KSTAR: An algorithm to predict patient-specific kinase activities from phosphoproteomic data

Sam Crowl[1,†],Benjamin Jordan[1,†], Cynthia Ma[2], Kristen M. Naegle[1,*]

1 University of Virginia, Department of Biomedical Engineering and the Center for Public Health Genomics, Charlottesville, VA, 22903
2 Department of Medicine and Siteman Cancer Center, Washington University in St. Louis, St. Louis, Missouri 63108

*** kmn4mj@virginia.edu**
† authors contributed equally to this work

Abstract

Kinase inhibitors are one of the largest classes of FDA-approved drugs and are major targets in oncology. Although kinase inhibitors have played an important role in improving cancer outcomes, major challenges still exist, including the development of resistance and failure to respond to treatments. Improvements for tumor profiling of kinase activity would be an important step in improving treatment outcomes and identifying effective kinase targets. Here, we present a graph- and statistics-based algorithm, called KSTAR, which harnesses the phosphoproteomic profiling of human cells and tissues by predicting kinase activity profiles from the observed phosphorylation of kinase substrates. The algorithm is based on the hypothesis that the more active a kinase is, the more of its substrates will be observed in a phosphoproteomic experiment. This method is error- and bias-aware in its approach, overcoming the challenges presented by the variability of phosphoproteomic pipelines, the limited information about kinase-substrate relationships, and limitations of global kinase-substrate predictions, such as training set bias and high overlap between predicted kinase networks. We demonstrate that the predicted kinase activities: 1) reproduce physiologically-relevant expectations and generates novel hypotheses within cell-specific experiments, 2) improve the ability to compare phosphoproteomic samples on the same tissues from different labs, and 3) identify tissue-specific kinase profiles. Finally, we apply the approach to complex human tissue biopsies in breast cancer, where we find that KSTAR activity predictions complement current clinical standards for identifying HER2-status – KSTAR can identify clinical false positives, patients who will fail to respond to inhibitor therapy, and clinically defined HER2-negative patients that might benefit from HER2-targeted therapy. KSTAR will be useful for both basic biological understanding of signaling networks and for improving clinical outcomes through improved clinical trial design, identification of new and/or combination therapies, and for identifying the failure to respond to targeted kinase therapies.

Introduction

Kinases make up the largest fraction of FDA-approved drugs for oncology [17], a reflection of their importance in oncogenesis and cancer progression. These drugs are also one of the most prevalent examples of precision medicine. For example, patients with BCR-ABL leukemias, HER2-overexpressing breast cancers, or EGFR-driven lung cancers benefit immensely from kinase-targeted therapies (often as adjuvants to chemotherapy, radiation, and/or surgery). Unfortunately, the selection and success of kinase inhibitors is often hampered by development of resistance mutations [2, 44], failure to respond to

treatment [33], and a limited set of current targets [17] – most new molecular entities target kinases for which inhibitors already exist. Hence, there is a need in the field of oncology to identify whether a patient might benefit from a kinase-inhibitor therapy and which kinase target(s) would be most beneficial.

In recent years, due to advances in proteomics, large-scale monitoring of protein phosphorylation is closer to being used within clinical profiling of tumor biopsies [16, 22, 23, 33, 45]. Since phosphorylation is a direct consequence of active kinases, or inactive phosphatases, this measurement might hold the key to better precision medicine. However, generating kinase activities from this data requires overcoming several challenges, including: 1) data sparsity or missing data – in shotgun phosphoproteomics (i.e. discovery-based approaches) lack of detection of a phosphorylation site may not be evidence that it is not present in the sample and 2) there is an extreme paucity of data regarding the direct connection between phosphorylation sites and their kinases [27]. Despite these challenges, phosphorylation is still more closely connected to kinase activity than commonly used proxies, such as mRNA, which rarely correlates with protein expression [7], or kinase expression and kinase phosphorylation, since kinase activity is regulated by a myriad of complex mechanisms [21, 24, 26, 28, 37, 43]. Hence, an approach that can harness the power of phosphoproteomic data, but can also account for the challenges presented via sampling sparsity and kinase-substrate relationships, might significantly improve our basic understanding of cell signaling and guide improved therapy selection for cancer patients.

Here, we present a novel algorithm that uses statistical and graph-theoretic approaches to infer the likely kinase activities from large-scale phosphoproteomic data, based on kinase-substrate prediction networks. Our approach hypothesizes that, given a network of kinase-substrate relationships, the more phosphorylated substrates observed from a kinase's network, the more likely it is to be active. This approach is unique from existing approaches – it does not require quantification of phosphorylation sites and it handles the kinase-substrate predictions in an error- and bias-aware manner. We wished to create an algorithm that relied only on the identification of phosphorylation sites from a sample, versus requiring relative quantification such as is required for the popular Kinase Set Enrichment Analysis (KSEA) approach, because: 1) even in the large-scale consortium samples that are related to a pooled sample, 75-95% of sites must be discarded as a result of coverage issues and 2) it is preferable in a clinical pipeline to avoid the requirement of a matched healthy tissue biopsies or a global pooled reference standard. Additionally, we wished to avoid using kinase-substrate algorithm predictions as thresholded networks, which is commonly done in downstream applications such as KSEA and INKA [3], since we found in recent work that [42] thresholding networks isolates phosphorylation sites (so most phosphorylation sites in an experiment would need to be dropped from analysis), many substrates exhibit high study bias – i.e. well studied phosphorylation sites are more likely to be well-connected to kinases, and there is high kinase network overlap that makes it difficult to discriminate between large groups of kinases. Hence, our approach operates on these networks in a new way to account for these issues and implements multiple statistical approaches to robustly convert the evidence observed in an experiment into an activity score. We refer to this method as KSTAR, which is short for Kinase-Substrate Transfer to Activity Relationships.

Ultimately our algorithm's kinase activity score is a reflection of "net kinase activity", since substrate phosphorylation is a reflection of the balance between kinase and phosphatase activity. In this work, we explain the details of our algorithmic approach and explore experiments to test whether the inferred kinase score changes with kinase activity. We demonstrate that we can predict: 1) increases in expected kinases as a result of network stimulation, 2) decreases in expected kinase activities as a result of kinase inhibition, and 3) tissue-specific kinase activity profiles, which are significantly more robust that the phosphoproteomic profiles, even across samples collected in different labs on different proteomic pipelines. Finally, we apply KSTAR to breast cancer biopsy-derived phosphoproteomic data and find that kinase activity profiles predicted by the algorithm can help identify misclassified HER2-positive breast cancer patients and identify clinically diagnosed HER2-negative patients that might respond to HER2-targeted therapy.

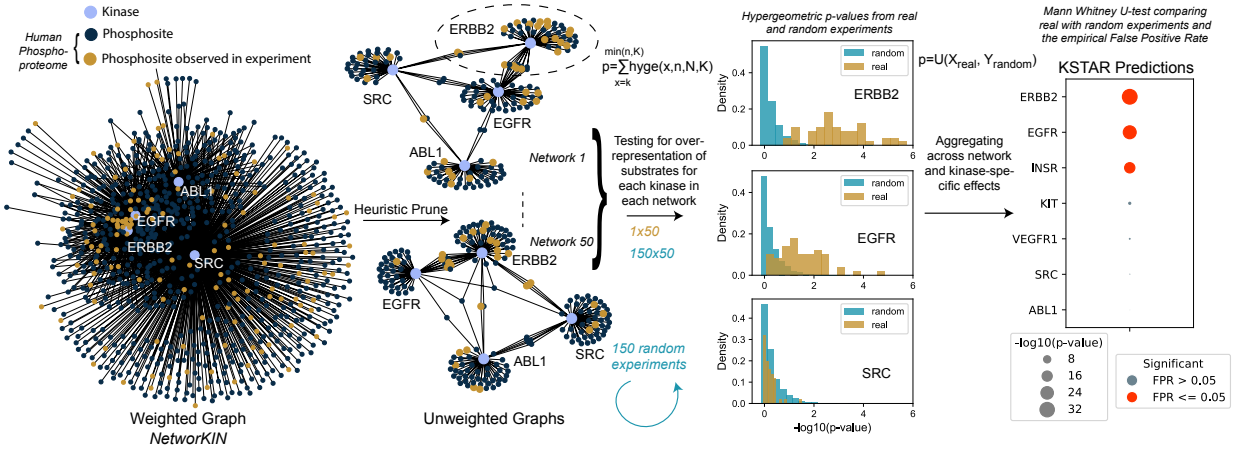

Figure 1. Overview of KSTAR algorithm. First, we heuristically prune dense and highly overlapping weighted kinase-substrate prediction graphs from NetworKIN [14] into many sparse, binary graphs. Statistical enrichment is calculated for an experiment that has a defined set of phosphorylation sites for every kinase across all networks using a hypergeometric distribution. We generate and calculate enrichment in 150 random experiments using the same approach. Next, we use the Mann-Whitney U test to measure the likelihood that the enrichment p-values in the real experiment are more significant than the random experiments, giving us a final p-value, which accounts for the underlying enrichment of substrates in a network, aggregates that information across the different network configurations, and controls for the kinase- and experiment-specific behavior of enrichment that occurs by random chance. We measure the false positive rate by measuring the distribution-based test for a random experiment against the remaining 149 random experiments, repeating this for 100 times. Finally, the numerical KSTAR "score" (the -log10 transformation of the Mann Whitney U-test) is presented in graphical format where the dot size is larger when there is more evidence phosphorylation sites are coordinately sampled from a kinase network. The FPR is indicated by "Significance" of having less than a specific empirical FPR.

Results

KSTAR algorithm

The KSTAR algorithm (Fig. 1) is based on the hypothesis that the more active a kinase is, the more of its substrates will be observed in a phosphoproteomic experiment. At its core, KSTAR is an algorithm that takes, as input, a set of phosphorylation sites observed in a mass spectrometry experiment, maps them onto global kinase substrate prediction graphs of KinPred [42], and converts the experimental input into a statistically robust KSTAR 'score' for each kinase, which increases with increasing representation of substrates from that kinase's network. When an experiment contains multiple conditions with relative quantification across all sites, the reported abundances are converted to binary evidence using a threshold relevant to the biological problem in question (S1 Table). We found that KSTAR predictions are fairly robust to changes in the threshold used (S1 Figure).

