

Combined kinome inhibition states are predictive of cancer cell line sensitivity to kinase inhibitor combination therapies

Chinmaya U. Joisa^{1,2}, Kevin A. Chen³, Samantha Beville², Timothy Stuhlmiller², Matthew E. Berginski², Denis Okumu², Brian T. Golitz², Michael P. East², Gary L. Johnson², Shawn M. Gomez^{1,2,*}

¹Joint Department of Biomedical Engineering, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA and North Carolina State University, Raleigh, NC, USA

²Department of Pharmacology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

³Department of Surgery, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

* Corresponding Author: Shawn M. Gomez (smgomez@unc.edu)

Protein kinases are a primary focus in targeted therapy development for cancer, owing to their role as regulators in nearly all areas of cell life. Recent strategies targeting the kinome with combination therapies have shown promise, such as trametinib and dabrafenib in advanced melanoma, but empirical design for less characterized pathways remains a challenge. Computational combination screening is an attractive alternative, allowing in-silico filtering prior to experimental testing of drastically fewer leads, increasing efficiency and effectiveness of drug development pipelines. In this work, we generated combined kinome inhibition states of 40,000 kinase inhibitor combinations from kinobeads-based kinome profiling across 64 doses. We then integrated these with transcriptomics from CCLE to build machine learning models with elastic-net feature selection to predict cell line sensitivity across nine cancer types, with accuracy $R^2 \sim 0.75-0.9$. We then validated the model by using a PDX-derived TNBC cell line and saw good global accuracy ($R^2 \sim 0.7$) as well as high accuracy in predicting synergy using four popular metrics ($R^2 \sim 0.9$). Additionally, the model was able to predict a highly synergistic combination of trametinib and omipalisib for TNBC treatment, which incidentally was recently in phase I clinical trials. Our choice of tree-based models for greater interpretability allowed interrogation of highly predictive kinases in each cancer type, such as the MAPK, CDK, and STK kinases. Overall, these results suggest that kinome inhibition states of kinase inhibitor combinations are strongly predictive of cell line responses and have great potential for integration into computational drug screening pipelines. This approach may facilitate the identification of effective kinase inhibitor combinations and accelerate the development of novel cancer therapies, ultimately improving patient outcomes.

Keywords: Kinase signaling, precision medicine, systems biology, drug response prediction.

1. Introduction

Protein kinases, which serve as the primary conduits for information transfer within cells, are often implicated as key drivers in cancer development and have become a cornerstone in current targeted therapies [1]. The rapid expansion of kinase inhibitor therapies as an oncology drug class is exemplified by the FDA's approval of nearly 60 such therapies over the past 20 years [2]. Despite their initial promise, kinase-targeting monotherapies frequently give rise to resistance [3], in part due to the dynamic nature of the kinase network, i.e., the “kinome,” which has been shown to reprogram and respond to the inhibition of single kinases by upregulating expression of partner pathways [4–6]. This necessitates the development of novel strategies to effectively target the kinome and harness the vast array of potential drug targets it offers.

One emerging strategy to counteract resistance involves the design of combination therapies, which perturb multiple targets with two or more drugs. These targets may be either known compensatory pathway partners, referred to as "horizontal pathway inhibition," or multiple targets within the same pathway, known as "vertical pathway inhibition" [7]. This approach has recently gained traction with the FDA approval of the combination of trametinib and dabrafenib for treating advanced melanoma [8]. This combination therapy "vertically" targets both BRAF and MEK within the RAF-MEK-ERK (MAPK) pathway, demonstrating the potential effectiveness of this strategy. However, this method of empirical design of combination therapies is not feasible for less characterized kinase pathways, and the sheer number of possible combinations of potential kinase targets (2^{500}) prevents brute-force screening or drug design.

To circumvent this issue, computational screening offers an appealing alternative, enabling the prediction of effective drug combinations in-silico prior to testing a reduced pool of potential combinations in-vitro. This method streamlines the drug development process, and when combined with patient-specific genomic profiling, can also enable personalized drug combination selection to potentially achieve resistance-proof responses in patients.

In recent years, a variety of computational approaches have been developed to predict combination therapy responses for cancer drug screening [9,10]. Most of these methods primarily rely on drug structure characteristics and cancer-specific baseline genomic profiling to predict effective drug combinations, spurred by advancements in the high-throughput acquisition of these data types. For example, a high-dimensional tensor-based modeling strategy used similar data and achieved impressive accuracy (Overall $R^2 \sim 0.8$) in predicting response to combination therapies, validated experimentally [11]. This approach and others employ intricate neural network architectures that, while capable of producing high performing models, can be challenging to interpret, posing a barrier to the broader understanding of their underlying mechanisms. Tree-based machine learning models on the other hand, although simpler and sometimes less powerful, are generally considered interpretable depending on the type of data fed to them [12]. Notably, drug-protein interactions, which are intuitively central to the process of phenotype reversal, have been relatively underexplored in these computational approaches. In part, the minimal amount of drug-target information leveraged in current response prediction efforts is because of the sheer amount of data generated by genomics and molecular fingerprinting, generating thousands of features for each measurement, while drug target data has been generally sparse with only a few annotated targets per drug. However, recent advances in technology to profile the interactions of clinical drugs with all the members of the kinome represent an unprecedented ability to measure drug-target information across ~ 500 proteins simultaneously in a quantitative manner [13,14]. The breadth, density, and ease of acquisition of this data, often measured at multiple dose points, is ideal for integration into machine learning models that can leverage diverse data types for drug response prediction.

Specifically, recent advances in proteomics techniques have facilitated the large-scale characterization of drug-kinase interactions, providing valuable information on the extent to which the entire kinome is inhibited by specific drugs or drug combinations. A landmark paper in 2017 used a mass spectrometry-based assay that used promiscuous kinase-binding compounds immobilized on beads to measure the binding competition between any given inhibitor and any

given kinase (henceforth called the “kinobeads” assay) [15]. Using this assay, the kinome-wide binding profiles for ~230 clinical kinase inhibitors at eight doses each were elucidated using cancer cell lysates, forming the largest in-cell drug-target binding database publicly available at this time. The data generated from these assays allow interrogation of how clinical and investigational drugs interact with the entire kinome on an unprecedented scale. By analyzing the degree of inhibition of all kinases simultaneously for a given inhibitor, we can treat this as characterizing the degree of departure from the “baseline kinome state”, thus moving through drug-induced alteration of multiple kinase activities to a new “kinome inhibition state”. Given the degree to which modulation of the kinome alters cellular state and downstream behavior, these baseline kinome states and kinome inhibition states can be directly connected to various measured cellular phenotypes. We have recently demonstrated this idea by showing that kinome inhibition state is significantly predictive of cancer cell responses to kinase inhibitor monotherapies when integrated with cancer-specific information, such as baseline transcriptomics, using tree-based machine learning models [16].

