Published in final edited form as:

Cancer Res. 2017 June 01; 77(11): 2810–2821. doi:10.1158/0008-5472.CAN-16-2460.

Tissue-specific signaling networks rewired by major somatic mutations in human cancer revealed by proteome-wide discovery

Junfei Zhao^{1,#}, Feixiong Cheng^{2,3,#}, and Zhongming Zhao^{1,4,*}

¹Center for Precision Health, School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston, TX 77030, USA

²Center for Cancer Systems Biology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA 02215, USA

³Center for Complex Networks Research, Northeastern University, Boston, MA 02115, USA

⁴Human Genetics Center, School of Public Health, The University of Texas Health Science Center at Houston, Houston, TX 77030, USA

Abstract

Massive somatic mutations discovered by large cancer genome sequencing projects provide unprecedented opportunities in the development of precision oncology. However, deep understanding of functional consequences of somatic mutations and identifying actionable mutations and the related drug responses currently remain formidable challenges. Dysfunction of protein post-translational modification plays critical roles in tumorigenesis and drug responses. In this study, we proposed a novel computational oncoproteomics approach, named kinome-wide network module for cancer pharmacogenomics (KNMPx), for identifying actionable mutations that rewired signaling networks and further characterized tumorigenesis and anticancer drug responses. Specifically, we integrated 746,631 missense mutations in 4,997 tumor samples across 16 major cancer types/subtypes from The Cancer Genome Atlas into over 170,000 carefully curated non-redundant phosphorylation sites covering 18,610 proteins. We found 47 mutated proteins (e.g., ERBB2, TP53, and CTNNB1) that had enriched missense mutations at their phosphorylation sites in pan-cancer analysis. In addition, tissue-specific kinase-substrate interaction modules altered by somatic mutations identified by KNMPx were significantly associated with patient survival. We further reported a kinome-wide landscape of pharmacogenomic interactions by incorporating somatic mutation-rewired signaling networks in 1,001 cancer cell lines via KNMPx. Interestingly, we found that cell lines could highly reproduce oncogenic phosphorylation site mutations identified in primary tumors, supporting the confidence in their associations with sensitivity/resistance of inhibitors targeting EGF, MAPK, PI3K, mTOR, and Wnt signaling pathways. In summary, our KNMPx approach is powerful for identifying

* Address correspondence to: Zhongming Zhao, Ph.D., School of Biomedical Informatics, The University of Texas Health Science Center at Houston, 7000 Fannin St. Suite 820, Houston, TX 77030, Phone: 713-500-3631, zhongming.zhao@uth.tmc.edu.

#These authors contributed equally to this work.

The authors declare no potential conflicts of interest.

oncogenic alterations via rewiring phosphorylation-related signaling networks and drug sensitivity/resistance in the era of precision oncology.

Keywords

Pan-cancer; Phosphorylation; Signaling Networks; Somatic Mutations; Oncoproteomics; Pharmacogenomics

INTRODUCTION

More than 46 million somatic mutations have been accumulated through several large-scale cancer genome sequencing projects such as The Cancer Genome Atlas (TCGA) and the International Cancer Genome Consortium projects (1,2). Currently, how to assess the impact of these somatic mutations in the process of tumorigenesis and disease progression is a main challenge. Considering the existing observations that many somatic mutations promote tumorigenesis by rewiring protein signaling networks (3), one potential strategy is to incorporate somatic mutations with the protein structure information and investigate them at functional sites (e.g., phosphorylation sites) (4–8). Phosphorylation-dependent signaling network is fundamental in cellular physiology and its dysfunction plays a critical role in tumorigenesis. Recently, the rapid advancements of technologies (e.g., mass spectrometry) have made a wealth of kinase-substrate interactions with the specific phosphorylation sites available (9). This provides us with a unique opportunity to interrogate the impact of somatic mutations on kinase-substrate phosphorylation sites to determine their pathophysiological roles in cancer and prioritize potentially actionable mutations (10).

Previous studies have suggested that many infrequent but specific somatic mutations may affect cancer hallmark processes and lead to tumorigenesis (11). However, a systematic investigation of the somatic mutations in phosphorylation networks and pathways governing cell signaling has not been done yet. Meanwhile, the human kinome has become one of the most important classes of drug targets in the pharmaceutical industry (12,13). So far, more than 20 drugs targeting one or more kinases have been approved for clinical use in a variety of cancers, including lung, breast, melanoma, colorectal, pancreatic, and prostate cancers (14). Moreover, as of 2012, more than 500 kinase inhibitors have been investigated as therapeutic agents, approximately a third of which are undergoing clinical trials (15,16). Frustratingly, while effective, patients treated with these kinase inhibitors eventually develop resistance, and the prolonged survivals are typically only a few months (17,18). One primary reason for resistance is that kinases are extensively involved in complex biological mechanisms through adaptive crosstalk or feedback within cellular networks (19). This suggests that it is imperative not only to investigate somatic mutations on individual kinase-substrate protein relationships but also to acquire in-depth knowledge of the interconnectivity of these networks and the biochemical specificity of phospho-catalytic events (20,21). So far, a systematic investigation of signaling networks disturbed by somatic mutations at the kinome-wide has not been done yet. There is an urgent need to develop novel computational approaches for kinome-wide identification of their rewiring signaling networks altered by somatic mutations at functional sites (e.g. phosphorylation sites) so that

the findings, especially the top candidates, can expedite cancer pharmacogenomics studies in the emerging era of precision oncology (10,22).

In this study, we incorporated the somatic missense mutations into the protein phosphorylation sites to decipher the functional role of somatic mutations. We found a significant higher mutation rate at phosphorylation sites in comparison with the whole protein sequences in all the 16 cancer types that we examined. Our analysis revealed that mutations at protein phosphorylation sites were more likely deleterious than those at non-phosphorylation sites. We examined over 740,000 missense mutations in nearly 5000 tumor-normal matching samples across 16 cancer types from TCGA. We found 47 statistically significantly mutated proteins (SMPs) that have enriched missense mutations at their phosphorylation sites in this pan-cancer analysis. We further characterized the distribution of phosphorylation site mutations in kinase-substrate network and developed a new algorithm to detect significantly tissue-specific kinase-substrate interaction modules altered by somatic mutations. Functional annotation analysis of these modules revealed the tissue-specific phosphorylation-related mechanisms of somatic mutations in cancer. We found that the significantly mutated phosphorylation sites or proteins and their perturbed network modules were highly associated with patient survival and anticancer drug responses. Finally, we reported a kinome-wide landscape of pharmacogenomic interactions through integrating somatic mutation-rewired signaling networks in approximately 5,000 primary tumors and approximately 1,000 cancer cell lines using a statistical framework.