The first key development of KSTAR was to avoid the use of published thresholded networks. Instead, a "heuristic pruning" approach was implemented that decreases the graph density and creates many possible representations of sparse kinase-substrate networks (Fig. 1) according to a set of rules that maximizes phosphoproteome coverage and minimizes kinase overlap. In this work, all results are generated using the kinase-substrate graph of NetworKIN [14], with serine/threonine and tyrosine graphs being treated independently, since they are non-overlapping. Starting with the full graph of all calculated edges between available kinases and all phosphorylation sites known in the human proteome, the algorithm removes an edge in each step according to the following rules: 1) lower likelihood edges are more likely to be removed, 2) edges that are connected to "hub" substrates and/or "hub kinases" are

more likely to be removed so that a substrate only provides evidence for its most likely kinases, increasing discrimination of evidence between kinases, 3) edge removal is unlikely or zero once a substrate and/or kinase hits a minimum target to ensure that substrates and kinases are not "isolated" from the network due to having only low probability edges, and 4) substrates that represent high "study bias" – as defined by the number of compendia they are documented in [42] – are more likely to be pruned, since we observed significant kinase- and experiment-specific false positive rates without this constraint (S2 Figure). The graph pruning finishes when a fixed number of substrates are connected to each kinase and the distribution of substrate study bias is the same for all kinases and reflects the phosphoproteome study bias distribution. The example graphs shown in Fig. 1 depicts the success of this approach – on average, KSTAR networks reduce the overlap of evidence between EGFR and ERBB2 for sites observed in an experiment from 80% in a thresholded network to 18% overlap.

The first statistical test in KSTAR measures overrepresentation of substrates in each kinase network, controlling for the likelihood that the observation happened by random chance. Since we use binary evidence (a substrate was observed) with binary networks, we can use a well-defined theoretical probability function to estimate this overrepresentation – the hypergeometric distribution, where we calculate the probability of having observed k or more substrates of a kinase in an experimental dataset with n phosphorylation sites, when the full graph (all sites in the human phosphoproteome) has N total phosphosites and a kinase has K total substrates. This approach has the benefit of scaling with the size of the dataset and the distribution of edges in the graph. The output of this calculation is a p-value for each graph (i.e. a set of 50 p-values for every kinase).

We next wished to control for the variability of the resulting hypergeometric p-values across the different network configurations and the likelihood of making the same observations in a random experiment exhibiting similar properties as the experiment under consideration. To do this, we create 150 random experiments, where we randomly draw the same number of phosphorylation sites as the real experiment from the background phosphoproteome, such that the distribution of substrate study bias is the same as the real experiment. For each random experiment, we calculate enrichment in each kinase network as we did for the real experiment, producing a "random" distribution of p-values. This random distribution can then be compared to the real distribution using the Mann Whitney U-test, a non-parametric, distribution-based test which estimates the probability that the ranked p-values of the real experiment are more significant than those of the random experiments. Fig. 1 shows examples of distributions that arise from this process. For EGFR, Fig. 1 indicates that some random networks with real data give low -log10(p-values), yet the preponderance of evidence across all networks is more significant than what was observed in the random experiments, resulting in a small U-test p-value and a large KSTAR 'score' (the -log10(MW U-test p-value)). In contrast, it is clear that the p-values observed in SRC networks of the real experiment are no more significant than what occurs in random experiments, resulting in small a KSTAR score. Finally, we measure a false positive rate for the observed Mann Whitney p-value by measuring how often that p-value or more significant is observed in an empirical null model (repeatedly calculating the Mann Whitney p-values for treating a random experiment as the real experiment and comparing it to the remaining random experiments). The KSTAR score must be interpreted with the false positive rate (FPR) as we have observed some experiments result in large kinase-specific KSTAR scores that occur by random chance, due to the particular composition of the experiment. Hence, the final KSTAR score is a statistically powered value that ultimately reflects the fundamental measurement of the proposed hypothesis – an activity score that increases when more evidence is observed across many different possible network architectures and is more significant than what is observed by random chance.

KSTAR correctly infers expected tissue-specific kinases, kinase activation, and kinase inhibition

We wished to explore if the KSTAR score correlates with kinase activity. To explore this, we used KSTAR to predict kinase activities across a range of experiments that covered fundamentally different tissue types and perturbed kinase activities through stimulation or inhibition of kinases. Here, we

summarize a set of key kinase predictions from KSTAR, where full activity predictions and their false positive rates are presented in S3 Figure. We find that KSTAR predicts expected activity changes in response to either activation or inhibition and introduces interesting new hypotheses. We also found from these experiments that using the base kinase-substrate prediction network from NetworKIN [14] results in predictions consistent with physiological expectation, but using kinase-substrate networks from other prediction algorithms (GPS [39] and PhosphoPICK [29]), did not result in physiologically relevant hypotheses. Therefore, we have used NetworKIN as the foundation for KSTAR.

Activation of kinases in diverse cell models

Growth factor stimulation in epithelial cells: We used two phosphoproteomic studies of a human mammary epithelial cell line (184A1) stimulated with EGF to explore whether KSTAR accurately predicts the onset of EGFR activation and other related EGFR-network kinases. One of the studies additionally includes a HER2-overexpressing model of 184A1 cells (24H cells, expressing 600,000 vs. 20,000 copies of HER2/ERBB2) [41] and measured response to EGF- and HRG-stimulation, where EGF drives EGFR:ERBB2 dimers and HRG drives ERBB2:ERBB3 dimers. Despite the different phosphoproteomic pipelines used in the two studies the KSTAR predictions are extremely similar for parental 184A1 cells stimulated with EGF (S3 Figure and Fig. 2A), suggesting no basal EGFR or ERBB2 activity exists after serum starvation and is followed by rapid onset and peak activation between 4 and 10 minutes post-stimulation. Notably, cytosolic kinases that are not expected to be in the epithelial lineage are not predicted to be active in any condition (S3 Figure). In the HER2-perturbation experiment there are predictions of both dynamics and activity differences between cell lines and growth factors (Fig 2A), consistent with the expected biology of the cells. For example, there is basal EGFR activity in HER2-overexpressing cells, slower dynamics of ERBB2 activation by HRG, and lower activity of both receptors from HRG-treatment – all of which are consistent with the effects of HER2-overexpression, the increased migration of 24H cells in the absence of stimulation, and the maximal migration upon EGF-addition [41]. Finally, and somewhat surprisingly, we see maximal statistical significance of ERBB2 in the EGF stimulation of both cell lines, with some indication that HER2-overexpression sustains that activity longer in the 24H cells. The failure to see additional increases in ERBB2 activity in HER2-overexpressing cells may be a result of our statistical limit, as no additional gains can be seen beyond this amount of evidence that is already present in the parental cells stimulated by EGF. Taken together, the specific kinases predicted active, their associated scores, and their patterns of activation suggest that KSTAR predictions are capable of reproducing expected biology and that the use of evidence in experiments alone (i.e. the removal of quantitative fold changes) is useful for making comparative hypotheses between conditions.

TCR activation of Jurkat cells: To explore a cell lineage that is expected to have a distinctly different fingerprint of kinase activities than epithelial cells, we turned to datasets available in hematopoietic lineages, including several datasets in K562 (chronic myeloid leukemia) (S4 Figure) and Jurkat cells (S3 Figure). The Jurkat cell experiment by Chylek et al. [6] captured the fast dynamics of TCR activation at 5, 15, 30 and 60 seconds after stimulation. Importantly, the KSTAR predictions result in robust tissue-specific signaling expectations (Fig. 2B), predicting fast and robust (statistically saturated) activity of cytosolic kinases downstream of TCR, including LCK, FYN, HCK, BTK, and ITK. Slower to reach maximal detectable activity include YES1, BLK, and FGR. Additional hypotheses suggested by KSTAR predictions include a slower (30-second) onset of the RTK activity of VEGFR2 and NTRK1. Most importantly, these tissue-specific kinases were not predicted to be active in the epithelial experiments and epidermal growth factor-specific kinases are not predicted to be active in Jurkat/TCR signaling. Hence, KSTAR predictions are consistent with tissue- and signaling-specific expectations.

KSTAR correctly predicts inhibition of tyrosine and serine/threonine kinases

BCR-ABL inhibition by dasatinib: The oncogenic BCR-ABL fusion protein drives some chronic myeloid leukemias (CML) and represents an important target for treatment. Unlike many kinase inhibitors, inhibition of BCR-ABL can initiate cell death pathways within the first hour of treatment.