In this work, we show that by combining the inhibition states of two kinase inhibitors, we can generate a hypothetical “combined” inhibition state for an untested inhibitor combination. In this manner, we can rationally use all combinatorial kinome inhibition states to sample all possible kinase target combinations, hypothetically including all pathway partners. By integrating these inhibition states with cancer-specific baseline transcriptomics, we demonstrate that the combined inhibition state can predict the sensitivity of cancer cell lines to inhibitor combination treatments from the NCI-ALMANAC dataset using interpretable machine learning models. We further validate these models experimentally by examining novel inhibitor combinations in a PDX-derived triple-negative breast cancer (TNBC) cell line. By focusing on dual-inhibitor drug-kinase interactions combined with cancer-specific baseline genomic profiling, we can enhance computation combination drug screening pipelines with combinatorial kinase targeting. Furthermore, this approach lays the foundation for the rational design and a priori prediction of combination kinase inhibitor treatments for patients with the potential to ultimately reduce single kinase inhibitor resistance acquisition by prior rational targeting of partner pathways and associated kinases.

2. Results

2.1. Creating a Set of Combined Kinome Inhibition States Representing Current and Potential Kinase Inhibitor Combination Therapies

In this work, we have focused on a specific set of 200 kinase inhibitors characterized using the kinobeads assay [15]. These inhibitors were profiled in-cell for their interactions with ~500 kinases and kinase-interacting proteins, across eight doses. From this data, as described previously (insert citation), we extracted monotherapy “kinome inhibition states”, denoting the degree to which they inhibit each kinase in the kinome at eight doses on a scale of 0-1 (0 is complete inhibition and 1 is no inhibition of a given kinase). We next tested different methods to approximate the kinome inhibition state of a kinase inhibitor combination. Intuitively, this can be thought of as simply superimposing two individual monotherapy inhibition states, but for the few cases where different inhibitors target the same kinase, we found ways to accurately reflect the resulting effect on the kinome. Here, we tested combining monotherapy kinome inhibition state vectors through addition,

multiplication, truncated multiplication (excluding kinase inhibition values >1). All three methods were compared for downstream model performance.

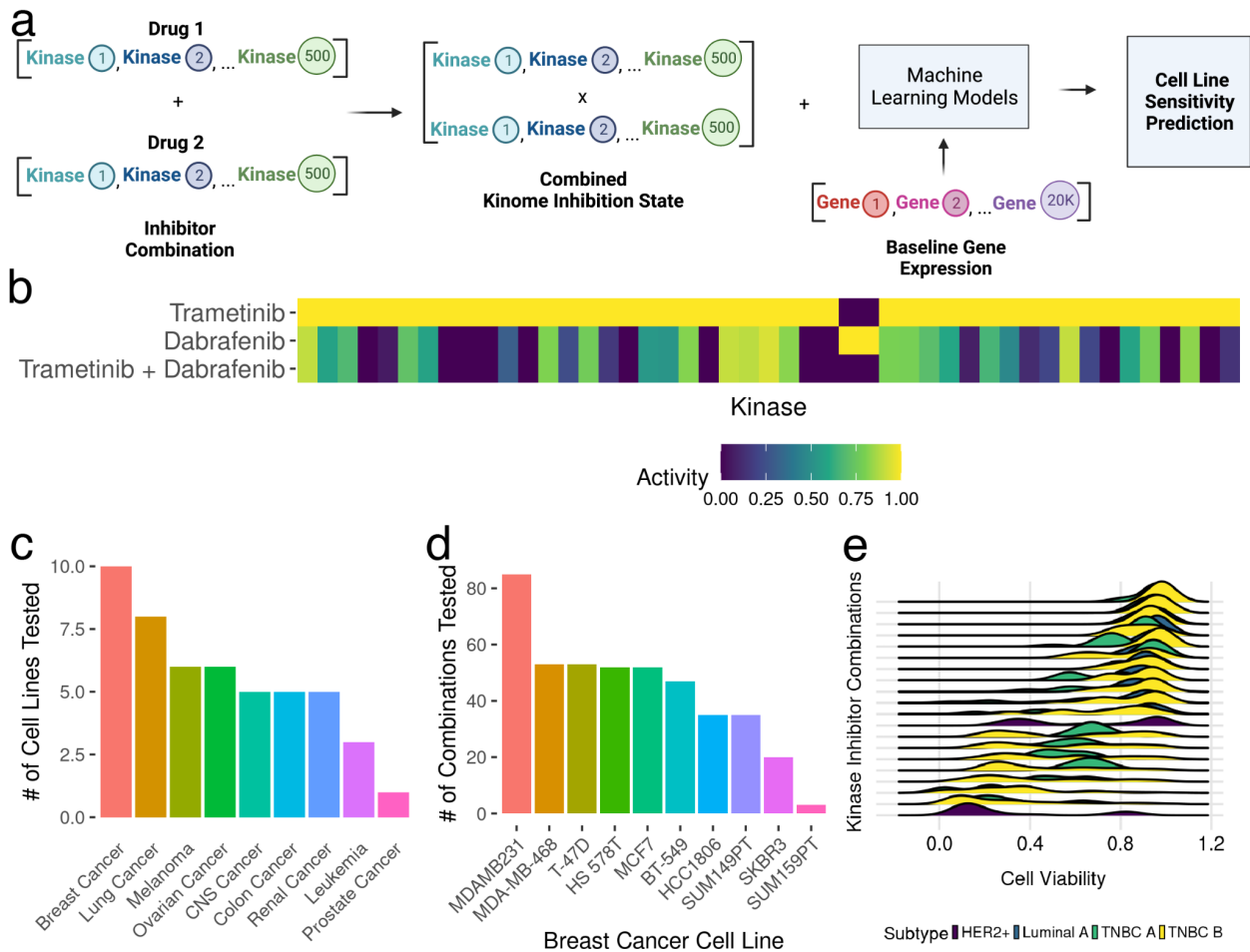


Figure 1. Kinome inhibition State Combination Modeling and Data Overview. (a) Schematic of modeling pipeline. (b) Heatmap showing the inhibition state of individual kinase inhibitors (row 1 and 2), and the hypothetical “combined” inhibition state for the two inhibitors (row 3) (c) Bar plot showing number of cell lines tested per cancer type in training data set (d) Bar plot showing number of unique combinations tested per cell line for the breast cancer subset of the training data set (e) Ridge plots showing cell viability (x-axis) variation for a random subset of different kinase inhibitor combinations (y-axis) in the NCI-ALMANAC data for breast cancer cell lines. Different breast cancer subtypes are represented with differing colors.

