Published in final edited form as:

J Proteomics. 2013 October 8; 91: . doi:10.1016/j.jprot.2013.07.023.

Quantitative Phosphoproteomic Profiling of Human Non-Small Cell Lung Cancer Tumors

Devin K. Schweppe¹, James R. Rigas^{2,3}, and Scott A. Gerber^{1,2,4,*}

¹Department of Genetics, Geisel School of Medicine, Lebanon, NH 03756

²Norris Cotton Cancer Center, Geisel School of Medicine, Lebanon, NH 03756

³Department of Medicine, Geisel School of Medicine, Lebanon, NH 03756

⁴Department of Biochemistry, Geisel School of Medicine, Lebanon, NH 03756

Abstract

Non-small cell lung cancer (NSCLC) is the leading cause of cancer-related deaths worldwide. Within the molecular scope of NSCLC, a complex landscape of dysregulated cellular signaling has emerged, defined largely by mutations in select mediators of signal transduction, including the epidermal growth factor receptor (EGFR) and anaplastic lymphoma (ALK) kinases. Consequently, these mutant kinases become constitutively activated and targets for chemotherapeutic intervention. Encouragingly, small molecule inhibitors of these pathways have shown promise in clinical trials or are approved for clinical use. However, many protein kinases are dysregulated in NSCLC without genetic mutations. To quantify differences in tumor cell signaling that are transparent to genomic methods, we established a super-SILAC internal standard derived from NSCLC cell lines grown *in vitro* and labeled with heavy lysine and arginine, and deployed them in a phosphoproteomics workflow. We identified 9019 and 8753 phosphorylation sites in two separate tumors. Relative quantification of phosphopeptide abundance between tumor samples allowed for the determination of specific hubs and pathways differing between each tumor. Sites downstream of Ras showed decreased inhibitory phosphorylation (Raf/Mek) and increased activating phosphorylation (Erk1/2) in one tumor versus another. In this way, we were able to quantitatively access oncogenic kinase signaling in primary human tumors.

Keywords

phosphoproteomics; lung cancer; SILAC; quantitative proteomics

Introduction

Lung cancer is the leading cause of cancer related death in the world. An estimated 156,940 Americans (28% of all cancer deaths) in 2011 and 1.4 million people worldwide in 2008 die annually from lung cancer¹. The majority of diagnosed lung cancers (85%) are classified as non-small cell in origin (non-small cell lung carcinoma, NSCLC); adenocarcinoma and squamous-cell carcinoma are the most common NSCLC histological subtypes. Currently,

© 2013 Elsevier B.V. All rights reserved.

*To whom correspondence should be addressed: scott.a.gerber@dartmouth.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

late detection and poor prognosis are common in NSCLC, leading to high mortality rates. For example, more than 60% of NSCLC patients are diagnosed with locally advanced or metastatic disease, which significantly reduces the overall 5-year survival rate to 15%. These clinical difficulties are due at least in part to an incomplete knowledge of the molecular mechanisms underlying NSCLC and therefore an inability specifically target and treat patients². Within the molecular scope of the disease, a variety of cellular processes have been implicated, including cellular proliferation pathways downstream of KRAS, EGFR, ALK, HER2, c-Met, and c-Raf^{3, 4}. Hyperactivation of these kinase pathways has been shown to play a role in cell growth, tumor invasion and metastasis³⁻⁵. The action and variety of these and other kinase pathways in NSCLC makes understanding global signal transduction in individual human tumors an important goal of translational oncology.

Quantitative shotgun proteomics is now widely recognized as a potent technology for discovery-based analyses of complex biological systems. Specifically, bottom-up shotgun LC-MS/MS platforms, in combination with a variety of quantitative workflows, has made it possible to identify thousands of potential affecter/affected proteins involved in diverse biological processes, as well as to establish their degree of enrichment or perturbation^{4, 6}.

Many technologies have been developed for the quantitative analysis of cellular proteins and peptides, including chemical labeling methods (e.g. iTRAQ⁷ and ICAT⁸), internal standards (e.g. AQUA^{9, 10}), and label free analyses (e.g. spectral counting¹¹). In addition, stable isotope labeling of amino acids in cell culture (SILAC) makes use of metabolic labeling of essential amino acids to create isotopic peptide pairs that can be compared and quantified using mass spectrometry¹². SILAC has been employed in a diverse set of reports, some of which investigate relative protein abundances, while others focus on the abundance differences of post-translationally modified peptides, including phosphopeptides^{5, 6, 13, 14}. In this light, SILAC coupled to modern LC-MS/MS platforms provides a strong technical framework for the determination of dynamic phosphorylation changes in cultured human cells.

To extend the SILAC methodology for use in quantifying primary tissues, Ishihama *et al.* developed an internal standard of cell culture derived proteins that had been labeled with heavy amino acids which they termed “culture derived isotope tags (CDIT)” to quantify differences between proteins in primary murine brain tissue and neuronal cell lines¹⁵. Relative differences between different tissue samples were quantified as a ratio of ratios, using a common, SILAC-derived internal standard (Figure 1). Later, additional cell lines in culture were added to improve the robustness of peptide quantification for orphan analytes in breast carcinoma tumors that were not represented (or present at very low intensity) in a single SILAC breast cancer cell line¹⁶. Recently, Monetti *et al.* investigated the phosphorylation profile of murine livers in response to insulin using a single heavy-labeled liver hepatocyte cell line as an internal standard¹⁷. In the context of cancer, quantitative proteomics has been applied to clinical samples including B-cell lymphoma subtype¹⁸, colonic adenoma¹⁹, and vulvar squamous cell carcinoma²⁰ proteomes, as well as the breast cancer N-glycoproteome²¹ and phosphoproteome²². Yet, the application of these technologies to primary lung cancer tumors remains undone.