MATERIALS AND METHODS

The detailed descriptions for Experimental Procedures are provided in supplemental file.

Collection of somatic mutations

We downloaded the tumor-normal pairwise somatic mutation data for patients from three sources: (1) Elledge Lab website at Harvard University (http://elledgelab.med.harvard.edu/?page_id=689; accessed in March 2014); (2) Sanger website: (<ftp://ftp.sanger.ac.uk/pub/cancer/AlexandrovEtAl>, December 16, 2013); and (3) COSMIC: Catalogue of Somatic Mutations in Cancer (v69). To reduce redundancy and ensure the quality of somatic mutation data, in this study, we only focused on the somatic mutations in TCGA tumor-normal matched samples from aforementioned three datasets (see supplemental Materials and Methods).

Proteome-wide phosphorylation sites

We collected over 170,000 carefully curated non-redundant phosphorylation sites covering 18,610 proteins from the PhosphoSitePlus (23) and dbPTM3 databases (24).

Construction of kinase-substrate interaction network

We compiled high-quality kinase-substrate interactions (KSI) from two sources: (i) literature-derived low-throughput experiments, and (ii) high-throughput experiments (e.g., protein microarrays and mass spectrometry-based approaches) (25). As a result, we

compiled 7,346 unique KSI pairs connecting 379 kinases and 1,961 non-kinase substrate proteins (see supplemental Methods and Materials).

Collection of somatic mutation and drug response data in 1,001 cancer cell lines

We downloaded putative somatic mutations for 1,001 cancer cell lines and natural log half maximal inhibitory concentration (IC₅₀) and the area under the dose-response curve (AUC) values for all screened cell line/drug combinations from the Genomics of Drug Sensitivity in Cancer (GDSC, <http://www.cancerrxgene.org/>).

Significance test of phosphorylation site mutations

All flanking regions with 7-residues window of phosphorylation sites as described previously (5) were tested for the recurrence of mutations. We assumed that the observed number of mutations for a given region followed a binomial distribution, binomial (T, p_{gi}), where T was the total number of mutations observed in one gene and p_{gi} was the estimated mutation rate for the region of interest in gene g_i under the null hypothesis that the region was not recurrently mutated. Using $length(g_i)$ to represent the length of protein product of gene g_i for each phosphorylation site, we computed the P value – the probability of observing more than k mutations around this phosphorylation site out of T total mutations observed in this gene – using the equation below:

$$P(X \geq k) = 1 - P(X < k) = 1 - \sum_{x=0}^{k-1} \binom{T}{x} p_{gi}^x (1-p_{gi})^{T-x} \quad (1)$$

in which $p_{gi} = \frac{window_size}{length(g_i)}$. Finally, we set the minimal P value across all the phosphorylation sites in a specific protein as the representative P value of its coding gene g_i denoted $P(g_i)$.

Highly mutated kinase–substrate module detection

We aim to integrate information from both kinase-substrate interaction network (KSIN) and co-expression networks to identify cancer-specific modules that are enriched for phosphorylation site mutations and highly co-expressed in the according tissue. Firstly, for each cancer type, we used the expression data of matched normal tissues to calculate the co-expression value for each kinase-substrate interaction and used cutoff P -value < 0.05 to construct co-expressed KSIN. For each gene, we assigned a score $S(g_i) = -\log(P(g_i))$ based on its respective P -value. Then, given one module $M = \{g_1, g_2, \dots, g_n\}$, we could define the

objective scoring function: $Obj(M) = \frac{\sum_{g \in M} (S(g) - \mu)}{\sqrt{n}}$, in which μ is the mean value of the scores for all the genes considered.

We proposed four following steps to detect densely mutated kinase-substrate modules in co-expressed KSIN:

1. Start from one of the substrates with significantly mutated phosphorylation sites and its upstream kinase and downstream substrate as the seed.

2. Select one gene from the neighbors of the current module, and use it to expand the current module. Continue this expansion until the objective score stops growing.
3. Repeat Step 2) 10,000 times and 10,000 modules are generated accordingly.
4. Construct a consensus module for each cancer type by merging the top 5% modules with the highest objective scores.

For this ensemble of the genes in the consensus module, we calculated how many times each gene appears in different suboptimal solutions and finally assigned a “confidence score” for each gene.

ANOVA model

For each drug, we constructed a drug-response vector consisting of IC_{50} values from treatment of cell lines. Then, drug-response vector was modeled as a linear combination of the tissue of origin of the cell lines, screening medium, growth properties and the status of a genomic feature. In this study, considering the data sparsity, we only performed pan-cancer analysis. A genomic feature-drug pair was tested only if the final drug-response vector contained at least 3 positive cell lines and at least 3 negative cell lines. Effect size was quantified through the Cohen's d using the difference between two means divided by a pooled standard deviation for the data. The resulting P -values were corrected by Benjamini-Hochberg method (26).

Statistics and network visualization

All the statistical analyses were performed by the R package (v3.2.3, <http://www.r-project.org/>).

RESULTS

Framework of kinome-wide network module for cancer pharmacogenomics

In this study, we developed a novel computational oncoproteomics approach, namely kinome-wide network module for cancer pharmacogenomics (KNMPx), to identify novel actionable mutations that have rewired signaling networks and to further characterize tumorigenesis and anticancer drug responses. As shown in Figure 1, KNMPx includes three components: (1) identify SMPs that harbor the enriched missense mutations at their phosphorylation sites (Figure 1A); (2) detect significantly tissue-specific rewired signaling network modules through integrating somatic mutations-altered phosphorylation sites into tissue-specific KSIN using a novel network-based statistical approach (Figure 1B); (3) build a kinome-wide landscape of pharmacogenomic interactions by integrating somatic mutation-rewired signaling networks in 5,000 primary tumors and 1,001 cancer cell lines using a statistical framework (Figure 1C). In the present KNMPx approach, we investigated a total of 746,631 somatic mutations in 4,997 tumor samples across 16 major cancer types/subtypes from TCGA. In this study, we only used missense mutations – the single nucleotide base substitutions leading to amino acid changes. Among the 16 cancer types, acute myeloid leukemia (LAML) had the lowest mean mutation rate (8 missense mutations per sample),

while uterine corpus endometrial carcinoma (UCEC) had the highest mutation rate (497 per sample). Overall, the average mutation rate across the 16 cancer types is 160.