After combining the individual inhibition states, we were left with a dataset describing all possible pairwise combinations of ~220 kinase inhibitors. These ~45,000 combinations represent the kinome inhibition states of existing clinical therapies (example), therapies currently in clinical trials (example), as well as potential therapies. Together, they interrogate a search space that includes nearly every known kinase on the phylogenetic tree (Fig S1).

2.2. Connecting Inhibited Kinome States with Cancer Cell Line Combination Sensitivities

Next, we linked the data set describing kinase inhibitor combinations to their cell sensitivity phenotypes in the large-scale ALMANAC drug combination screen. The ALMANAC screen contains cell sensitivity data for 53 kinase inhibitor combinations, over ~200 unique dose combinations for 45 cell lines across 9 cancer types. Additionally, previous high-throughput combination screens conducted in our lab in breast cancer offered data for 56 inhibitor combinations in four cell lines. Ideally, we would like exact matches between the dose at which kinome inhibition state is profiled and the dose at which cell sensitivity was measured. However, there are very few exact matches between the datasets. To overcome this, we found the nearest dose (6 exact matches, 14 nearest matches at maximum differing by 1 μ M) at which kinome inhibition was profiled for each cell sensitivity measurement and connected the two datasets using these dose matches.

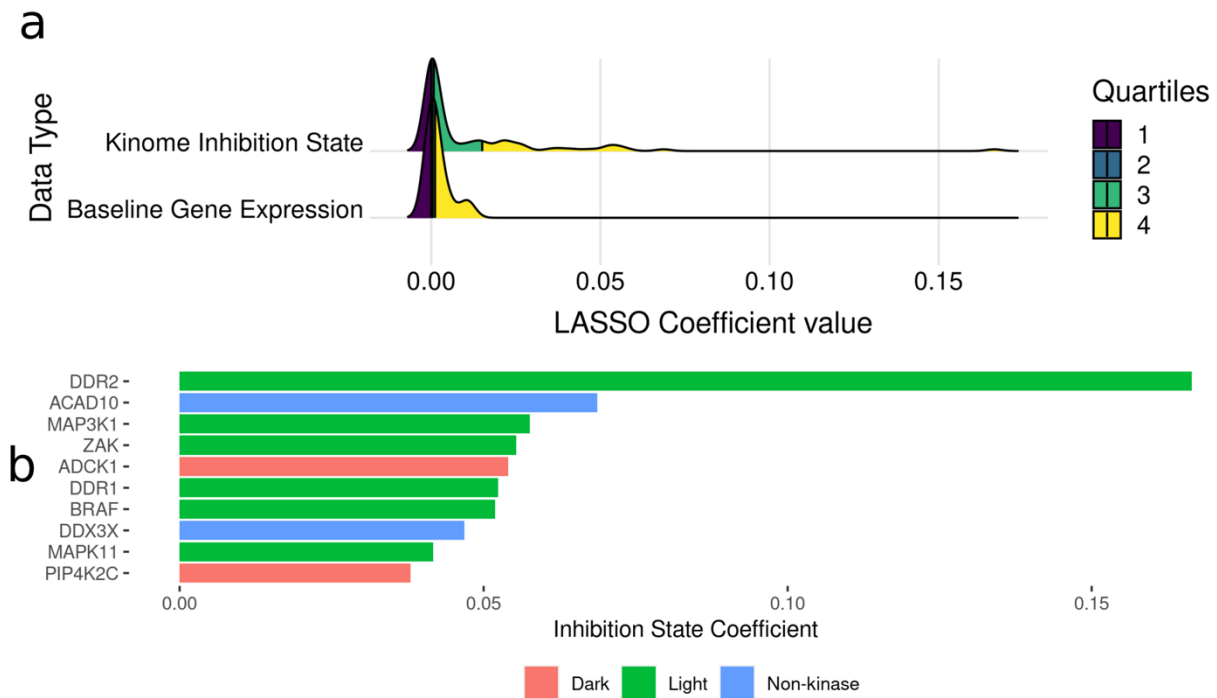


Figure 2. Feature Selection using an Elastic-net Regression Model against Cancer Cell Line Sensitivity. (a) Ridge plot showing the distribution of LASSO coefficient sizes as a metric for feature importance, for each feature type (b) Horizontal bar plot showing kinases with the largest elastic-net coefficient values, coloured by whether they are defined as “understudied” (Dark) or “well-characterized” (Light).

Additionally, we added cell line specific information to the dataset to complement the drug-specific kinome inhibition states. The CCLE database contains baseline transcriptomics data for ~1500 cancer cell lines, and almost all of the cell lines included in our data set were represented. Using this, we further added baseline gene expression into the dataset, now containing kinase inhibitor combinations, their inhibition state of the kinome, the cell line sensitivity to their treatment,

as well as that cell line's baseline gene expression. In this way, the dataset connects the kinome inhibition states of inhibitor combinations to their cell sensitivity phenotypes.

The collected dataset represents a total of eight major cancer types, with the majority having ~7 cell lines represented each, while breast cancer had the most representation (11 cell lines). To ensure that the machine learning model downstream could find cancer-specific linkages between the kinome and cell sensitivity, we split the dataset into eight individual cancer type datasets and conducted all modeling on each data split in parallel.