To better understand misregulated cellular signaling in primary NSCLC tissues, we developed and deployed a CDIT/Super-SILAC-based strategy to interrogate the differences in phosphorylation signatures exhibited by two tumors resected from patients diagnosed with NSCLC. Our strategy establishes that the phosphoproteomes of primary human lung tumors can be mapped and quantified using heavy internal standard proteomes derived from multiple human lung cancer cell lines. For each tumor, we sequenced roughly 9,000 phosphorylation sites on more than 3,200 proteins. Through quantifying the differences in

phosphorylation between them, we observed a number of disparate phosphoproteome signaling network signatures (e.g. ErbB, Raf/MEK/ERK, and ATR/Chk2) in the two tumors. Importantly, many of these kinase networks are known to be frequently misregulated in NSCLC. In addition, quantitative analysis of the unphosphorylated supernatants from phosphopeptide isolation demonstrated that the magnitude of phosphorylation changes observed in tumors cannot be explained by only protein abundance differences, suggesting that the tumor phosphoproteome may contain considerably greater dynamic information than the proteome alone. Taken together, these approaches and data expand our ability to identify and understand kinase signaling in human cancers.

Materials and Methods

Tumor harvesting and tissue culture

Resected tumors from de-identified patients previously diagnosed with non-small cell lung cancer were acquired through an IRB-approved protocol. Tumors were snap frozen within 30 minutes of surgical resection and stored at -80°C until an alysis (see below), as done previously^{23, 24}. The mass of both tumors was approximately 0.2 grams wet weight at the onset of the analytical pipeline. Three quarters of the tumor mass was used for mass spectrometry, and the remainder was saved for SDS-PAGE and immunoblotting.

NCI-H23, H1975, and H2170 (internal standard (IS) mix) were obtained from ATCC (Manassas, VA) and cultured in arginine, lysine, and leucine-free RPMI Medium 1640 (Gibco) containing 75 mg/L of both $^{13}\text{C}_6$, $^{15}\text{N}_2$ lysine and $^{13}\text{C}_6$, $^{15}\text{N}_4$ arginine (Cambridge Isotope Labs), 15mg/L proline (Fisher), 50 mg/L leucine (Fisher), 1mM pyruvate (CellGro), 10% dialyzed FBS (HyClone) and 1% penicillin/streptomycin (HyClone). Cells were grown for at least 6 doublings in heavy media (median incorporation levels for H23, H1975 and H2170 were 98%, 95%, and 98%, respectively; Supplementary Figure 2), harvested by trypsinization from ~90% confluent plates followed by washing twice in DPBS (CellGro) prior to being snap frozen.

Phosphopeptide sample preparation and SCX separation

Cells were resuspended in lysis buffer (8M urea, 150mM NaCl, 50mM Tris-Cl pH 8.7, phosphatase inhibitors (1mM sodium tartrate, 1mM sodium orthovanadate, 1mM sodium fluoride, 1mM sodium molybdate, and 1mM α -glycero-phosphate), and Roche complete EDTA-free protease inhibitor cocktail), followed by sonication (3 $\times$ 10s, rest 1min on ice), reduction (2mM DTT, 50°C for 20 minutes) and alkylation (after cooling to RT, 8mM iodoacetamide for 45 minutes, followed by 4mM DTT quench for 15 minutes). After a 5-fold dilution in 50mM Tris-Cl pH 8.7, lysates were digested overnight with sequencing grade trypsin (Promega) at a 1:200 ratio (w/w) of trypsin:protein at 37°C . The peptide samples were then acidified to ~pH 2.5 using trifluoroacetic acid (Burdick-Jackson), desalted on a C₁₈ SepPak column (Waters; 100mg sorbent; wash: 100% methanol; equilibration: 3% methanol/0.1% TFA; elution: 60% methanol/0.1% formic acid) and lyophilized to dryness. Peptides from all three IS cell lines were then mixed 1:1:1.

Tumors (A and B) were individually resuspended in lysis buffer, dounce homogenized, spun at 3000 rpm for 15 min at 4°C to remove insoluble tumor masses, sonicated, reduced, alkylated, and digested as above. The mixed IS peptides were added 1:1 to each tumor lysate (4.8mg of IS peptides added to 4.8mg of tumor peptides for both tumors).

Lyophilized peptide mixtures were resuspended in SCX buffer A (5mM KH_2PO_4 , 30% acetonitrile [ACN], pH 2.65), spun down at 13000 for 10 min to remove insoluble particulates, and separated on a gradient from 0 to 21% buffer B (5mM KH_2PO_4 , 350mM KCl, 30% ACN, pH 2.65) on a 9.1mm polysulfoethyl aspartamide SCX column (PolyLC).

The eluate was collected in 24 fractions; each fraction was lyophilized, resuspended in 0.1% trifluoroacetic acid, and desalted on an Oasis C₁₈ solid-phase extraction (SPE) plate (Waters) as above and dried by vacuum centrifugation.

Each desalted SCX fraction was then enriched for phosphopeptides in a batch TiO₂ format using 5µm Titansphere beads (GL Sciences). Briefly, desalted peptide fractions were dried and resuspended in binding buffer (2M lactic acid/50% ACN). Peptide fractions were added to TiO₂ beads (375µg) pre-equilibrated in binding buffer and vortexed at room temperature for 1 hour, followed by washing twice with binding buffer and once with 50% ACN/0.1% TFA. Unbound peptides were dried by vacuum centrifugation, desalted on an SPE plate as above, dried by vacuum centrifugation and saved for proteomic analysis and protein quantification. Phosphopeptide elution from TiO₂ beads was carried out with two ten minute incubations with 50mM KH₂PO₄ in ammonia water, pH 10.1. Eluates were neutralized with 50% ACN/5% formic acid, dried by vacuum centrifugation and desalted on an SPE plate as above. Finally, samples were resuspended in 5% ACN, 2% formic acid and analyzed by LC-MS/MS.