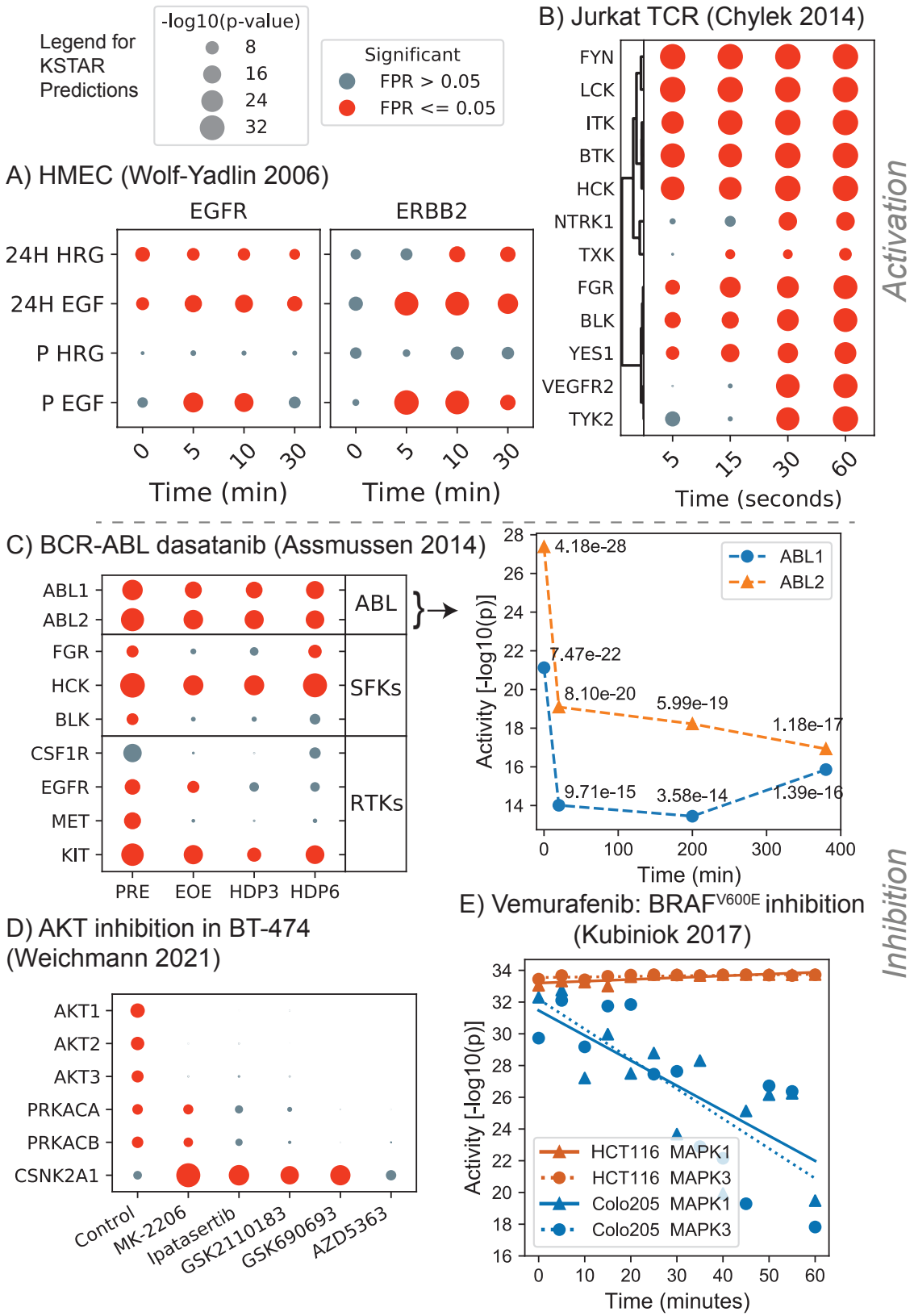

Figure 2. KSTAR applied to diverse cell models of kinase activation and inhibition. Full KSTAR results for data in this figure available in S3 Figure). Panel titles give the reference for the publication study of the phosphoproteomic data. All KSTAR predictions use the same legend for score size and significance as given above panel A. (caption continued on next page)

Figure 2. A) Predicted activation patterns of HMEC cell lines (P for parental 184A1 and 24H for HER2 overexpressing 184A1) in response to EGF and HRG stimulation. **B)** Predicted activation patterns of TCR stimulation in Jurkat cells shows early and robust activation of TCR-specific kinases. **C)** Predicted kinase patterns in response to inhibition of BCR-ABL inhibition by dasatinib in K562 cells with a detailed plot of significance changes for the ABL family kinases demonstrating a decrease, but continued activity of the oncogene. Kinase activity decreases in receptor tyrosine kinases (RTKs) correspond with findings of the original publication [1] as do changes in the off-target interactions with Src family kinases (SFKs). **D)** AKT inhibition by five inhibitors, all competitive ATP inhibitors, except MK-2206 an allosteric inhibitor of AKT demonstrate robust inhibition of all AKT homologs and interesting increases in CSNK2A1. **E)** Vemurafenib treatment, targeting the BRAF^{V600} mutation found in Colo205 colorectal cancer cells, but not the HCT116 cell line, demonstrates a decrease in MAPK activity specific to BRAF mutation, although still statistically significant MAPK activity.

This occurs even though ABL activity has been shown to be recovered within 4-8 hours after drug washout, depending on the study [1, 13, 34, 35]. To validate the predicted tyrosine kinase activities from KSTAR, we compared the kinase activity profiles of the CML cell line K562 in response to treatment with the ABL-inhibitor dasatinib from a study by Asmussen et al., who profiled phosphopeptide abundance before treatment, at the time of drug washout (EOE), and 3 and 6 hours post drug washout (HDP3 and HDP6, respectively) [1]. KSTAR predictions for ABL1 and ABL2 activity show a decrease following treatment, although the activity levels remain significant across all time points (Fig. 2C). Also, as expected, this activity is partially recovered after drug washout and the dynamic patterns of ABL kinases are mirrored in the Src family kinases (SFKs) BLK, HCK, and FGR, known alternate targets of dasatinib [44]. Finally, we note that there is robust down-regulation of a number of receptor tyrosine kinases, which is consistent with the conclusions of the original study [1] and others [13], which suggests dasatinib treatment in K562 cells is dependent upon on the elimination of ABL-GF-R interactions more than the elimination of ABL activity.

AKT inhibition: To this point, we have demonstrated the utility of KSTAR for predicting tyrosine kinase activities. To validate serine/threonine kinase predictions, we profiled the response of the breast cancer cell BT-474 to five different clinically relevant AKT inhibitors (Fig. 2D), based on data obtained from Wiechmann et al. [40]. BT-474 is a cell line that overexpresses ERBB2, which commonly leads to increased AKT activity. Upon applying KSTAR to this dataset, a complete elimination of AKT activity was predicted across all five inhibitors, suggesting that each inhibitor is highly effective at targeting AKT. While AKT3 was not identified by chemical proteomics in Wiechmann et al., our results suggests that AKT3 is active basally in BT-474 cells and that AKT3 activity is equally affected by treatment with these AKT inhibitors. We validated that the coordinated AKT1/2/3 predictions are not due to indiscriminate kinase networks – no AKT kinase shared more than 20% of substrates in KSTAR networks with any other AKT kinase. Hence, the basal activity predictions and predicted decrease of AKT family kinases in response to drug comes from independent networks.

Fig. 2D highlights two interesting predictions from KSTAR. First, AKT inhibition appears to increase CSNK2A1 (Casein Kinase 2) activity in four of the five drugs. There is evidence that there is complex interconnections between CSNK2A1 and AKT kinases [32] and these predictions suggest that there might be an unintended increase in casein kinase activity as a result of AKT inhibition. Second, we observed interesting patterns in PRKACA/B, which shows decreasing activity with the ATP competitive inhibitors, but not the PH-domain binding allosteric inhibitor MK-2206. Hence, these predictions suggest the competitive ATP inhibitors bind to PRKACA/B kinases, and the greatest decreases predicted by KSTAR in PRKACA/B (GSK2110183, GSK690693, and AZD5363) were shown to bind PRKACA/B by chemoproteomics [40]. Hence, KSTAR predictions accurately predict AKT inhibition and help identify cross-talk at the network level (CK2) and the inhibitor level (PRKACA/B).

Vemurafenib inhibition of BRAF^{V600E}: While the previous studies help to demonstrate how KSTAR effectively predicts changes to kinases directly impacted by stimulation or inhibition, KSTAR is currently limited by the kinases that are present in NetworKIN and some commonly studied cancer-specific kinases do not currently have predictions, such as AXL, DDR2, and RAF. In order to

determine whether KSTAR is still applicable in cases where direct kinase targets may be absent from predictions, we profiled the response of colorectal cancer cell lines to RAF inhibition by vemurafenib, using a time-resolved dataset generated by Kubiniok et al [18]. Vemurafenib is commonly used to target cells that harbor a $BRAF^{V600E}$ mutation, such as the colorectal cancer cell line Colo205. However, it has achieved limited clinical utility, in part, because tumors that contain a RAS mutation or overactive RTKs often see an adverse response where the MAP-ERK pathway is activated rather than inhibited [12, 18]. HCT116 cell line harbors a mutation in KRAS, and provides an example of a cancer cell that exhibits this paradoxical response and was also treated in the study by Kubiniok et al.

KSTAR predicts a clear decrease in MAPK1/3 activity over time in Colo205 cells, and a small increase in ERK activity is observed in HCT116, consistent with the paradoxical effect of RAS mutation in HCT116 (Fig. 2E). However, the basal activity values limits the ability to observe significant increases in activity due to the statistical limit. The increase in activity becomes more apparent when the kinase activity profiles are transformed via quantile normalization across samples (S3 Figure). Importantly, pre-treatment MAPK1/3 activity levels are comparable between HCT116 and Colo205, indicating that differences are a result of treatment and not basal activity. While the ultimate goal might be to expand NetworKIN and KSTAR to include these kinases in predictions, these results indicate that KSTAR is able accurately predict kinase activity profiles affected by missing kinases.

Kinase activity profiles are more robust than phosphoproteomic data

Due to the high complexity and stochastic nature of mass spectrometry measurements, obtaining reproducible results across studies and across laboratories can often be difficult. Multiple inter-laboratory studies have demonstrated that differences in instrumentation and pipelines can greatly impact the peptides identified in a single run, with reproducibility being highest between technical replicates from the same instrument [4, 36]. In addition, the choice of phosphopeptide enrichment method can further increase variability and lead to differences in the type of phosphopeptides identified [8]. Here, we sought to determine whether kinase activity profiles obtained from KSTAR can robustly identify similarities and differences between samples, even in cases where there is low overlap in the phosphopeptides identified by mass spectrometry. To do so, we obtained a total of 11 different phosphotyrosine datasets from 7 different studies and 5 different labs. Of these 11 datasets, 7 profiled the phosphoproteome of non-small cell lung carcinoma (NSCLC) cell lines with activating mutations in EGFR (H3255 and HCC827) [3, 11, 25, 31, 46] and 4 datasets profiled K562 cells, a chronic myeloid leukemia (CML) cell line containing the BCR-ABL fusion protein [1, 3, 8].

Based on the Jaccard similarity between sites identified in each dataset, most experiments exhibit low overlap with the other experiments (Fig 3A). The datasets with the highest overlap all stem from either the same study, the same lab, or both. In the only case where both K562 cells and NSCLC cells were profiled by the same study (dataset 6 and 10, [3]), the NSCLC sample shared the highest site similarity with the corresponding K562 sample rather than NSCLC samples from other studies. Overall, it is difficult to identify a clear pattern of separation between the NSCLC cell lines and K562 cells with the phosphoproteomic datasets alone. However, kinase activity profiles are capable of identifying shared profiles amongst all cells of the same type and reducing perceived similarity between studies from different cell types – datasets from the same cancer type were either moderately or strongly correlated ($r \geq 0.47, p < 0.05$), while those from different cancer types were either uncorrelated or weakly negatively correlated ($r < 0.15$). Hence, the transfer of phosphoproteomic data into kinase activity profiles greatly improves the comparability of independent phosphoproteomic experiments.