2.3. Elastic-Net Feature Selection Reveals Kinome Inhibition States to be Most Informative

In our collected dataset, kinome inhibition states and baseline gene expression together represent ~20,000 variables or “features” that could affect the phenotype of cell sensitivity to kinase inhibitors. It is both practically prohibitive and ineffective to build models using all available features, and so keeping in mind computational efficiency we sought to filter down the dataset to include only the most informative features. To accomplish this “feature selection”, we built our machine learning pipeline starting with an elastic-net regression [17] model built against the outcome of cell sensitivity. This generated coefficients for each feature, with the absolute value of the feature coefficient directly proportional to its predictive value for the outcome. We ensured non-informative features were not included in modeling by only considering features with non-zero coefficients. We fit the model on the entire dataset to visualize a snapshot of the feature coefficients globally. This revealed overwhelmingly larger coefficients for kinome inhibition states compared to baseline gene expression (Fig 2a), thus indicating that kinome inhibition states were globally more informative for cell sensitivity prediction compared to baseline gene expression.

For downstream model building, the data set was split into a training and testing set five times (five-fold cross validation). For the training set data to not have any influence on the test set (to prevent data leakage), the elastic net model is fit on only the training data, and features are selected within each fold. Parameters for the elastic net model and hyperparameters for the tested model types were also tuned this way.

2.4. Machine Learning Models Can Predict Cancer Cell Line Sensitivity to Combination Therapies by Integrating Kinome Inhibition States and Baseline Transcriptomics

After data set preparation and feature selection, we built machine learning models that can predict cell sensitivity to kinase inhibitor combinations. For each cancer type, three machine learning model types were tested: random forest, boosted trees (xgboost) and deep neural networks. Xgboost performed the best for all cancer types, with type-specific performance largely dependent on abundance of data in the training set (Fig 3b). The most abundant cancer type (breast) had the best performing model with an R^2 score of 0.93 (Fig 3b) while the lowest performing model was prostate cancer with $R^2 = 0.73$. Given our previous lab experience with breast cancer, we chose the breast cancer model for downstream experiments and validation.

Additionally, since the best-performing model was tree-based gradient boosting, we were able to further analyze the model using computed feature importance to find the most informative

features in the data set based on the feature importance metric. Similar to the feature selection output, we saw much higher feature importance scores overall for kinome inhibition states when compared to baseline gene expression, and several kinases implicated in breast cancer dysfunction had high importance scores, such as MAP2K1/2 and EGFR(Fig. 3c).

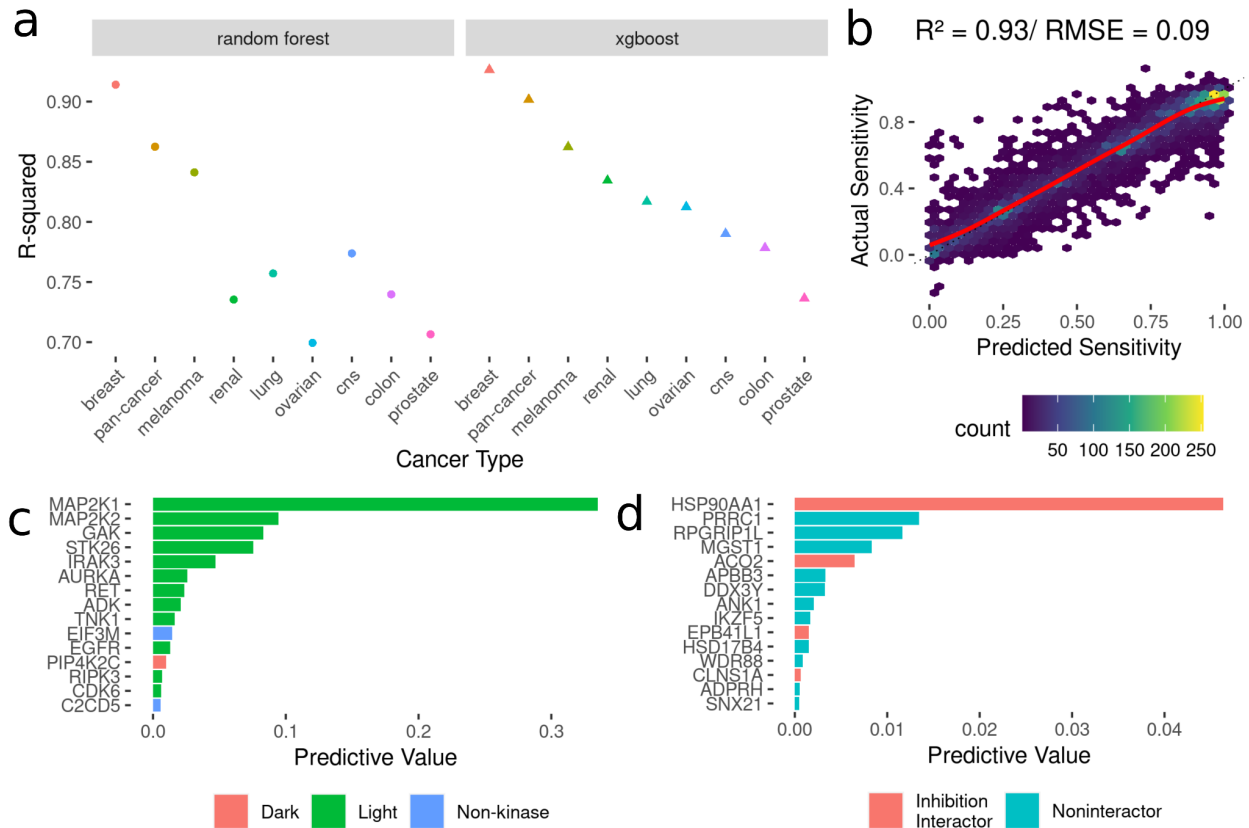


Figure 3. Development of Models to Predict Cancer Cell Line Sensitivities to Kinase Inhibitor Combination Therapies from Kinome Inhibition States. (a) Model performance metrics (R-squared) for Random Forest (dots) and XGBoost (triangles). (b) Scatter Plot of predicted sensitivity values from the best-performing model vs actual sensitivity values. The red line indicates a smooth fit through the data points. (c) Horizontal bar plot showing model importance of individual kinase inhibition states by importance values. (d) Horizontal bar plot showing model importance of individual baseline gene expression by importance values.

2.5. Experimental Validation of Model Predictions in a PDX-Derived Triple Negative Breast Cancer Cell Line was Successful.

We demonstrated that machine learning models using the kinome inhibition states of combination therapies along with cell-specific baseline gene expression could robustly predict cell sensitivity in multiple cancer types. However, to see if these predictive models could extend to real-world experiments, we experimentally validated 35 kinase inhibitor combinations in a PDX-tumor derived cell line(Fig 4A).

