCI¹⁹, a cell-cell communication tool specifically designed for SCST data, we infer the ligand-receptor (L-R) interactions involving tumor cells in the zoom-in FOV (**Fig. 4a**). The aggregated L-R interactions that occur between tumor cells and adjacent cells are shown in the chord diagram (**Fig. 5a**), with the chord width indicating the interaction strength. Of note, we observe that macrophage, fibroblast, and CD4 and CD8 memory T cells interact strongly with both sensitive cells and resistant tumor cells. NK cells uniquely crosstalk with sensitive but not resistant tumor cells.

The involved L-R pairs that contribute to the cellular crosstalk with tumor cells are further presented in **Fig. 5b**. The gradient colors represent the interaction strength of each L-R pair. Specifically, CD4 and CD8 memory T cells are involved in more L-R interactions with resistant than sensitive cells. *MMP9-CD44* is uniquely involved in the interactions between NK and sensitive cells, which is supported by previous study^{20,21}. Moreover, fibroblast expressed *DCN* (ligand) shows stronger interactions with resistant tumor cells' MET (receptor). *DCN* has been reported to interact antagonistically with the *MET* factor (c-Met)^{22,23}, and play roles in cancer development and metastasis²⁴. Other L-R interactions including *HGF²⁵-MET²⁶*, *VCAN²⁷-ITGB1²⁸*, and *VCAN²⁷-CD44²⁹* also play a crucial role in cancer cells forming resistant state against drugs.

SpaRx can be used to explore optimal drug combinations. For example, in addition to cisplatin, SpaRx also identifies the spatially differential cellular response to the other lung cancer drug, docetaxel. As in **Fig. 5c**, some tumor cells that are resistant to one drug appear sensitive to the other drug. Tumor cells sensitive to each of the two drugs are complementary, with Jaccard similarity as 0.381. This result suggests that the combined therapy of

cisplatin and docetaxel may overcome resistance and improve therapeutics, which has also been confirmed in clinical trials for patients with unresectable NSCLC³⁰⁻³².

SpaRx uncovers an orderly pattern of resistant tumor cells

Next we apply SpaRx to the Vizgen MERSCOPE liver cancer SCST data (**Fig. 6a**). In this case, most tumor cells (green colored cells) are confined in three regions with clear boundaries, with some infiltrating tumor cells within the hepatocytes. Specifically, the tumor region on the left (region-1) is surrounded by Kupffer cells, and the other two tumor cell regions (region-2 and region-3) at the right are surrounded by hepatoblasts. SpaRx is applied to predict the tumor cells' response to a typical liver cancer drug, mitoxantrone. As shown in **Fig. 6b**, SpaRx uncovers tumor cells' response to mitoxantrone, with both sensitive and resistant cells revealed. Interestingly, sensitive cells mostly present at the outer area, whereas resistant cells majorly locate in the inner area of each tumor region. Such orderly patterns of resistant cells shared by three tumor regions indicate that those resistant cells may share similar molecular characteristics.

To investigate the underlying differences between the resistant and sensitive cells, differentially expressed gene (DEG) analysis is performed for each tumor region. For example, the DEGs of resistant cells at region-1 are shown in **Fig. 6c**, among them *VTN*³³ and *VEGFA*³⁴ are overexpressed. Vitronectin (encoded by *VTN*) has been reported to protect cancer cells from drug-induced apoptosis³³. *VEGFA* decreases the sensitivity of cancer cells to chemotherapy by suppressing *VEGFA*-mediated autophagy³⁴. These over-expressed genes identified in resistant cells may serve as resistance biomarkers and potential therapeutic targets. More importantly, the DEGs of both resistant and sensitive cells across three tumor regions are largely in common (**Fig. 6d**), further confirming that these regions share similar molecular mechanisms for mitoxantrone resistance. Enrichment analysis of these shared DEGs (**Fig. 6e**) among resistant cells (R1, R2, and R3) reveals signaling pathways that are potentially responsible for mitoxantrone resistance, including the focal adhesion-induced PI3K-AKT signaling. In contrast, interleukin signaling and cytokine signaling pathways are enriched in sensitive cells.

DISCUSSION

The spatial heterogeneity in cells and their microenvironment play critical roles in the treatment of complex diseases such as cancers³⁵ and Alzheimer's diseases (AD)³⁶. For example, tumor microenvironment is crucial for tumor cell metastasis³⁷ and drug resistance³⁸. Recent emerging single-cell spatial technologies utilizing molecular imaging for targeted gene profiling, provide deep insights into the spatial cellular ecosystems³⁹⁻⁴¹. These state-of-art technologies help resolve the cellular heterogeneous response to drugs, the intercellular communications that contribute to drug resistance, and how tumor ecosystem acquires drug resistance.

In this work, we have developed a novel SpaRx model that leverages the pharmacogenomics knowledgebase with SCST data to systematically reveal spatial complexity of therapeutic response. As to our knowledge, SpaRx is the first method to incorporate the large-scale pharmacogenomics profiles with SCST data, to accurately predict the heterogeneous cellular response to drugs. SpaRx is able to reveal the spatial cell variability in drug response and uncover the underlying biological mechanisms for drug resistance. For example, based on the lung cancer SCST data, we observe the multitude interactions related to tumor cell resistance and identify the spatially adjacent cell interactions that may alter tumor cells' sensitivity to cisplatin. In addition to cancers, SpaRx also holds the promises for repurposing anti-cancer drugs for complex diseases such as AD, which is also known for its complexity and heterogeneity. Collectively, SpaRx is anticipated to reveal the mechanisms of drug resistance, prioritize tailored drugs for complex diseases, and provide clues for drug repositioning.

Given the advantages of SpaRx, there are several aspects that SpaRx can be improved. First, current single-cell spatial technologies are still not able to detect sufficient number of genes, which may limit the potentials of SpaRx in some degree. Future advances in spatial technologies that captures more genes and less dropouts will help enhance the SpaRx model. Second, with the rapid development of single-cell spatial omics technologies⁴², SpaRx can also be improved through incorporating spatial multi-omics. Though the current version of SpaRx that enables

the predictions of drug responses based on SCST data, SpaRx can be improved by utilizing new data types, e.g., single-cell spatial ATAC-seq profiles⁴³, thus to further unveil the underlying mechanisms such as the upstream cis-regulatory elements and associated transcription factors involved in drug resistance.

MATERIALS AND METHODS

Data sources and preparation

1) The GDSC¹⁰ and CCLE⁹ cell-line-based drug screening database. The gene expressions of cell lines and the sensitivity profile (IC50) of drugs are downloaded from the GDSC and CCLE database. The binary drug responses for each cell line are obtained from previous studies^{11,13}. Here for GDSC and CCLE database, cell lines and drug information without missing values are retained and integrated based on overlapped drug compounds. Collectively, we obtained 1,280 cell lines with drug response information across 80 shared drug compounds.

2) The NanoString CosMx lung-13 SCST data⁵ and the Vizgen MERSCOPE liver cancer-1 SCST data⁶.

Benchmarking data

The benchmarking datasets are generated from the collected 1,280 cell lines. Here we randomly select a proportion of cell lines (p) as the source domain, whereas the remaining cell lines ($1-p$) are used to synthesize the target domain. For the target domain, the gene expression profiles of the remaining cell lines ($1-p$) are further mixed randomly to mimic tumor cell complexity. These mixed data is then downsampled to assure the total counts are comparable to the single-cell level, which allows the synthesized gene expression data in the target domain mimic that of single tumor cells in SCST data. Specifically, to mimic the complexity of tumor cells, we select two to ten cell lines from the remaining cell lines ($1-p$), then combine their transcriptomic profiles as one tumor cell profile. In order to better mimic real tumor cell, if the total counts of the resulting tumor data exceed 2,000, we downsample it accordingly. In this way, the synthesized gene expression data serving as the target domain is more likely to resemble the real heterogenous tumor cells in SCST data.

