Pathway activation model for personalized prediction of drug synergy

Quang Thinh Trac[1†], Yue Huang[2†], Tom Erkers[3], Päivi Östling[3, 4], Anna Bohlin[5], Albin Österroos[6], Mattias Vesterlund[3], Rozbeh Jafari[3], Ioannis Siavelis[3], Helena Bäckvall[3], Santeri Kiviluoto[3], Lukas M. Orre[3], Mattias Rantalainen[1], Janne Lehtiö[3], Sören Lehmann[5, 6], Olli Kallioniemi[3, 4], Yudi Pawitan[1], Trung Nghia Vu[1*]

***For correspondence:** trungnghia.vu@ki.se (TNV)

[†]These authors contributed equally to this work

[1]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden.; [2]Department of Health Statistics, School of Public Health, Weifang Medical University, Weifang, Shandong, China.; [3]Department of Oncology Pathology, Karolinska Institutet, Science for Life Laboratory, Stockholm, Sweden.; [4]Institute for Molecular Medicine Finland, University of Helsinki, Helsinki, Finland.; [5]Department of Medicine Huddinge, Karolinska Institutet, Unit for Hematology, Karolinska University Hospital Huddinge, Stockholm, Sweden.; [6]Department of Medical Sciences, Hematology, Uppsala University Hospital, Uppsala, Sweden.

Abstract Targeted monotherapies for cancer often fail due to inherent or acquired drug resistance. By aiming at multiple targets simultaneously, drug combinations can produce synergistic interactions that increase drug effectiveness and reduce resistance. Computational models based on the integration of omics data have been used to identify synergistic combinations, but predicting drug synergy remains a challenge. Here, we introduce DIPx, an algorithm for personalized prediction of drug synergy based on biologically motivated tumor- and drug-specific pathway activation scores (PASs). We trained and validated DIPx in the AstraZeneca-Sanger (AZS) DREAM Challenge dataset using two separate test sets: Test Set 1 comprised the combinations already present in the training set, while Test Set 2 contained combinations absent from the training set, thus indicating the model's ability to handle novel combinations. The Spearman correlation coefficients between predicted and observed drug synergy were 0.50 (95% CI: 0.47–0.53) in Test Set 1 and 0.26 (95% CI: 0.22–0.30) in Test Set 2, compared to 0.38 (95% CI: 0.34–0.42) and 0.18 (95% CI: 0.16–0.20), respectively, for the best performing method in the Challenge. We show evidence that higher synergy is associated with higher functional interaction between the drug targets, and this functional interaction information is captured by PAS. We illustrate the use of PAS to provide a potential biological explanation in terms of activated pathways that mediate the synergistic effects of combined drugs. In summary, DIPx can be a useful tool for personalized prediction of drug synergy and exploration of activated pathways related to the effects of combined drugs.

Introduction

Targeted therapies such as specific inhibitors are the most promising class of cancer drugs, but often fail or achieve only temporary remission due to inherent or acquired resistance. Theoretically, by aiming at multiple targets simultaneously, drug combinations can produce a synergistic

interaction that increases drug effectiveness and reduces resistance and the chances of relapse (***Medicine, 2017***; ***Pemovska et al., 2018***; ***Plana et al., 2022***). This is illustrated in the combination of a BRAF inhibitor dabrafenib with a MEK inhibitor trametinib, which suppresses paradoxical reactivation and resistance observed in patients with BRAF-mutated melanoma treated with dabrafenib alone (***Zhong et al., 2022***; ***Banzi et al., 2016***). This recently approved combination has been shown to improve progression-free and overall survival rates (***Subbiah et al., 2023***).

The discovery of effective drug combinations has traditionally relied on expert knowledge and understanding of known biological mechanisms (***Li et al., 2015***). However, this expert-based approach has limited scope to come up with novel combinations. Furthermore, ideally, novel combinations are clinically tested, but it is practically impossible to test all reasonable combinations in a clinical setting. Computational models of drug synergy have shown some potential for personalized prediction of synergistic combinations (***Güvenç Paltun et al., 2021***; ***Wu et al., 2022***; ***Kong et al., 2022***). These models are typically based on the integration of patient-specific molecular data, such as mutation profiles, gene expression, and drug response information (***Güvenç Paltun et al., 2021***). For example, TAJI, the best performing method in the AstraZeneca-Sanger DREAM Challenge, uses these diverse data types to predict drug synergy (***Li et al., 2018***). The drug combinations predicted to be effective will expand the therapeutic options while maintaining the same level of adverse effects profile. However, despite the advantages offered by modern machine learning methodologies and the availability of large-scale datasets, prediction of synergistic combinations and validating computational models remains challenging. For example, drug screening protocols often vary across studies, and there is a limited overlap in tested drugs and cell lines, complicating the external validation of these models. Additionally, the reliance on 'black-box' machine learning approaches hinders the exploration of underlying molecular mechanisms driving synergistic combinations.

To address this limitation, several studies have introduced statistical and computational approaches to infer the mechanisms of action of synergistic combinations within cancer signaling pathways. For example, Liu et al. proposed TranSynergy, a drug synergy prediction model that uses the interaction between drug target genes in a protein-protein interaction (PPI) network (***Liu and Xie, 2021***). However, TranSynergy only relies on target gene information, neglecting information on upstream and downstream activities of the targets and their differential contributions to synergy. More recently, Tang et al. developed SynPathy, a deep learning model for drug synergy prediction that incorporates drug-pathway associations (***Tang and Gottlieb, 2022***). SynPathy calculates pathway enrichment scores as a measure of the distance between target genes of each drug in a combination and pathway genes in the PPI network. These pathway enrichment scores, along with chemical structure information, are then combined to fit the model and infer pathway importance scores for each combination. More recently, Wu et al. introduced ForSyn (***Wu et al., 2023***), a deep forest-based method. Although ForSyn implemented a gene enrichment analysis to identify cancer-related pathways, it does not directly identify them through prediction.

Here we present a Drug synergy Interaction Prediction (DIPx) based on tumor- and drug-specific pathway activation scores (PASs). PASs are biologically motivated features that provide potentially relevant information on the underlying mechanisms of synergistic combinations. We trained and validated DIPx using the AstraZeneca-Sanger (AZS) DREAM Challenge dataset (***Menden et al., 2019***), and compared its performance with the best performing method in the Challenge. Furthermore, we assessed the generalizability of the model by validating it on the ONeil dataset (***O'Neil et al., 2016***), and provided illustrations of pathways that could mediate the synergistic combinations found by DIPx. DIPx is publicly available at https://www.github.com/tracquangthinh/DIPx.

Results

A pathway based drug synergy prediction model

Figure 1 provides an overview of DIPx, which uses gene expression, mutation profiles, and drug synergy data from the AZS dataset to train and validate its prediction model. The test set comprised