Mass spectrometry and peptide matching

Each SCX fraction was analyzed in technical duplicate by separation on an in-house packed polymer-fritted trap column (1.5 cm, 100 µm inner diameter (ID), Reprosil C₁₈ 5µm 200 Å pore (Dr. Maisch, Ammerbuch, Germany)) and eluted by split flow onto an in-house pulled (Sutter P-2000, Sutter Instruments, San Francisco, CA) 100µm ID) analytical column packed with 3µm C₁₈ material (Reprosil), using a 70 minute gradient from 3 to 29% Buffer B (Buffer A: 3% ACN, 0.125% formic acid; Buffer B: 95% ACN, 0.125% formic acid) at a flow rate of ~400nl/min. The LTQ-Orbitrap (control software v 2.5.5, build 4, 06/20/2008) was run as follows: one Orbitrap survey scan (AGC Orbitrap target value: 700K; R = 60K; max. ion time 800 ms; mass range: 400 to 1400 m/z; lock mass: 445.120029) followed by ten data-dependent tandem mass spectra on each of the top ten most abundant precursor ions (isolation width: 1.6 m/z; CID relative collision energy: 35%; MS1 signal threshold: 12,500; AGC LTQ target value: 3,500; maximum MS/MS ion time: 125 ms; dynamic exclusion: repeat count of 1, maximum exclusion list size of 500, 24 seconds wide in time, +/- 20 ppm wide in m/z; doubly- and triply-charged precursors only; no neutral loss dependent or multi-stage activation methods)²⁵. Spectra were searched against a concatenated forward and reversed human SwissProt/UniProt database (downloaded September, 2010; containing 74,338 reverse & forward proteins) using SEQUEST^{26, 27} with a mass tolerance of +/- 1 Da and requiring fully tryptic peptides with up to three missed cleavages, carbamidomethylcysteine as a fixed modification and oxidized methionine and phosphorylated serine/threonine/tyrosine as variable modifications. Search results were filtered to a false discovery rate (FDR) of <1% based on reverse hit matches by multiplying the number of reverse hits by 2 and dividing this number by the total number of non-redundant peptide identifications²⁸. Peptide filtering involved applying a narrow mass measurement cutoff to each LC-MS/MS run based on the idiosyncratic mass measurement accuracy distribution for that run (typically +/- 1.5ppm) post-search²⁹⁻³¹. Additional criteria included a delta XCorr (dCn) cutoff of 0.08, and XCorr greater than 1.5 for both +2 and +3 charge state peptides (the full dataset can be found in Supplementary Table 1). SILAC peptide ratios were calculated using an in-house modified version of Xpress³² and log₂ normalized. Sample distributions were zero-shifted based on the median log₂ values^{16, 33}.

Antibodies and immunoblotting

Anti-PLK1 antibody for Western blotting was purchased from Sigma, and phospho-specific antibodies were purchased from Cell Signaling Technologies (eEF2, T56; ERK1/2, S202/204) and Abcam (MCM2, S27). Tumor samples were diced and lysed in sample buffer

containing phosphatase inhibitors (as above), and protease inhibitors (as above), followed by reduction with 10mM DTT. Samples were separated via SDS-PAGE using a 4–12% gradient gel (Invitrogen), transferred to nitrocellulose and blotted with the above antibodies. HRP-conjugated secondary antibody (Jackson ImmunoResearch) binding was visualized by chemiluminescence (Pierce).

Motif generation and clustering

Three subsets of the merged Tumor A and B overlapping datasets were taken based on the quartiles of the quantified modsite distributions (Up in A: Q3 to maximum Ratio-of-ratio; overlap: Q1 to Q3; up in B: minimum to Q1). Phosphorylation site motifs from each subset were generated using default values in Motif-x³⁴. Motifs were then clustered using ClustalW2 set to default values (Protein weight matrix: Gonnet; Gap open: 10; Gap extension: 0.20; Gap distances: 5; No end gaps: no; Iteration: none; Numiter: 1; Clustering: NJ; Format: Aln w/numbers; Order: aligned)³⁵.

Results and Discussion

The highly dynamic phosphoproteome represents a snapshot of all steady-state cellular signal transduction cascades, including those that have been implicated in many canonical transformative processes such as MEK/ERK, p53, ErbB2, and AKT signaling³⁶ as well as other potential networks yet to be discovered. Herein, we sought to specifically and quantitatively explore the phosphoproteome of primary human lung tumor samples. To this end we deployed a super-SILAC-like scheme to elucidate these global phosphosignaling networks.

For quantification, our SILAC internal standard (IS) consisted of two adenocarcinoma cell lines (NCI-H23 and H1975) and one squamous cell line (NCI-H2170). Multiple lung cancer cell lines were employed to enhance the coverage of the heavy internal standard for the tumor phosphoproteome (Figure 1; Supplementary Figure 1). These cell lines were selected to represent tumors affected by pathways frequently found to be dysregulated in NSCLC, including p53 disruption^{37, 38} (H23, H1975, H2170) and constitutive activation of Ras^{38, 39} (H23) and EGFR^{38, 40} (H1975), as well as for their relative ease in preferentially incorporating heavy lysine and arginine, without unwanted conversion to other amino acids^{41, 42}. Analysis of heavy-only proteins from SDS-PAGE separated cell lysates from each labeled cell line yielded a median incorporation rate of 95 – 98% (Supplementary Figure 2). Validation of the extent of heavy labeling is a key component of the super-SILAC workflow, given that some analytes may be present at significantly lower abundances in tumors than in cell culture, and that incomplete labeling thus compresses the dynamic range with which such analytes can be quantified.