Moreover, four benchmarking scenarios are included. 1) Different numbers of cell lines in the source data. We choose different proportions of cell lines as the source domain, i.e., $p = 10\%$, 30% , 50% , respectively. The remaining cell lines ($1 - p$) are used to generate the target domain. 2) Different levels of noises. For both source and target domain based on the setting of $p = 30\%$, we add extra noises randomly sampled from normal distributions $N(0, \sigma^2)$, with standard deviation σ as 1, 1.5, and 2, respectively. 3) Different levels of dropouts. For both source and target domain based on the setting of $p = 30\%$, dropouts are simulated by replacing the gene expression values with zeros, to ensure the proportions of zeros among all gene expression values are 70%, 80%, and 90%, respectively. 4) Different numbers of genes. Based on the setting of $p = 30\%$, the number of genes in the cell line profiles is reduced to 2k, 853, and 500 genes. The 2k genes are randomly selected, while the 853 and 500 genes are selected based on the RNA panels used in the NanoString CosMx and the Vizgen MERSCOPE data, respectively.

SpaRx model

Source domain. The GDSC and the CCLE data are used as the source domain, denoted as $D_s = \{(x_i^s, y_i^s)\}_{i=1}^{n_s}$. Here s represents the source domain (the cell-line based drug response profiles), $x_i^s \in R^{d_s \times n_s}$ represents the gene expression, with i representing a cell line, d_s representing the number of genes, and n_s denoting the number of cell lines. The cell-line similarity graph A_s is constructed using mutual nearest neighbors (MNN)⁴⁴, with the number of mutual nearest neighbor as k . In the graph A_s , each node v_i represents a cell line i , and if two nodes v_i and v_j are connected, it means that the corresponding gene expression profiles x_i^s and x_j^s are similar.

Target domain. The SCST data is used as the target domain. Each of the SCST data is represented by $D_t = \{(x_i^t)\}_{i=1}^{n_t}$, $x_i \in R^{d_t \times n_t}$, where t denotes the target domain (the SCST data), d_t denotes the number of genes, n_t represents the number of cells, and i represents a cell in the SCST data. A spatial cell graph A_t is constructed according to cell locations using k -nearest neighbors. If two cells i and j are spatially adjacent, then the corresponding nodes v_i and v_j are connected in A_t .

The SpaRx model uses cell lines in the source domain and cells in the target domain as samples, gene expressions as features, and drug responses as outcomes. SpaRx is composed of three components: 1) feature extractor to extract gene expression features from the source and the target domain, 2) drug response predictor for both cell lines and single cells, and 3) global and drug-specific discriminators. The final output of the SpaRx is the predicted drug responses of each cell in the target SCST domain.

1) Feature Extractor: The shared feature extractor T_r is composed of multi-head graph transformer¹⁵ layers to project the graph representation of the cellular transcriptomics data to a latent space in which cells that demonstrate similar responses to treatments are close to each other. Briefly, for a cell line i from the GDSC or the CCLE data or a cell i from the SCST data, the propagation of the graph transformer from the l layer to the $l + 1$ layer is defined as:

$$\mathbf{h}_i^{(l+1)} = \text{ReLU}\left(W_1^{(l)}\mathbf{h}_i^{(l)} + \sum_{v_j \in \mathcal{N}(v_i)} \alpha_{i,j} \mathbf{V}_j^{(l)}\right) \quad [1]$$

, where i represents either a cell line (x_i^s) in the source domain or a cell (x_i^t) in the target domain, j represents a neighbor cell line or cell in their corresponding graph (i.e., $v_j \in \mathcal{N}(v_i)$), the rectified linear unit (ReLU)⁴⁵ is used as the nonlinear gated activation function. When $l = 0$, $\mathbf{h}_i^{(l)} = x_i^s$ for the cell line data and $\mathbf{h}_i^{(l)} = x_i^t$ for the SCST data. The attention module is defined as: $\alpha_{i,j} = \text{softmax}\left(\frac{\langle Q_i^{(l)}, K_j^{(l)} \rangle}{\sum_{v_j \in \mathcal{N}(v_i)} \langle Q_i^{(l)}, K_j^{(l)} \rangle}\right)$, where:

$$\text{query: } Q_i^{(l)} = W_Q^{(l)}\mathbf{h}_i^{(l)} + b_Q^{(l)}$$

$$\text{key: } K_j^{(l)} = W_K^{(l)}\mathbf{h}_j^{(l)} + b_K^{(l)}$$

$$\text{value: } V_j^{(l)} = W_V^{(l)}\mathbf{h}_j^{(l)} + b_V^{(l)}$$

and $\langle Q, K \rangle \equiv \exp\left(Q^T K / \sqrt{\dim(\mathbf{h}_i^{(l)})}\right)$. The multi-head attentions are concatenated. In this way, we obtain the latent representation, $z_s = \{\mathbf{h}_i^l\}_{i=1}^{n_s} = T_r(x_i^s, A_s)$, as the extracted features for source domain, and $z_t = \{\mathbf{h}_i^l\}_{i=1}^{n_t} = T_r(x_i^t, A_t)$ as the extract features for target domain, respectively. The feature extractor T_r is shared by both source and target domain.

2) Drug response predictor: The predictor (F_y), a fully connected classifier, is designed to classify the drug response results using latent features from the feature extractor T_r . It is trained by minimizing the differences between the predicted source labels and the source domain labels of drug response (ground truth labels) by the cross-entropy loss, which is formulated as:

$$L_y = -\frac{1}{n_s} \sum_{x_i \in D_s} \left(P_{x_i \in S} \log F_y(T_r(x_i, A)) + P_{x_i \in R} \log F_y(T_r(x_i, A)) \right) \quad [2]$$

, where $P_{x_i \in S}$ is the probability of x_i belonging to drug sensitive, $P_{x_i \in R}$ is the probability of x_i belonging to drug resistance. F_y is the response predictor and T_r is the feature extractor. A represents the graph A_s in source domain. Here the drug response predictor is shared by both source and target domain.

3) Discriminators. A global discriminator G_d is trained to align the latent representations of source and target domain. Here the loss of the global discriminator is formulated as:

$$L_g = \frac{1}{n_s + n_t} \sum_{x_i \in D_s \cup D_t} L_g^{x_i} = \frac{1}{n_s + n_t} \sum_{x_i \in D_s \cup D_t} L_d(G_d(T_r(x_i, A)), d_i) \quad [3]$$

, where L_d denotes the cross-entropy, G_d denotes the global discriminator, T_r is the feature extractor, and d_i is the domain label for the input x_i ($d_i = 0$ for source domain, $d_i = 1$ for target domain). A represents the graph A_s when $x_i \in D_s$ and A_t when $x_i \in D_t$.

Drug-specific discriminators (g_d^S and g_d^R): g_d^S and g_d^R are used to match the latent representations from source and target domains under drug-sensitive and drug-resistant category, respectively. Both drug-specific

discriminators are trained to minimize the differences in the latent representations of source and target domain under each drug category. The output of the drug response predictor ($F_y(T_r(x_i))$) is used to show the probability of being included into each drug category. The loss for each discriminator is calculated using cross-entropy:

$$L_c = \frac{1}{n_s+n_t} \sum_{x_i \in D_s \cup D_t} L_c^{S,x_i} + \frac{1}{n_s+n_t} \sum_{x_i \in D_s \cup D_t} L_c^{R,x_i}$$

$$= \frac{1}{n_s+n_t} \sum_{x_i \in D_s \cup D_t} L_d^S \left(g_d^S \left(\hat{y}_i^S T_r(x_i, A) \right), d_i \right) + L_d^R \left(g_d^R \left(\hat{y}_i^R T_r(x_i, A) \right), d_i \right) \quad [4]$$

, where g_d^S , L_d^S and g_d^R , L_d^R are the drug-specific discriminator loss and its cross-entropy loss associated with drug categories, respectively. \hat{y}_i^S and \hat{y}_i^R is the predicted probability of the input x_i belonging to drug-sensitive or drug-resistant category, i.e., $F_y(T_r(x_i))$. d_i is the label for the input x_i ($d_i = 0$ for source domain, $d_i = 1$ for target domain). A represents the graph A_s when $x_i \in D_s$ and A_t when $x_i \in D_t$.