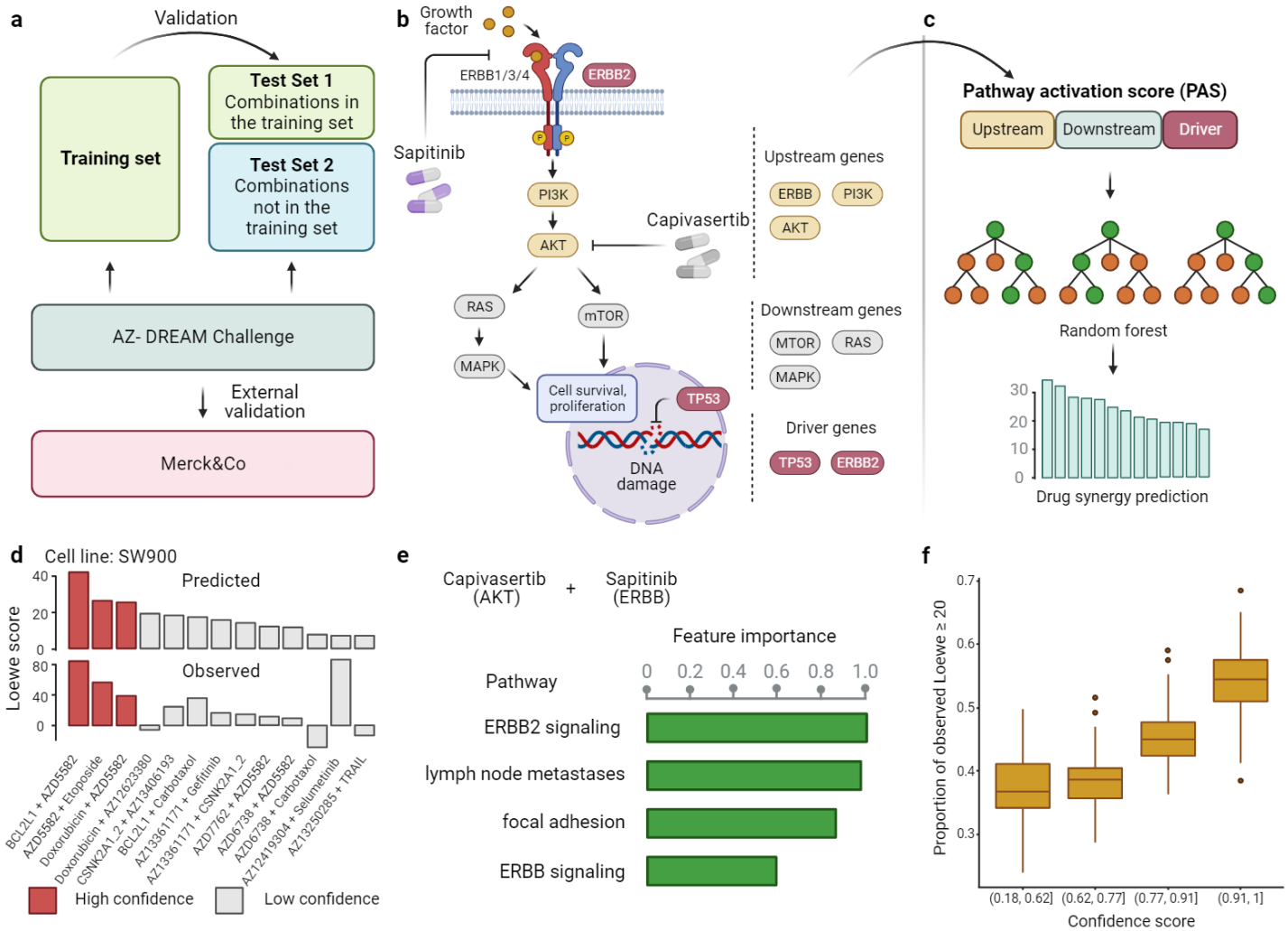

Figure 1. Overview of DIPx. **a**) The AZS omics data were used to train and validate the model. The test set was split into two subsets: Test Set 1 contained combinations found in the training set, while Set 2 comprised combinations not found in the training set. The model was also externally validated using the ONeil dataset. **b**) A cartoon illustration of the ERBB pathway in a breast cancer cell line treated with the combination of Capivasertib + Sapitinib. Capivasertib targets the AKT gene, whereas Sapitinib targets the ERBB genes. Pathway genes were classified into upstream and downstream genes relative to the position of the target genes in the network. **c**) The drug synergy prediction model was trained using pathway activation scores (PAS) of the upstream, downstream and driver genes. **d**) The predicted and observed Loewe scores of a cell line achieved the median Spearman correlation in Test Set 1 of the AZS dataset. The color of each bar shows the confidence score information with the threshold of 0.75. **e**) The main pathways that contribute to the prediction of the synergy of the Capivasertib + Sapitinib combination. **f**) The proportion of validated high synergistic predictions (Loewe score ≥ 20) increases with higher confidence scores. The x-axis presents four groups defined by quartiles of confidence scores.

two subsets: (i) Test Set 1 includes combinations from the training set, and (ii) Test Set 2 includes combinations absent from the training set. Together, both sets assess the generalizability of the prediction for new patients and new combinations. The analysis involved a total of 75 cell lines tested in 910 combinations in the AZS dataset. DIPx was also validated using an external dataset, as shown in Figure 1a.

Figure 1b illustrates the ERBB signaling pathway in relation to the Capivasertib + Sapitinib combination, where the genes belonging to the pathway are classified into upstream and downstream genes relative to the position of the target genes: AKT targeted by Capivasertib and ERBB targeted by Sapitinib. Putative driver mutations were identified in each sample based on a well-characterized list of frequently mutated genes in cancer; see Section 4.3. We first calculate the PAS of the upstream and downstream part of the pathway relative to the driver genes; see the

Methods section for details. The PAS values are then combined to train a random forest regression model. Given a new drug combination experiment, DIPx predicts the Loewe score for drug synergy, as shown in Figure 1c.

Figure 1d presents the predicted and observed synergies for the SW900 lung cancer cell line, which has a median correlation of 0.50 among the cell lines in Test Set 1; each bar in the figure represents a drug combination. The best predicted combinations include BCL2L1 + AZD5582, AZD5582 + etoposide and doxorubicin + AZD5582, with predicted Loewe scores of 42.34, 26.60, and 25.72, respectively, and high confidence scores of 1.0, 0.90, and 0.82, respectively. A combination with Loewe score greater than 20 is considered highly synergistic (***Menden et al., 2019***). Although the combination of doxorubicin + AZ12623380 is predicted to have high synergy, it is a low confidence prediction with a confidence score of 0.33. Indeed, the observed Loewe synergy score for this combination is near zero.

The use of PAS allows DIPx to infer the potential biological mechanisms of synergistic drug combinations. Figure 1e shows pathways with the highest contribution to prediction of drug synergy of the Capivasertib + Sapitinib combination: these include the ERBB-related pathways (ERBB2 signaling pathway, ERBB pathway), and tumor-related pathway (lymph-node metastates, focal adhesion).

Figure 1f demonstrates the association between the confidence scores and the validation of predictions. The x-axis represents the confidence scores grouped into quartiles, while the y-axis displays the proportion of validated high synergy (Loewe score ≥ 20). Predictions with higher confidence scores are expected to exhibit a greater level of validation. Indeed, in this figure, the proportion of high synergistic predictions that are validated in the combination of Test Set 1 and 2 of the AZS dataset increases as the confidence score rises.

Validation and comparisons in the AZS dataset

We evaluated the performance of DIPx in the AZS test sets and compared it with TAJI, which was the best performing method in the AZS DREAM Challenge (***Li et al., 2018***). TAJI was trained using both monotherapy drug-response and molecular data. Since DIPx uses only molecular data, to make a fair comparison, we trained TAJI using only molecular features and referred to it as TAJI-M. The extra information from the use of monotherapy data in TAJI is rather small, approximately 10% increase in the overall Spearman correlation, and, of course, we could also use such data in DIPx, so it is more convenient and informative to focus the comparisons on prediction based on molecular data alone. For instance, this allows us to compare DIPx with TAJI-M on the prediction of combinations that contain un-trained drug(s), which is not possible with TAJI.

Figure 2a shows the correlation between the predicted and observed Loewe scores of 963 experiments in Test Set 1 ($r = 0.5$, 95% CI: 0.47–0.53), where each experiment represents a combination drug A + drug B tried on cell line C, yielding one data point. In comparison, TAJI-M gives $r = 0.38$ (95% CI: 0.34–0.42). We also bootstrapped the training set (n = 100 times) and for each bootstrap replicate calculated the Spearman correlation between the predicted and observed scores of all experiments. As illustrated in Figure 2b, DIPx achieved stable Spearman correlations across all bootstrap replicates, which are significantly higher than that of TAJI-M. The bootstrap distribution actually indicates that the Spearman correlation from DIPx is negatively biased, while from TAJI it is slightly positively biased. This means that the gap between the bias-corrected estimates of the Spearman correlations from DIPx and TAJI-M would be even larger; see the Method section for a theoretical explanation.

Furthermore, we compared the performance of DIPx and TAJI-M across all cell lines in Test Set 1 using a Spearman correlation between the predicted and observed synergy scores, as shown in Figure 2c. A majority of the cell lines (63%) are below the diagonal line, indicating that DIPx outperforms TAJI-M in predicting synergy scores for these cell lines.

We also compared the performance of DIPx and TAJI-M in Test Set 2. As expected, the prediction performance of both methods was worse in Test Set 2 than in Test Set 1 since Test Set 2 consists