Two tumors were resected from separate patients diagnosed with NSCLC adenocarcinoma and were subjected to our quantitative phosphoproteomic pipeline (Figure 1). These tumors were snap frozen immediately after resection and stored at –80 °C until lysis to maintain potentially labile phosphorylation events^{23, 24}. The two tumors were then lysed and the lysates were standardized using a protein concentration assay, followed by digestion with trypsin. Immediately following digestion, an equivalent amount (by mass) of heavy internal standard protein digest was mixed with each tumor sample, followed by desalting of the combined sample/internal standard mixture. Aliquots of these digests were then separated by strong cation exchange (SCX) chromatography into 24 fractions, enriched for phosphopeptides (TiO₂), and analyzed by reverse phase LC-MS/MS in technical duplicates on an LTQ-Orbitrap. The supernatants from these TiO₂ enrichments were also retained, desalted and analyzed as above to establish unmodified protein abundance differences, as well.

After filtering to a peptide false discovery rate (FDR) of $< 1\%$ (the protein FDR was estimated at 1.5% and 1.1% for Tumor A and B, respectively), the individual Tumor A and Tumor B datasets consist of 9019 and 8753 total unique phosphorylation sites on 3212 and 3323 proteins, respectively with a consistent distribution of single, double, and triple phosphorylated peptides (Figure 2A, B; Table 1, 2; Supplementary Table 2). Of these sites, ~25% of each tumor dataset was not previously annotated in Uniprot, Phospho.ELM or PhosphositePlus^{43–45} (3138 and 3043 sites in Tumor A and B, respectively). We identified these previously unreported sites with primary database search characteristics that are indistinguishable from the rest of our data (Supplementary Table 3), lending credibility to their peptide assignments. In addition, this subset of phosphorylation sites was enriched for pathways and functions associated with tumor invasion (i.e. actin cytoskeletal regulation, integrin signaling, and leukocyte extravasation signaling; Supplementary Figure 3), suggesting that a possible reason for the lack of prior annotation could be discrepancies between primary tumor phosphoproteomes (this dataset) and cell line based analyses (Uniprot, Phospho.ELM, PhosphoSite Plus). Individual phosphorylation sites for which localization of the specific phosphorylation site was not possible were considered ambiguous and were condensed into unique identifiers (“ModSites”) to increase confidence in identification of near-neighbor phosphorylation sites and to improve the rate of peptide quantification⁶. Merging all quantifiable phosphorylation sites collapsed the datasets down to 7198 ModSites and 7080 ModSites for Tumor A and B, respectively (Table 1). After quantification, individual $\log_2$ light (tumor)-to-heavy (IS) ModSite ratios were zero-corrected using the median $\log_2$ ratio value of the distribution of ratios for each tumor (Figure 2C,D)^{16, 33}. Within our median-normalized datasets, 80% and 74% of the phosphopeptide identifications showed ratio differences of ten-fold or less between the internal standard and Tumor A or Tumor B, respectively, allowing for accurate quantification of the majority of identified phosphopeptides. Technical replicate injections of all SCX fractions from both tumors allowed us to visualize the quantitative precision of the approach by plotting the average tumor-to-IS ratio for those phosphopeptides identified and quantified in both technical replicates for each tumor (Supplementary Figure 4). In addition, each dataset showed a pseudo-normal distribution of $\log_2$ ratios, with modest tailing at low tumor-to-IS ratios, indicative of a population of phosphopeptides in tumors with increased phosphorylation relative to cell lines. We hypothesize that these tumor-specific modifications could exist due to phosphorylation that is attenuated or ablated in cell culture, or is unique to the tumor micro environment.

We found that 4764 ModSites (or ~66% of each dataset) existed in both tumors (Table 1), allowing us to calculate ratios-of-ratios and quantitatively determine site-specific differences between the two tumor phosphoproteomes (Figure 3A). To profile the tumors, we separated the ModSite distribution into three subsets based on the quartiles of the ratio-of-ratios of Tumor A/Tumor B (i.e. minimum A/B ratio (−13.17) to quartile 1 (Q1) (−1.42): increased phosphorylation in Tumor B over Tumor A; Q1 to Q3 (−1.42 to 0.75) represents A/B ratios that are approximately equal; Q3 to maximum A/B ratio (0.75 to 9.46): increased phosphorylation in Tumor A over Tumor B). Based on these quartiles and the representative peptides for each serine-centered modsite (no noticeable motif based differences were found for threonine centered motifs), we then constructed phosphorylation site motifs using Motif-X³⁴ for each subset. Using ClustalW2, we aligned and clustered the motifs based on their sequence (Figure 4). Clustering showed a clear discrimination of subpopulations of phosphorylation site motifs between each tumor (e.g. Tumor B: acidic residues downstream of the phosphorylation site; Tumor A: arginines in the −2, −3, and further upstream sites) and a third population of motifs that were common between the tumors. We suggest that these tumor-specific profiles of linear phosphorylation site motifs may provide additional context when determining the kinases involved in cellular dysregulation of different tumors.

Next, using commercial signaling pathway analysis software, we identified a strong enrichment of factors within the overlap of the two datasets known to play a role in cancer development and progression, either through promotion of cellular proliferation or anti-apoptotic pathways (Figure 3B). Within these pathways, we were able to identify a diverse population of phosphorylation sites, including sites found on Cdk1, p53 and p21, AKT, ATR and Raf (Figure 5, Supplementary Figure 5), many of which are known to have specific regulatory functions from prior published literature reports (Supplementary Table 4). We were then able to use this information to tease out specific regulatory modules, as exemplified by the increased inhibitory phosphorylation of the cell cycle phosphatase Cdc25 (S216) in Tumor B relative to Tumor A, which correlated with increased phosphorylation of an important Cdc25 substrate, Cdk1 in the same sample (Figure 5). Within the tumor overlap, we found differential phosphorylation that was observed from initial membrane signal (e.g. ErbB2, ITG5A, c-Met), through multi-functional hubs (e.g. AKT, Raf, ATR, Pi3K) and on to direct transcriptional regulators (e.g. Mef2D, TP53, ELK3, ATF7) (Figure 5). In addition, we were able to track potential tumor-specific epigenetic alterations to the proteome (Supplementary Table 5). This is highlighted by the elevation of H2AX Ser139/140 and HDAC4 Ser246 phosphorylation in Tumor A relative to Tumor B. These two protein phosphorylation events have been specifically implicated in the cellular DNA damage response^{46–48}. Finally, the phosphoproteins within this dataset are known to regulate a large number of transformative and oncogenic pathways including cell cycle (e.g. CDK1, Cdc25, p21, Rb1), cell adhesion (e.g. Fyn and Fak), cellular proliferation (e.g. HER2, Pi3K, AKT), and apoptotic pathways (e.g. p53, Bad, p21).