Loss function. Given the three major components, feature extractors, domain discriminators, and label classifier in our model, the final learning objective is formulated as:

$$L_T = L_y - L_t, \text{ where } L_t = (1 - \omega)L_g + \omega L_c \quad [5]$$

In the loss function, ω is a dynamic adversarial factor which balances the relative weight of the global and the drug-specific discriminator loss. During the training, ω is dynamically updated according to the losses of the three discriminators: $\omega = \frac{d_g}{d_g + \frac{1}{2}(d_c^S + d_c^R)}$, where d_g , d_c^S , and d_c^R are the proxy \mathcal{A} -distances⁴⁶ between the source and the target domains for the three domain discriminators, respectively. Specifically, for the discriminator G_d , $d_g(D_s, D_t) = 2 \left(1 - 2 * \frac{1}{n_s+n_t} \sum_{x_i \in D_s \cup D_t} L_g^{x_i} \right)$; for the discriminator g_d^S , $d_c^S(D_s, D_t) = 2 \left(1 - 2 * \frac{1}{n_s+n_t} \sum_{x_i \in D_s \cup D_t} L_c^{S,x_i} \right)$; and for the discriminator g_d^R , $d_c^R(D_s, D_t) = 2 \left(1 - 2 * \frac{1}{n_s+n_t} \sum_{x_i \in D_s \cup D_t} L_c^{R,x_i} \right)$, where n_s represents the number of cell lines in source domain, and n_t is the number of cells in target domain.

The SpaRx model is trained using the stochastic gradient descent (SGD) optimizer. In the SpaRx model, the parameters including the number of adjacent neighbors or mutual nearest neighbors in graph construction, as well as the latent dimensions in graph transformer layers, are determined through grid-based hyper-parameter fine tuning. The hyperparameters used in the final model are: for the SGD optimizer, momentum=0.9 and weight decay= $5e^{-5}$; the learning rate is set to $1e^{-3}$; gradient clip threshold at 5; the number of graph transformer layers is 2, with the dimensions of 512 and 64, respectively. After the model training, SpaRx accurately predicts the drug response labels of cells in the spatial data and uncovers the spatially heterogeneous responses to different types of drugs.

Benchmarking methods and comparison measurement

To evaluate the performance of SpaRx, we compare it with four deep learning models, including SpaRx-GAT, SpaRx-GCN, scDEAL¹³, SCAD¹⁴, and four machine learning methods including random forest (RF), support vector machines (SVM), lightGBM and XGBoost. SpaRx-GAT is built based on the SpaRx model, with the feature extractor as GAT¹⁶ layers, rather than the graph transformer¹⁵. SpaRx-GCN uses GCN¹⁷ layers as the feature extractor. scDEAL¹³ and SCAD¹⁴ are proposed to predict single cell response to cancer drugs by integrating large-scale bulk cell-line data and scRNA-seq data. To evaluate the performance of each model, we use the F1 score to assess the agreement between the predicted drug response and the ground truth. The F1 ranges from 0 to 1 referring to the increasing match between the predicted drug response with ground truth. With TP denoting true positive, FP representing false positive, and FN representing false negative, F1 score is calculated by $F1 = \frac{2*TP}{2*TP+(FP+FN)}$. Additional metrics including precision, recall, AUROC, and AUPR are included for comprehensive evaluation of SpaRx and benchmarking methods.

Identifying surrounding microenvironment, L-R interactions, and adjacent cell communications

When characterizing the surrounding microenvironment, we have summed up the adjacent (with 5 nearest neighbors) cells (by each cell type) around resistant/sensitive cells. After dividing the total number of cells within

each cell type, we obtain the percentage of different cell types in the microenvironment of resistant/sensitive cells. To identify L-R interactions, our previous tool spaCI¹⁹ is used here for SCST data. With the L-R interactions, we further characterize the adjacent cell communications with interaction strength. Specifically, for an L-R interaction pair, we define its interaction strength as the multiplication of their average expression values among adjacent cells, where the top and bottom 10% expressions of the ligand and the receptor are ignored. The interaction strength of all identified L-R pairs is then summarized as the interaction strength between two cell types. Thus, the higher value of the interaction strength, the stronger the two cell types adjacently interact.

DATA AVAILABILITY

NanoString CosMx SMI data: The single-cell spatial dataset (Lung-13), profiled by CosMx SMI on Formalin-Fixed Paraffin-Embedded (FFPE) samples of the non-small-cell lung cancer (NSCLC) tissue⁵, is available from <https://nanosttring.com/products/cosmx-spatial-molecular-imager/ffpe-dataset/>. **Vizgen MERSCOPE data:** We includes the Vizgen MERFISH liver cancer 1 dataset that contains a MERFISH measurement of a 500 gene panel. Data is downloaded from <https://info.vizgen.com/merscope-ffpe-solution>, which includes the list of detected transcripts, gene counts per cell matrix, and additional spatial cell metadata. The gene expression profiles of GDSC and CCLE cell lines are downloaded from <https://www.cancerrxgene.org/> and <https://depmap.org/portal/>. The gene expression profile data of TCGA lung cancer patients, including lung adenocarcinoma and lung squamous cell carcinoma patients, are downloaded from the UCSC Xena database (<http://xena.ucsc.edu/>). The corresponding response information to cisplatin is retrieved from previous studies⁴⁷, where responders (including complete response and partial response) and non-responders (including stable disease and progressive disease), are characterized according to the RECIST standard⁴⁸.

CODE AVAILABILITY

SpaRx is provided as a Python package available at <https://github.com/QSong-github/SpaRx>, with detailed functions for the general applicability on different SCST data.

COMPETING INTERESTS

The authors declare no competing interests.

FUNDING

QS is supported in part by the Bioinformatics Shared Resources under the NCI Cancer Center Support Grant to the Comprehensive Cancer Center of Wake Forest University Health Sciences (P30CA012197). QS is also supported by the American Cancer Society Institutional Research Grant Pilot. JS is partially financially supported by the Indiana University Precision Health Initiative and the Indiana University Melvin and Bren Simon Comprehensive Cancer Center Support Grant from the National Cancer Institute (P30CA082709). JS is also supported by R01LM013771.

MATERIALS & CORRESPONDENCE

Correspondence and requests for materials should be addressed to QS or JS.

FIGURE LEGENDS

Fig. 1: Schematic overview of the SpaRx method. **a**, Concept design of the SpaRx model. With the pharmacogenomic profiles serving as the source data, SpaRx enables to predict the cells' response to drugs in single-cell spatial transcriptomics data via graph-based adversarial adaptation. **b**, SpaRx comprises three components (feature extractors, drug-response predictor, and discriminators) to leverage source domain and predict cellular response in target domain.

Fig. 2: SpaRx accurately identifies drug response in target domain. **a**, Performance of SpaRx and other methods (SpaRx-GAT, SpaRx-GCN, RF, and SVM) are measured by the F1 scores across 80 different drug

compounds. Each point represents the F1 score of SpaRx versus an alternative method on one type of drug. **b**, Boxplot of F1 scores based on different sizes of source data.

Fig. 3: Performance evaluation in different benchmarking scenarios. **a**, Accuracy of identifying ground truth labels for the scenario with extra noise (noise level = 1) in source and target domain. Dashed lines refer to 15% percentile of SpaRx's F1 scores. **b**, Boxplot of F1 scores over three scenarios with extra noise (noise level = 1, 1.5, 2) in source and target data. **c**, Accuracy of identifying ground truth labels for the scenario with extensive dropouts (dropout rate = 70%) in source and target domain. **d**, Boxplot of F1 scores over three scenarios with different dropout levels (dropout rate = 70%, 80%, 90%) in source and target domain. **e**, Accuracy of identifying ground truth labels for the scenario with limited number of genes (number of genes = 2k) in source and target domain. **f**, Boxplot of F1 scores over three scenarios with limited number of genes (number of genes = 2k, 853, 500) in source and target domain.

Fig. 4: Spatial heterogeneity of cellular response to cisplatin. **a**, Spatial visualization of single-cell spatial data from lung cancer tissue. Eight Field Of Views (FOVs) are included. Different colors denote different types of cells. **b**, Spatial visualization of cellular response to cisplatin. Red and Cyan colors denote resistant (R) and sensitive (S) tumor cells to cisplatin. **c**, Surrounding microenvironment of resistant (R) and sensitive (S) cells in each FOV. **d**, Heatmap quantitatively delineates the constituents around sensitive and resistant cells. **e**, Summarization across eight FOVs for the cell type distributions surrounding resistant and sensitive tumor cells to cisplatin. **f**, Characterization of the infiltrated cell types in responders and non-responders to cisplatin in TCGA lung cancer patient samples.