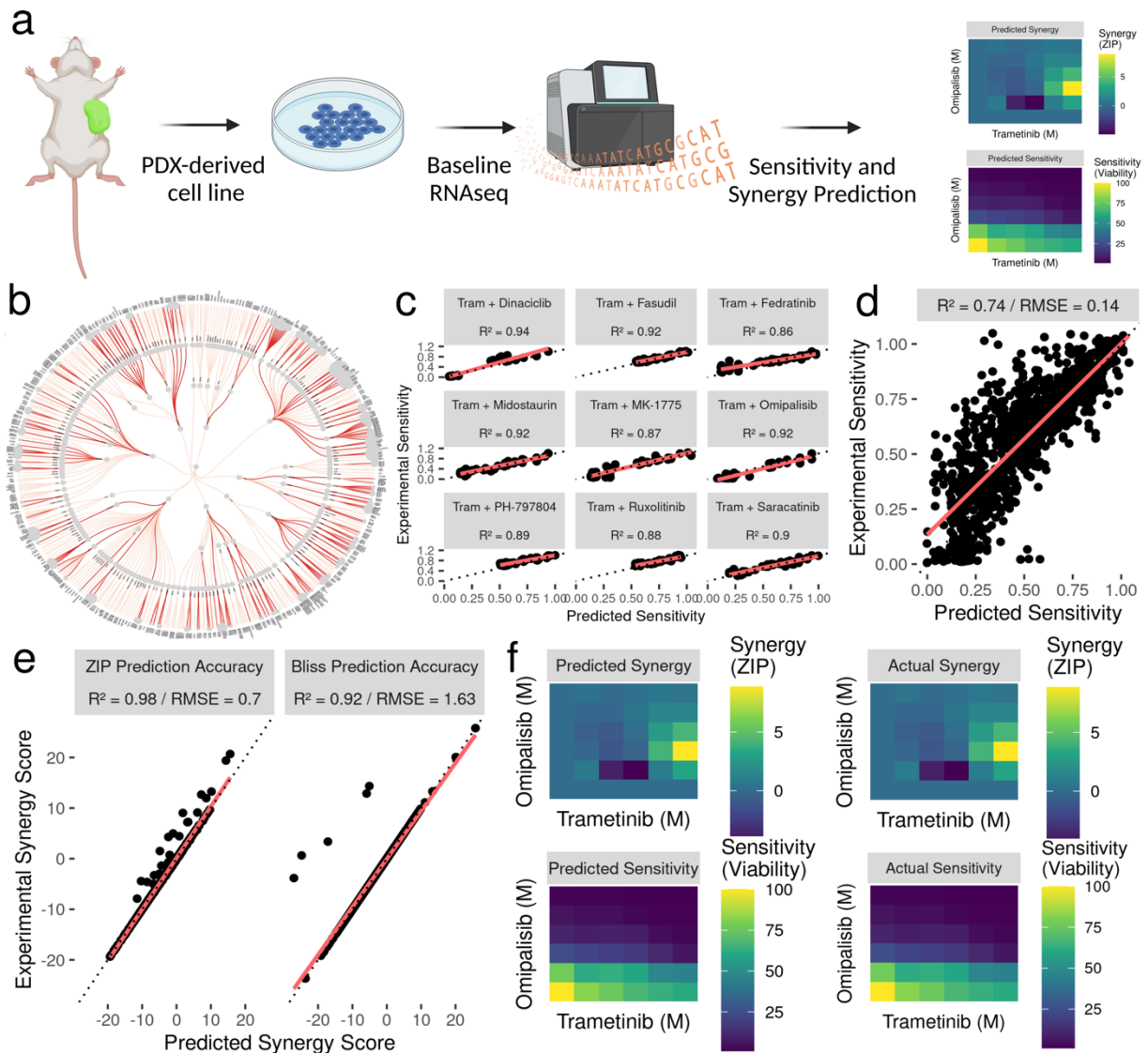


Figure 4. Experimental Validation of Model through a Trametinib Combination Screen in the WHIM12 Patient-Derived TNBC Cell Line. (a) Schematic showing experimental validation pipeline for the WHIM12 PDX-derived cell line. (b) Kinome phylogenetic map showing diversity of kinome targeted (red = inhibited by a validated kinase inhibitor combination). (c) Grid of scatter plots showing accuracy of top nine tested combinations. For all scatter plots, the dashed line indicates where perfect predictions would lie and the red line shows a linear fit through the data. Quantitative accuracy is represented by the R-squared score. (d) Scatter plot showing the global accuracy of model. (e) Grid of scatter plots showing accuracy of model predicted synergy scores compared to experimentally measured synergy scores across two metric types (ZIP, Bliss). (f) Grid of heatmap plots showing comparison of predicted vs experimentally measured sensitivity and synergy for the highly synergistic trametinib / omipalisib combination.

High-throughput cell line drug screens have been widely documented to suffer from a lack of reproducibility and poor translation to more complex samples like patient tumors. We sought to test

whether our model of cell sensitivity in breast cancer, trained on 11 well-characterized immortalized cell lines, could effectively predict cell sensitivity in a PDX (Patient-Derived Xenograft) derived cell line. We chose the WHIM12 PDX-derived cell line, which was generated from a highly chemo-resistant TNBC tumor [18]. Previous experiments in the lab had conducted a drug combination screen in the WHIM12 cell line, out of which 35 kinase inhibitors were tested in combination with trametinib. Complementary baseline gene expression data was also generated through RNAseq. Using these in-house data, we were able to input the unseen WHIM12 gene expression into the trained model and predict the cell sensitivity outcomes of the conducted drug combination screen. We achieved robust prediction accuracy (Global $R^2 = 0.74$ / RMSE = 0.14) in predicting exact cell viability in response to treatment with 35 kinase inhibitor combinations, across 64 dose combinations (Fig 4c, d).

2.6. Model Predictions Reveal Known Synergy in trametinib/omipalisib Combination

The model predictions in the WHIM12 cell line were further interrogated for potential synergy. We generated synergy scores for all 35 combinations at each of the 64 dose points using the R package SynergyFinder [19] based on four different metrics: Zero-Interaction Potency [10] (ZIP), Bliss Independence [20], Highest Single-Agent (HSA), and Loewe Additivity [21]. Additionally, we generated similar synergy scores using the actual experimental data generated for validation as a comparison. We found a high degree of similarity (Global $R^2 \sim 0.94$ / RMSE ~ 0.5) between predicted and actual synergy, with trametinib + omipalisib as our most synergistic predicted combination, with a ZIP score of ~ 8 at certain dose combinations (Fig 4e, f). This is significant as the model predictions were in a TNBC PDX-derived line, and the trametinib/omipalisib combination represents the popular strategy of simultaneously targeting the MAPK and PI3K pathways [22].