Pan-cancer analysis of phosphorylation site mutations

We mapped the 746,631 missense mutations into 15,401 human proteins harboring phosphorylation sites. Then, we obtained 8,057 unique missense mutations directly at the phosphorylation sites across 4,331 unique human proteins. After considering the somatic mutations at the seven immediate flanking residues of each phosphorylation site similarly applied in a previous study, we obtained 97,277 phosphorylation site mutations affecting 12,552 proteins (including 5,547 protein having recurrently mutated phosphorylation sites – at least 3 mutations affecting the same site).

We first examined the 4,331 proteins bearing mutations directly at the phosphorylation sites. We compared mutation frequency in the whole protein sequences with that at the phosphorylation sites across 16 cancer types. We found a significantly higher mutation rate at the direct phosphorylation sites than that of the whole protein sequence in 9 cancer types among the total 16 cancer types (Figure 2A). For the instance of breast invasive carcinoma (BRCA), the average mutation rate is 13.8 per million amino acids at phosphorylation sites. This compared with 5.5 mutations per million amino acids at the whole protein level. The mutation rate at the phosphorylation sites is nearly three times of that at the whole protein level ($P = 1.0 \times 10^{-61}$, Wilcoxon rank-sum test). On the sample level, in pan-cancer analysis, more than 50% of samples harbored somatic mutation at phosphorylation site in at least one protein (Table S1). Especially in lung squamous cell carcinoma (LUSC) and skin cutaneous melanoma (SKCM), the percentage of samples having phosphorylation site mutations is up to 80%. These observations indicated that phosphorylation sites were significantly altered by somatic missense mutations.

Phosphorylation site mutations tend to be deleterious mutations

We tested the hypothesis that somatic missense mutations at phosphorylation sites have more deleterious impact than missense mutations at other sites. We measured the functional impact scores of mutations using two popular and complementary tools: SIFT (27) and PolyPhen-2 (28). We examined the cumulative distribution of SIFT and PolyPhen-2 scores for phosphorylation site mutations (direct position and seven flanking residues) and non-binding site mutations (amino acid positions excluding the binding sites and seven flanking residues). The distribution is shown in Figure 2B and 2C. We defined the mutations with SIFT scores < 0.95 or PolyPhen-2 score > 0.909 as deleterious mutations based on previous studies (4). We found that phosphorylation site mutations were more likely to be deleterious than non-phosphorylation site mutations when they were evaluated using both SIFT ($P = 1.47 \times 10^{-6}$, Fisher's exact test, Fig. 2B) and PolyPhen-2 ($P = 0.0038$, Fig. 2C) scores.

Landscape of the significantly mutated phosphorylation sites across 16 cancer types

By analyzing 746,631 missense mutations from 4,997 tumors across 16 cancer types from TCGA using KNMPx, we found that more than 50% tumor samples harbored phosphorylation-associated single nucleotide variants (SNVs), and further identified 47 proteins that harbored significantly mutated phosphorylation sites (false discovery rate $q <$

0.3) including several well-known cancer proteins: BRAF: p.T599 ($q = 2.2 \times 10^{-308}$), TP53: p.S269 ($q = 2.2 \times 10^{-45}$), and EGFR: p.T290 ($q = 3.2 \times 10^{-7}$) (Figure 3A). We next evaluated the predictions using two benchmark cancer gene sets. Specifically, we examined the enrichment of these 47 SMPs using two well-curated cancer gene datasets: (1) 547 experimentally validated cancer genes from Cancer Gene Census (29) (accessed on November 5, 2014, denoted as CGC genes), and (2) a curated list of 693 significantly mutated genes (SMGs) in cancer from TCGA publications and several other publications (details were provided in Table S2 and described in our previous study (7)). We found that the genes harboring significantly mutated phosphorylation sites were significantly enriched in CGC ($P = 1.1 \times 10^{-19}$, observed overlap = 29, expected = 3, hypergeometric test) and SMGs ($P = 1.3 \times 10^{-25}$, observed overlap = 37, expected = 3), respectively.

Next, we prioritized potential SMPs for each individual cancer type (Figure 3B, Figure S1, Table S3). In total, we identified 74 SMPs with $q < 0.3$ by integrating results in 16 cancer types, including 29 CGC proteins ($P = 4.7 \times 10^{-26}$, hypergeometric test) and 37 published SMG products ($P = 1.5 \times 10^{-37}$). Using the benchmark classification from Vogelstein (30), these 74 SMPs included 9 tumor suppressor proteins (TP53, APC, PIK3R1, ATM, SMAD4, SOX9, VHL, CREBBP, and ARID1A) and 5 oncogene proteins (BRAF, CTNNB1, EGFR, KIT, and ERBB2). In the case of BRCA, we found 2 SMPs: ERBB2: Y772 ($q = 0.00043$), TP53: S269 ($q = 0.0071$). Both are well-known breast cancer related proteins. For ERBB2, which encodes HER2, four mutations (V777L, D769Y, D769H, and I767M) may alter its phosphorylation site of p.Y772. Three of them (V777L, D769Y and D769H) were functionally characterized as activating mutations in *HER2* gene amplification-negative breast cancer (31). Interestingly, R273H on TP53 altering its phosphorylation site p.S269 was recently reported as an oncogenic mutation in the mouse model (32). In the analysis of UCEC, we found 9 SMPs with $q < 0.3$ including CTNNB1: p.S37 ($q = 1.7 \times 10^{-93}$), CCND1: p.T286 ($q = 0.00049$), PIK3R1: p.T576 ($q = 0.0024$), and TP53: p.S269 ($q = 0.0047$). CTNNB1 (Catenin beta-1) is the key downstream component of the canonical Wnt signaling pathway, which plays crucial roles in the regulation of diverse cell behaviors, including cell fate, proliferation, survival, differentiation, migration and polarity (33). In BLCA, there were 4 SMPs: AHNK: p.S1943 ($q = 0.0012$), TP53: p.T155 ($q = 0.012$), ILF3: p.S482 ($q = 0.044$) and ARID1A: p.S607 ($q = 0.1$). AHNK is originally identified as a nuclear phosphoprotein in human neuroblastomas and skin epithelial cells. A recent study reported that AHNK could mediate negative regulation of cell growth and act as novel tumor suppressor through potentiation of TGF β signaling (34). ILF3 promotes breast tumor progression by regulating sustained urokinase-type plasminogen activator expression (35). Altogether, we demonstrated that KNMPx is a useful approach for identifying cancer-associated proteins harboring enriched somatic mutations at their phosphorylation sites.