To verify that the kinase profiles that improved similarity within cell types and discrimination between cell types was connected to the underlying biology, we explored the top-ranked kinases across each cell type (Fig 3B). On average, EGFR was ranked as the most active kinase across all of the NSCLC cell lines, which is consistent with these lines carrying activating EGFR mutations. In K562 cells, the most active kinases were hematopoietic kinases HCK and BTK. Only MET appears as a highly active kinase in both cell types. We found that using only these top ranked kinases was sufficient to separate the cancer types by hierarchical clustering (Fig 3C). Further, the two NSCLC cell lines tended to cluster separately with this subset of kinases, which was not observed when using the entire kinase

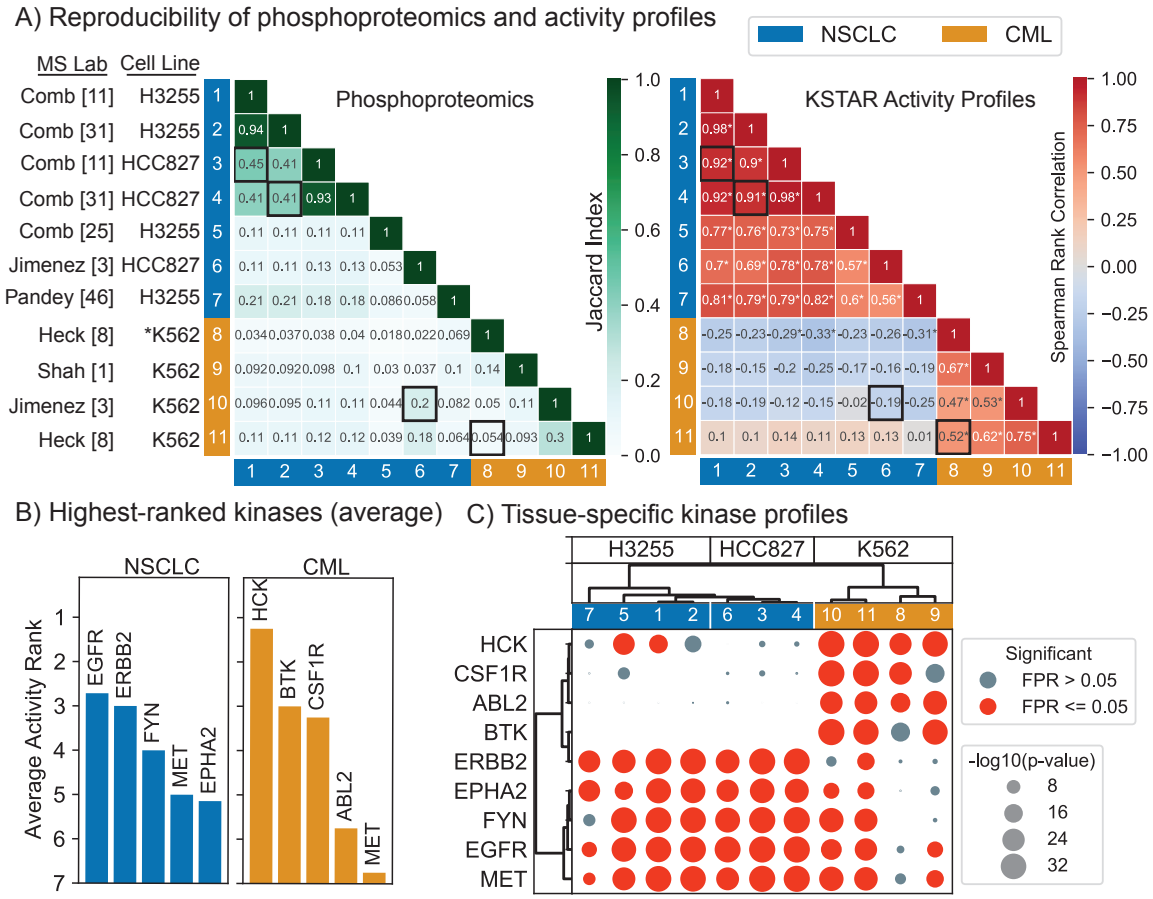

Figure 3. Tissue-specific profiles of kinase activities across independent studies of non-small cell lung carcinoma (NSCLC) and chronic myeloid leukemia (CML) cell lines. A) Comparison of the phosphoproteomic results obtained by each study (left) and the kinase activity profile predicted by KSTAR (right). We used similarity metrics to match the different data types – Jaccard similarity for phosphoproteomics and Spearman's rank correlation for kinase activity profiles. The ordering of the experiments in each heatmap is based on hierarchical clustering of the full kinase activity profile (S4 Figure). **B)** Kinases with the highest average activity ranking in NSCLC and CML cell lines. A rank of 1 indicates the most active kinase and a rank of 50 indicates the least active. For each study, kinases were sorted by their Mann Whitney p-values to obtain the experiment-specific ranking, and then the average rank across experiments was calculated for each kinase. **C)** Kinase activity profiles for top-ranked kinases. Both the kinases and experiments were sorted using hierarchical clustering with ward linkage. Full KSTAR results for data in this figure available in S4 Figure.

activity profile (S4 Figure). Overall, these results suggest that KSTAR is able to robustly predict kinase activity profiles that define a particular tissue type and results in robust identification of samples coming from similar tissues, which is not possible from phosphoproteomics directly.

Kinase predictions in breast cancer

We next wished to ask if KSTAR predictions from bulk biopsy of human tumors might be informative for identifying patient-specific kinase activity profiles. We selected breast cancer in order to compare HER2/ERBB2 activity with clinical diagnosis of HER2-overexpression. We applied KSTAR to the CPTAC dataset of 107 breast cancer patients, using data from the consortium as published in Mertins et al. [22]. Figure 4A focuses on ERBB2-activity predictions and the clinical status of the patient tumor. To our surprise, some of the most ERBB2-active predictions occur in HER2-negative patients. Also, some HER2-positive patients demonstrate extremely low levels of ERBB2-activity. Of the 22 HER2+ patients, we predict 10 of them are ERBB2-active and 13 are ERBB2-inactive. We predict the vast majority of HER2- patients lack ERBB2-activity (62/74), but there are 12 HER2- patients that we predict have evidence of ERBB2-activity. This difference between HER2-status and ERBB2-activity is feasible, since overexpression of HER2 may not necessarily lead to functional ERBB2 receptors at the cell surface and that breast epithelial cells contain some level of ERBB2, which could have activity without the requirement of amplification or overexpression. However, from this data, which lacks longitudinal patient information, it is not possible to know if KSTAR predictions are accurately identifying the roughly 15-30% of HER2-positive patients that will not be responsive to HER2-targeted therapies, or instead demonstrates issues with inferring tumor activity from complex bulk biopsies.

In order to test whether the departure of ERBB2-activity from HER2-status is clinically predictive of ERBB2-therapy response, we sought out additional datasets. These additional datasets include the phosphoproteomics of patient-derived xenografts (PDX) of breast cancer tumors [16], predominantly from HER2-negative patients, and the phosphoproteomics of HER2-positive breast cancer patients with data both before and 48-72 hours after starting a combined chemotherapy and HER2-targeted therapy [33]. Although small, each of these datasets included response data to HER2-targeted therapy, allowing us to query whether KSTAR predictions are correlated with response.

PDX models of breast cancer

The PDX dataset [16] measured the phosphoproteomics of 25 breast cancer PDX models, which were predominantly HER2-negative. Of the three HER2-positive PDX models, we predict two of them are ERBB2-active and represent the highest activity levels of HER2/ERBB2 that we see across the entire dataset. We predict the third HER2-positive model as having globally low tyrosine kinase activity, including insignificant levels of ERBB2-activity. On the other hand, we predict 31.8% of the HER2-negative PDX models have evidence of ERBB2-activity, double the rate we observed in the CPTAC predictions. The researchers in this study selected 4 PDX models to treat with the EGFR/HER2 inhibitor lapatinib. They selected two HER2-positive models, WHIM35 and WHIM8 (both of which we predict are the models we predict are ERBB2-active), and two HER2-negative models, WHIM14 and WHIM6. The authors were surprised to find that WHIM14 demonstrated significant response to lapatinib treatment. However, KSTAR predicts that the tumor has significant levels of both EGFR and ERBB2 activity, similar to WHIM35 and WHIM8, which also respond to lapatinib treatment. Hence, KSTAR predictions of EGFR and ERBB2 basal activity correspond with lapatinib response, including the surprising result of a HER2-negative tumor responding to HER2-specific therapy. This data suggests that kinase activity predictions from phosphoproteomic data may be capable of identifying patients who could benefit from HER2-targeted therapy despite being diagnosed in current clinical standards as HER2-negative. Together with the the CPTAC analysis, it suggests there are around 16 to 31% of HER2-negative patients that might benefit from HER2-targeted therapy.