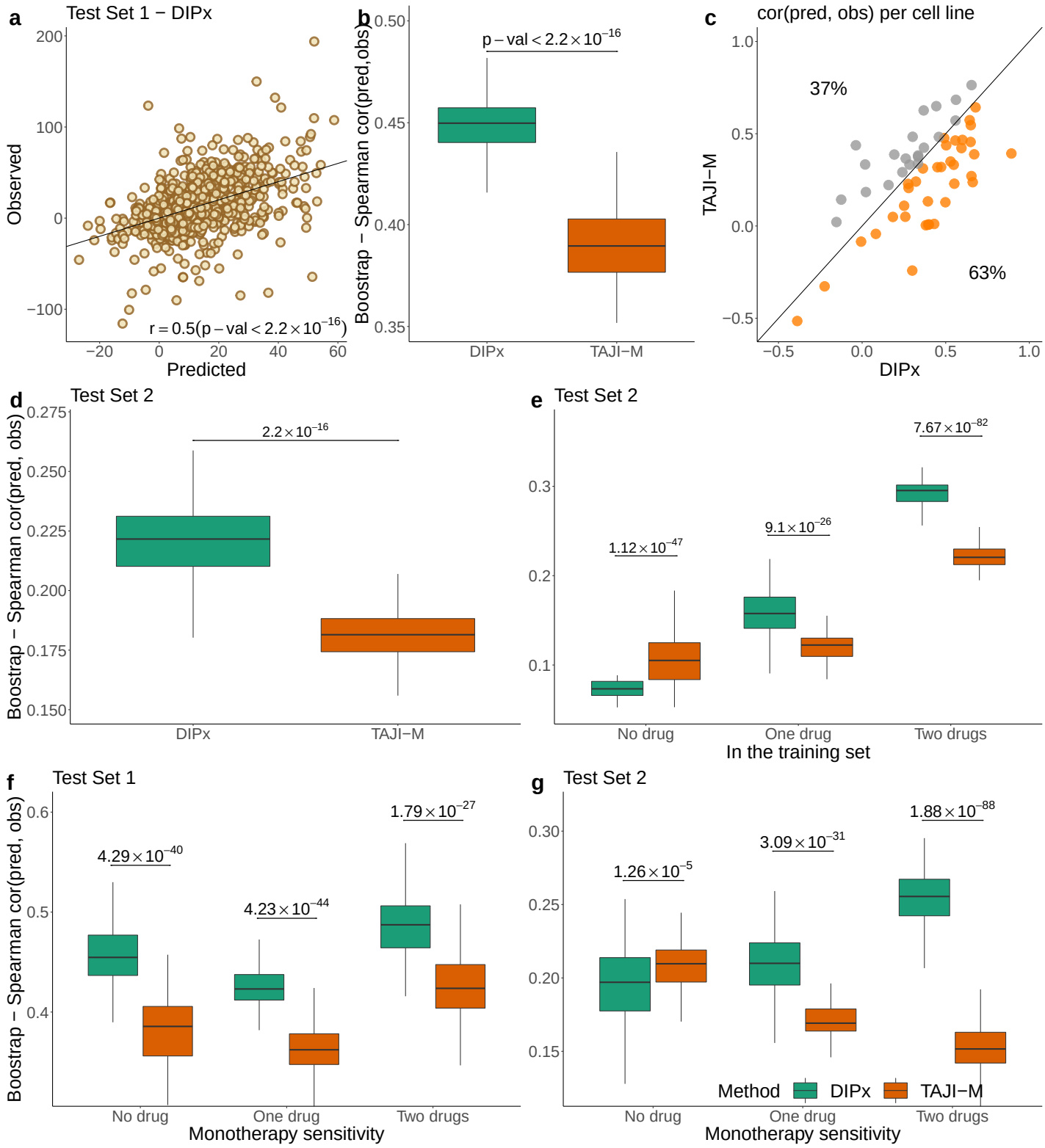

Figure 2. Performance of DIPx in the AZS dataset. This includes both Test Set 1 (panels a, b, c, f) and Test Set 2 (panels d, e, g). **a**) Comparison of predicted vs observed synergy scores for all experiments in Test Set 1. **b**) Comparison of DIPx vs TAJI-M in terms of the correlation between predicted and observed synergy scores from all experiments in Test Set 1. Each boxplot shows the results based on 100 bootstrap replicates of the training set. **c**) Comparison of DIPx and TAJI-M performance across cell lines in Test Set 1. Each point represents the correlation between the predicted and observed synergy for a given cell line. **d**) Comparison of DIPx vs TAJI-M in Test Set 2. Each boxplot displays the correlations between the predicted and observed values obtained from 100 bootstrap replicates of the training set. **e**) Comparison of performance between DIPx and TAJI-M in Test Set 2 in relation to the number of drugs in common (x-axis) between the combinations in the test set and the training set. **f**) and **g**) DIPx vs TAJI-M in three groups classified by monotherapy sensivitity of two drugs in a combination in Test Set 1 (**f**) and Test Set 2 (**g**).

of new combinations absent from the training set. The Spearman correlation of the observed vs predicted synergy using DIPx is 0.26 (95% CI: 0.22–0.30), which is greater than 0.18 (95% CI: 0.16–0.20) using TAJI-M. Figure 2d show that this result is stable across 100 bootstrap replications of the training set. A similar downward bias for DIPx is observed in the bootstrap distribution.

To investigate the effect of unseen combinations on prediction performance, we divided each combination (drug A + drug B) in Test Set 2 into one of three groups based on the number of individual drugs present in the training set: (i) neither drug A nor drug B in the training set ("no drug"), (ii) either drug A or drug B in the training set ('one drug'), (iii) and both drugs A and B in the training set ('two drugs'), as shown in Figure 2e. Overall, both DIPx and TAJI-M showed improved performance as the number of drugs present in the training set increased. For experiments in which both drugs were not in the training set ($n = 262$), TAJI-M achieved a median correlation of 0.11, while DIPx performed worse with a median correlation of –0.03. For experiments with at least one drug in the training set ($n = 2,499$), both methods showed improved performance with median correlations of 0.16 and 0.12 for DIPx and TAJI-M, respectively. When both drugs in an experiment were present in the training set ($n = 4,370$), DIPx achieved a median correlation of 0.30, which was better than TAJI-M's performance ($r = 0.22$).

Monotherapy drug response profiles have been shown to correlate with synergistic effects and contribute to improving prediction performance, e.g., in TAJI (***Li et al., 2018***). Here, we compared the performance of DIPx and TAJI-M in relation to monotherapy sensitivity as measured by the IC50 value. We categorized each experiment in the AZS test sets into three groups according to the monotherapy response. Briefly, we first calculated the median sensitivity to monotherapy for each drug A (T_A) across all experiments. Measuring the response of a cell line to drug A in an experiment by S_A, the drug is considered sensitive if $S_A \geq T_A$. We then compared the synergy of a combination of drugs A and B in relation to the monotherapy sensitivity to both drugs, only one drug, or neither drug.

In Test Set 1, DIPx outperformed TAJI-M in all three groups of monotherapy sensitivity, with the highest performance in the group sensitive to both drugs (median $r = 0.48$, P-value $< 1.79 \times 10^{-27}$), see Figure 2f. In Test Set 2, TAJI-M performed slightly better in the group with no sensitive drug (median $r = 0.21$ vs $r = 0.20$ by DIPx, P-value $< 1.26 \times 10^{-5}$). Interestingly, we found that, while the performance of DIPx improved as the number of monotherapy-sensitive drugs in a combination increased, the performance of TAJI-M decreased, see Figure 2g.

External validation of DIPx in the ONeil dataset

We used a similar computational approach to evaluate the prediction performance of DIPx in relation to the sensitivity of the constituent monotherapies and the impact of unseen combinations in the ONeil dataset. As shown in Figure 3a, the performance of DIPx improved with an increasing number of monotherapy-sensitive drugs in the combination, consistent with the results of Test Set 2 of the AZS data. The highest Spearman correlation between the predicted and observed scores was seen in combinations with two sensitive drugs (median $r = 0.11$). In relation to the number of drugs in a combination present in the training set, DIPx achieved better performance for combinations with none or one drug in the training set (middle box plot, Figure 3 - figure supplement 1). Poor performance for combinations with two drugs in the training set could be due to the limited number of drug combinations (42/583). We also analyzed the prediction performance of DIPx across the 29 cell lines from 6 different cancer tissues (Figure 3b). Colon cancer (yellow boxplots) and lung cancer cell lines (purple boxplots) showed better validation compared to cell lines from breast, ovarian, melanoma, and prostate cancers.

Inference of the mechanism of action based on PAS

The use of PAS in DIPx allows us to infer the potential mechanisms of action of drug combinations while maintaining the prediction performance of the model. For instance, in Test Set 1 of the AZS data, DIPx suggests the involvement of ERBB2 signaling pathways in the Capivasertib + Sapi-

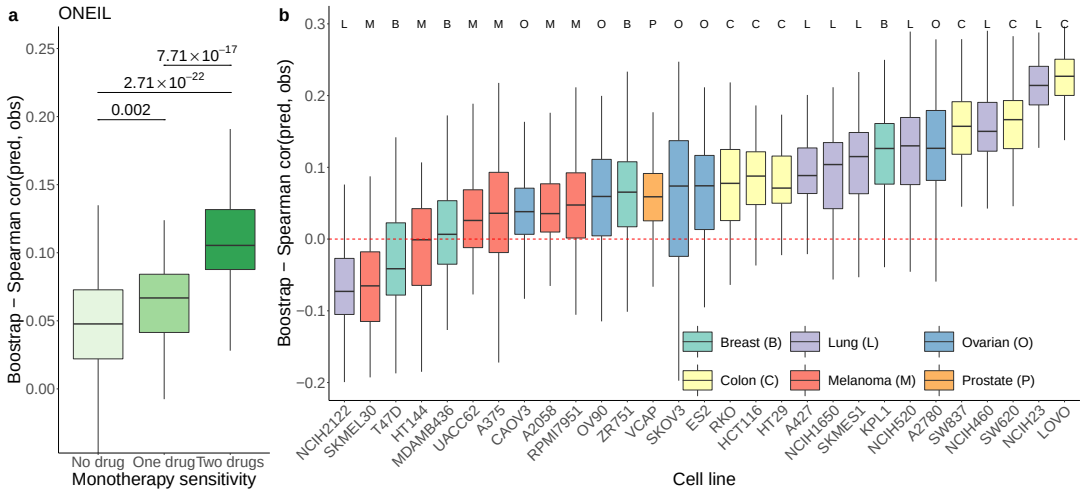

Figure 3. Prediction performance of DIPx in the ONeil dataset. **a**) monotherapy sensitivity, **b**) 29 cell lines from 6 cancer tissues. The y-axis in all box plots shows the Spearman correlation between predicted and observed values in 100 bootstrap replicates.