Fig. 5: Roles of cell-cell communications in tumor cell resistance. **a**, Summary chord diagram of the identified cell-cell communication network. The chord width is proportional to the interaction strength across different cell types. **b**, Heatmap shows the identified L-R interactions between major cell types and resistant (R) or sensitive (S) cells. **c**, Spatial visualization of cellular response to cisplatin, docetaxel, and their combinations.

Fig 6: SpaRx reveals drug resistance pattern in liver tumor tissue. **a**, Spatial visualization of the single-cell spatial data from liver tumor tissue. **b**, Spatial visualization of cellular response to mitoxantrone. Red and blue colors denote resistant and sensitive tumor cells to mitoxantrone. **c**, Volcano plot shows the differentially expressed genes (DEGs) of resistant versus sensitive tumor cells in tumor region-1. **d**, Overlaps of DEGs across three tumor regions. **e**, Enriched pathways based on overlapped DEGs across three tumor regions.

Supplementary Fig. 1: Performance of SpaRx and additional benchmarking methods. **a**, Performance of SpaRx and other methods (SCAD, scDEAL, LightGBM, and XGBoost) are measured by the F1 scores across 80 different drug compounds. Each point represents the F1 score of SpaRx versus an alternative method on one type of drug. **b**, Boxplot of F1 scores in different scenarios, including different source data sizes, noise levels, dropout levels, and number of genes.

Supplementary Fig. 2: AUROC scores of benchmarking methods in different scenarios.

Supplementary Fig. 3: AUPR scores of benchmarking methods in different scenarios.

Supplementary Fig. 4: Precision values of benchmarking methods in different scenarios.

Supplementary Fig. 5: Recall values of benchmarking methods in different scenarios.

Supplementary File 1: Comprehensive evaluation of SpaRx with benchmarking methods.

REFERENCES

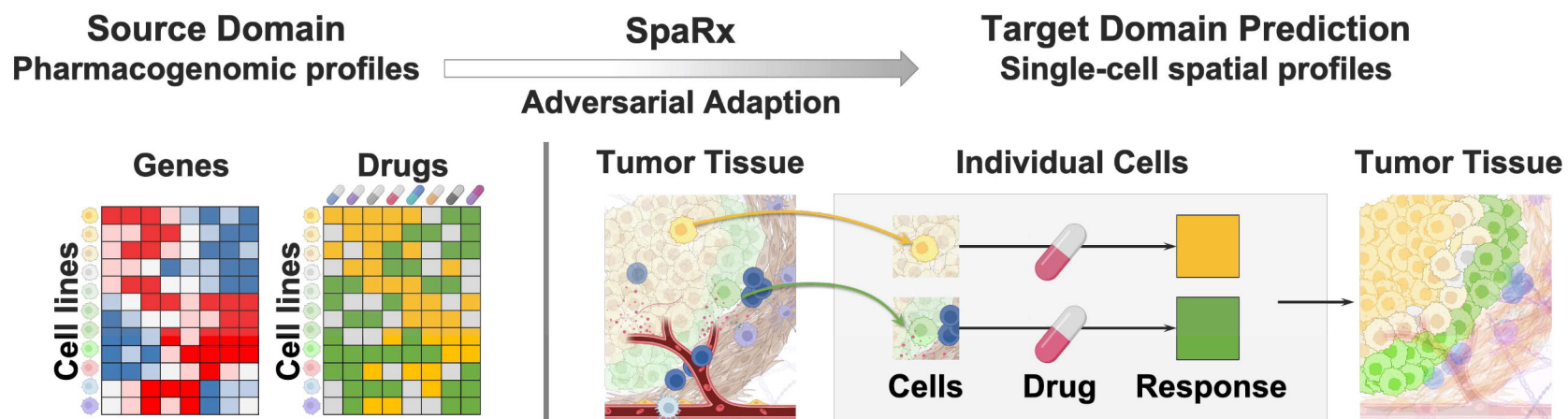
- 1 Nirmal, A. J. *et al.* The Spatial Landscape of Progression and Immunoediting in Primary Melanoma at Single-Cell Resolution. *Cancer Discovery* **12**, 1518-1541 (2022). <https://doi.org/10.1158/2159-8290.Cd-21-1357>
- 2 Luca, B. A. *et al.* Atlas of clinically distinct cell states and ecosystems across human solid tumors. *Cell* **184**, 5482-5496.e5428 (2021). [https://doi.org:https://doi.org/10.1016/j.cell.2021.09.014](https://doi.org/https://doi.org/10.1016/j.cell.2021.09.014)
- 3 El-Sayes, N., Vito, A. & Mossman, K. Tumor Heterogeneity: A Great Barrier in the Age of Cancer Immunotherapy. *Cancers* **13**, 806 (2021).
- 4 Kemper, K. *et al.* Intra- and inter-tumor heterogeneity in a vemurafenib-resistant melanoma patient and derived xenografts. *EMBO Mol Med* **7**, 1104-1118 (2015). <https://doi.org/10.15252/emmm.201404914>
- 5 He, S. *et al.* High-plex imaging of RNA and proteins at subcellular resolution in fixed tissue by spatial molecular imaging. *Nat Biotechnol* **40**, 1794-1806 (2022). <https://doi.org/10.1038/s41587-022-01483-z>
- 6 Fang, R. *et al.* Conservation and divergence of cortical cell organization in human and mouse revealed by MERFISH. *Science* **377**, 56-62 (2022). <https://doi.org/10.1126/science.abm1741>
- 7 Zhang, Q. *et al.* The spatial transcriptomic landscape of non-small cell lung cancer brain metastasis. *Nature Communications* **13**, 5983 (2022). <https://doi.org/10.1038/s41467-022-33365-y>
- 8 Dagogo-Jack, I. & Shaw, A. T. Tumour heterogeneity and resistance to cancer therapies. *Nat Rev Clin Oncol* **15**, 81-94 (2018). <https://doi.org/10.1038/nrclinonc.2017.166>
- 9 Barretina, J. *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603-607 (2012). <https://doi.org/10.1038/nature11003>
- 10 Yang, W. *et al.* Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic acids research* **41**, D955-D961 (2012).
- 11 Iorio, F. *et al.* A Landscape of Pharmacogenomic Interactions in Cancer. *Cell* **166**, 740-754 (2016). <https://doi.org/10.1016/j.cell.2016.06.017>
- 12 Gambardella, G. *et al.* A single-cell analysis of breast cancer cell lines to study tumour heterogeneity and drug response. *Nature Communications* **13**, 1714 (2022). <https://doi.org/10.1038/s41467-022-29358-6>
- 13 Chen, J. *et al.* Deep transfer learning of cancer drug responses by integrating bulk and single-cell RNA-seq data. *Nature Communications* **13**, 6494 (2022). <https://doi.org/10.1038/s41467-022-34277-7>
- 14 Zheng, Z. *et al.* Enabling Single-Cell Drug Response Annotations from Bulk RNA-Seq Using SCAD. *Advanced Science* **10**, 2204113 (2023). [https://doi.org:https://doi.org/10.1002/advs.202204113](https://doi.org/https://doi.org/10.1002/advs.202204113)
- 15 Vaswani, A. *et al.* Attention is all you need. *Advances in neural information processing systems* **30** (2017).
- 16 Veličković, P. *et al.* Graph attention networks. *arXiv preprint arXiv:1710.10903* (2017).
- 17 Kipf, T. N. & Welling, M. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907* (2016).
- 18 Newman, A. M. *et al.* Robust enumeration of cell subsets from tissue expression profiles. *Nature Methods* **12**, 453-457 (2015). <https://doi.org/10.1038/nmeth.3337>
- 19 Tang, Z., Zhang, T., Yang, B., Su, J. & Song, Q. spaCI: deciphering spatial cellular communications through adaptive graph model. *Briefings in Bioinformatics* **24** (2022). <https://doi.org/10.1093/bib/bbac563>