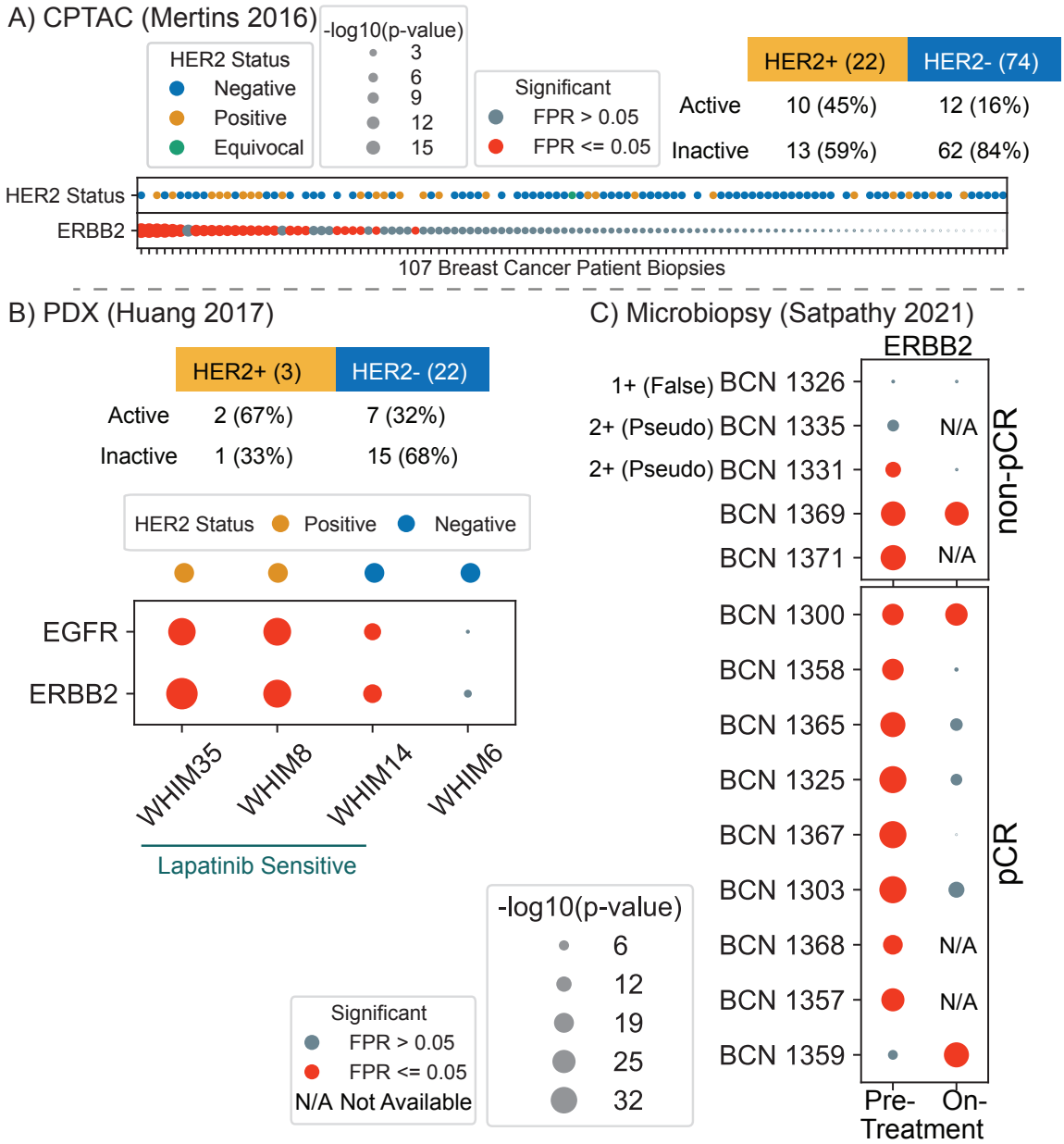

Figure 4. KSTAR applied to breast cancer biopsies in three studies. HER2 is used when referring to clinical diagnosis and ERBB2-activity for ERBB2/HER2 activity predictions. **A)** KSTAR predictions of ERBB2-activity for the 107 breast cancer patients in the CPTAC dataset [22] and their HER2 status is given. The table gives the total number of HER2-positive and HER-negative patients and the KSTAR predictions for ERBB2/HER2 activity. **B)** Predictions of EGFR and ERBB2 activities for the patient derived xenograft (PDX) models published in Huang et al. [16] subset that were treated with lapatinib (EGFR/HER2 targeted therapy), where WHIM14 is a HER2-negative tumor that showed a surprising response to lapatinib treatment. The table reports the HER2-status of all 25 PDX tumors and KSTAR ERBB2/HER2-activity predictions. **C)** The ERBB2-activity predictions for tumor biopsies of patients enrolled in a HER2-positive study by Satpathy et al. [33]. Five patients were non-pathologically complete responders (non-pCR) and the remainder were pathologically complete responders (pCR). Biopsies were taken pre-treatment and most patient's also had an on-treatment biopsy taken with phosphoproteomic profiling. The first three non-responders were reclassified for HER2-status upon additional analysis in Satpathy et al. and results are shown as one false positive and two classified as "Pseudo-positives". Full KSTAR results for data in this figure available in S5 Figure.

HER2-positive patient study and response to treatment

In the study by Satpathy et al. [33], the researchers developed a method to section a biopsy for a suite of molecular analyses, including microscaled proteomics and phosphoproteomics. This study consisted of 14 clinically-identified HER2-positive patients. Microscaling the analyses allowed the researchers the ability to perform replicates and analyses on new biopsies taken 48 to 72 hours following the commencement of a combination of chemotherapy and HER2-targeted therapy. Thirteen of the patients were treated with docitaxel and the combination of trastuzumab and pertuzumab, while one patient (BCN 1369) received paclitaxel and trastuzumab alone. In the study, 9 of the 14 patients were diagnosed as pathologically complete responders (pCR) and 5 patients were non-pathologically complete responders (non-pCR). This study provides an excellent opportunity to explore the reproducibility of KSTAR predictions across replicates, the relationship between therapy response and predictions of kinase activity pre-treatment, and the connection between response and ERBB2-activity, including on-treatment biopsy data, although complicated by the combination of treatment with chemotherapy.

From kinase activity predictions, we focused on the prediction of ERBB2 activity across replicates and samples, where we found consistent predictions across all replicates of a patient sample (S5 Figure), which confirms that both the replication of the data and predictions from that data are robust. Of all 14 patients that have been clinically diagnosed as HER2-positive, we predict 3 of the patients (BCN1326, BCN 1335, and BCN 1359) lack evidence of ERBB2-activity in the pre-treatment phosphoproteomic data and an additional patient, BCN1331 with the next lowest activity and an FPR rate right at significance of 0.05. Three of these patients are non-pCR, suggesting that lack of ERBB2-activity basally may be a reason for non-response to chemotherapy combined with targeted HER2-therapy (Fig. 4C). In the original study, the researchers re-analyzed patient data for the non-pCR patients and found that BCN1326 was a HER2 false-positive, having no indication of ERBB2-amplification or increased protein expression and it is also the patient we predict has the lowest ERBB2-activity. Patients BCN1331 and BCN1335, also with low KSTAR scores for ERBB2-activity, were re-labeled by the research team as"pseudo-HER2-positive" having evidence of HER2 copy number increases, which did not translate to the protein levels. Their work demonstrates the need for additional clinical analyses for accurate diagnostics, compared to standard of care. Unfortunately, it also required detailed analyses of mRNA and protein-levels guided by known patient outcomes. Alternatively, using phosphoproteomics alone, KSTAR predictions identified these patients as having significantly lower to no ERBB2-activity compared to the remaining patients, indicating that KSTAR predictions may be clinically useful.

KSTAR predictions from on-treatment samples are also promising, suggesting that six of the seven patients with pre-treatment ERBB2-activity are predicted to demonstrate a decrease, below significance, of ERBB2-activity on-treatment (Fig 4C). There are two patients where predictions suggest there is no response to ERBB2-targeted therapy, BCN 1369 (non-pCR) and BCN 1300 (pCR). These data suggest that these patients are not benefiting from the HER2-therapy arm. Additionally, these two patients are most similar according to their tyrosine kinase activity predictions (S5 Figure), suggesting that resistance to HER2-therapy might be encoded by their particular pattern of tyrosine kinase activities. BCN 1371, the last non-pCR patient, which has no on-treatment sample, has a similar kinase activity profile to the HER2- non-responders BCN 1369 and BCN 1300. Similarly, BCN 1357 and BCN 1368, which have no on-treatment phosphoproteomic data, cluster most similarly with BCN 1358 and BCN 1368, respectively, both of which show robust decreases in ERBB2-activity on-treatment. Where these similarities between pre-treatment patient responses and clinical outcomes are seen in tyrosine kinase activity profiles, they are do not occur in serine/threonine kinase activity profiles (S5 Figure).

The final ERBB2-predicted negative patient, BCN1359, represents a unique case, as we predict it has low pre-treatment activity that increases on therapy. Satpathy et al. did not redefine the HER2-status of this patient and their analyses show this patient also uniquely undergoes an increase in ERBB2 mRNA and protein upon treatment. A unique trait of BCN1359's pre-treatment kinase profile is EGFR-activity, in the absence of ERBB2-activity. This status is rare as we have typically observed, across the breast epithelial and breast cancer samples we have profiled, that EGFR and ERBB2 activity are often jointly active or inactive. Ultimately, BCN1359 went on to become a pathologically complete responder. However, KSTAR predictions suggest that response may have been a product predominantly of the

chemotherapy component of the treatment.

Discussion

The basic premise of KSTAR is the hypothesis that the more active a kinase is, the more phosphorylation sites it will produce, and the more likely those sites will be observed in a phosphoproteomic experiment. Converting this premise into a quantitative value required new approaches, such as heuristic network pruning and control of kinase- and experiment-specific false positive rates. We learned important lessons about using the underlying network predictions that suggest typical approaches that use the thresholded networks presents serious issues with respect to study bias, an effect we were able to address, but not completely overcome (e.g. we sometimes see high activity scores that also relate to high false positive rates, making it difficult to interpret the prediction). The application of KSTAR to diverse experiments supports the idea that the KSTAR score is a reflection of kinase activity – the score increases with increasing kinase activities and decreases as a result of direct kinase inhibition. We found that these continuous-valued scores, up to the limit of statistical saturation, are capable of identifying the most active kinases within tissues and differences in kinase activities between samples. Therefore, KSTAR scores appear to be a direct reflection of kinase activity, supporting the central premise of the approach. Perhaps the most surprising result occurred when we transitioned from controlled cell line data to complex cancer biopsy data, where we found that KSTAR was able to identify clinically-relevant kinase activity profiles. The relatively high mismatch between patient HER2-status in the HER2/ERBB2-activity predictions in the CPTAC data could easily be a reflection of failure to accurately predict HER2-activity, either as a failure of the approach or due to complications from immune/stroma contamination within the biopsy. However, the additional breast cancer experiments we profiled also involved complex tissue biopsies. Astoundingly, KSTAR predictions accurately identified that a PDX model of a HER2-negative tumor, WHIM14, is basally ERBB2-active, explaining the originally surprising result that it responded to lapatinib – a drug that would never be prescribed for that patient under current clinical standards. If KSTAR is capable of identifying ERBB2 activity that results from regulation beyond that of gene amplification, then our data suggests that between 16% and 32% of HER2-negative patients might benefit from a HER2-targeted therapy.