3. Methods

Data Sources. The kinome profiling data set from the kinobeads assay was downloaded from the supplementary materials of Klaeger et al. 2017 [15]. For cancer cell line sensitivity to kinase inhibitor combinations, data was downloaded from (1) NCI-ALMANAC: cell sensitivity data was downloaded from the NCI wiki database (<https://wiki.nci.nih.gov/display/NCIDTPdata/NCI-ALMANAC>) and (2) Supplementary materials of previous lab combination screens published in Beville et al. 2019 [28] and Stuhlmiller et al. 2015 [29]. The CCLE gene expression set (“CCLE_expression.csv”) was downloaded from the DepMap portal (<https://depmap.org/portal/download/all/>) to create the set of cancer cell lines and their gene expression characteristics. In-house baseline gene expression data for the PDX-derived WHIM12 line was downloaded from the GEO repository for the Zawitowski et al. paper [26] (GSE87424).

Data Preprocessing. The scripts implementing these descriptions are all available through github.

Klaeger et al. Kinobead Kinase Inhibition Profiles: As previously described [16], we read the values from the supplemental data table into R and produced a filtered list of kinase and kinase

interactor relative intensity values. We imputed missing values with the default “no interaction” value of 1 and truncated likely outlier values to the 99.99 percentile (3.43).

Creating the Combination Inhibition State Data Set: To create a “combined” inhibition state of a given kinase inhibitor combination, we sought to superimpose the inhibition states of two individual states at specific doses. There were eight doses measured for each individual inhibitor, thus there were 64 possible combinations for each combination. We took the monotherapy kinome inhibition states from the Klaeger et al. set and computed a “combined” inhibition state for each kinase, based on three different combination schemes:

1. Simple Multiplicative: The simple conditional probability rule assumes two independent events (A and B in Eq. 1). Since the default “no interaction” inhibition value is 1, for kinases that are not targeted by both inhibitors simultaneously, the “combined” inhibition state (C') value is simply either one in monotherapy.
2. Truncated Multiplicative: A minority of measured kinase inhibition states (~1%) have values > 1 in the Klaeger et al. dataset, a possible artifact from the mass spectrometry measuring process. To avoid inflating those values, all >1 values were truncated at 1 and simple multiplication was performed as described above (Eq. 2).
3. Addition: All kinase inhibition states were inverted into “Percent Inhibition” values (A' and B'), where 0 denotes no inhibition and 100 denotes complete inhibition. Then, when two inhibition states were combined, they were added together and truncated at a max value of 100 (Eq. 3).

$$\text{Eq. 1.} \quad C' = A * B$$

$$\text{Eq. 2.} \quad C' = \min(1, A) * \min(1, B)$$

$$\text{Eq. 3.} \quad C' = \min(100, A' + B')$$

All three methods were tested in downstream modeling, resulting in minor variation. Truncated multiplied vectors were slightly more predictive (R^2 score of ~0.01 greater) so we used that scheme for all downstream modeling. In this way, we were able to compute hypothetical “combined” inhibition states for all possible combinations of ~220 inhibitors, altogether comprising ~2,000,000 combined inhibition states.

Dataset of Cancer Cell Line Sensitivity to Kinase Inhibitor Combinations: The cell sensitivity dataset from NCI-ALMANAC and previous lab publications were filtered to contain only kinase inhibitor small molecules, then summarized over replicates and converted to cell viability (1 = fully viable cell and 0 = full cell death). Relevant cancer types were annotated and individual cancer type datasets were subsetted for downstream cancer type-specific modeling.

Matching of Kinase Inhibitors between Inhibition State Dataset and Cell Line Sensitivity Dataset: The drug names from each dataset were read into R, and the package Webchem [30] was used to retrieve PubChem compound IDs (cid's). The two sets of drug names were then matched based on these reference IDs, with a total of ~100 matches between the two sets.

Baseline Gene Expression from CCLE: Data was preprocessed as described before [31] from the “CCLE_expression.csv” file. Cell line names were matched manually between CCLE and the NCI naming scheme. All cell lines represented in NCI-ALMANAC had a match in the CCLE database.

String: The STRING database [32] was processed as described previously [31] to annotate kinases and kinase interacting genes.

Modeling Techniques: To assess our models we used a random 5-fold cross validation strategy. We implemented Elastic-net regression using the glmnet engine [33] for the feature selection scheme [17]. We compared the performance of three model types using this strategy: random forest using the ranger engine [34] and gradient boosting using the XGBoost engine [35]. Model performance was assessed by the R-squared value between predicted and actual outcome within the cross-validation scheme. For each model type and for the feature selection model, we tuned sets of 20 hyperparameters to find the best possible performer as follows: (a) Elastic-net: Penalty (0 - 0.1), Regularization (0.1-1) (b) Random Forest: Trees (100 - 2000) (c) XGBoost: Trees (100 - 1000), Tree Depth (4 - 30). After final model selection, we fit the model on the entire dataset and then made predictions on the experimental validation data.

All of the data and code written to support this paper is available through github (https://github.com/gomezlab/kinotype_combination_prediction).

Experimental Validation. 6x6 dose combination screens were performed in the WHIM12 cell line as described in Beville et al. 2019 [28]. Briefly, cells were seeded in 384-well plates and dosed with drug after 24h. The screening library was tested for growth inhibition alone or in combination with Trametinib across 6 doses: 10 nmol/L, 100 nmol/L, 300 nmol/L, 1 μ mol/L, 3 μ mol/L, and 10 μ mol/L. 0.1% DMSO was included as the control for growth inhibition on each plate. Plates were incubated at 37°C for 96 hours and lysed using CellTiter-Glo Reagent (Promega, catalog. no. G7570). Luminescence was measured using a PHERAstar FS instrument and growth inhibition was calculated relative to DMSO-treated wells.