- 20 Yosef, G., Hayun, H. & Papo, N. Simultaneous targeting of CD44 and MMP9 catalytic and hemopexin domains as a therapeutic strategy. *Biochemical Journal* **478**, 1139-1157 (2021). <https://doi.org/10.1042/bcj20200628>
- 21 Albini, A. & Noonan, D. M. Decidual-Like NK Cell Polarization: From Cancer Killing to Cancer Nurturing. *Cancer Discov* **11**, 28-33 (2021). <https://doi.org/10.1158/2159-8290.CD-20-0796>
- 22 Neill, T. *et al.* Decorin antagonizes the angiogenic network: concurrent inhibition of Met, hypoxia inducible factor 1alpha, vascular endothelial growth factor A, and induction of thrombospondin-1 and TIMP3. *The Journal of biological chemistry* **287**, 5492-5506 (2012). <https://doi.org/10.1074/jbc.M111.283499>
- 23 Gubbiotti, M. A., Vallet, S. D., Ricard-Blum, S. & Iozzo, R. V. Decorin interacting network: A comprehensive analysis of decorin-binding partners and their versatile functions. *Matrix Biol* **55**, 7-21 (2016). <https://doi.org/10.1016/j.matbio.2016.09.009>
- 24 Sofeu Feugaing, D. D., Götte, M. & Viola, M. More than matrix: the multifaceted role of decorin in cancer. *Eur J Cell Biol* **92**, 1-11 (2013). <https://doi.org/10.1016/j.ejcb.2012.08.004>
- 25 Grugan, K. D. *et al.* Fibroblast-secreted hepatocyte growth factor plays a functional role in esophageal squamous cell carcinoma invasion. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 11026-11031 (2010). <https://doi.org/10.1073/pnas.0914295107>
- 26 Dong, Y., Xu, J., Sun, B., Wang, J. & Wang, Z. MET-Targeted Therapies and Clinical Outcomes: A Systematic Literature Review. *Molecular Diagnosis & Therapy* **26**, 203-227 (2022). <https://doi.org/10.1007/s40291-021-00568-w>
- 27 Chang, M. Y. *et al.* Versican is produced by Trif- and type I interferon-dependent signaling in macrophages and contributes to fine control of innate immunity in lungs. *Am J Physiol Lung Cell Mol Physiol* **313**, L1069-L1086 (2017). <https://doi.org/10.1152/ajplung.00353.2017>
- 28 Wantoch von Rekowski, K. *et al.* The Impact of Integrin-Mediated Matrix Adhesion on Cisplatin Resistance of W1 Ovarian Cancer Cells. *Biomolecules* **9** (2019). <https://doi.org/10.3390/biom9120788>
- 29 Yang, G. H. *et al.* Osteopontin combined with CD44, a novel prognostic biomarker for patients with hepatocellular carcinoma undergoing curative resection. *Oncologist* **13**, 1155-1165 (2008). <https://doi.org/10.1634/theoncologist.2008-0081>
- 30 Schiller, J. H. *et al.* Comparison of four chemotherapy regimens for advanced non-small-cell lung cancer. *N Engl J Med* **346**, 92-98 (2002). <https://doi.org/10.1056/NEJMoa011954>
- 31 Kaya, A. O. *et al.* Concomitant chemoradiotherapy with cisplatin and docetaxel followed by surgery and consolidation chemotherapy in patients with unresectable locally advanced non-small cell lung cancer. *Med Oncol* **27**, 152-157 (2010). <https://doi.org/10.1007/s12032-009-9186-z>
- 32 Katayama, H. *et al.* Preoperative concurrent chemoradiotherapy with cisplatin and docetaxel in patients with locally advanced non-small-cell lung cancer. *Br J Cancer* **90**, 979-984 (2004). <https://doi.org/10.1038/sj.bjc.6601624>
- 33 Ding, Y. & Shen, Y. Notch increased vitronectin adhesion protects myeloma cells from drug induced apoptosis. *Biochemical and Biophysical Research Communications* **467**, 717-722 (2015). [https://doi.org:https://doi.org/10.1016/j.bbrc.2015.10.076](https://doi.org/https://doi.org/10.1016/j.bbrc.2015.10.076)
- 34 Li, X. *et al.* Inhibition of VEGFA Increases the Sensitivity of Ovarian Cancer Cells to Chemotherapy by Suppressing VEGFA-Mediated Autophagy. *Onco Targets Ther* **13**, 8161-8171 (2020). <https://doi.org/10.2147/ott.S250392>

- 35 Baghban, R. *et al.* Tumor microenvironment complexity and therapeutic implications at a glance. *Cell Communication and Signaling* **18**, 59 (2020). <https://doi.org/10.1186/s12964-020-0530-4>
- 36 Guo, T. *et al.* Molecular and cellular mechanisms underlying the pathogenesis of Alzheimer's disease. *Molecular Neurodegeneration* **15**, 40 (2020). <https://doi.org/10.1186/s13024-020-00391-7>
- 37 Quail, D. F. & Joyce, J. A. Microenvironmental regulation of tumor progression and metastasis. *Nat Med* **19**, 1423-1437 (2013). <https://doi.org/10.1038/nm.3394>
- 38 Maacha, S. *et al.* Extracellular vesicles-mediated intercellular communication: roles in the tumor microenvironment and anti-cancer drug resistance. *Molecular Cancer* **18**, 55 (2019). <https://doi.org/10.1186/s12943-019-0965-7>
- 39 Moses, L. & Pachter, L. Museum of spatial transcriptomics. *Nature Methods* **19**, 534-546 (2022). <https://doi.org/10.1038/s41592-022-01409-2>
- 40 Lewis, S. M. *et al.* Spatial omics and multiplexed imaging to explore cancer biology. *Nature Methods* **18**, 997-1012 (2021). <https://doi.org/10.1038/s41592-021-01203-6>
- 41 Moffitt, J. R., Lundberg, E. & Heyn, H. The emerging landscape of spatial profiling technologies. *Nature Reviews Genetics* (2022). <https://doi.org/10.1038/s41576-022-00515-3>
- 42 Vickovic, S. *et al.* SM-Omics is an automated platform for high-throughput spatial multi-omics. *Nature Communications* **13**, 795 (2022). <https://doi.org/10.1038/s41467-022-28445-y>
- 43 Deng, Y. *et al.* Spatial profiling of chromatin accessibility in mouse and human tissues. *Nature* **609**, 375-383 (2022). <https://doi.org/10.1038/s41586-022-05094-1>
- 44 Haghverdi, L., Lun, A. T., Morgan, M. D. & Marioni, J. C. Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nat Biotechnol* **36**, 421-427 (2018).
- 45 Agarap, A. F. Deep learning using rectified linear units (relu). *arXiv preprint arXiv:1803.08375* (2018).
- 46 Ben-David, S., Blitzer, J., Crammer, K. & Pereira, F. Analysis of representations for domain adaptation. *Advances in neural information processing systems* **19** (2006).
- 47 Ding, Z., Zu, S. & Gu, J. Evaluating the molecule-based prediction of clinical drug responses in cancer. *Bioinformatics* **32**, 2891-2895 (2016). <https://doi.org/10.1093/bioinformatics/btw344>
- 48 Eisenhauer, E. A. *et al.* New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* **45**, 228-247 (2009). <https://doi.org/10.1016/j.ejca.2008.10.026>