In addition to the HER2-negative patients that show predicted ERBB2-activity, KSTAR was able to identify clinically-defined HER2-positive patients that do not demonstrate significant basal kinase activity, explaining 60% of the non-pathologically complete responders. Also, with the combination of on-treatment biopsy samples, KSTAR predicts that three of those patients are not benefiting from the HER2-therapy. If true, then this suggests that phosphoproteomics, in combination with KSTAR, might prove to be significantly more powerful at identifying patients that will benefit from HER2-therapy than current clinical approaches. Additionally, the HER2-active patients that showed no decrease in activity on-treatment displayed similar kinase activity profiles, suggesting that these profiles might be useful for understanding the mechanistic underpinning of failure to respond. Hence, we envision that KSTAR can be useful for a range of approaches, from the basic sciences to the clinic, including clinical trial design.

Materials and Methods

Data Preparation and Mapping

All phoshoproteomic datasets used in this study were downloaded from the relevant journal site (typically found as supplementary data or tables in Excel). We identified the datasheet columns (a protein accession and a peptide indicating the phosphorylation site or an amino acid position) that allowed us to map the specific protein and phosphorylation site to the central ontology developed in KinPred [42]. If Uniprot accessions were not available, we used Uniprot web services to map to the Uniprot accession. KSTAR has a mapping function that matches the protein record and site information to the central KinPred ontology and appends two confirmed columns in that same ontology. Failure to map the protein or the site results in exclusion of that peptide from the mapped experiment, which

happens in about 1 to 10% of peptides in any given dataset, an event that is logged. Also logged by mapping is if the site position is altered from the dataset to the ontology in cases where the reference protein is different than the database representation used by the authors of the study. The mapped files and their error and warning logs are provided in Supplementary Data available on FigShare (`https://doi.org/10.6084/m9.figshare.14919726`).

In the experiments where relative quantification of phosphosites across multiple conditions was available, we used a relevant threshold to include or exclude specific sites as evidence in that particular experimental condition, while still ensuring that every experiment contained a sufficient number of sites to be statistically powerful (generally 50 or more sites for tyrosine, 1000 or more sites for serine and threonine). These thresholds, their meaning, and the resulting evidence sizes for experiments for all datasets are given in S1 Table and these experiments are stored in KSTAR outputs as binary evidence used to generate kinase activity predictions (also available in FigShare Supplementary Data).

Comparing Phosphoproteomics Datasets

When comparing different phosphoproteomics datasets, all data was first mapped to KinPred, as described above. After mapping, the set of phosphosites identified in each dataset was obtained, where any site present in a sample, regardless of relative quantification, was considered in the set. The similarity of two sets of sites were compared using the Jaccard index, defined by dividing the intersection of the two sets (number of sites identified in both experiments) by the union (the total number of unique sites identified across both experiments). To compare these same experiments based on their KSTAR profiles, we used Spearman rank correlation, as implemented in the SciPy Python package [38]

KSTAR algorithm implementation and details

All parts of the KSTAR algorithm, including pruning, mapping, activity inference, and plotting, were implemented using Python3. The KSTAR code is highly customizable regarding the networks used, the limits on when to stop network pruning, the number of random experiments, and the number of decoy runs to use in FPR calculation. However, here we used the same parameters for all analyses as described and all data and code are provided on open source platforms (GitHub and Figshare), described below.

The 50 heuristically pruned networks were generated according to the rules discussed in Algorithm Overview, where tyrosine kinases had a finishing limit of 2000 substrates per kinase and serine/threonine kinases had a finishing limit of 5000 substrates per kinase. The final networks also guaranteed that each kinase had the same proportion of substrates based on "study bias" according to the total number of compendia substrates are annotated by and as defined by KinPred [42]. Substrates can be found in none or any of the five compendia included in ProteomeScout [20]: phosphoELM [10], HPRD [30], dbPTM [19], or PhosphoSitePlus [15] The networks used were downloaded from KinPred version 1.0 and based on the February 2020 reference phosphoproteome. The final networks

We generated 150 unique random experiments for every experiment under consideration and such that the distribution of "study bias class" was the same as the real experiment, where study bias was defined as having low, medium, or high study bias based on being annotated in 0, 1-2, or 3-5 compendia. Unlike network generation, which uses the exact number of compendia annotations in the definition of study bias, random sampling groups the categories into classes, in order to guarantee true random sampling from the high study bias categories, since the number of total substrates annotated by 5 compendia is small. The false positive rate

Statistical tests, including the hypergeometric test and the one-sided Mann Whitney U test, were performed using functions found within the stats module of the SciPy Python package [38]. For cases where datasets were large, such as for the two global phosphoproteomics studies [18, 40] or for large serine/threonine networks, the KSTAR activity predictions were performed using a highly parallel version of the original code, implemented using the Nextflow software package [9].

Data and Code Availability

All data and results used throughout this paper, including the processed and mapped datasets, the random experiments generated for each run, and the final KSTAR results have all been deposited in KSTAR FigShare (`https://doi.org/10.6084/m9.figshare.14919726`). KSTAR can be implemented using the Python package found here: https://github.com/NaegleLab/KSTAR, which also includes Jupyter Notebooks of all analyses and plots for every dataset in this paper.

Supporting Information

S1 Figure

Impact of thresholding decision on the estimated kinase activity. We used phosphoproteomic data from untreated samples in K562 cells from Palma et al. [8] for pTyr sensitivity analysis and untreated BT-474 samples from Weichmann et al. [40] for pSer/pThr sensitivity analysis. KSTAR analysis was performed on these datasets with run multiple times using progressively smaller evidence sizes using abundance estimates from these datasets to simulate the effect of using increasing threshold stringencies.

S2 Figure

Demonstration that phosphorylation site sampling from resources causes kinase- and experiment-specific false positive rates. Supplement gives methods and shows the activity predictions for datasets randomly sampled from the following compendia: **A)** phosphoELM [10], **B)** HPRD [30], **C)** dbPTM [19], or **D)** PhosphoSitePlus [15]. **E)** Demonstration that phosphorylation sites in ProteomeScout [20] is very different than what is observed in an example phosphoproteomic dataset [16]. These results lead to constraining networks so that all kinases have the same distribution of study bias in their connected substrates (reducing kinase-specific false positive rates seen here) and generating random datasets that have the same distribution of study bias as the experiment being used (experiment- and kinase-specific controls).

S3 Figure

Full KSTAR predictions for all control experiments in Figure 2.

S4 Figure

Full KSTAR predictions for all NSCLC and CML cell lines used in Figure 3.

S5 Figure

Full KSTAR predictions for all Breast Cancer biopsy experiments in Figure 4.

S1 Table

Table outlining experiments, the thresholds used to define binary evidence within them, and the resulting sizes of evidence.

Acknowledgments

Research reported in this publication was supported by the National Cancer Institute of the National Institutes of Health under Award Number R21CA231853. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We

wish to thank the open source Python community, the researchers behind NetworKIN, and the authors of all studies we used in this work. We also wish to thank members of the UVA Systems Biology Journal Club for critical reading and feedback on this manuscript. Finally, we wish to thank the cancer patients for their participation in the clinical trials that resulted in phosphoproteomic profiling of breast tumor biopsies.

References

1. J. Asmussen, E. a. Lasater, C. Tajon, J. Oses-Prieto, Y.-W. Jun, B. S. Taylor, A. Burlingame, C. S. Craik, and N. P. Shah. MEK-Dependent Negative Feedback Underlies BCR-ABL-Mediated Oncogene Addiction. *Cancer discovery*, pages 200–215, dec 2013.

2. R. Barouch-bentov. Mechanisms of Drug-Resistance in Kinases. *Expert Opin Investig Drugs*, 20(2), 2011.

3. R. Beekhof, C. van Alphen, A. A. Henneman, J. C. Knol, T. V. Pham, F. Rolfs, M. Labots, E. Henneberry, T. Y. Le Large, R. R. de Haas, S. R. Piersma, V. Vurchio, A. Bertotti, L. Trusolino, H. M. Verheul, and C. R. Jimenez. INKA, an integrative data analysis pipeline for phosphoproteomic inference of active kinases. *Molecular systems biology*, 15(4):e8250, 2019.

4. A. W. Bell, E. W. Deutsch, C. E. Au, R. E. Kearney, R. Beavis, S. Sechi, T. Nilsson, J. J. Bergeron, T. A. Beardslee, T. Chappell, G. Meredith, P. Sheffield, P. Gray, M. Hajivandi, M. Pope, P. Predki, M. Kullolli, M. Hincapie, W. S. Hancock, W. Jia, L. Song, L. Li, J. Wei, B. Yang, J. Wang, W. Ying, Y. Zhang, Y. Cai, X. Qian, F. He, H. E. Meyer, C. Stephan, M. Eisenacher, K. Marcus, E. Langenfeld, C. May, S. A. Carr, R. Ahmad, W. Zhu, J. W. Smith, S. M. Hanash, J. J. Struthers, H. Wang, Q. Zhang, Y. An, R. Goldman, E. Carlsohn, S. van der Post, K. E. Hung, D. A. Sarracino, K. Parker, B. Krastins, R. Kucherlapati, S. Bourassa, G. G. Poirier, E. Kapp, H. Patsiouras, R. Moritz, R. Simpson, B. Houle, S. LaBoissiere, P. Metalnikov, V. Nguyen, T. Pawson, C. C. Wong, D. Cociorva, J. R. Yates, M. J. Ellison, A. Lopez-Campistrous, P. Semchuk, Y. Wang, P. Ping, G. Elia, M. J. Dunn, K. Wynne, A. K. Walker, J. R. Strahler, P. C. Andrews, B. L. Hood, W. L. Bigbee, T. P. Conrads, D. Smith, C. H. Borchers, G. A. Lajoie, S. C. Bendall, K. D. Speicher, D. W. Speicher, M. Fujimoto, K. Nakamura, Y. K. Paik, S. Y. Cho, M. S. Kwon, H. J. Lee, S. K. Jeong, A. S. Chung, C. A. Miller, R. Grimm, K. Williams, C. Dorschel, J. A. Falkner, L. Martens, and J. A. Vizcaíno. A HUPO test sample study reveals common problems in mass spectrometry-based proteomics. *Nature Methods*, 6(6):423–430, 2009.