Figure 3—figure supplement 1. Prediction performance of DIPx in the ONeil dataset, grouped by unseen combinations.

tinib combination, as illustrated by the top pathways depicted in Figure 1e and marked yellow in Figure 4a. This combination therapy has shown promise in overcoming resistance to anti-ERBB2 monotherapy in HER2+ breast cancer (***Fujimoto et al., 2020***), and ERBB2 has been identified as a key biomarker associated with synergistic responses to this combination in the AZS DREAM Challenge study (***Menden et al., 2019***).

Figure 4a further shows the distribution of feature importance versus PAS for all pathways for Capivasertib + Sapitinib combination. The feature importance value is calculated using the permutation method of Ishwaran et al. (***Ishwaran and Lu, 2019***). Our focus is on pathways with high feature importance (e.g., the top 5%) as well as highly activated (top 5% PAS). Therefore, the top-right section of Figure 4a is the interesting region. We present additional examples to further demonstrate the capabilities of DIPx. Figure 4b gives the top pathways of MEDI3622, an ADAM17 inhibitor, in combinations with AKT inhibitors including Capivasertib and MK-2206. These ADAM17 + AKT combinations target multiple parts of the PI3K/AKT pathway through ERBB activation (***Menden et al., 2019***), which aligns with the potential pathway candidates suggested by DIPx.

One of the key strengths of DIPx is its ability to infer potential mechanisms of both known and novel drug combinations, even in cases where limited biological or clinical information is available. This capability is particularly valuable for new combinations that have not been included in the training set. For instance, in Figure 4c, we present the key pathways identified for the Selumetinib + MK-2206 combination from Test Set 2 of the AZS data. We observe the involvement of RAS signaling, with Selumetinib targeting MEK and MK-2206 targeting AKT, as shown in Figure 4-figure supplement 2. A recent clinical study has used Selumetinib + MK-2206 to target downstream components of the RAS pathway (***Chung et al., 2017***).

If the drugs in a combination have the same target, the efficacy of the combination is likely similar to that of each individual drug at higher doses, i.e., they will only have an additive effect. So it seems reasonable to hypothesize that a synergistic combination is more likely to occur when the two drugs have different targets (***Chen et al., 2015***). But how should the targets be related to each other? To investigate this, we examine the pathways suggested by DIPx. First, we choose a synergistic combination of BCL2L1 + AZD5582 in the SW900 cell line for further illustration. The contour plot of the BCL2L1 + AZD5582 inhibition in the SW900 cell line is illustrated in Figure 4 - figure supplement 3a. We first collected the top 15 pathways (ranked by feature importance) for

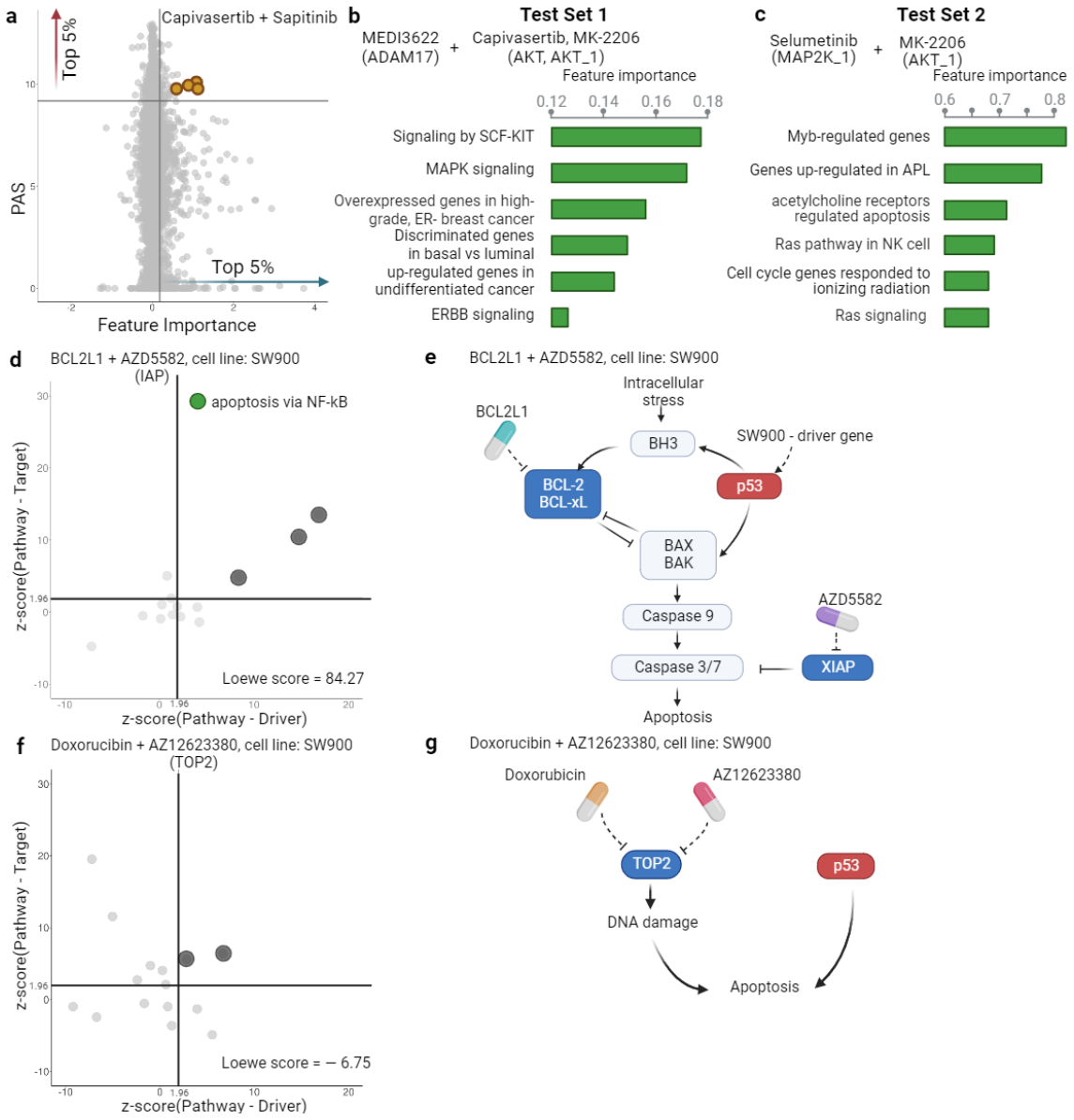

Figure 4. Inference of pathway importance scores in the AZS dataset. **a**) Scatter plot showing feature importance (x-axis) vs PAS (y-axis) for the Capivasertib + Sapitinib combination. Pathways with high PAS and feature importance (top 5%) are of particular interest. **b, c**) Top pathways contributing to the prediction of the combinations in Test Set 1 (b) and Test Set 2 (c). For each pathway, the bar plots show its feature importance. **d, f**) Functional interaction between the pathway vs driver genes (x-axis) and the pathway vs target genes (y-axis) of the top pathways suggested by DIPx in the SW900 cell line treated with synergistic combination BCL2L1 + AZD5582 (d) and the non-synergistic combination Doxorubicin + AZ12623380 (f). The z-score from network enrichment analysis (NEA) is a measure of functional interaction between two gene sets. A higher z-score indicates a stronger interaction compared to a random permutation of the network. The upper right quadrant (z-score > 1.96) represents pathways that are potentially interesting. **e, g**) Cartoon illustration of the potential pathways mediated by the synergistic combination of BCL2L1 + AZD5582 (e) and the non-synergistic combination Doxorubicin + AZ12623380 (g).

Figure 4—figure supplement 2. A cartoon illustration of the RAS pathway mediated by the Selumetinib + MK-2206 combination

Figure 4—figure supplement 3. Observed vs predicted inhibition in the SW900 cell line treated by BCL2L1 + AZD5582 and Doxorubicin + AZ12623380 combinations

Figure 4—figure supplement 4. Functional interaction between driver genes, target genes, and top pathways suggested by DIPx in the SW900 cell line treated with BCL2L1 + AZD5582.