In a more targeted approach, we correlated known kinase targets of Polo-like kinase 1 (PLK1) with those sites identified in the overlap of the tumor datasets. PLK1 has been implicated in the transformation of NSCLC tumors based on increased mRNA levels of the kinase that correlate with patient survival^{49, 50}. Taking advantage of our recent study on the mitotic substrates of this kinase and relevant literature sources we identified substrate sites targeted by PLK1⁶ in the two tumors. Within the distribution of possible PLK1 phosphorylation sites, there was an enrichment for substrates in Tumor B (Figure 6A); Western blotting of tumor lysates demonstrated a higher PLK1 expression level in Tumor B versus Tumor A (Figure 6B). Based on these data, we considered a well-characterized subset of PLK1 substrate sites (including APC1, NudC, Cdc25C, Cdc23 and KIF23) and found a consistent increase in phosphorylation of these sites in Tumor B relative to Tumor A (Figure 6C). Additionally, analysis of the quartile based phosphosite subsets using Ingenuity Pathway Analysis (IPA) showed Tumor B contained an enrichment for the mitotic roles of PLK1 that was not seen in Tumor A (Supplementary Figure 6). A kinase-specific profile such as this may indicate an increased role of PLK1 in the maintenance of a given tumor, thereby possibly differentiating the efficacy of chemotherapeutic interventions targeting Plk1 in such tumors⁵¹.

To define whether our tumor phosphopeptide ratios were indicative of differences in phosphorylation events or protein abundances, we identified and quantified ~3000 proteins in the proteomes (non-phosphorylated peptides) of both Tumor A and Tumor B. We then normalized individual phosphorylation site ratios to their corresponding protein abundance and found that protein normalization had no obvious effect on the distribution of phosphorylation site ratios for the two tumors or on the average phosphorylation site ratio for the dataset as a whole (p-value > 0.2) (Supplementary Figure 7A,B, Supplementary Table 6). Consistently, we found that the average Tumor A:Tumor B ratio-of-ratios for individual phosphorylation sites differed by ~5-fold more compared to the corresponding protein abundance differences (Figure 6D, Supplementary Figure 7C). Taken together, these observations indicate that the differences in phosphorylation site abundance that we

observed in these tumors are primarily due to changes in phosphorylation site occupancy, and not protein abundance.

Finally, we performed Western blotting to validate the quantification of phosphopeptides within our tumor datasets using commercial phospho-specific antibodies. Western blots using antibodies for phosphorylation sites within ERK2 (S185/187), MCM2 (S27) and Ef2 (S56) recapitulated our median adjusted IS ratios from the tumor datasets (Figure 6D–F). Notably, S185/S187 on ERK2 illustrated an interesting site disproportion between tumors, as this mass spec and Western blot data highlighted a marked disparity in the activating phosphorylation sites of ERK2 (Figure 5, Figure 6D–F). ERK kinase activity has been implicated in a diverse set of oncogenic pathways including cellular proliferation and differentiation, apoptosis, inflammation, transformation and drug resistance^{3, 52–54}, making the ERK pathway an attractive target for small-molecule therapeutic intervention⁵². Thus, the increased Erk2 S185/S187 phosphorylation in Tumor A compared to Tumor B provides another potential feature of a therapeutically relevant phosphoproteomic signature between these two tumors (IPA network analysis showed enrichment for ERK/MEK signaling in Tumor A over Tumor B; Supplementary Figure 7). With the ability to identify and quantify specific pathway differences, such as ERK and PLK1 signaling, we believe that this template analysis could be more broadly applicable to determining the efficacy and off-target effects of clinical targeted chemotherapeutics that are known to directly impact cellular phosphorylation signaling in a patient-specific way (i.e. kinase inhibitors such as gefitinib and PLX4032)³.

Conclusions

Using a super-SILAC based methodology combined with rapid clinical capture of human tumor samples, we have validated a robust platform to quantitatively compare global phosphoproteome differences between primary human tumors. Our analysis was able to capture and quantify ~9000 phosphorylation sites from two NSCLC tumor samples derived from NSCLC tumor resections. By applying prior knowledge of kinase-specific substrate phosphorylation sites (a specific kinase “substrate-ome”) to these quantitative primary tumor data, we were able to correlate substrate phosphorylation levels with upstream kinase expression for the oncogenic kinase PLK1, which is overexpressed in the majority of human lung adenocarcinomas. Interestingly, we found that the magnitude of changes in phosphorylation site occupancy was much larger than could be explained by protein abundance differences alone, suggesting that considerable biomedically relevant information may be contained in tumor-specific phosphoproteomic signatures. More broadly, knowledge of global phosphorylation signals and the potential oncogenic kinases mediating these signal transduction events offers a promising avenue to improve and expand current drug treatment regimens in the fight against cancer.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank A.N. Kettenbach, J.M. Gilmore, J.A. Milloy and B.K. Faherty for technical assistance and thoughtful discussion concerning data analysis. We thank A. Kisselev, M. Havrda, and M. Israel for antibodies. The authors would like to acknowledge a pre-doctoral fellowship from the National Institutes of Health, National Institute of General Medical Sciences (T32-GM008704) to D.K.S. and funding from the National Cancer Institute (R01-CA155260) to S.A.G.