4. Discussion

Kinase inhibitors are one of the fastest growing drug classes for cancer therapy, with ~62 FDA approved in total against neoplasms [2]. With 500 potential druggable targets, there is significant interest in streamlining the kinase inhibitor screening process. We have previously introduced [16,23,24] the idea that the full spectrum of a given inhibitor's effect on the kinome as measured by recent advances in kinobead-competition/MS technology [15] can be represented as a "kinome inhibition state", i.e. a vector representing the effect of a given inhibitor on the kinome as a whole.

In this work, we have extended this idea to represent the kinome inhibition state of a combination of inhibitors, using a multiplicative probability model to "combine" the inhibition states of two given kinase inhibitors. By generating these "combined" inhibition states, we can vastly expand the search space targeted by inhibitor monotherapies, sampling all possible combinations of currently available therapies. To accomplish this, we used publicly available drug-kinome interaction data to generate snapshots of the combined effect of a combination therapy on the protein kinome. We then linked these kinome inhibition states of inhibitor combinations to cancer cell sensitivity phenotypes to combination treatment, creating a framework for predicting the efficacy of combination therapies in different cancer types.

We fit tree-based machine learning models on this linked data set to robustly predict precise cancer cell line sensitivity and synergy for untested kinase inhibitor combinations therapies and validate those predictions in complex patient derived samples. gradient-boosted tree models were highly accurate across cancer types (R^2 0.75-0.93), comparable to two recent neural-network driven

attempts to predict cell line response to drug combinations [9,11]. We chose to validate our model predictions in the PDX-derived WHIM12 line, reasoning that PDX-derived cell lines retain many of the molecular and genetic features of the xenografted original tumors. We were able to show that the models performed robustly on novel gene expression data ($R^2 \sim 0.74$), representing its ability to extend to complex and clinical-adjacent samples compared to well-characterized cell line data.

One of the strengths of tree-based models is that they are considered to be interpretable through feature importance computation [12,25]. Using this, we were able to investigate the “black box” and query which specific kinase inhibition states and baseline genes were most predictive of cell sensitivity. We found that for the breast cancer model, the inhibition of the kinases MAP2K1/2 were the most informative by far. This is intuitive considering the most abundant kinase inhibitor in the dataset is the allosteric MEK inhibitor trametinib, but it must be noted that MEK inhibition is always only just one half of the kinome targeting in the combination. There has been increasing clinical interest recently in targeting the PI3K and MAPK pathways [22], and our lab has shown before that MEK1/2 inhibition in TNBC by trametinib induces widespread transcriptional adaptation, and that there is potential for clinical efficacy in complementary kinome targeting with trametinib [26]. Since our model’s sensitivity predictions can effectively simultaneously predict synergy, our top synergy prediction for breast cancer according to the ZIP metric was trametinib and omipalisib, which we were able to validate experimentally in the WHIM12 line. This indicates that from the breast cancer screening data, the model was able to learn that targeting the complementary PI3K and MAPK pathways is effective and synergistic in TNBC.

Interestingly, the predicted high-synergy combination of trametinib/omipalisib was recently in phase I clinical trials for advanced solid tumors but failed due to patient intolerability [27]. This highlights some limitations of our modeling approach. Ideally, kinome inhibition state would be one of many different drug modalities included for response prediction, and we plan to further expand these models in the future by considering toxicity, drug structure and cancer-describing multi-omic data types not limited to baseline gene expression. Additionally, in this proof-of-concept study we utilized multiplicative probability models to generate the “combined” inhibition state of two inhibitors on the kinome, by assuming that the inhibition of a given kinase is mutually exclusive from that of other kinases. We know that kinases function physiologically as part of complex signaling networks, and their inhibition may have downstream effects on other kinases and signaling pathways. To address this limitation, future models will incorporate more biologically representative schemes to hypothesize combined kinome inhibition states.

In summary, through this work we demonstrate the development of a framework for predicting the efficacy of combination therapies in different cancer types using just kinome-drug interactions and baseline gene expression. We generated the combined “kinome inhibition state” and linked these states to cancer cell sensitivity phenotypes. First, we were able to show that a given combination therapy’s cancer-agnostic interaction with the kinome was far more informative than baseline genomics in predicting downstream response. This is intuitive fundamentally, as drug-protein interactions are the primary means of drug effect on physiology, but this type of data is still underutilized in computational screening approaches. We then used machine learning models to predict cell line sensitivity and synergy for untested kinase inhibitor combination therapies and validated those predictions experimentally in complex patient derived samples.

Acknowledgements

We would like to thank UNC Research Computing for access to the computational resources necessary for this work. We would like to thank Michael P. East for his help with data compilation. This work was supported by grants through the National Institutes of Health (Grant #s CA274298, CA233811, CA238475, DK116204)

This is a preprint of an article submitted for consideration in Pacific Symposium on Biocomputing © 2024 [copyright World Scientific Publishing Company] [psb.stanford.edu]

References

1. Kothari V, Wei I, Shankar S, Kalyana-Sundaram S, Wang L, Ma LW, et al. Outlier kinase expression by RNA sequencing as targets for precision therapy. *Cancer Discov.* 2013;3: 280–293.
2. Roskoski R Jr. Properties of FDA-approved small molecule protein kinase inhibitors: A 2023 update. *Pharmacol Res.* 2023;187: 106552.
3. Jiang L, Li L, Liu Y, Lu L, Zhan M, Yuan S, et al. Drug resistance mechanism of kinase inhibitors in the treatment of hepatocellular carcinoma. *Front Pharmacol.* 2023;14: 1097277.
4. Duncan JS, Whittle MC, Nakamura K, Abell AN, Midland AA, Zawistowski JS, et al. Dynamic reprogramming of the kinome in response to targeted MEK inhibition in triple-negative breast cancer. *Cell.* 2012;149: 307–321.
5. Engelman JA, Zejnullahu K, Mitsudomi T, Song Y, Hyland C, Park JO, et al. MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. *Science.* 2007;316: 1039–1043.
6. Chandarlapaty S. Negative feedback and adaptive resistance to the targeted therapy of cancer. *Cancer Discov.* 2012;2: 311–319.
7. Yesilkanal AE, Johnson GL, Ramos AF, Rosner MR. New strategies for targeting kinase networks in cancer. *J Biol Chem.* 2021;297: 101128.
8. Atkinson V, Sandhu S, Hospers G, Long GV, Aglietta M, Ferrucci PF, et al. Dabrafenib plus trametinib is effective in the treatment of BRAF V600-mutated metastatic melanoma patients: analysis of patients from the dabrafenib plus trametinib Named Patient Program (DESCRIBE II). *Melanoma Res.* 2020;30: 261–267.
9. Kuenzi BM, Park J, Fong SH, Sanchez KS, Lee J, Kreisberg JF, et al. Predicting Drug Response and Synergy Using a Deep Learning Model of Human Cancer Cells. *Cancer Cell.* 2020;38: 672–684.e6.
10. Yadav B, Wennerberg K, Aittokallio T, Tang J. Searching for Drug Synergy in Complex Dose-Response Landscapes Using an Interaction Potency Model. *Comput Struct Biotechnol J.* 2015;13: 504–513.