Fig. 1

a



b

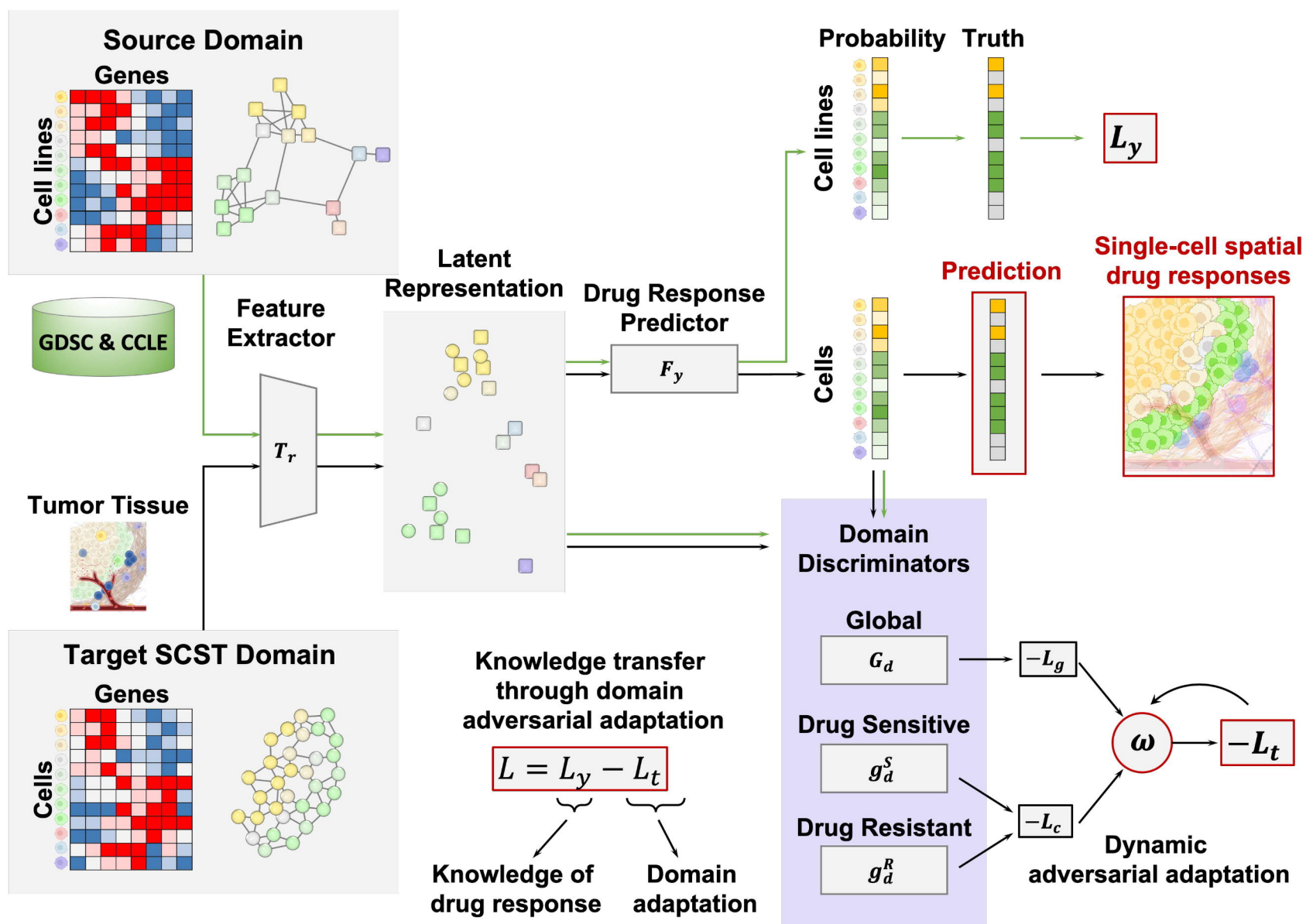


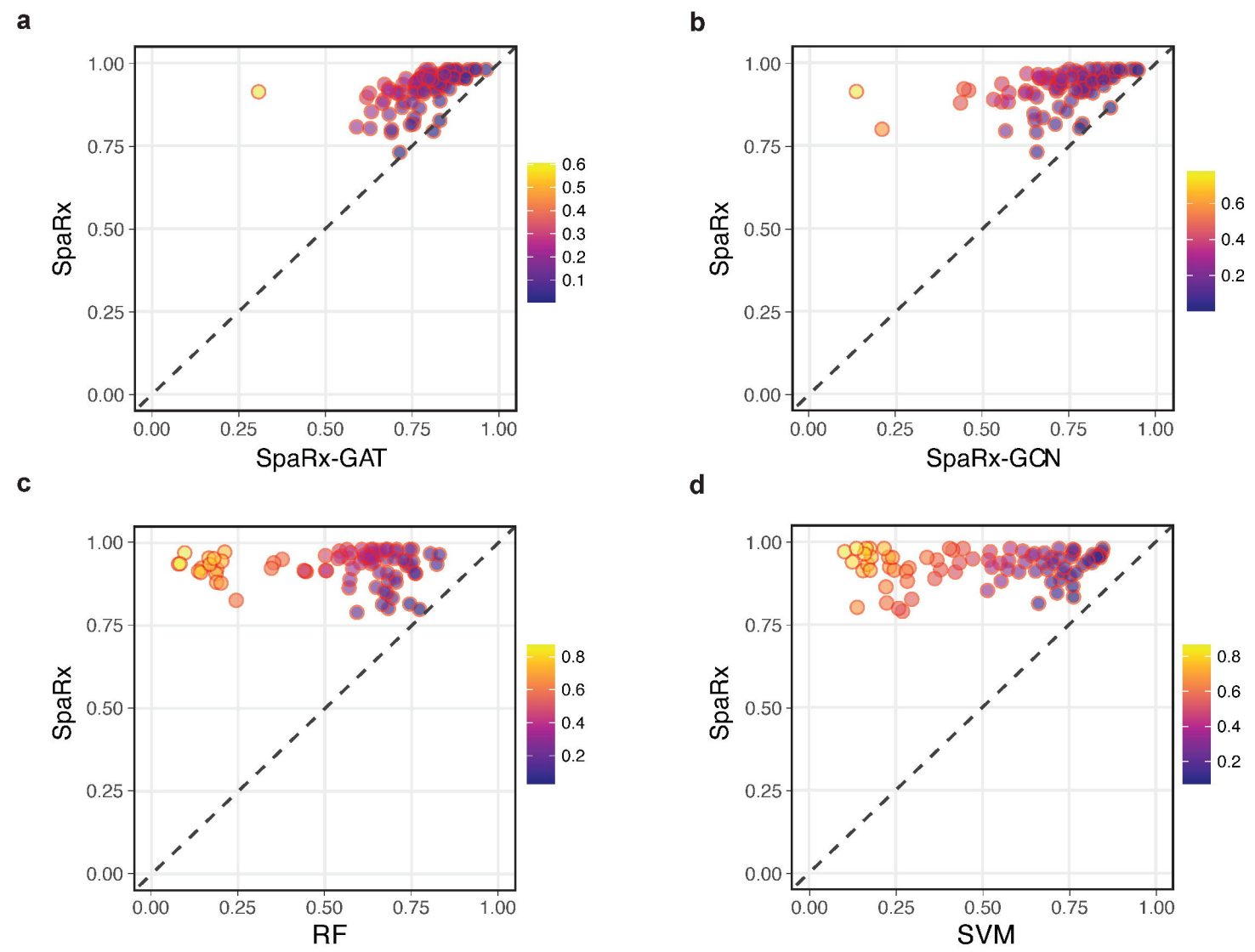
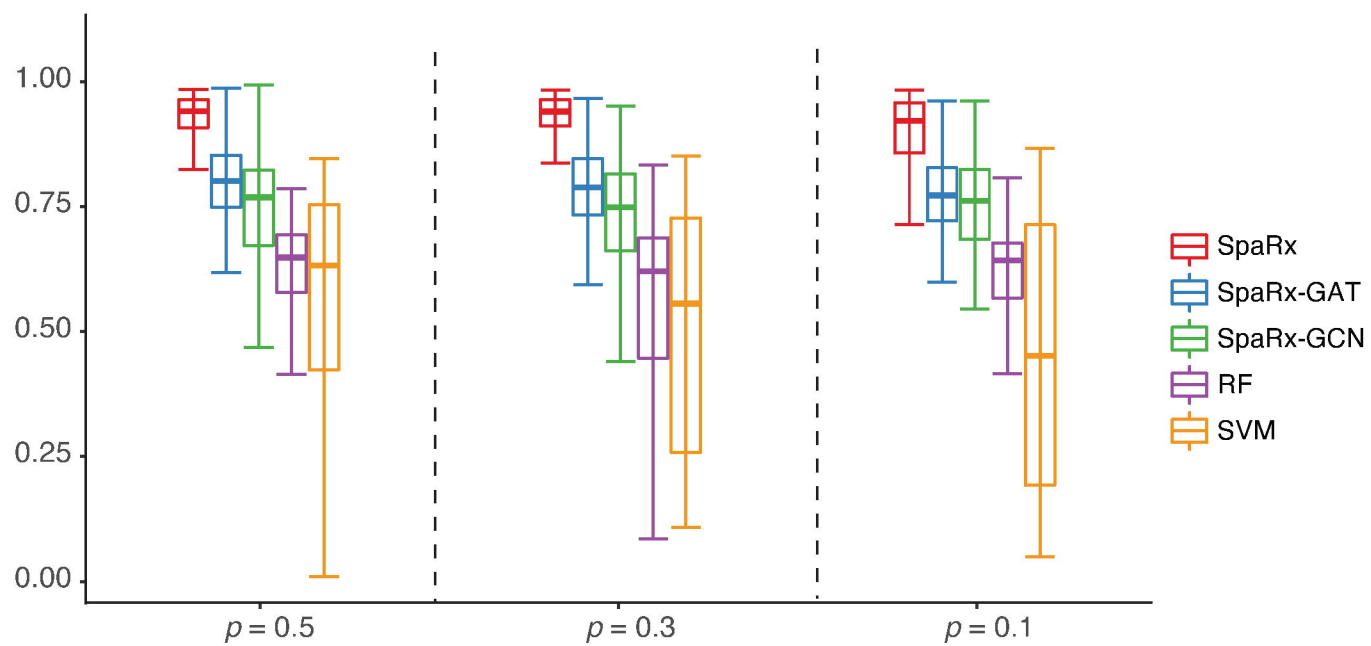
Fig. 2**E**