References

1. Siegel R, Ward E, Brawley O, Jemal A. Cancer statistics, 2011: the impact of eliminating socioeconomic and racial disparities on premature cancer deaths. *CA Cancer J Clin.* 2011; 61(4): 212–36. [PubMed: 21685461]
2. Herbst RS, Heymach JV, Lippman SM. Lung cancer. *N Engl J Med.* 2008; 359(13):1367–80. [PubMed: 18815398]
3. Roberts PJ, Der CJ. Targeting the Raf-MEK-ERK mitogen-activated protein kinase cascade for the treatment of cancer. *Oncogene.* 2007; 26(22):3291–310. [PubMed: 17496923]
4. Bose R, Molina H, Patterson AS, Bitok JK, Periaswamy B, Bader JS, Pandey A, Cole PA. Phosphoproteomic analysis of Her2/neu signaling and inhibition. *Proc Natl Acad Sci U S A.* 2006; 103(26):9773–8. [PubMed: 16785428]
5. Hammond DE, Hyde R, Kratchmarova I, Beynon RJ, Blagoev B, Clague MJ. Quantitative analysis of HGF and EGF-dependent phosphotyrosine signaling networks. *J Proteome Res.* 2010; 9(5): 2734–42. [PubMed: 20222723]
6. Kettenbach AN, Schweppe DK, Faherty BK, Pechenick D, Pletnev AA, Gerber SA. Quantitative phosphoproteomics identifies substrates and functional modules of aurora and polo-like kinase activities in mitotic cells. *Sci Signal.* 2011; 4(179):rs5. [PubMed: 21712546]
7. Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S, Purkayastha S, Juhasz P, Martin S, Bartlett-Jones M, He F, Jacobson A, Pappin DJ. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics.* 2004; 3(12):1154–69. [PubMed: 15385600]
8. Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol.* 1999; 17(10):994–9. [PubMed: 10504701]
9. Gerber SA, Rush J, Stemman O, Kirschner MW, Gygi SP. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc Natl Acad Sci U S A.* 2003; 100(12):6940–5. [PubMed: 12771378]
10. Kettenbach AN, Rush J, Gerber SA. Absolute quantification of protein and post-translational modification abundance with stable isotope-labeled synthetic peptides. *Nat Protoc.* 2011; 6(2): 175–86. [PubMed: 21293459]
11. Liu H, Sadygov RG, Yates JR. 3rd, A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal Chem.* 2004; 76(14):4193–201. [PubMed: 15253663]
12. Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics.* 2002; 1(5):376–86. [PubMed: 12118079]
13. Brewis IA, Brennan P. Proteomics technologies for the global identification and quantification of proteins. *Adv Protein Chem Struct Biol.* 2011; 80:1–44. [PubMed: 21109216]
14. Ong SE, Foster LJ, Mann M. Mass spectrometric-based approaches in quantitative proteomics. *Methods.* 2003; 29(2):124–30. [PubMed: 12606218]
15. Ishihama Y, Sato T, Tabata T, Miyamoto N, Sagane K, Nagasu T, Oda Y. Quantitative mouse brain proteomics using culture-derived isotope tags as internal standards. *Nat Biotechnol.* 2005; 23(5):617–21. [PubMed: 15834404]
16. Geiger T, Cox J, Ostasiewicz P, Wisniewski JR, Mann M. Super-SILAC mix for quantitative proteomics of human tumor tissue. *Nat Methods.* 2010; 7(5):383–5. [PubMed: 20364148]
17. Monetti M, Nagaraj N, Sharma K, Mann M. Large-scale phosphosite quantification in tissues by a spike-in SILAC method. *Nat Methods.* 2011; 8(8):655–8. [PubMed: 21743459]
18. Deeb SJ, D'Souza RC, Cox J, Schmidt-Supprian M, Mann M. Super-SILAC allows classification of diffuse large B-cell lymphoma subtypes by their protein expression profiles. *Mol Cell Proteomics.* 2012; 11(5):77–89. [PubMed: 22442255]
19. Wisniewski JR, Dus K, Mann M. Proteomic workflow for analysis of archival formalin fixed and paraffin embedded clinical samples to a depth of 10,000 proteins. *Proteomics Clin Appl.* 2012