The distribution of mutated phosphorylation sites on kinase-substrate network

We next investigated tissue-specific phosphorylation site mutations in kinase-substrate network. For each cancer type, the Pearson correlation coefficient (PCC) of each kinase-substrate pair in KSIN was calculated based on the normalized gene expression data from RNASeqV2 of the matched normal tissue in TCGA. Here, we defined a putative somatic mutation-altered kinase-substrate interaction if the corresponding substrate harbored any

phosphorylation site mutations. Accordingly, we found that in 9 cancer types the kinase-substrate interactions affected by phosphorylation site mutations were significantly shifted towards higher correlation coefficients when compared to the distributions expected for the randomly selected kinase-substrate interactions that are not affected by phosphorylation site mutations ($P < 0.05$ in 9 of 12 cancer types that had matched normal tissues, Kolmogorov-Smirnov test). The most significant patterns were observed in LUSC ($P = 5.9 \times 10^{-6}$), UCEC ($P = 0.0026$), and lung adenocarcinoma (LUAD) ($P = 0.0028$), which is consistent with the reports that phosphorylation site mutations have been frequently observed in these three cancer types. Altogether, the clustering of phosphorylation site mutations in highly correlated kinase-substrate interactions that we observed might suggest tissue-specific patterns of phosphorylation site mutations in KIN. Standing on this feature, we next integrated tissue-specific co-expression information in order to search for the highly mutated modules in kinase-substrate network.

Identifying the consensus mutated module in tissue-specific kinase-substrate network

We hypothesized that phosphorylation site mutations rewired kinase-substrate interaction networks and further associated with tumorigenesis or anticancer drug responses (36). As shown in KNMPx framework (Figure 1B), we searched the consensus mutant modules via 4 steps as described in Experimental Procedure section. We identified one consensus module for each cancer type, resulting in a total of 16 consensus modules (Table S4). We then combined 16 consensus mutant modules which included 330 genes for follow up pan-cancer analysis. We found that these 330 genes were significantly enriched in CGC ($P = 8.4 \times 10^{-24}$, hypergeometric test) and the previously reported SMGs ($P = 3.1 \times 10^{-31}$, hypergeometric test). In addition, genes in the consensus mutant modules for all the 16 individual cancer types were also significantly enriched in CGC (P values ranging from 2.8×10^{-11} to 0.017, hypergeometric test) and previously reported SMGs (P values ranging from 1.3×10^{-21} to 0.0035, hypergeometric test). These enrichment analyses revealed the functional relevance of consensus mutant modules identified by our KNMPx approach. We further examined the correlation of expression of genes in the consensus modules with the patient survival information using LUAD and ovarian serous cystadenocarcinoma (OV) as two examples. We found that the expression of genes in consensus modules was highly correlated with the patient survival (Figure S2). For examples, high expression of several genes in consensus modules was significantly associated with poor survival, such as *CD44* ($P = 0.01$) and *LCK* ($P = 0.044$) in OV, and *BAD* ($P = 0.012$) and *RGPD4* ($P = 0.048$) in LUAD.

In UCEC, the final consensus module (mutated in 90/248 samples) consisted of 11 genes including 2 well-known cancer genes (*CTNNB1* and *TP53*) but also the genes with mutations that were too rare to be significant by single-gene tests (Figure 4A). Most of these genes have literature evidence supporting their roles in cancer. For instance, we found Bruton's tyrosine kinase (BTK), a key component of B cell receptor (BCR) signaling and functions as an important regulator of cell proliferation and cell survival in various B cell malignancies, has three somatic mutations around its phosphorylation site p.S21. A previous study has reported that Btk phosphorylation at p.S21 creates a binding site for the prolyl isomerase Pin1, which modulates Btk activity in a cell cycle-dependent manner (37). We observed significant mutual exclusivity across the 11 genes in the module which was largely

due to the previously reported mutual exclusivity between *TP53* and *CTNNB1* mutations in UCEC (Figure 4B) (38). For example, 66 missense mutations on *CTNNB1* in UCEC around site p.S37 (Figure 4C and 4D) might alter β -catenin signaling in colon cancer (39). A previous study has suggested that the mutations within a pathway tended to exhibit mutual exclusivity (40). Enrichment analysis also showed that 7 of them (*GSK3B*, *FYN*, *RASSF1*, *TP53*, *BTB ABL1*, and *MAP2*) belong to developmental process annotated in Gene Ontology ($P = 7.21 \times 10^{-3}$, <http://geneontology.org/>). These observations suggested that the phosphorylation site mutations in these 11 genes might define a subtype of UCEC.

Kinome-wide landscape of pharmacogenomic interactions

Although it is well known that phosphorylation is frequently involved in the binding of drug inhibitors to target proteins, there has not been a systematic survey of the mutations at the phosphorylation sites against drug sensitivity. Here, we analyzed putative somatic mutations across 1,001 human cancer cell lines and how their phosphorylation site mutations could correlate with drug sensitivity or resistance to the 265 drugs examined.

First, we examined the phosphorylation site mutation profiles from cancer cell lines and for each cancer cell line, we evaluated the extent of recapitulating the mutational profile observed in the corresponding primary tumor genomes. To this end, we applied the same method in analyzing primary tumor data to map the somatic mutations in cell lines to protein structures. Because kidney chromophobe (KICH) did not have the corresponding cell line data, we analyzed 15 cancer types only. Among the 277,073 putative somatic mutations revealed in these cancer cell lines, 29,821 were at the phosphorylation sites (located at the direct phosphorylation sites or their 7-residue flanking regions) and half of them (15,626/29,821) also occurred in primary tumor. When comparing the 47 SMPs that we identified in primary tumor pan-cancer analysis, the correlation between the frequency of phosphorylation site mutations in cell lines and primary tumors was greater than 0.5 (from 0.53 in LAML to 0.97 in skin cutaneous melanoma [SKCM]) for all the 13 cancer types that had at least 10 samples with both variant and drug response data in primary tumors and the corresponding cell lines with the similar origins (Figure 5). For example, thyroid carcinoma (THCA) primary tumors shows correlation of 0.70 with THCA cell lines and 0.95 with SKCM cell lines. Put together, this result confirmed that a large panel of cell lines could capture the phosphorylation site mutation patterns of primary tumors with the similar origin.