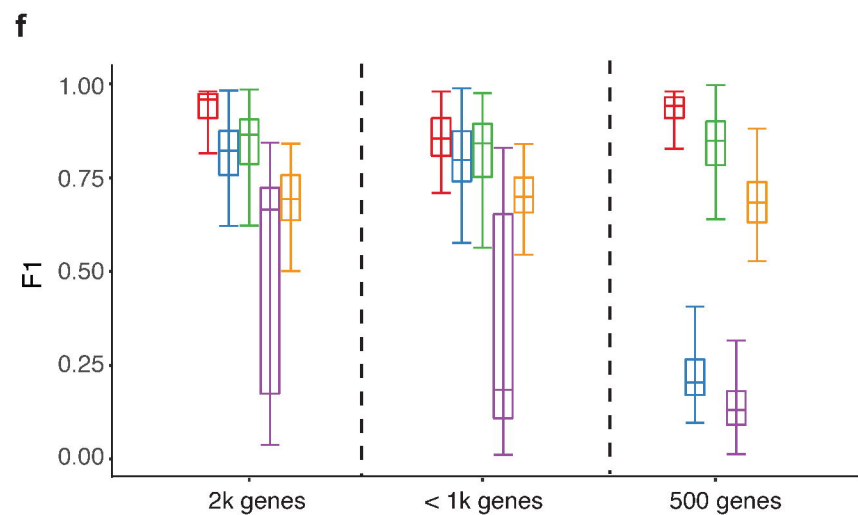
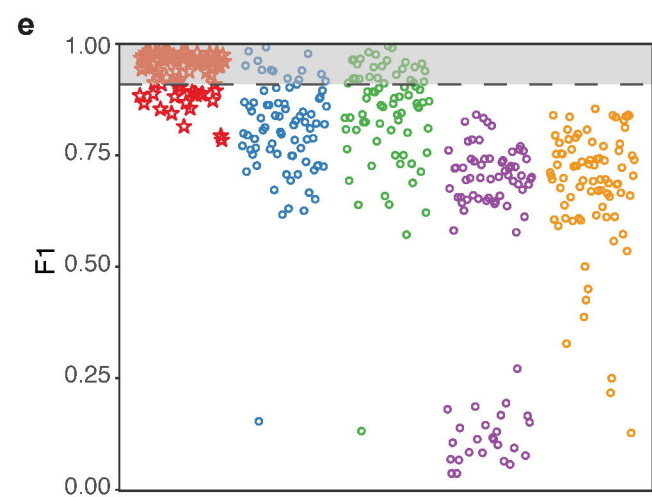
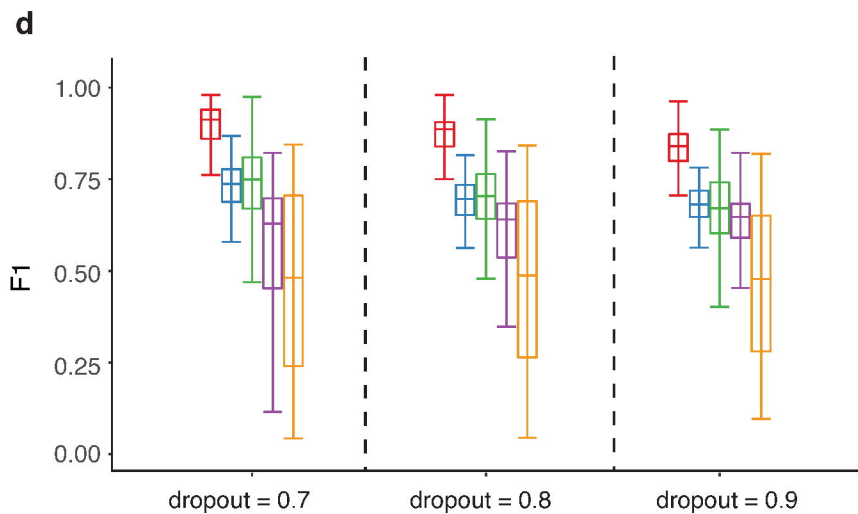
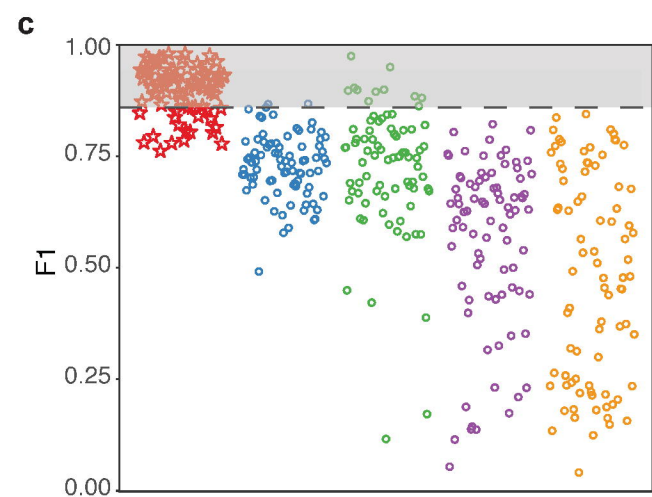
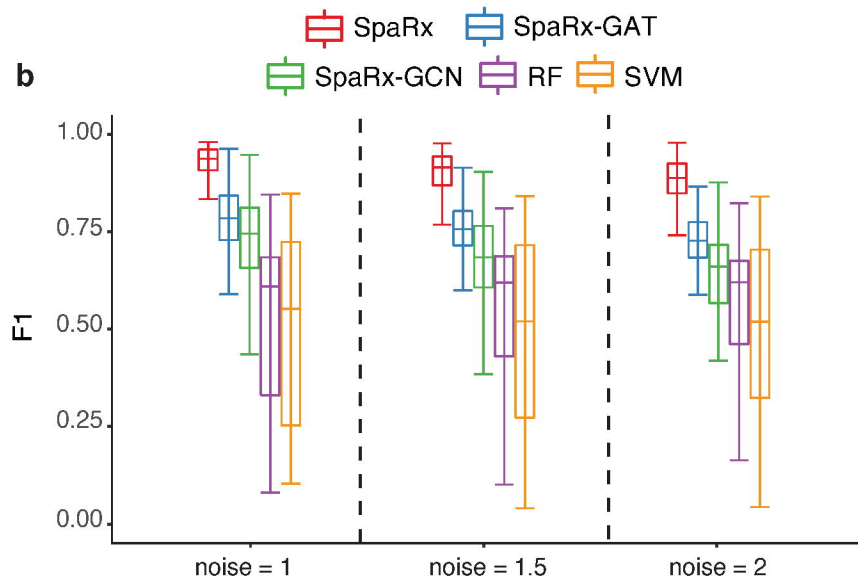
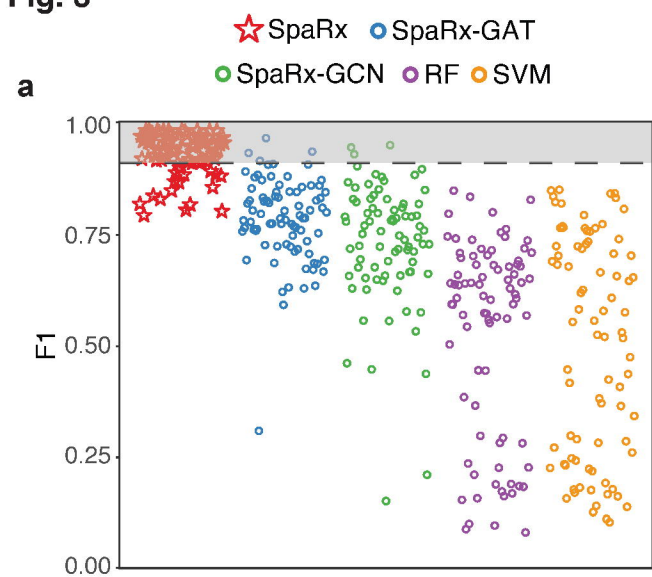
Fig. 3