11. Julkunen H, Cichonska A, Gautam P, Szedmak S, Douat J, Pahikkala T, et al. Leveraging multi-way interactions for systematic prediction of pre-clinical drug combination effects. *Nat Commun.* 2020;11: 1–11.
12. Izza Y, Ignatiev A, Marques-Silva J. On Tackling Explanation Redundancy in Decision Trees. *jair.* 2022;75: 261–321bussy.
13. Cann ML, McDonald IM, East MP, Johnson GL, Graves LM. Measuring Kinase Activity-A Global Challenge. *J Cell Biochem.* 2017;118: 3595–3606.
14. Koleti A, Terryn R, Stathias V, Chung C, Cooper DJ, Turner JP, et al. Data Portal for the Library of Integrated Network-based Cellular Signatures (LINCS) program: integrated access to diverse large-scale cellular perturbation response data. *Nucleic Acids Res.* 2018;46: D558–D566.
15. Klaeger S, Heinzlmeir S, Wilhelm M, Polzer H, Vick B, Koenig P-A, et al. The target landscape of clinical kinase drugs. *Science.* 2017;358. doi:10.1126/science.aan4368
16. Berginski ME, Joisa CU, Golitz BT, Gomez SM. Kinome inhibition states and multiomics data enable prediction of cell viability in diverse cancer types. *PLoS Comput Biol.* 2023;19: e1010888.
17. Zou H, Hastie T. Regularization and Variable Selection via the Elastic Net. *J R Stat Soc Series B Stat Methodol.* 2005;67: 301–320.
18. Li S, Shen D, Shao J, Crowder R, Liu W, Prat A, et al. Endocrine-therapy-resistant ESR1 variants revealed by genomic characterization of breast-cancer-derived xenografts. *Cell Rep.* 2013;4: 1116–1130.
19. Ianevski A, Giri AK, Aittokallio T. SynergyFinder 2.0: visual analytics of multi-drug combination synergies. *Nucleic Acids Res.* 2020;48: W488–W493.
20. Bliss CI. THE TOXICITY OF POISONS APPLIED JOINTLY1. *Ann Appl Biol.* 1939;26: 585–615.
21. Chou TC, Talalay P. Quantitative analysis of dose-effect relationships: the combined effects of multiple drugs or enzyme inhibitors. *Adv Enzyme Regul.* 1984;22: 27–55.
22. Lee J, Liu H, Pearson T, Iwase T, Fuson J, Lalani AS, et al. PI3K and MAPK Pathways as Targets for Combination with the Pan-HER Irreversible Inhibitor Neratinib in HER2-Positive Breast Cancer and TNBC by Kinome RNAi Screening. *Biomedicines.* 2021;9. doi:10.3390/biomedicines9070740
23. Berginski ME, Jenner MR, Joisa CU, Herrera Loeza SG, Golitz BT, Lipner MB, et al. Kinome state is predictive of cell viability in pancreatic cancer tumor and stroma cell lines. *bioRxiv.* 2021. p. 2021.07.21.451515. doi:10.1101/2021.07.21.451515

24. Joisa CU, Chen KA, Berginski ME, Golitz BT, Jenner MR, Herrera Loeza SG, et al. Integrated Single-Dose Kinome Profiling Data is Predictive of Cancer Cell Line Sensitivity to Kinase Inhibitors. *bioRxiv*. 2022. p. 2022.12.06.519165. doi:10.1101/2022.12.06.519165
25. Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, et al. Explainable AI for Trees: From Local Explanations to Global Understanding. *arXiv [cs.LG]*. 2019. Available: <http://arxiv.org/abs/1905.04610>
26. Zawistowski JS, Bevill SM, Goulet DR, Stuhlmiller TJ, Beltran AS, Olivares-Quintero JF, et al. Enhancer Remodeling during Adaptive Bypass to MEK Inhibition Is Attenuated by Pharmacologic Targeting of the P-TEFb Complex. *Cancer Discov*. 2017;7: 302–321.
27. Grilley-Olson JE, Bedard PL, Fasolo A, Cornfeld M, Cartee L, Razak ARA, et al. A phase Ib dose-escalation study of the MEK inhibitor trametinib in combination with the PI3K/mTOR inhibitor GSK2126458 in patients with advanced solid tumors. *Invest New Drugs*. 2016;34: 740–749.
28. Bevill SM, Olivares-Quintero JF, Sciaky N, Golitz BT, Singh D, Beltran AS, et al. GSK2801, a BAZ2/BRD9 Bromodomain Inhibitor, Synergizes with BET Inhibitors to Induce Apoptosis in Triple-Negative Breast Cancer. *Mol Cancer Res*. 2019;17: 1503–1518.
29. Stuhlmiller TJ, Miller SM, Zawistowski JS, Nakamura K, Beltran AS, Duncan JS, et al. Inhibition of Lapatinib-Induced Kinome Reprogramming in ERBB2-Positive Breast Cancer by Targeting BET Family Bromodomains. *Cell Rep*. 2015;11: 390–404.
30. Szöcs E, Stirling T, Scott ER, Scharmüller A, Schäfer RB. webchem: An R Package to Retrieve Chemical Information from the Web. *J Stat Softw*. 2020;93: 1–17.
31. Berginski ME, Joisa CU, Golitz BT, Gomez SM. Kinome Inhibition States and Multiomics Data Enable Prediction of Cell Viability in Diverse Cancer Types. *bioRxiv*. 2022. p. 2022.04.08.487646. doi:10.1101/2022.04.08.487646
32. Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res*. 2021;49: D605–D612.
33. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw*. 2010;33: 1–22.
34. Wright MN, Ziegler A. ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. *J Stat Softw*. 2017;77: 1–17.
35. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. *arXiv [cs.LG]*. 2016. Available: <http://arxiv.org/abs/1603.02754>