Figure 4—figure supplement 5. Functional interaction between driver genes, target genes, and top pathways suggested by DIPx in the SW900 cell line treated with Doxorubicin + AZ12623380.

this BCL2L1 + AZD5582 combination suggested by DIPx. The full list of these pathways is shown in Figure 4 - figure supplement 4. Figure 4d illustrates the functional interaction between the genes of the top 15 pathways and the driver genes of the SW900 cell line (x-axis); the target genes of the combination BCL2L1 + AZD5582 (y-axis). To assess the strength of this interaction, we used the network enrichment analysis (NEA) (***Alexeyenko et al., 2012***), which provides z-score, an enrichment score, indicating the degree of interaction. A higher z-score reflects a stronger interaction between the two gene sets. The top pathways exhibiting high functional interaction with both the driver genes and target genes (z-score > 1.96) are particularly notable, located in the upper right quadrant of Figure 4d. In particular, the apoptosis pathway via NF-kB (highlighted in green) has the highest pathway-target interaction among these pathways. Figure 4e shows the cartoon illustration of the pathway in which the drug BCL2L1 targets BCL-xL and AZD5582 targets XIAP. This suggests an explanation for the observed synergy between the two drugs. Thus, it appears that in this case we get synergy when the drugs target different parts of a driving pathway, either directly or via other functional interactions.

As a negative control, we examine the non-synergistic combination Doxorubicin + AZ12623380, which targets the same gene TOP2; see Figure 4f and g and Figure 4 - figure supplement 3b. We similarly obtain 15 top-ranking pathways according to feature importance, but now we do not expect to see anything obviously relevant to the SW900 cell line (more details in Figure 4 - figure supplement 5). Some pathways that have a high functional interaction with the target genes (upper-left quadrant) have little interaction with the drivers. There are no clearly outlying points in the upper-right quadrant; the two pathways near the boundary are (i) Shen_Smarca2_targets_up, containing genes whose expression negatively correlated with the expression of the SMARCA2 gene in prostate cancer samples, discovered in relation to androgen-induced proliferation in the prostate; and (ii) Kokkinakis_Methione _deprivation_48hr_up, which contains up-regulated genes in melanoma cell-line MEWO cells after 48h of methionine deprivation. They do not appear to be relevant for the lung cancer cell line SW900.

PAS captures the functional interaction of drug targets

As we discuss above, to get synergy, the two drugs in a combination theoretically should not have the same target. However, there is of course no guarantee that two drugs that do not share target genes can produce synergy. In Figure 5a, using the AZS data, we compare the observed drug synergy of combinations of two drugs that share some target genes vs those that do not share any target genes. No significant differences were observed (p-value > 0.72), suggesting that non-overlapping drugs in terms of their targets do not necessarily result in improved drug synergy.

However, we also observed synergy when the two drugs target different genes in the same pathway. More generally, we hypothesize that synergistic effects occur when the targets have functional interaction. As before, the functional interaction is assessed using NEA (***Alexeyenko et al., 2012***), where a higher z-score value indicates a stronger functional interaction between the two drugs. Figure 5b shows the observed drug synergy (y-axis) in the AZS data for the four groups defined by the quartile values of the z-scores (x-axis). It indicates that combinations with higher functional interaction are more likely to achieve higher drug synergy, with the highest z-score group ($z \in (2.97, 29.3]$) exhibiting the most favorable drug synergy (median Loewe score = 10.34).

However, when added to the prediction model, the functional interaction z-score did not improve the prediction of synergy (data not shown). Statistically, this can happen if PAS already captures the functional interaction information. To show this, using the AZS training data, we trained a prediction model using PAS as the feature and the functional interaction z-score as the output. We then evaluated the performance of the model in the test set. As shown in Figure 5c, we observed a significant correlation between the predicted and observed z-scores, with a Spearman correlation coefficient of 0.46. This explains why the functional interaction does not give additional predictive power in our model.

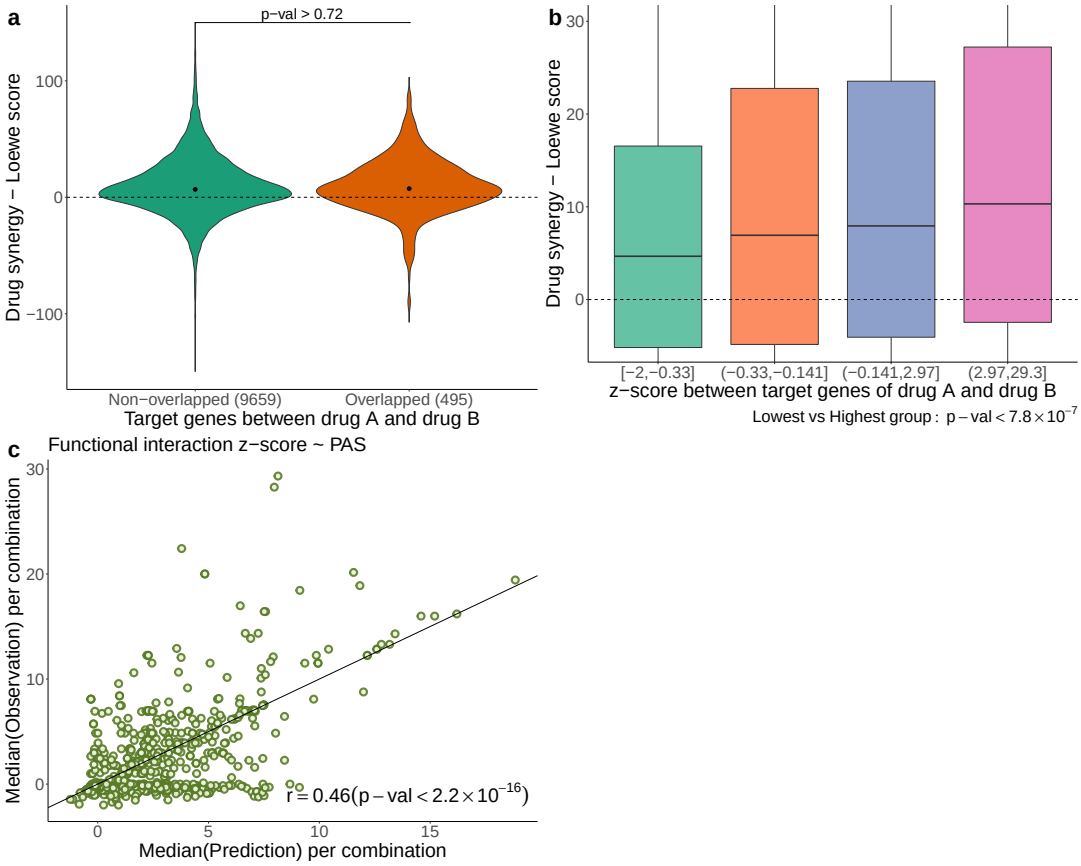

Figure 5. a) Comparison of drug synergy between combinations (drug A + drug B) with vs without overlapping target genes. The numbers in parentheses show the sample sizes of each group. **b**) Drug synergy between four groups in relation to increased functional interaction between the target genes of the two drugs. **c**) Comparison between the observed functional interaction (z-score in the network enrichment analysis) and the predicted z-score by PAS.

Discussion

We have developed and validated DIPx, an advanced computational model that incorporates gene expression and mutation profiles to predict synergistic drug combinations. DIPx performs well against the best performing method in the AstraZeneca-Sanger DREAM Challenge. Through the use of tumor- and patient-specific pathway activation scores, DIPx also provides valuable information on the potential underlying pathways associated with an observed synergistic drug interaction. In addition to rigorous validation using the AZS dataset, DIPx is further validated on the independent ONeil dataset. This comprehensive validation ensures the robustness and reliability of DIPx in predicting drug synergy across different cancers.

The recent availability of large-scale drug combination assay data has allowed the development of realistic prediction models for drug synergy. These datasets offer a substantial number of samples encompassing hundreds of combinations, allowing for extensive validation studies. However, it is important to note that these datasets were generated using different protocols and drug screening techniques. For instance, the AZS data used a 5-by-5 concentration matrix, while the study by ONeil et al. used a 4-by-4 format. In addition, there is limited overlap in the cell lines used among the datasets. These differences pose challenges to the proper validation of prediction methods (***Menden et al., 2019***).

A particular strength of our study is that we use the best-performing method in the Challenge as a benchmark. This is a convenient but robust benchmarking, as there were 160 teams that participated in the Challenge (73 teams submitted in the final round). Altogether, these teams used practically all of the commonly used machine learning tools; see the summary in Menden et al(***Menden et al., 2019***). Another strength is our use and validation of the confidence score metric, which captures the statistical uncertainty in the predicted synergy by a single number. This is more convenient for clinical interpretation than the standard prediction interval, because there is a target level for which a combination is considered synergistic, so the score measures our confidence in achieving the target.

Despite promising results, our study has several limitations. First, the use of cell lines as training and validation samples from the AZS and ONeil datasets may not fully capture the heterogeneity present in actual tumors. Second, the computation of PAS relies solely on the primary target genes of the drug combinations, potentially disregarding valuable information from non-primary targets. There could also be off-targets that we do not know about. This limitation might lead to the loss of information about the broader effects of drug combinations. Third, cancer is a heterogeneous disease that occurs in many tissues. Even within a single tissue, cancer exhibits distinct (molecular) subtypes with varying biological mechanisms and clinical outcomes. Since DIPx was developed using pan-cancer datasets, it may not be optimal for tissue-specific predictions.