Next, we used ANOVA to identify phosphorylation site mutations as markers associated with drug responses. As shown in Figure 6A, we identified 70 statistically significant pairwise interactions between phosphorylation site mutations and drug responses by combining data in 1,001 cell lines (false discovery rate [FDR] < 30% and Cohen's $d > 0.8$). As shown in Figure 6A and 6B, high sensitivity (low IC_{50} value) for four BRAF inhibitors: Dabrafenib ($P = 1.7 \times 10^{-57}$), PLX4720 ($P = 1.0 \times 10^{-26}$), SB590885 ($P = 5.2 \times 10^{-21}$), and AZ628 ($P = 9.4 \times 10^{-8}$), was significantly associated with BRAF p.T599 mutations. In addition, five MEK inhibitors, Bicalutamide ($P = 1.7 \times 10^{-57}$), RDEA119 ($P = 1.2 \times 10^{-14}$), PD-0325901 ($P = 1.2 \times 10^{-12}$), Trametinib ($P = 8.8 \times 10^{-9}$), and CI-1040 ($P = 1.2 \times 10^{-7}$), had a higher response to NRAS p.Y64 mutated cell lines in comparison with wild-type cell lines. Interestingly, one hotspot mutation site, TP53 p.S269 also lead to a higher response to

Bicalutamide ($P = 2.2 \times 10^{-15}$) in TP53 p.S269 mutated cell lines compared to wild-type ones. Studies of multiple-target receptor tyrosine kinase inhibitors have suggested high efficiency in targeted cancer therapies (41). For example, Linifanib, which has high binding affinity with PDGF receptor beta, KDR, and colony stimulating factor 1 receptor, is effective orally in VEGF-induced uterine edema and corneal angiogenesis (42). Figure 6A and 6B indicated that a higher response of Linifanib was significantly associated with several phosphorylation site mutations on energy homeostasis enzyme, creatine kinase, M-type (CKM) ($P = 7.4 \times 10^{-8}$) and a cell metabolism gene, *SLC25A1* ($P = 3.1 \times 10^{-7}$). Furthermore, Quizartinib (AC220), a FDA-approved FLT3 inhibitor for the treatment of acute myeloid leukemia, was found significantly associated with a higher response in cancer cell lines harboring the rewired phosphorylation network by several proteins, such as MKNK1 p.S39 ($P = 1.7 \times 10^{-9}$), PKP2 p.Y161 ($P = 1.6 \times 10^{-8}$), CKM p.T313 ($P = 2.5 \times 10^{-7}$), and USP37 p.S170 ($P = 2.2 \times 10^{-7}$), in comparison with wild-type groups. This pattern of multiple targets is similar to the multiple-target RTK inhibitors like Linifanib. Put together, we identified several promising actionable mutations with their mechanisms on altering phosphorylation sites. These markers may be useful for further characterizing drug sensitivity to MAPK signaling and ERK signaling inhibitors (Figure 7).

We next explored potential mechanisms involved in the resistance of PI3K signaling inhibitors. For example, a PDK1 inhibitor, OSU-03012, showed potential for treating various cancers, such as multiple myeloma (43), brain tumor (44), and thyroid cancer. Figure 6B revealed that somatic mutations altering ATP8A1 p.S1126 might correlate with the high risk of resistance of OSU-03012 ($P = 7.6 \times 10^{-8}$). In addition, QL-VIII-58, a mTOR inhibitor, presented high risk of resistance with somatic mutations altering ABLIM3 p.T276 ($P = 1.1 \times 10^{-7}$). FH535, a Tcf/beta-catenin signaling inhibitor, showed high potential for the treatment of several cancer types, such as lung, liver, and colon cancers. Figure 6B displayed that mutations around APBB1IP p.S652 were significantly correlated with the resistance of FH535 in pan-cancer analysis. A previous study has suggested that APBB1IP plays an important role in Rap1 signaling pathway that is closely involved with beta-catenin signaling (45). Collectively, rewired signaling network altered by phosphorylation site mutations is an important mechanism in mediating the resistance of cancer therapies for various signaling pathways, such as PI3K signaling and Wnt/beta-catenin signaling (Figure 7).

DISCUSSION

In this study, we developed an oncoproteomics-based framework to perform systematic investigation of the somatic mutations by altering phosphorylation sites and the related signaling networks in approximately 5,000 tumor samples across 16 major cancer types. After physically locating a total of 746,631 somatic mutations in protein sequences, we identified 8,057 unique missense mutations directly located at the direct phosphorylation sites and 97,277 mutations in the seven residues flanking regions around phosphorylation sites. In our comparison of the mutational load distribution of missense mutations at phosphorylation sites versus other sites of the protein sequences, we found a higher mutation rate at phosphorylation sites than that of the remaining protein sites in nine cancer types we examined. Furthermore, we found that missense mutations located around phosphorylation sites were more likely to be deleterious when compared to other residues. Based on the

significance test implemented in our kinome-wide network module for cancer pharmacogenomics, KNMPx, we reported 47 proteins that harbored significantly mutated phosphorylation sites in pan-cancer analysis and 74 proteins in 16 individual cancer type analysis. Enrichment analysis indicated that these proteins were significantly enriched in two well-studied cancer gene data sets.