20. Sandberg A, Lindell G, Kallstrom BN, Branca RM, Danielsson KG, Dahlberg M, Larson B, Forshed J, Lehtio J. Tumor proteomics by multivariate analysis on individual pathway data for characterization of vulvar cancer phenotypes. *Mol Cell Proteomics*. 2012; 11(7):M112 016998. [PubMed: 22499770]
21. Boersema PJ, Geiger T, Wisniewski JR, Mann M. Quantification of the N-glycosylated secretome by super-SILAC during breast cancer progression and in human blood samples. *Mol Cell Proteomics*. 2012
22. Narumi R, Murakami T, Kuga T, Adachi J, Shiromizu T, Muraoka S, Kume H, Kodera Y, Matsumoto M, Nakayama K, Miyamoto Y, Ishitobi M, Inaji H, Kato K, Tomonaga T. A strategy for large-scale phosphoproteomics and SRM-based validation of human breast cancer tissue samples. *J Proteome Res*. 2012
23. Albanell J, Lonardo F, Rusch V, Engelhardt M, Langenfeld J, Han W, Klimstra D, Venkatraman E, Moore MA, Dmitrovsky E. High telomerase activity in primary lung cancers: association with increased cell proliferation rates and advanced pathologic stage. *J Natl Cancer Inst*. 1997; 89(21):1609–15. [PubMed: 9362159]
24. Petty WJ, Dragnev KH, Memoli VA, Ma Y, Desai NB, Biddle A, Davis TH, Nugent WC, Memoli N, Hamilton M, Iwata KK, Rigas JR, Dmitrovsky E. Epidermal growth factor receptor tyrosine kinase inhibition represses cyclin D1 in aerodigestive tract cancers. *Clin Cancer Res*. 2004; 10(22):7547–54. [PubMed: 15569985]
25. Kettenbach AN, Gerber SA. Rapid and reproducible single-stage phosphopeptide enrichment of complex peptide mixtures: Application to general and phosphotyrosine-specific phosphoproteomics experiments. *Anal Chem*. 2011
26. Eng JK, McCormack AL, Yates JR III. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *Journal of the American Society for Mass Spectrometry*. 1994; 5(11):976–989.
27. Faherty BK, Gerber SA. MacroQUEST: efficient candidate-centric searching and high-resolution correlation analysis for large-scale proteomics data sets. *Anal Chem*. 2010; 82(16):6821–9. [PubMed: 20684545]
28. Elias JE, Gygi SP. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat Methods*. 2007; 4(3):207–14. [PubMed: 17327847]
29. Beausoleil SA, Villen J, Gerber SA, Rush J, Gygi SP. A probability-based approach for high-throughput protein phosphorylation analysis and site localization. *Nature Biotechnology*. 2006; 24(10):1285–1292.
30. Hsieh EJ, Hoopmann MR, Maclean B, Maccoss MJ. Comparison of Database Search Strategies for High Precursor Mass Accuracy MS/MS Data. *J Proteome Res*. 2010; 9(2):1138–43. [PubMed: 19938873]
31. Villen J, Gygi SP. The SCX/IMAC enrichment approach for global phosphorylation analysis by mass spectrometry. *Nat Protoc*. 2008; 3(10):1630–8. [PubMed: 18833199]
32. Han DK, Eng J, Zhou H, Aebersold R. Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. *Nat Biotech*. 2001; 19(10):946–951.
33. Geiger T, Wisniewski JR, Cox J, Zanivan S, Kruger M, Ishihama Y, Mann M. Use of stable isotope labeling by amino acids in cell culture as a spike-in standard in quantitative proteomics. *Nat Protoc*. 2011; 6(2):147–57. [PubMed: 21293456]
34. Schwartz D, Gygi SP. An iterative statistical approach to the identification of protein phosphorylation motifs from large-scale data sets. *Nat Biotechnol*. 2005; 23(11):1391–8. [PubMed: 16273072]
35. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. Clustal W and Clustal X version 2.0. *Bioinformatics*. 2007; 23(21):2947–8. [PubMed: 17846036]
36. Kosako H, Nagano K. Quantitative phosphoproteomics strategies for understanding protein kinase-mediated signal transduction pathways. *Expert Rev Proteomics*. 2011; 8(1):81–94. [PubMed: 21329429]

37. Takahashi T, Carbone D, Nau MM, Hida T, Linnoila I, Ueda R, Minna JD. Wild-type but not mutant p53 suppresses the growth of human lung cancer cells bearing multiple genetic lesions. *Cancer Res.* 1992; 52(8):2340–3. [PubMed: 1559236]
38. Forbes SA, Bindal N, Bamford S, Cole C, Kok CY, Beare D, Jia M, Shepherd R, Leung K, Menzies A, Teague JW, Campbell PJ, Stratton MR, Futreal PA. COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res.* 2011; 39(Database issue):D945–50. [PubMed: 20952405]
39. Zundelovich A, Elad-Sfadia G, Haklai R, Kloog Y. Suppression of lung cancer tumor growth in a nude mouse model by the Ras inhibitor salirasib (farnesylthiosalicylic acid). *Mol Cancer Ther.* 2007; 6(6):1765–73. [PubMed: 17541036]
40. Sordella R, Bell DW, Haber DA, Settleman J. Gefitinib-sensitizing EGFR mutations in lung cancer activate anti-apoptotic pathways. *Science.* 2004; 305(5687):1163–7. [PubMed: 15284455]
41. Bendall SC, Hughes C, Stewart MH, Doble B, Bhatia M, Lajoie GA. Prevention of amino acid conversion in SILAC experiments with embryonic stem cells. *Mol Cell Proteomics.* 2008; 7(9): 1587–97. [PubMed: 18487603]
42. Schmidt F, Strozynski M, Salus SS, Nilsen H, Thiede B. Rapid determination of amino acid incorporation by stable isotope labeling with amino acids in cell culture (SILAC). *Rapid Commun Mass Spectrom.* 2007; 21(23):3919–26. [PubMed: 17979108]
43. Hornbeck PV, Chabra I, Kornhauser JM, Skrzypek E, Zhang B. PhosphoSite: A bioinformatics resource dedicated to physiological protein phosphorylation. *Proteomics.* 2004; 4(6):1551–61. [PubMed: 15174125]
44. Reorganizing the protein space at the Universal Protein Resource (UniProt). *Nucleic Acids Res.* 40(Database issue):D71–5. [PubMed: 22102590]
45. Dinkel H, Chica C, Via A, Gould CM, Jensen LJ, Gibson TJ, Diella F. Phospho.ELM: a database of phosphorylation sites--update 2011. *Nucleic Acids Res.* 39(Database issue):D261–7. [PubMed: 21062810]
46. Basile V, Mantovani R, Imbriano C. DNA damage promotes histone deacetylase 4 nuclear localization and repression of G2/M promoters, via p53 C-terminal lysines. *J Biol Chem.* 2006; 281(4):2347–57. [PubMed: 16293626]
47. Stewart GS, Wang B, Bignell CR, Taylor AM, Elledge SJ. MDC1 is a mediator of the mammalian DNA damage checkpoint. *Nature.* 2003; 421(6926):961–6. [PubMed: 12607005]
48. Foster ER, Downs JA. Histone H2A phosphorylation in DNA double-strand break repair. *FEBS J.* 2005; 272(13):3231–40. [PubMed: 15978030]
49. Wolf G, Elez R, Doerner A, Holtrich U, Ackermann H, Stutte HJ, Altmannsberger HM, Rubsamen-Waigmann H, Strebhardt K. Prognostic significance of polo-like kinase (PLK) expression in non-small cell lung cancer. *Oncogene.* 1997; 14(5):543–9. [PubMed: 9053852]
50. Strebhardt K, Ullrich A. Targeting polo-like kinase 1 for cancer therapy. *Nat Rev Cancer.* 2006; 6(4):321–30. [PubMed: 16557283]
51. Raab M, Kappel S, Kramer A, Sanhaji M, Matthess Y, Kurunci-Csacsko E, Calzada-Wack J, Rathkolb B, Rozman J, Adler T, Busch DH, Esposito I, Fuchs H, Gailus-Durner V, Klingenspor M, Wolf E, Sanger N, Prinz F, Angelis MH, Seibler J, Yuan J, Bergmann M, Knecht R, Kreft B, Strebhardt K. Toxicity modelling of Plk1-targeted therapies in genetically engineered mice and cultured primary mammalian cells. *Nat Commun.* 2011; 2:395. [PubMed: 21772266]
52. Abrams SL, Steelman LS, Shelton JG, Wong EW, Chappell WH, Basecke J, Stivala F, Donia M, Nicoletti F, Libra M, Martelli AM, McCubrey JA. The Raf/MEK/ERK pathway can govern drug resistance, apoptosis and sensitivity to targeted therapy. *Cell Cycle.* 2010; 9(9):1781–91. [PubMed: 20436278]
53. McCubrey JA, Steelman LS, Chappell WH, Abrams SL, Wong EW, Chang F, Lehmann B, Terrian DM, Milella M, Tafuri A, Stivala F, Libra M, Basecke J, Evangelisti C, Martelli AM, Franklin RA. Roles of the Raf/MEK/ERK pathway in cell growth, malignant transformation and drug resistance. *Biochim Biophys Acta.* 2007; 1773(8):1263–84. [PubMed: 17126425]
54. Shaul YD, Seger R. The MEK/ERK cascade: from signaling specificity to diverse functions. *Biochim Biophys Acta.* 2007; 1773(8):1213–26. [PubMed: 17112607]