5. P. Casado, J.-C. Rodriguez-Prados, S. C. Cosulich, S. Guichard, B. Vanhaesebroeck, S. Joel, and P. R. Cutillas. Kinase-Substrate Enrichment Analysis Provides Insights into the Heterogeneity of Signaling Pathway Activation in Leukemia Cells. *Science Signaling*, 6(268):rs6–rs6, 2013.

6. L. A. Chylek, V. Akimov, J. Dengjel, K. T. G. Rigbolt, B. Hu, W. S. Hlavacek, and B. Blagoev. Phosphorylation Site Dynamics of Early T-cell Receptor Signaling. *PLoS ONE*, 9(8):e104240, 2014.

7. R. De Sousa Abreu, L. O. Penalva, E. M. Marcotte, and C. Vogel. Global signatures of protein and mRNA expression levels. *Molecular BioSystems*, 5(12):1512–1526, 2009.

8. S. Di Palma, A. Zoumaro-Djayoon, M. Peng, H. Post, C. Preisinger, J. Munoz, and A. J. R. Heck. Finding the same needles in the haystack? A comparison of phosphotyrosine peptides enriched by immuno-affinity precipitation and metal-based affinity chromatography. *Journal of Proteomics*, 91:331–337, 2013.

9. P. DI Tommaso, M. Chatzou, E. W. Floden, P. P. Barja, E. Palumbo, and C. Notredame. Nextflow enables reproducible computational workflows. *Nature Biotechnology*, 35(4):316–319, 2017.

10. F. Diella, S. Cameron, C. Gemünd, R. Linding, A. Via, B. Kuster, T. Sicheritz-Pontén, N. Blom, and T. J. Gibson. Phospho.ELM: a database of experimentally verified phosphorylation sites in eukaryotic proteins. *BMC bioinformatics*, 5:79, jun 2004.

11. A. Guo, J. Villén, J. Kornhauser, K. A. Lee, M. P. Stokes, K. Rikova, A. Possemato, J. Nardone, G. Innocenti, R. Wetzel, Y. Wang, J. MacNeill, J. Mitchell, S. P. Gygi, J. Rush, R. D. Polakiewicz, and M. J. Comb. Signaling networks assembled by oncogenic EGFR and c-Met. *Proceedings of the National Academy of Sciences of the United States of America*, 105(2):692–697, 2008.

12. S. J. Heidorn, C. Milagre, S. Whittaker, A. Nourry, I. Niculescu-Duvas, N. Dhomen, J. Hussain, J. S. Reis-Filho, C. J. Springer, C. Pritchard, and R. Marais. Kinase-Dead BRAF and Oncogenic RAS Cooperate to Drive Tumor Progression through CRAF. *Cell*, 140(2):209–221, 2010.

13. D. K. Hiwase, D. L. White, J. A. Powell, V. A. Saunders, S. A. Zrim, A. K. Frede, M. A. Guthridge, A. F. Lopez, R. J. D'Andrea, L. B. To, J. V. Melo, S. Kumar, and T. P. Hughes. Blocking cytokine signaling along with intense Bcr-Abl kinase inhibition induces apoptosis in primary CML progenitors. *Leukemia*, 24(4):771–778, 2010.

14. H. Horn, E. M. Schoof, J. Kim, X. Robin, M. L. Miller, F. Diella, A. Palma, G. Cesareni, L. J. Jensen, and R. Linding. KinomeXplorer: An integrated platform for kinome biology studies, 2014.

15. P. V. Hornbeck, J. M. Kornhauser, S. Tkachev, B. Zhang, E. Skrzypek, B. Murray, V. Latham, and M. Sullivan. PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic acids research*, 40(Database issue):D261–70, jan 2012.

16. K. L. Huang, S. Li, P. Mertins, S. Cao, H. P. Gunawardena, K. V. Ruggles, D. R. Mani, K. R. Clauser, M. Tanioka, J. Usary, S. M. Kavuri, L. Xie, C. Yoon, J. W. Qiao, J. Wrobel, M. A. Wyczalkowski, P. Erdmann-Gilmore, J. E. Snider, J. Hoog, P. Singh, B. Niu, Z. Guo, S. Q. Sun, S. Sanati, E. Kawaler, X. Wang, A. Scott, K. Ye, M. D. McLellan, M. C. Wendl, A. Malovannaya, J. M. Held, M. A. Gillette, D. Fenyö, C. R. Kinsinger, M. Mesri, H. Rodriguez, S. R. Davies, C. M. Perou, C. Ma, R. Reid Townsend, X. Chen, S. A. Carr, M. J. Ellis, and L. Ding. Proteogenomic integration reveals therapeutic targets in breast cancer xenografts. *Nature Communications*, 8(14868):1–16, 2017.

17. M. S. Kinch. An analysis of FDA-approved drugs for oncology. *Drug Discovery Today*, 19(12):1831–1835, 2014.

18. P. Kubiniok, H. Lavoie, M. Therrien, and P. Thibault. Time-resolved phosphoproteome analysis of paradoxical RAF activation reveals novel targets of ERK. *Molecular and Cellular Proteomics*, 16(4):663–679, 2017.

19. C.-T. Lu, K.-Y. Huang, M.-G. Su, T.-Y. Lee, N. A. Bretaña, W.-C. Chang, Y.-J. Chen, Y.-J. Chen, and H.-D. Huang. dbPTM 3.0: an informative resource for investigating substrate site specificity and functional association of protein post-translational modifications. *Nucleic acids research*, 41(Database issue):D295–305, jan 2013.

20. M. K. Matlock, A. S. Holehouse, and K. M. Naegle. ProteomeScout: A repository and analysis resource for post-translational modifications and proteins. *Nucleic Acids Research*, 2015.

21. B. J. Mayer, H. Hirai, and R. Sakai. Evidence that SH2 domains promote processive phosphorylation by protein-tyrosine kinases. *Current biology : CB*, 5(3):296–305, 1995.

22. P. Mertins, D. R. Mani, K. V. Ruggles, M. A. Gillette, K. R. Clauser, P. Wang, X. Wang, J. W. Qiao, S. Cao, F. Petralia, E. Kawaler, F. Mundt, K. Krug, Z. Tu, J. T. Lei, M. L. Gatza, M. Wilkerson, C. M. Perou, V. Yellapantula, K.-l. Huang, C. Lin, M. D. McLellan, P. Yan, S. R. Davies, R. R. Townsend, S. J. Skates, J. Wang, B. Zhang, C. R. Kinsinger, M. Mesri,

H. Rodriguez, L. Ding, A. G. Paulovich, D. Fenyö, M. J. Ellis, and S. A. Carr. Proteogenomics connects somatic mutations to signalling in breast cancer. *Nature*, 534(7605):55–62, 2016.

23. P. Mertins, L. C. Tang, K. Krug, D. J. Clark, M. A. Gritsenko, L. Chen, K. R. Clauser, T. R. Clauss, P. Shah, M. A. Gillette, V. A. Petyuk, S. N. Thomas, D. R. Mani, F. Mundt, R. J. Moore, Y. Hu, R. Zhao, M. Schnaubelt, H. Keshishian, M. E. Monroe, Z. Zhang, N. D. Udeshi, S. R. Davies, R. R. Townsend, D. W. Chan, R. D. Smith, H. Zhang, T. Liu, and S. A. Carr. Reproducible workflow for multiplexed deep-scale proteome and phosphoproteome analysis of tumor tissues by liquid chromatography-mass spectrometry. *Nature Protocols*, 13(7):1632–1661, 2018.

24. M. a. Meyn, M. B. Wilson, F. a. Abdi, N. Fahey, A. P. Schiavone, J. Wu, J. M. Hochrein, J. R. Engen, and T. E. Smithgall. Src family kinases phosphorylate the Bcr-Abl SH3-SH2 region and modulate Bcr-Abl transforming activity. *Journal of Biological Chemistry*, 281(41):30907–30916, 2006.

25. A. Moritz, Y. Li, A. Guo, J. Villén, Y. Wang, J. MacNeill, J. Kornhauser, K. Sprott, J. Zhou, A. Possemato, J. M. Ren, P. Hornbeck, L. C. Cantley, S. P. Gygi, J. Rush, and M. J. Comb. Akt - RSK - S6 kinase signaling networks activated by oncogenic receptor tyrosine kinases. *Science Signaling*, 3(136):1–12, 2010.

26. P. J. Müller, K. T. G. Rigbolt, D. Paterok, J. Piehler, J. Vanselow, E. Lasonder, J. S. Andersen, F. Schaper, and R. M. Sobota. Protein tyrosine phosphatase SHP2 / PTPN11 mistargeting as a consequence of SH2-domain point mutations associated with Noonan Syndrome and leukemia. *Journal of Proteomics*, 84:132–147, 2013.

27. E. J. Needham, B. L. Parker, T. Burykin, D. E. James, and S. J. Humphrey. Illuminating the dark phosphoproteome. *Sci Signal*, 12(eaau8645):1–18, 2019.

28. H. Park, M. I. Wahl, D. E. Afar, C. W. Turck, D. J. Rawlings, C. Tam, a. M. Scharenberg, J. P. Kinet, and O. N. Witte. Regulation of Btk function by a major autophosphorylation site within the SH3 domain. *Immunity*, 4(5):515–25, may 1996.

29. R. Patrick, C. Horin, B. Kobe, K. A. L. Cao, and M. Bodén. Prediction of kinase-specific phosphorylation sites through an integrative model of protein context and sequence. *Biochimica et Biophysica Acta - Proteins and Proteomics*, 2016.