Recently, Creixell et al. identified six types of network-attacking mutations (NAMs), including the changes in kinase and SH2 modulation, network rewiring, and the genesis and extinction of phosphorylation sites (19). They showed that signaling networks are both dynamically (e.g., mutant molecular logic gates) and structurally rewired (e.g., both constitutive activation and inactivation of kinase and SH1 domains) in cancer cells. They further demonstrated that the amino acids in the kinase domain encoding substrate specificity form sparse networks of non-conserved residues spanning distant regions (20). For example, they found that inter-residue allostery play important roles in specificity and an evolutionary decoupling of kinase activity and specificity. However, those two studies did not provide novel network algorithms and show the relationships of NAMs with clinical investigations (survival analysis) and drug responses. Compared to previous studies (19,20), this study has two significant contributions. (1) Building on our observation of the clustering of phosphorylation site mutations in highly co-expressed kinase-substrate interactions, we proposed a new network algorithm to search for cancer-related modules consisting of a set of genes that are enriched with phosphorylation site mutations and tissue-specific kinase-substrate network. We demonstrated that network module genes were significantly enriched in well-known cancers for both individual cancer and pan-cancer analyses, as well as were associated with patient survival. (2) We reported a kinome-wide landscape of cancer pharmacogenomic interactions through integrating somatic mutation-rewired signaling networks in approximately 4,997 primary tumors and 1,001 cancer cell lines using a statistical framework. Surprisingly, we found that mutation profiles in cell lines could highly reproduce the oncogenic phosphorylation site mutations identified in primary tumors (Figure 5). We further demonstrated that many of these associations with drug sensitivity or resistance of various target therapeutic agents, including inhibitors targeting EGF, MAPK, PI3K, mTOR, and Wnt signaling pathways (Figures 6 and 7). All dataset and predicted results in this study are free available at website: <https://bioinfo.uth.edu/mutPhosphoProt>. To our best knowledge, this is the first study of systematic oncoproteomics analysis for investigating phosphorylation site mutations mediating tumorigenesis and anticancer drug efficacy at the kinome-level. This study, although based on computational and statistical approaches, identified many promising mutation candidates for drug development in the rapidly evolving precision oncology.

Similar to many published computational works, our results might be influenced by potential data bias, data incompleteness, and confounding factors. For example, we only observed higher mutation rates at phosphorylation sites in nine of the 16 cancer types. This may be biased due to the incompleteness of the current phosphorylation data and the lack of sufficient number of samples. To overcome the potential bias from the phosphorylation sites, we further investigated the tissue-specific kinase-substrate network models which may be potentially altered by phosphorylation site mutations. Although two tumors may not have any phosphorylation site mutations in common, they may share the biological networks

affected by these mutations. By investigating phosphorylation site mutations in the context of network, we might be able to identify some rarely mutated proteins in the subnetworks linking to the well-characterized cancer-related proteins. This approach increased our power to detect potential biomarkers at the protein and network levels (46).

Another limitation is that this study only focused on missense mutations which alter phosphorylation sites by amino acid substitution and excluded other types of mutations, such as nonsense mutations, insertions/deletions (indels) or gene fusion events. All these mutations may also play important roles in altering phosphorylation sites and leading to tumor initiation and progression. Moreover, our present study mainly focused on the distribution of phosphorylation site mutations. We did not deeply characterize the potential impact of mutations by gain-of-function or loss-of-function at the kinome-level. Integrating the functional impact of phosphorylation site mutations using high-throughput phenotyping assays with our present analysis framework will be one promising direction for our next-step work (47). In addition to phosphorylation signaling, dephosphorylation signaling by phosphatases also play critical roles in various diseases, such as cancer (48). In our future work, we will investigate both phosphorylation signaling by kinases and dephosphorylation signaling by phosphatases altered by somatic mutations, which may mediate tumorigenesis and anticancer drug responses.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding support. This work was partially supported by National Institutes of Health grant R01LM011177 (Z. Zhao). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

The authors would like to thank Dr. Pora Kim for developing the mutPhosphoProt website.

Abbreviations

CGC	Cancer Gene Census
FDR	false discovery rate
IC₅₀	half maximal inhibitory concentration
KSI	kinase-substrate interactions
KNMPx	kinome-wide network module for cancer pharmacogenomics
NAMs	network-attacking mutations (NAMs)
PCC	Pearson correlation coefficient
SMGs	significantly mutated genes
SMPs	significantly mutated proteins

SNVs	single nucleotide variants
TCGA	The Cancer Genome Atlas

References

1. Hudson TJ, Anderson W, Aretz A, Barker AD, Bell C, Bernabé RR, et al. International network of cancer genome projects. *Nature*. 2010; 464:993–8. [PubMed: 20393554]
2. Weinstein JN, Collisson EA, Mills GB, Shaw KRM, Ozenberger BA, Ellrott K, et al. The cancer genome atlas pan-cancer analysis project. *Nat Genet*. 2013; 45:1113–20. [PubMed: 24071849]
3. Pawson T, Warner N. Oncogenic re-wiring of cellular signaling pathways. *Oncogene*. 2007; 26:1268–75. [PubMed: 17322911]
4. Vuong H, Cheng F, Lin CC, Zhao Z. Functional consequences of somatic mutations in cancer using protein pocket-based prioritization approach. *Genome Med*. 2014; 6:81. [PubMed: 25360158]
5. Reimand J, Bader GD. Systematic analysis of somatic mutations in phosphorylation signaling predicts novel cancer drivers. *Mol Syst Biol*. 2013; 9:637. [PubMed: 23340843]
6. Takeda K, Takemoto C, Kobayashi I, Watanabe A, Nobukuni Y, Fisher DE, et al. Ser298 of MITF, a mutation site in Waardenburg syndrome type 2, is a phosphorylation site with functional significance. *Hum Mol Genet*. 2000; 9:125–32. [PubMed: 10587587]
7. Zhao J, Cheng F, Wang Y, Arteaga CL, Zhao Z. Systematic prioritization of druggable mutations in ~5000 genomes across 16 cancer types using a structural genomics-based approach. *Mol Cell Proteomics*. 2016; 15:642–56. [PubMed: 26657081]
8. Shen Q, Cheng F, Song H, Lu W, Zhao J, An X, et al. Proteome-scale investigation of protein allosteric regulation perturbed by somatic mutations in 7,000 cancer genomes. *Am J Hum Genet*. 2017; 100:5–20. [PubMed: 27939638]
9. Aebersold R, Mann M. Mass-spectrometric exploration of proteome structure and function. *Nature*. 2016; 537:347–55. [PubMed: 27629641]
10. Fleuren ED, Zhang L, Wu J, Daly RJ. The kinome 'at large' in cancer. *Nat Rev Cancer*. 2016; 16:83–98. [PubMed: 26822576]
11. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011; 144:646–74. [PubMed: 21376230]
12. Cohen P, Tcherpakov M. Will the ubiquitin system furnish as many drug targets as protein kinases? *Cell*. 2010; 143:686–93. [PubMed: 21111230]
13. Cohen P. Protein kinases—the major drug targets of the twenty-first century? *Nat Rev Drug Discovery*. 2002; 1:309–15. [PubMed: 12120282]
14. Knight ZA, Lin H, Shokat KM. Targeting the cancer kinome through polypharmacology. *Nat Rev Cancer*. 2010; 10:130–7. [PubMed: 20094047]
15. Knox C, Law V, Jewison T, Liu P, Ly S, Frolkis A, et al. DrugBank 3. 0: a comprehensive resource for 'omics' research on drugs. *Nucleic Acids Res*. 2011; 39:D1035–D41. [PubMed: 21059682]
16. Cohen P, Alessi DR. Kinase drug discovery—what's next in the field? *ACS Chem Biol*. 2012; 8:96–104. [PubMed: 23276252]
17. Hata AN, Niederst MJ, Archibald HL, Gomez-Caraballo M, Siddiqui FM, Mulvey HE, et al. Tumor cells can follow distinct evolutionary paths to become resistant to epidermal growth factor receptor inhibition. *Nat Med*. 2016; 22:262–9. [PubMed: 26828195]
18. Gharwan H, Groninger H. Kinase inhibitors and monoclonal antibodies in oncology: clinical implications. *Nat Rev Clin Oncol*. 2016; 13:209–27. [PubMed: 26718105]
19. Creixell P, Schoof EM, Simpson CD, Longden J, Miller CJ, Lou HJ, et al. Kinome-wide decoding of network-attacking mutations rewiring cancer signaling. *Cell*. 2015; 163:202–17. [PubMed: 26388441]
20. Creixell P, Palmeri A, Miller CJ, Lou HJ, Santini CC, Nielsen M, et al. Unmasking determinants of specificity in the human kinome. *Cell*. 2015; 163:187–201. [PubMed: 26388442]