Last but not least, prediction of previously untrained combinations remains a great challenge. The worst case is for combinations of drugs that were not previously trained, with the Spearman correlation only around 0.1. However, from a clinical perspective, it is perhaps more realistic to look for combinations among drugs previously trained in monotherapy or in other combinations. Improving the prediction for the combination of such drugs would be worthwhile.

Methods and Materials

Pathway activation score for drug combinations

Pathway activation scores (PASs) are the key features in DIPx. The PAS of pathway P in cell line C is calculated for each drug combination (drug A + drug B) and pathway P. Genes in pathway P are grouped into three subgroups: (a) G_u, which includes all the target genes of drugs A and B, as well as the upstream genes of pathway P; (b) G_d, which includes the downstream genes of pathway P; and (c) G_{dr}, which consists of all the driver genes of cell line C in pathway P. In the example of the ERBB pathway targeted by Capivasertib + Sapitinib. (Figure 1b), G_u consists of ERBB, PI3K, and also AKT; G_d contains MTOR, RAS and MAPK, while G_{dr} includes TP53 and ERBB2.

The score for upstream activity (PAS_u) is calculated by the sum of mRNA expression for genes in G_u. Similarly, the scores for the downstream activity (PAS_d) and the set of driver genes (PAS_{dr}) are calculated from G_d and G_{dr}. In practice, the genes of the N = 4,762 curated human pathways are provided from the MsigDB database (***Liberzon et al., 2015***). The target genes of the drugs are collected from the AZS dataset and extended from the DrugBank database (***Wishart et al., 2018***) and the ChEMBL database (***Zdrazil et al., 2024***). The extraction of the driver genes of the cell lines is described in the Datasets section.

A pathway based model for drug synergy prediction

The training features of DIPx consist of three components: upstream activity (PAS_u), downstream activity (PAS_d), and driver genes (PAS_{dr}), as shown in Figure 1b. The final training matrix has a size of K experiments by 14,286 PASs, where each row corresponds to a specific experiment (drug A + drug B, cell line C).

To address potential sparsity in the training matrix caused by pathways with no target or driver genes, we explored an alternative model with N = 4,762 additional features. Each feature corresponds to a pathway P and is calculated as S(g) * (w1 + w2), where S(g) represents the sum of mRNA expression for all genes in pathway P, and w1 and w2 denote the functional interactions between gene sets: (pathway genes ↔ target genes) and (pathway genes ↔ driver genes), respectively. The functional interactions were estimated using NEA and converted into normal probability scores for w1 and w2. The feature value is zero only when the pathway lacks both targets and driver genes, as well as any interactions with drug targets and driver genes. Additionally, we incorporated the NEA enrichment score between target genes and driver genes into the final matrix. Despite adding these new features, the alternative model did not exhibit any significant improvements in predictive power (data not shown).

For the predictor, we used the random forest algorithm implemented in the randomforestRSC package (with default parameters) in R version 4.0.4. During the development of DIPx, we experimented with various machine learning methods, such as the support vector machine (SVM) and the elastic net. However, we found that these other methods yielded comparable results and that tuning their parameters did not significantly improve prediction performance while requiring extensive additional computations (data not shown). The random forest algorithm in the randomforestRSC package also offers multiple options to calculate the importance of features. In this study, we used the permutation (or Breiman-Cutler) method (***Ishwaran and Lu, 2019***) to infer the importance of each PAS.

The confidence score (CS) is used to assess the statistical quality of synergy prediction; see Pawitan (2001, Section 5.6) (***Pawitan, 2001***) for the confidence concept in general. First, as previously defined for example in (***Menden et al., 2019***), a combination is considered synergistic if the Loewe score is greater than or equal to 20. For each sample s, we have the actual predicted synergy P_s. We then generate $N_b = 100$ bootstrap replicates of the training data and obtain the bootstrap predictions for the sample: $P^*_{s1}, \ldots, P^*_{s100}$. The CS of P_s is defined as follows:

$$CS(Ps) = \frac{\#(P^*_{si}s \geq 20)}{N_b}.$$

The bootstrap replicates are also used to evaluate the standard errors (se) of the Spearman correlation between the observed and predicted synergy scores in the test sets. The 95% confidence intervals are computed by the usual formula: $\hat{\rho} \pm 1.96$se, where $\hat{\rho}$ is the observed Spearman correlation. Though less frequently used, the bootstrap can also be used for bias correction (Pawitan, 2001, Section 5.2). Bias occurs if there is a nontrivial gap between the observed estimate and the mean of the bootstrap replications. Theoretically,

$$\text{Bias} = E_F(\hat{\rho}) - \rho,$$

where F is the underlying data distribution. So, the bias-corrected estimate should be

$$\hat{\hat{\rho}} = \hat{\rho} - \text{Bias}$$

In practice, the bias is estimated by

$$\begin{aligned} \widehat{\text{Bias}} &= E_{\hat{F}}(\hat{\rho}) - \hat{\rho} \\ &= \text{average}\{\hat{\rho}_1^* \dots \hat{\rho}_n^*\} - \hat{\rho}, \end{aligned}$$

where $\hat{\rho}_1^* \dots \hat{\rho}_n^*$ are the bootstrap replicates of $\hat{\rho}$. When the estimated bias is negative, as we observed for DIPx, the bias-corrected estimate is shifted upward. And vice versa, if the bias is positive, as observed for TAJI-M, the corrected estimate is shifted downward.

Datasets

AstraZeneca-Sanger (AZS) DREAM Challenge dataset The AZS DREAM Challenge is a rigorous competition in the effort to systematically develop and validate drug synergy prediction methods. Indicating the strong interest in the topic, 160 international teams (***Menden et al., 2019***) participated in the Challenge. It was organized into two subchallenges: i) Prediction for known (tested) combination and ii) Prediction for unknown (untested) drug combinations. The final dataset comprised 11,576 experiments from 85 cell lines and 910 combinations. The gene expression data of these cell lines was obtained from Affymetrix microarray (***Menden et al., 2019***). However, to ensure consistency between the AZS dataset and the Oneil dataset (***O'Neil et al., 2016***) (which did not provide gene expression profiles of cell lines), we utilized gene expression data from the Cancer Cell Line Encyclopedia (CCLE) cohort (***Ghandi et al., 2019***).

Out of the 85 cell lines, we identified 75 cell lines with available gene expression data in the CCLE cohort, resulting in a total of 10,154 experiments involving 910 combinations used in our study. Supplementary File S1 shows the list of 75 cell lines. For the validation of the prediction model, the data were split into a training set (n = 2060) and two test sets (n = 963 and 7131) according to subchallenges 1 and 2, respectively. The first test set contains experiments from 167 combinations (of 69 single drugs) that are also in the training set. The second test set includes experiments with 729 drug combinations that are not in the training set.

We collected gene expression data of 75 cell lines, measuring the transcript per million (TPM) of 37,222 genes, of the CCLE cohort downloaded from the DepMap Portal (***Tsherniak et al., 2017***). The gene expression data was logarithmically transformed to the base 2 scale for downstream analysis. Additionally, we obtained potential driver genes for these cell lines, including both mutations and fusion genes, from the DepMap Portal. The portal provides information on mutations in 1,637 protein-coding genes associated with cancer biology in a collection of 1,030 cell lines.

To filter the list of mutations, we focused on those occurring in at least 2.5% of the total cell lines. Subsequently, we extracted the list of mutations specific to the 75 cell lines under investigation. For fusion genes, we focused on those present in the Miltelman database (***Mitelman, 2022***) and occurring at least twice, considering them as relevant for our analysis. The final list of potential driver genes for the 75 cell lines can be found in Supplementary File S1. On average, each cell line had a median of 29 potential driver genes.

For the drug synergy data, we used a 5-by-5 concentration matrix provided by the Challenge. Drug synergy values were estimated using the Loewe reference model from Combenefit (***Di Veroli et al., 2016***).

ONeil dataset ONeil dataset is a large-scale drug synergy screening dataset from Merck&Co company (***O'Neil et al., 2016***). A total of 23,062 experiments with 583 unique drug combinations (38 monotherapy drugs) was carried out on 38 cancer cell lines by a 4-by-4 drug concentration matrix. Out of 38 cell lines, we found 29 cell lines with available gene expression data from the DepMap Portal. The detail of 29 cell lines is described in Supplementary File S1. The gene expression data of 37,222 genes from 29 cell lines, as well as the driver genes of these cell lines, were collected from the DepMap Portal using the same procedure as in the AZS dataset. The original release of this

dataset provides only the raw data on drug synergy. Here, we calculated the Loewe synergy score for each experiment using Combenefit (***Di Veroli et al., 2016***). In total, we obtained 16,907 experiments for 583 combinations in 29 cell lines for further analysis. Drug targets of 38 monotherapy drugs were collected from the DrugBank database (***Wishart et al., 2018***) and the ChEMBL database (***Zdrazil et al., 2024***).