Significance

Through the use of quantitative proteomics, we demonstrated the feasibility and coverage that large scale mass spectrometry can leverage for understanding kinase networks in cancer. By incorporating Super-SILAC based quantitation into a typical pathology workflow, we were able to access and compare tumors from multiple patients in this analysis with high accuracy and dynamic range. We analyzed tumors from patients diagnosed with non-small cell lung cancer and were able to detect comprehensive phosphorylation networks relaying through known hubs of oncogenesis in lung cancer. We hereby show that it is possible to track changes to phosphorylation networks across multiple tumors, opening up the possibility that drug susceptibility and patient-specific stratification can be implemented downstream of classical pathology.

Highlights

- Proteomes of two primary NSCLC tumors were quantified and compared.
- Differential kinase networks were assembled into tumor-specific profiles.
- PLK1 substrates and ERK activation were enriched in a patient-specific way.

Figure 1. Overview of phospho-Super-SILAC (IS) workflow

Whole protein lysates were isolated from resected primary human tumors, digested and mixed with heavy internal standard before phosphopeptide enrichment. LC-MS/MS was performed on an LTQ-Orbitrap and quantified phosphopeptide and peptide ratios which were compared to give a ratio-of-ratios between tumors for each identified peptide.

Figure 2. Primary IS phosphoproteomic analysis in NSCLC tumors

(A) Venn diagram of the distribution of unique identified proteins between Tumor A and B. (B) Same as A for unique phosphorylation sites. Tumor phosphorylation sites were compared to popular phosphorylation site databases to assess novelty. (C) Histograms of median-adjusted log₂ ratios for each tumor.

Figure 3. Comparison of primary tumor ratios and molecular processes
 (A) Histogram of the overlap of the ModSites used for quantification of ratio-of-ratios. (B) Ingenuity Pathway Analysis results for cellular processes with enriched representation within the intersection of the two tumor datasets.

Figure 4. Clustering of tumor-specific motif profiles

Hierarchical clustering based on sequence alignments of motifs from within the overlap of phosphopeptides identified in each tumor dataset were arrayed based on quartile distributions of the tumor ratio-of-ratios (>A: Q3 to max, A~B: Q1–3, >B: min to Q1). Example kinase motifs were adapted from literature derived linear motif predictions (Motif-X).

Figure 5. Representative tumor-observed oncogenic pathways

Representative pathway depicting proteins and sites identified and quantified (ratio-of-ratios) within the intersection of the Tumor A and B datasets. References for interactions and regulatory sites are in Supplementary Table 4, Supplementary Data.

Figure 6. PLK1 substrates, phosphopeptide and protein quantification, and western blot validation of site-specific pIS ratios from NSCLC tumors

(A) Tumor A/Tumor B ratio-of-ratio distribution for phosphorylation sites found within the Tumor overlap that correspond with sites previously found to be PLK1 substrate sites⁶. (B) Western blot for total PLK1 from tumor whole cell lysate. (C) Select well-known sites from within the PLK1 substratome, showing site-specific TumorA/TumorB ratio-of-ratios for each site. (D) Relative phosphopeptide and protein ratios for MCM2, EF2, and ERK2 compared between Tumor A and Tumor B. (E) Coomassie stained SDS-PAGE gel of a titration of equal loading amounts of whole cell lysate from each tumor. (F) Relative immunoblotting intensities for phospho-MCM2, EF2, and ERK2 in Tumor A and Tumor B.

Table 1

Overview of data output from phosphoproteomic workflow.

Tumor Sample	# Unique Proteins	# Unique Phosphorylation Sites	# Total Modsites	Overlap	% in Overlap
A	3212	9019	7198	4764	66.2
B	3323	8753	7080	4764	67.3

Table 2

Multiplicity of phosphorylation events per peptide in each tumor.

# Phosphorylation Events	Tumor	Occurrences
Single	A	7374
	B	7448
Double	A	1656
	B	1262
Triple	A	131
	B	139