21. Tokheim C, Bhattacharya R, Niknafs N, Gyax DM, Kim R, Ryan MC, et al. Exome-scale discovery of hotspot mutation regions in human cancer using 3D protein structure. *Cancer Res.* 2016; 76:3719–31. [PubMed: 27197156]
22. Shrager J, Tenenbaum JM. Rapid learning for precision oncology. *Nat Rev Clin Oncol.* 2014; 11:109–18. [PubMed: 24445514]
23. Hornbeck PV, Kornhauser JM, Tkachev S, Zhang B, Skrzypek E, Murray B, et al. PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res.* 2012; 40:D261–70. [PubMed: 22135298]
24. Lu C-T, Huang K-Y, Su M-G, Lee T-Y, Bretaña NA, Chang W-C, et al. DbPTM 3. 0: an informative resource for investigating substrate site specificity and functional association of protein post-translational modifications. *Nucleic Acids Res.* 2013; 41:D295–305. [PubMed: 23193290]
25. Newman RH, Hu J, Rho HS, Xie Z, Woodard C, Neiswinger J, et al. Construction of human activity-based phosphorylation networks. *Mol Syst Biol.* 2013; 9:655. [PubMed: 23549483]
26. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological).* 1995:289–300.
27. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.* 2003; 31:3812–4. [PubMed: 12824425]
28. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nat Methods.* 2010; 7:248–9. [PubMed: 20354512]
29. Forbes SA, Bindal N, Bamford S, Cole C, Kok CY, Beare D, et al. COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res.* 2011; 39:D945–50. [PubMed: 20952405]
30. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW. Cancer genome landscapes. *Science.* 2013; 339:1546–58. [PubMed: 23539594]
31. Bose R, Kavuri SM, Searleman AC, Shen W, Shen D, Koboldt DC, et al. Activating HER2 mutations in HER2 gene amplification negative breast cancer. *Cancer Discovery.* 2013; 3:224–37. [PubMed: 23220880]
32. Soussi T, Wiman K. TP53: an oncogene in disguise. *Cell Death Differ.* 2015; 22:1239–49. [PubMed: 26024390]
33. Anastas JN, Moon RT. WNT signalling pathways as therapeutic targets in cancer. *Nat Rev Cancer.* 2013; 13:11–26. [PubMed: 23258168]
34. Lee I, Sohn M, Lim H, Yoon S, Oh H, Shin S, et al. Ahnak functions as a tumor suppressor via modulation of TGFβ/Smad signaling pathway. *Oncogene.* 2014; 33:4675–84. [PubMed: 24662814]
35. Hu Q, Lu Y-Y, Noh H, Hong S, Dong Z, Ding H-F, et al. Interleukin enhancer-binding factor 3 promotes breast tumor progression by regulating sustained urokinase-type plasminogen activator expression. *Oncogene.* 2013; 32:3933–43. [PubMed: 22986534]
36. Reimand J, Wagih O, Bader GD. The mutational landscape of phosphorylation signaling in cancer. *Sci Reports.* 2013; 3:2651.
37. Yu L, Mohamed AJ, Vargas L, Berglöff A, Finn G, Lu KP, et al. Regulation of Bruton tyrosine kinase by the peptidylprolyl isomerase Pin1. *J Biol Chem.* 2006; 281:18201–7. [PubMed: 16644721]
38. Kandoth C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, et al. Mutational landscape and significance across 12 major cancer types. *Nature.* 2013; 502:333–9. [PubMed: 24132290]
39. Chan TA, Wang Z, Dang LH, Vogelstein B, Kinzler KW. Targeted inactivation of CTNNB1 reveals unexpected effects of beta-catenin mutation. *P Natl Acad Sci USA.* 2002; 99:8265–70.
40. Zhao J, Zhang S, Wu L-Y, Zhang X-S. Efficient methods for identifying mutated driver pathways in cancer. *Bioinformatics.* 2012; 28:2940–7. [PubMed: 22982574]
41. Knight ZA, Lin H, Shokat KM. Targeting the cancer kinome through polypharmacology. *Nat Rev Cancer.* 2010; 10:130–7. [PubMed: 20094047]