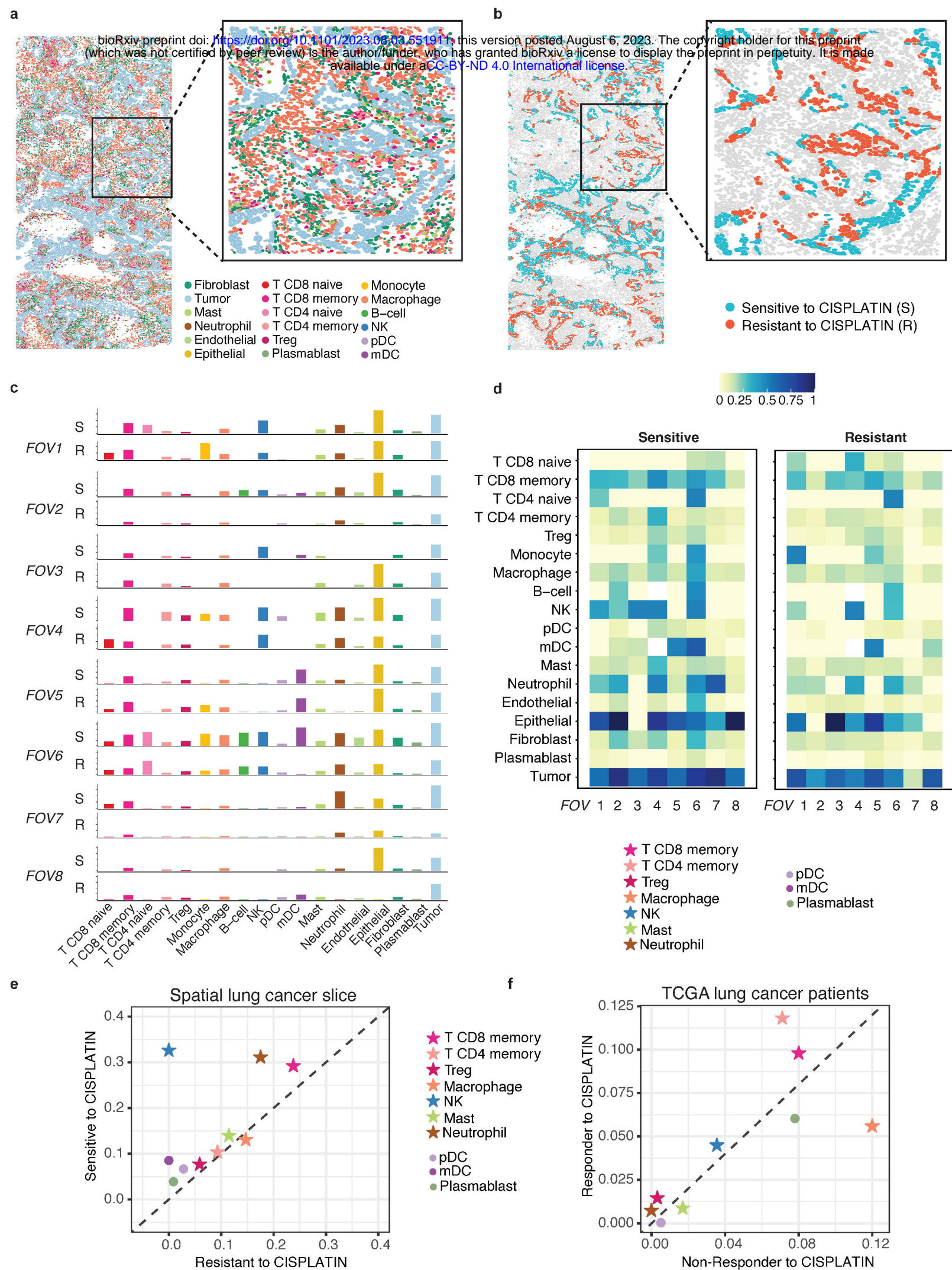
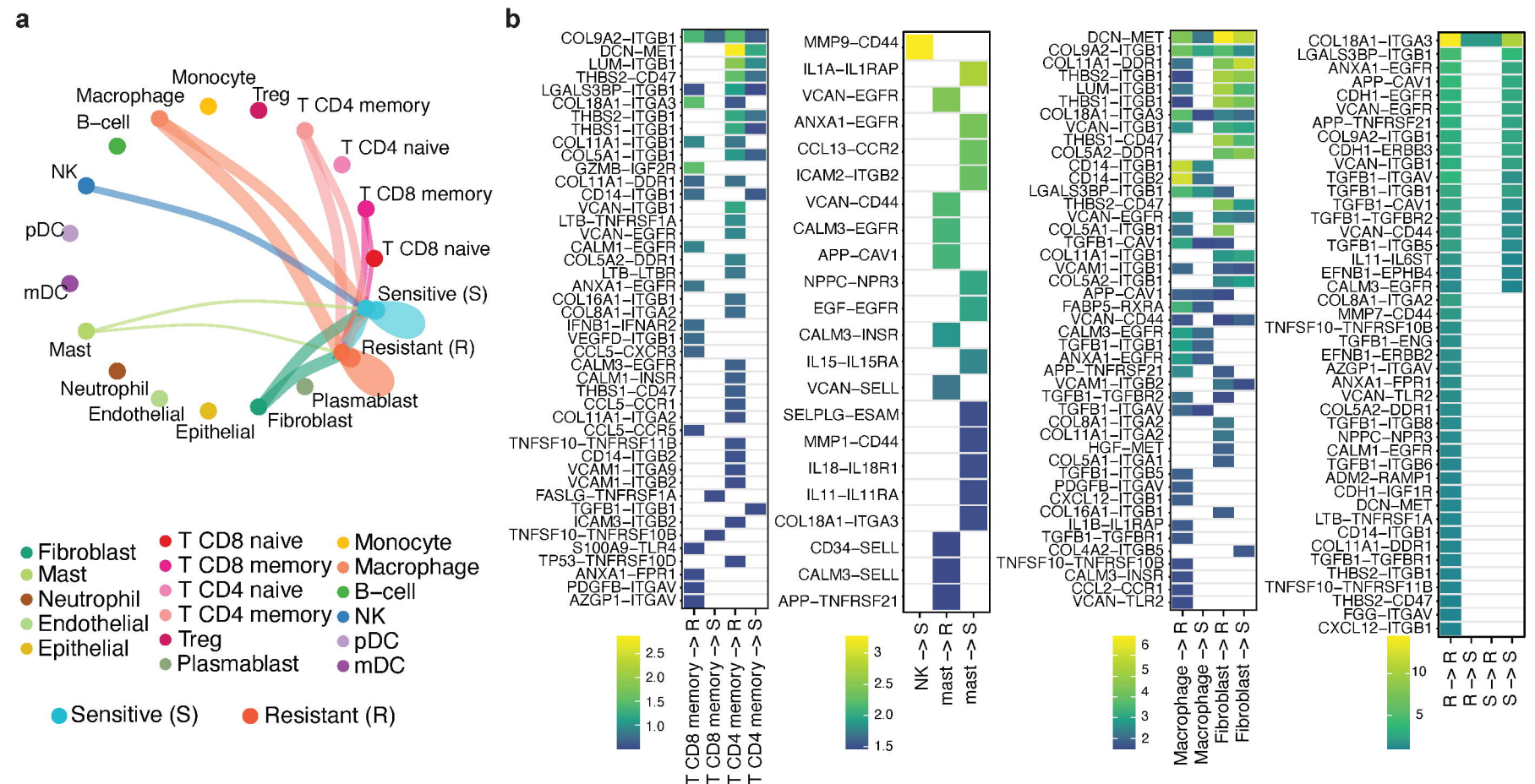
Fig. 4

Fig. 5**c****Cisplatin****Docetaxel****Combined**