Data Availability

The implementation of DIPx, and related data are publicly available in https://www.github.com/tracquangthinh/DIPx. Drug synergy data are available from their original studies: Synapse database at synapse.org/DrugCombinationChallenge for the AZS dataset (***Menden et al., 2019***), raw data from the supplementary data for the ONeil dataset (***O'Neil et al., 2016***).

Acknowledgments

This work was partially supported by funding from the Swedish Research Council (VR), Cancer-Fonden, and the Swedish Foundation for Strategic Research (SSF). The computations were enabled by resources provided by the National Academic Infrastructure for Supercomputing in Sweden (NAISS) at UPPMAX partially funded by the Swedish Research Council through grant agreement no. 2018-05973. We acknowledge the investigators of the AstraZeneca-Sanger DREAM Challenge for data access.

Conflict of interest statement.

The authors declare no competing interests.

References

Alexeyenko A, Lee W, Pernemalm M, Guegan J, Dessen P, Lazar V, Lehtiö J, Pawitan Y. Network enrichment analysis: extension of gene-set enrichment analysis to gene networks. BMC bioinformatics. 2012; 13:1–11.

Banzi M, De Blasio S, Lallas A, Longo C, Moscarella E, Alfano R, Argenziano G. Dabrafenib: a new opportunity for the treatment of BRAF V600-positive melanoma. OncoTargets and therapy. 2016; p. 2725–2733.

Chen D, Liu X, Yang Y, Yang H, Lu P. Systematic synergy modeling: understanding drug synergy from a systems biology perspective. BMC systems biology. 2015; 9:1–10.

Chung V, McDonough S, Philip PA, Cardin D, Wang-Gillam A, Hui L, Tejani MA, Seery TE, Dy IA, Al Baghdadi T, et al. Effect of selumetinib and MK-2206 vs oxaliplatin and fluorouracil in patients with metastatic pancreatic cancer after prior therapy: SWOG S1115 study randomized clinical trial. JAMA oncology. 2017; 3(4):516–522.

Di Veroli GY, Fornari C, Wang D, Mollard S, Bramhall JL, Richards FM, Jodrell DI. Combenefit: an interactive platform for the analysis and visualization of drug combinations. Bioinformatics. 2016; 32(18):2866–2868.

Fujimoto Y, Morita TY, Ohashi A, Haeno H, Hakozaki Y, Fujii M, Kashima Y, Kobayashi SS, Mukohara T. Combination treatment with a PI3K/Akt/mTOR pathway inhibitor overcomes resistance to anti-HER2 therapy in PIK3CA-mutant HER2-positive breast cancer cells. Scientific reports. 2020; 10(1):21762.

Ghandi M, Huang FW, Jané-Valbuena J, Kryukov GV, Lo CC, McDonald III ER, Barretina J, Gelfand ET, Bielski CM, Li H, et al. Next-generation characterization of the cancer cell line encyclopedia. Nature. 2019; 569(7757):503–508.

Güvenç Paltun B, Kaski S, Mamitsuka H. Machine learning approaches for drug combination therapies. Briefings in bioinformatics. 2021; 22(6):bbab293.

Ishwaran H, Lu M. Standard errors and confidence intervals for variable importance in random forest regression, classification, and survival. Statistics in medicine. 2019; 38(4):558–582.

Kong W, Midena G, Chen Y, Athanasiadis P, Wang T, Rousu J, He L, Aittokallio T. Systematic review of computational methods for drug combination prediction. Computational and structural biotechnology journal. 2022; 20:2807–2814.

Li H, Li T, Quang D, Guan Y. Network propagation predicts drug synergy in cancers. Cancer research. 2018; 78(18):5446–5457.

Li P, Huang C, Fu Y, Wang J, Wu Z, Ru J, Zheng C, Guo Z, Chen X, Zhou W, et al. Large-scale exploration and analysis of drug combinations. Bioinformatics. 2015; 31(12):2007–2016.

Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The molecular signatures database hallmark gene set collection. Cell systems. 2015; 1(6):417–425.

Liu Q, Xie L. TranSynergy: Mechanism-driven interpretable deep neural network for the synergistic prediction and pathway deconvolution of drug combinations. PLoS computational biology. 2021; 17(2):e1008653.

Medicine N. Rationalizing combination therapies. Nature Medicine. 2017; 23(10):1113. doi: 10.1038/nm.4426.

Menden MP, Wang D, Mason MJ, Szalai B, Bulusu KC, Guan Y, Yu T, Kang J, Jeon M, Wolfinger R, et al. Community assessment to advance computational prediction of cancer drug combinations in a pharmacogenomic screen. Nature communications. 2019; 10(1):2674.

Mitelman F. Mitelman database of chromosome aberrations and gene fusions in cancer. https://mitelmandatabaseisb-cgcorg. 2022; .

O'Neil J, Benita Y, Feldman I, Chenard M, Roberts B, Liu Y, Li J, Kral A, Lejnine S, Loboda A, et al. An unbiased oncology compound screen to identify novel combination strategies. Molecular cancer therapeutics. 2016; 15(6):1155–1162.

Pawitan Y. In all likelihood: statistical modelling and inference using likelihood. Oxford University Press; 2001.

Pemovska T, Bigenzahn JW, Superti-Furga G. Recent advances in combinatorial drug screening and synergy scoring. Current opinion in pharmacology. 2018; 42:102–110.

Plana D, Palmer AC, Sorger PK. Independent drug action in combination therapy: implications for precision oncology. Cancer discovery. 2022; 12(3):606–624.

Subbiah V, Kreitman RJ, Wainberg ZA, Gazzah A, Lassen U, Stein A, Wen PY, Dietrich S, de Jonge MJ, Blay JY, et al. Dabrafenib plus trametinib in BRAFV600E-mutated rare cancers: the phase 2 ROAR trial. Nature medicine. 2023; 29(5):1103–1112.

Tang YC, Gottlieb A. SynPathy: Predicting drug synergy through drug-associated pathways using deep learning. Molecular Cancer Research. 2022; 20(5):762–769.

Tsherniak A, Vazquez F, Montgomery PG, Weir BA, Kryukov G, Cowley GS, Gill S, Harrington WF, Pantel S, Krill-Burger JM, et al. Defining a cancer dependency map. Cell. 2017; 170(3):564–576.

Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, Sajed T, Johnson D, Li C, Sayeeda Z, et al. DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic acids research. 2018; 46(D1):D1074–D1082.

Wu L, Gao J, Zhang Y, Sui B, Wen Y, Wu Q, Liu K, He S, Bo X. A hybrid deep forest-based method for predicting synergistic drug combinations. Cell Reports Methods. 2023; 3(2).

Wu L, Wen Y, Leng D, Zhang Q, Dai C, Wang Z, Liu Z, Yan B, Zhang Y, Wang J, et al. Machine learning methods, databases and tools for drug combination prediction. Briefings in bioinformatics. 2022; 23(1):bbab355.

Zdrazil B, Felix E, Hunter F, Manners EJ, Blackshaw J, Corbett S, de Veij M, Ioannidis H, Lopez DM, Mosquera JF, et al. The ChEMBL Database in 2023: a drug discovery platform spanning multiple bioactivity data types and time periods. Nucleic acids research. 2024; 52(D1):D1180–D1192.

Zhong J, Yan W, Wang C, Liu W, Lin X, Zou Z, Sun W, Chen Y. BRAF inhibitor resistance in melanoma: mechanisms and alternative therapeutic strategies. Current Treatment Options in Oncology. 2022; 23(11):1503–1521.

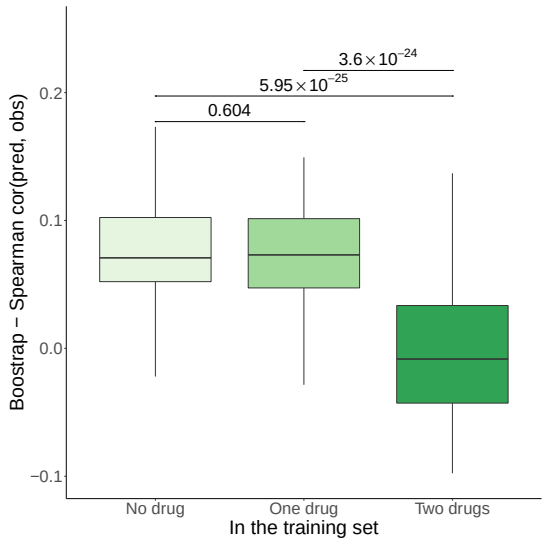

Figure 3—figure supplement 1. Prediction performance of DIPx in the ONeil dataset, grouped by unseen combinations in the training set (x-axist). The y-axis in all box plots shows the Spearman correlation between predicted and observed values in 100 bootstrap replicates.