42. Albert DH, Tapang P, Magoc TJ, Pease LJ, Reuter DR, Wei RQ, et al. Preclinical activity of ABT-869, a multitargeted receptor tyrosine kinase inhibitor. *Mol Cancer Therapeut.* 2006; 5:995–1006.
43. Zhang S, Suvannasankha A, Crean CD, White VL, Johnson A, Chen CS, et al. OSU-03012, a novel celecoxib derivative, is cytotoxic to myeloma cells and acts through multiple mechanisms. *Clin Cancer Res.* 2007; 13:4750–8. [PubMed: 17699852]
44. Booth L, Cruickshanks N, Ridder T, Chen CS, Grant S, Dent P. OSU-03012 interacts with lapatinib to kill brain cancer cells. *Cancer Biol Ther.* 2012; 13:1501–11. [PubMed: 22990204]
45. Tsai IC, Amack JD, Gao ZH, Band V, Yost HJ, Virshup DM. A Wnt-CKIvarepsilon-Rap1 pathway regulates gastrulation by modulating SIPA1L1, a Rap GTPase activating protein. *Dev Cell.* 2007; 12:335–47. [PubMed: 17336901]
46. Leiserson MD, Vandin F, Wu H-T, Dobson JR, Eldridge JV, Thomas JL, et al. Pan-cancer network analysis identifies combinations of rare somatic mutations across pathways and protein complexes. *Nat Genet.* 2015; 47:106–14. [PubMed: 25501392]
47. Berger AH, Brooks AN, Wu X, Shrestha Y, Chouinard C, Piccioni F, et al. High-throughput phenotyping of lung cancer somatic mutations. *Cancer Cell.* 2016; 30:214–28. [PubMed: 27478040]
48. Duan G, Li X, Kohn M. The human DEPhosphorylation database DEPOD: a 2015 update. *Nucleic Acids Res.* 2015; 43:D531–5. [PubMed: 25332398]

Figure 1. Workflow of kinome-wide network module identification for cancer pharmacogenomics

(A) Somatic mutation profiles, kinase phosphorylation spectrum, tissue-specific co-expression network, and drug response data were integrated for kinome-wide network module search in cancer pharmacogenomics. (B) Diagram of the algorithm to detect the consensus mutant kinase-substrate network modules. (C) Identifying new cancer proteins or network module genes harboring the enriched somatic mutations at their phosphorylation sites that are also associated with patient survival and drug responses.

Figure 2. Mutation frequencies and distribution patterns in 16 cancer types

(A) Missense mutation frequencies across 16 major cancer types. Distribution of missense mutation frequencies (number of mutations per 1 million residues) in whole protein sequences (All) versus phosphorylation sites in 16 major cancer types. (B, C) Cumulative distribution of deleterious mutations at direct phosphorylation sites, their seven immediate flanking residues, and outside positions. Cumulative frequencies of SIFT (B) and PolyPhen-2 scores (C) for direct phosphorylation sites (abbreviated as D), ± 7 flanking residues (D + 1-7 flanking positions), and outside positions. The 16 major cancer types are acute myeloid leukemia (LAML), bladder urothelial carcinoma (BLCA), breast invasive carcinoma (BRCA), colon and rectal adenocarcinoma (COAD/READ), glioblastoma multiforme (GBM), head and neck squamous cell carcinoma (HNSC), kidney chromophobe (KICH), kidney renal clear cell carcinoma (KIRC), lung adenocarcinoma (LUAD), lung

squamous cell carcinoma (LUSC), ovarian serous cystadenocarcinoma (OV), prostate adenocarcinoma (PRAD), skin cutaneous melanoma (SKCM), stomach adenocarcinoma (STAD), thyroid carcinoma (THCA), and uterine corpus endometrial carcinoma (UCEC).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 3. Manhattan plot of putative significantly mutated proteins identified in pan-cancer analysis (A) and individual cancer analysis (4 selected individual cancer types) (B)
Each dot represents a gene or its protein. Red dots represent the Cancer Gene Census (CGC) genes. Yellow dots represent the significantly mutated genes collected from literatures (Table S2). The horizontal red line denotes the false discovery rate at 0.05. The putative significantly mutated proteins identified in the remaining 12 cancer types were showed in Figure S1. Abbreviations of 16 cancer types in Figures 3 and 5 are provided in the legend of Figure 2.

Figure 4. The consensus mutant kinase-substrate network module identified for uterine corpus endometrial carcinoma (UCEC)

(A) The subnetwork for 11 genes in the final consensus module for UCEC. The size of nodes denotes the frequencies of genes identified in individual network modules quantified by the objective score (Figure 1B). (B) The mutual exclusivity of the phosphorylation site mutation profiles for the 11 genes in UCEC. (C) The three-dimensional model of phospho-degron motif of β-catenin (encoded by gene *CTNNB1*). PDB: 1P22 for β-TrCP1-Skp1-β-catenin complex was downloaded from PDB database (<http://www.rcsb.org/>). White-black: Skp1; purple: β-catenin phospho-degron motif; and the remaining color: β-TrCP1. Two phosphorylation sites: p.S33 and p.S37 were highlighted in red. (D) Distribution of missense mutations at p.S37 site in the 66 UCEC samples as shown in Figure 4B.

Figure 5. Representation of phosphorylation site mutations from primary tumors in cancer cell lines across 15 cancer types

Pearson correlation of phosphorylation site mutation frequency between cell lines and primary tumors for each cancer type. The number in parentheses refers to the sample size in each cancer type.

Figure 6. Kinome-wide landscape of pharmacogenomic interactions
(A) Pan-cancer ANOVA analysis for statistically significant interactions between phosphorylation site mutations and differential drug sensitivity. (B) Examples of significant pharmacogenomic interactions for mutant (Mut) vs. wild-type (WT) cell lines identified by ANOVA. The details for the remaining significant pharmacogenomic interactions identified by ANOVA are provided in Table S5. The color keys in A represent the P-value (-Log₁₀) of ANOVA analysis for the drug sensitivity (red) or drug resistance (blue) caused by phosphorylation site mutations in human cancer cell lines (see Methods).

Figure 7. A proposed model to illustrate the potential molecular mechanisms of kinome-wide landscape of pharmacogenomic interactions altered by phosphorylation site mutations as shown in Figure 6A

The FDA-approved or clinical investigational cancer drugs targeting 5 selected signaling pathways, including PI3K signaling, MAPK signaling, EFG signaling, mTOR signaling, and Wnt signaling, were illustrated. The data was provided in Figure 6 and Table S5.