30. T. S. K. Prasad, R. Goel, K. Kandasamy, S. Keerthikumar, S. Kumar, S. Mathivanan, D. Telikicherla, R. Raju, B. Shafreen, A. Venugopal, L. Balakrishnan, A. Marimuthu, S. Banerjee, D. S. Somanathan, A. Sebastian, S. Rani, S. Ray, C. J. Harrys Kishore, S. Kanth, M. Ahmed, M. K. Kashyap, R. Mohmood, Y. L. Ramachandra, V. Krishna, B. A. Rahiman, S. Mohan, P. Ranganathan, S. Ramabadran, R. Chaerkady, and A. Pandey. Human Protein Reference Database–2009 update. *Nucleic acids research*, 37(Database issue):D767–72, jan 2009.

31. K. Rikova, A. Guo, Q. Zeng, A. Possemato, J. Yu, H. Haack, J. Nardone, K. Lee, C. Reeves, Y. Li, Y. Hu, Z. Tan, M. Stokes, L. Sullivan, J. Mitchell, R. Wetzel, J. MacNeill, J. M. Ren, J. Yuan, C. E. Bakalarski, J. Villen, J. M. Kornhauser, B. Smith, D. Li, X. Zhou, S. P. Gygi, T. L. Gu, R. D. Polakiewicz, J. Rush, and M. J. Comb. Global Survey of Phosphotyrosine Signaling Identifies Oncogenic Kinases in Lung Cancer. *Cell*, 131(6):1190–1203, 2007.

32. M. Ruzzene, J. Bertacchini, A. Toker, and S. Marmiroli. Cross-talk between the CK2 and AKT signaling pathways in cancer, 2017.

33. S. Satpathy, E. J. Jaehnig, K. Krug, B. J. Kim, A. B. Saltzman, D. W. Chan, K. R. Holloway, M. Anurag, C. Huang, P. Singh, A. Gao, N. Namai, Y. Dou, B. Wen, S. V. Vasaikar, D. Mutch, M. A. Watson, C. Ma, F. O. Ademuyiwa, M. F. Rimawi, R. Schiff, J. Hoog, S. Jacobs,

A. Malovannaya, T. Hyslop, K. R. Clauser, D. R. Mani, C. M. Perou, G. Miles, B. Zhang, M. A. Gillette, S. A. Carr, and M. J. Ellis. Microscaled proteogenomic methods for precision oncology. *Nature Communications*, 11(1), 2020.

34. N. P. Shah, C. Kasap, C. Weier, M. Balbas, J. M. Nicoll, E. Bleickardt, C. Nicaise, and C. L. Sawyers. Transient Potent BCR-ABL Inhibition Is Sufficient to Commit Chronic Myeloid Leukemia Cells Irreversibly to Apoptosis. *Cancer Cell*, 14(6):485–493, 2008.

35. J. L. Snead, T. O'Hare, L. T. Adrian, C. A. Eide, T. Lange, B. J. Druker, and M. W. Deininger. Acute dasatinib exposure commits Bcr-Abl-dependent cells to apoptosis. *Blood*, 114(16):3459–3463, 2009.

36. D. L. Tabb, L. Vega-Montoto, P. A. Rudnick, A. M. Variyath, A. J. L. Ham, D. M. Bunk, L. E. Kilpatrick, D. D. Billheimer, R. K. Blackman, H. L. Cardasis, S. A. Carr, K. R. Clauser, J. D. Jaffe, K. A. Kowalski, T. A. Neubert, F. E. Regnier, B. Schilling, T. J. Tegeler, M. Wang, P. Wang, J. R. Whiteaker, L. J. Zimmerman, S. J. Fisher, B. W. Gibson, C. R. Kinsinger, M. Mesri, H. Rodriguez, S. E. Stein, P. Tempst, A. G. Paulovich, D. C. Liebler, and C. Spiegelman. Repeatability and reproducibility in proteomic identifications by liquid chromatography-tandem mass spectrometry. *Journal of Proteome Research*, 9(2):761–776, 2010.

37. S. S. Taylor and A. P. Kornev. Protein kinases: evolution of dynamic regulatory proteins. *Trends in biochemical sciences*, 36(2):65–77, feb 2011.

38. P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, A. Vijaykumar, A. P. Bardelli, A. Rothberg, A. Hilboll, A. Kloeckner, A. Scopatz, A. Lee, A. Rokem, C. N. Woods, C. Fulton, C. Masson, C. Häggström, C. Fitzgerald, D. A. Nicholson, D. R. Hagen, D. V. Pasechnik, E. Olivetti, E. Martin, E. Wieser, F. Silva, F. Lenders, F. Wilhelm, G. Young, G. A. Price, G. L. Ingold, G. E. Allen, G. R. Lee, H. Audren, I. Probst, J. P. Dietrich, J. Silterra, J. T. Webber, J. Slavič, J. Nothman, J. Buchner, J. Kulick, J. L. Schönberger, J. V. de Miranda Cardoso, J. Reimer, J. Harrington, J. L. C. Rodríguez, J. Nunez-Iglesias, J. Kuczynski, K. Tritz, M. Thoma, M. Newville, M. Kümmerer, M. Bolingbroke, M. Tartre, M. Pak, N. J. Smith, N. Nowaczyk, N. Shebanov, O. Pavlyk, P. A. Brodtkorb, P. Lee, R. T. McGibbon, R. Feldbauer, S. Lewis, S. Tygier, S. Sievert, S. Vigna, S. Peterson, S. More, T. Pudlik, T. Oshima, T. J. Pingel, T. P. Robitaille, T. Spura, T. R. Jones, T. Cera, T. Leslie, T. Zito, T. Krauss, U. Upadhyay, Y. O. Halchenko, and Y. Vázquez-Baeza. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, 17(3):261–272, 2020.

39. C. Wang, H. Xu, S. Lin, W. Deng, J. Zhou, Y. Zhang, Y. Shi, D. Peng, and Y. Xue. GPS 5.0: An Update on the Prediction of Kinase-specific Phosphorylation Sites in Proteins. *Genomics, Proteomics and Bioinformatics*, 2020.

40. S. Wiechmann, B. Ruprecht, T. Siekmann, R. Zheng, M. Frejno, E. Kunold, T. Bajaj, D. P. Zolg, S. A. Sieber, N. C. Gassen, and B. Kuster. Chemical Phosphoproteomics Sheds New Light on the Targets and Modes of Action of AKT Inhibitors. *ACS Chemical Biology*, 16(4):631–641, 2021.

41. A. Wolf-Yadlin, N. Kumar, Y. Zhang, S. Hautaniemi, M. Zaman, H.-D. Kim, V. Grantcharova, D. A. Lauffenburger, and F. M. White. Effects of HER2 overexpression on cell signaling networks governing proliferation and migration. *Molecular systems biology*, 2(1):1–15, jan 2006.

42. B. Xue, B. Jordan, S. Rizvi, and K. M. Naegle. KinPred: A unified and sustainable approach for harnessing proteome-level human kinase-substrate predictions. *PLOS Computational Biology*, 17(2):e1008681, 2021.

43. Q. Yan, T. Barros, P. R. Visperas, S. Deindl, T. a. Kadlecek, A. Weiss, and J. Kuriyan. Structural Basis for Activation of ZAP-70 by Phosphorylation of the SH2-Kinase Linker. *Molecular and cellular biology*, 33(11):2188–201, jun 2013.

44. K. Yang and L. wu Fu. Mechanisms of resistance to BCR-ABL TKIs and the therapeutic strategies: A review. *Critical Reviews in Oncology/Hematology*, 93(3):277–292, 2015.

45. H. Zhang, T. Liu, Z. Zhang, S. H. Payne, B. Zhang, J. E. McDermott, J. Y. Zhou, V. A. Petyuk, L. Chen, D. Ray, S. Sun, F. Yang, L. Chen, J. Wang, P. Shah, S. W. Cha, P. Aiyetan, S. Woo, Y. Tian, M. A. Gritsenko, T. R. Clauss, C. Choi, M. E. Monroe, S. Thomas, S. Nie, C. Wu, R. J. Moore, K. H. Yu, D. L. Tabb, D. Feny??, V. Vineet, Y. Wang, H. Rodriguez, E. S. Boja, T. Hiltke, R. C. Rivers, L. Sokoll, H. Zhu, I. M. Shih, L. Cope, A. Pandey, B. Zhang, M. P. Snyder, D. A. Levine, R. D. Smith, D. W. Chan, K. D. Rodland, S. A. Carr, M. A. Gillette, K. R. Klauser, E. Kuhn, D. R. Mani, P. Mertins, K. A. Ketchum, R. Thangudu, S. Cai, M. Oberti, A. G. Paulovich, J. R. Whiteaker, N. J. Edwards, P. B. McGarvey, S. Madhavan, P. Wang, D. W. Chan, A. Pandey, I. M. Shih, H. Zhang, Z. Zhang, H. Zhu, L. Cope, G. A. Whiteley, S. J. Skates, F. M. White, D. A. Levine, E. S. Boja, C. R. Kinsinger, T. Hiltke, M. Mesri, R. C. Rivers, H. Rodriguez, K. M. Shaw, S. E. Stein, D. Fenyo, T. Liu, J. E. McDermott, S. H. Payne, K. D. Rodland, R. D. Smith, P. Rudnick, M. Snyder, Y. Zhao, X. Chen, D. F. Ransohoff, A. N. Hoofnagle, D. C. Liebler, M. E. Sanders, Z. Shi, R. J. C. Slebos, D. L. Tabb, B. Zhang, L. J. Zimmerman, Y. Wang, S. R. Davies, L. Ding, M. J. C. Ellis, and R. R. Townsend. Integrated Proteogenomic Characterization of Human High-Grade Serous Ovarian Cancer. *Cell*, 166(3):755–765, 2016.

46. X. Zhang, T. Maity, M. K. Kashyap, M. Bansal, A. Venugopalan, S. Singh, S. Awasthi, A. Marimuthu, H. K. C. Jacob, N. Belkina, S. Pitts, C. M. Cultraro, S. Gao, G. Kirkali, R. Biswas, R. Chaerkady, A. Califano, A. Pandey, and U. Guha. Quantitative tyrosine phosphoproteomics of Epidermal growth factor receptor (EGFR) tyrosine kinase inhibitor-treated lung adenocarcinoma cells reveals potential novel biomarkers of therapeutic response. *Molecular and Cellular Proteomics*, 16(5):891–910, 2017.