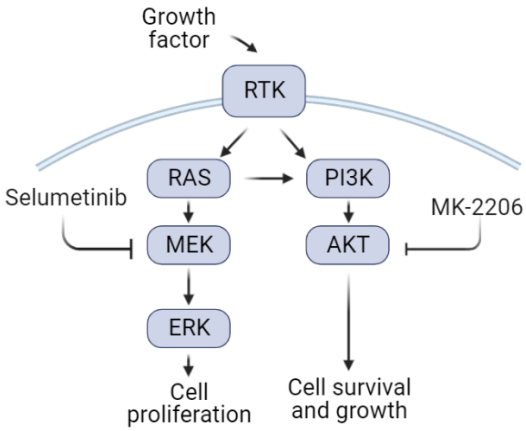

Figure 4—figure supplement 2. A cartoon illustration of the RAS pathway mediated by the Selumetinib + MK-2206 combination

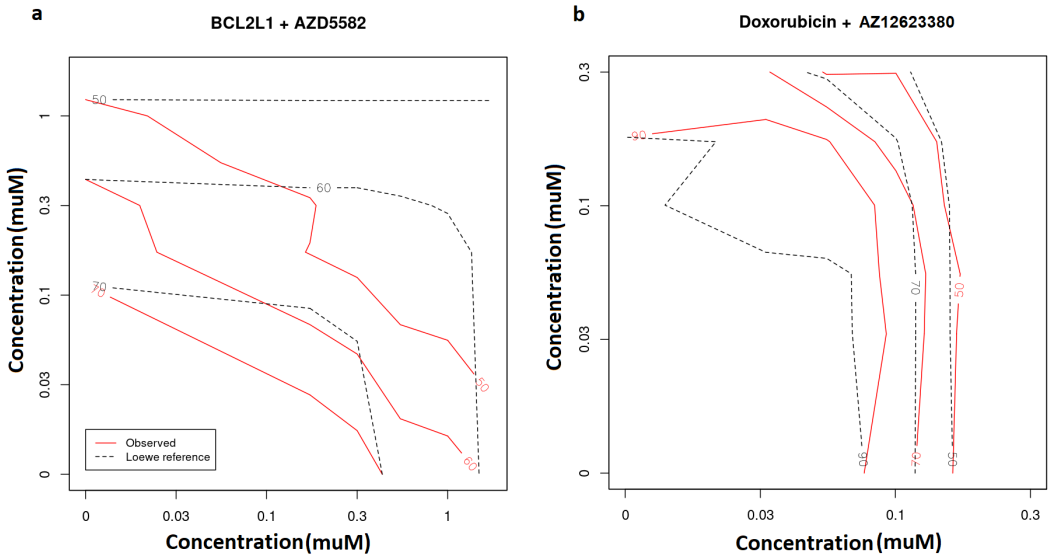

Figure 4—figure supplement 3. Observed (red lines) vs predicted inhibition (black, dash lines) from Loewe reference model in the SW900 cell line treated by the synergistic BCL2L1 + AZD5582 combination (a) and the non-synergistic Doxorubicin + AZ12623380 combination (b). The number in each line presents the percentage of inhibition..

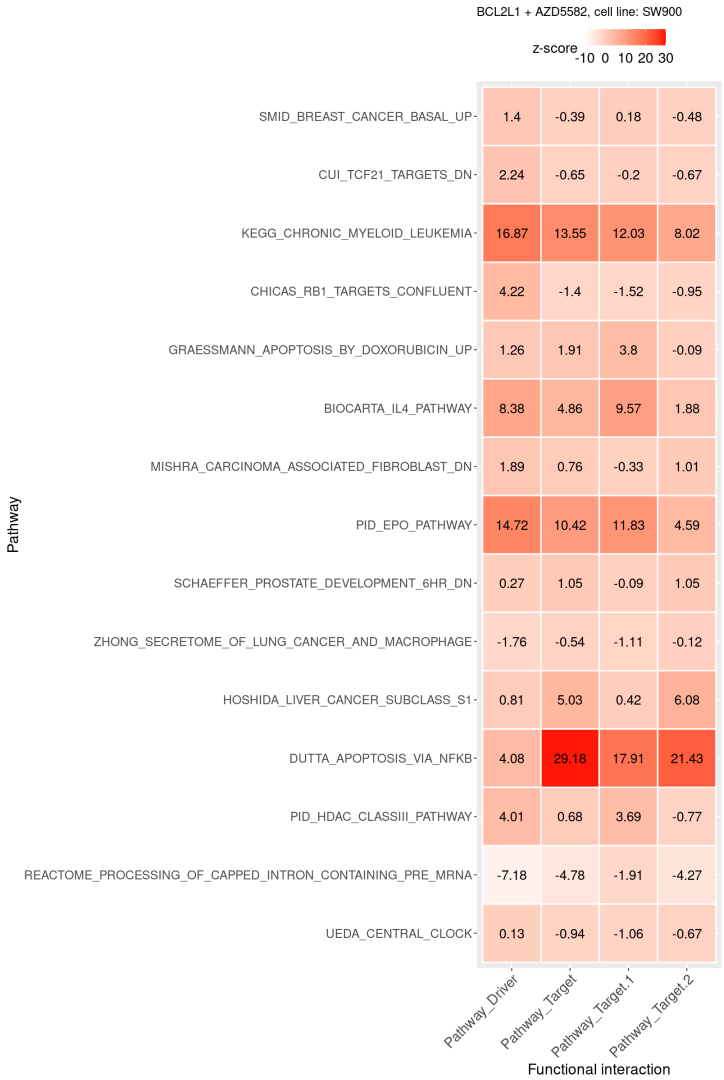

Figure 4—figure supplement 4. Functional interaction (x-axis) between the pathway vs driver genes (1st column), the pathway vs all target genes (2nd), the pathway vs BCL2L1 target genes (3th), and the pathway vs AZD5582 target genes (4th) of the top pathways suggested by DIPx in the SW900 cell line treated with synergistic combination BCL2L1 + AZD5582.

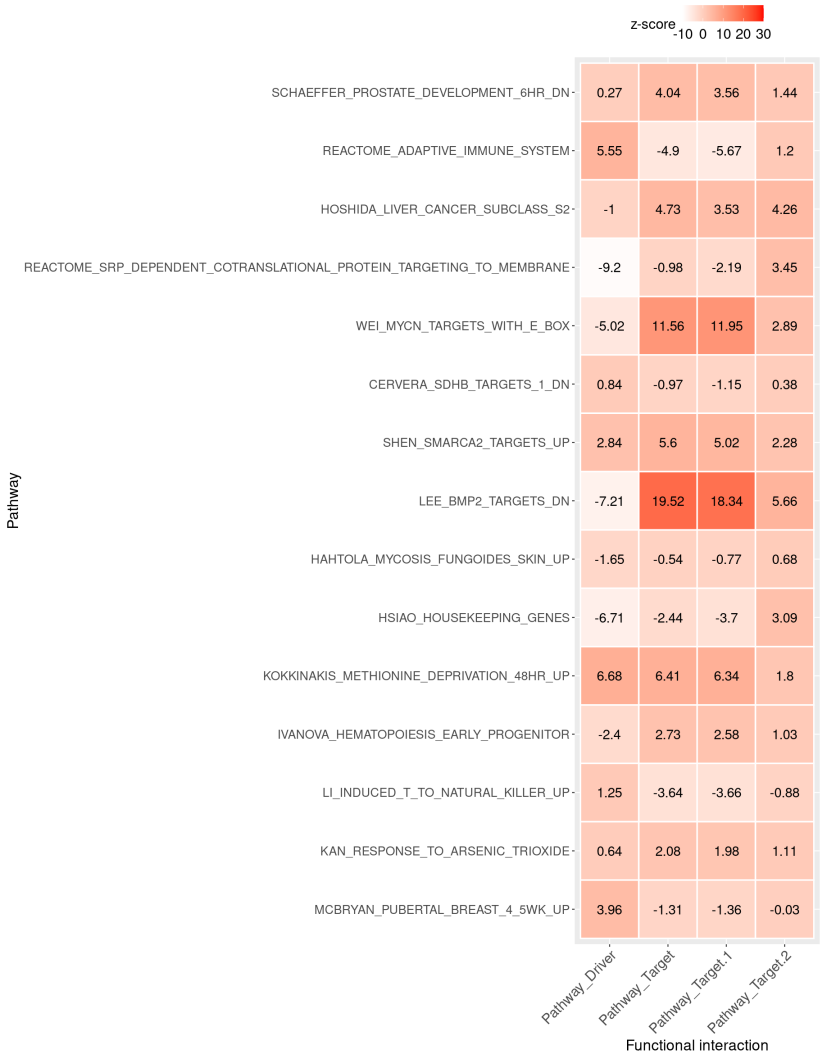

Figure 4—figure supplement 5. Functional interaction (x-axis) between the pathway vs driver genes (1st column), the pathway vs all target genes (2nd), the pathway vs Doxorubicin target genes (3th), and the pathway vs AZ12623380 target genes (4th) of the top pathways suggested by DIPx in the SW900 cell line treated with non-synergistic combination Doxorubicin + AZ12623380.