Published in final edited form as:

Curr Opin Genet Dev. 2010 February ; 20(1): 23–30. doi:10.1016/j.gde.2009.12.005.

QUANTIFYING ONCOGENIC PHOSPHOTYROSINE SIGNALING NETWORKS THROUGH SYSTEMS BIOLOGY

Amanda M. Del Rosario¹ and Forest M. White^{1,2}

¹Department of Biological Engineering, Massachusetts Institute of Technology, Building 56-787a, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

²David H. Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Building 56-787a, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

Abstract

Pathways linking oncogenic mutations to increased proliferative or migratory capacity are poorly characterized, yet provide potential targets for therapeutic intervention. As tyrosine phosphorylation signaling networks are known to mediate proliferation and migration, and frequently go awry in cancers, a comprehensive understanding of these networks in normal and diseased states is warranted. To this end, recent advances in mass spectrometry, protein microarrays, and computational algorithms provide insight into various aspects of the network including phosphotyrosine identification, analysis of kinase/phosphatase-substrates, and phosphorylation mediated protein-protein interactions. Here we detail technological advances underlying these systems level approaches and give examples of their application. By combining multiple approaches, it is now possible to quantify changes in the phosphotyrosine signaling network with various oncogenic mutations, thereby unveiling novel therapeutic targets.

Keywords

Systems biology; tyrosine phosphorylation; cellular signaling; mass spectrometry

INTRODUCTION

Numerous cellular responses are regulated by reversible phosphorylation of serine, threonine, and tyrosine residues. While tyrosine phosphorylation accounts for less than 1% of the phosphoproteome, it plays a disproportionately large role in disease as nearly 50% of the 90 human tyrosine kinases are implicated in cancer [1]. Aberrant regulation of tyrosine phosphorylation by gain- or loss-of-function mutations to kinases or phosphatases can generate alterations in cellular signaling that contribute to oncogenic malignancies [2]. In addition to direct effects on the network associated with mutation of tyrosine kinases and phosphatases, the oncogenic potential of other mutations (e.g. NF1, PTEN) is driven, in part, through alteration of tyrosine phosphorylation-mediated signaling networks. Quantitative

© 2009 Elsevier Ltd. All rights reserved.

Corresponding author: White, Forest (fwhite@mit.edu).

POTENTIAL CONFLICTS OF INTEREST

None.

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.

phosphoproteomic analysis can reveal the individual phosphorylation sites affected by each oncogenic mutation, but linking this information to functional consequence remains challenging, mostly due to our incomplete knowledge of signaling network connectivity and signal transduction through the cell. To gain a more comprehensive understanding of the tyrosine phosphorylation signaling network, the following information is required: 1) identification and quantification of tyrosine phosphorylation sites, including dynamic regulation of these sites under different conditions, 2) identification of kinases and phosphatases responsible for modification of these sites, and 3) analysis of phosphorylation dependent protein-protein interactions. This review will discuss the most promising recent technological advances in multiple assays that have begun to address each of these aspects as well as the limitations and most appropriate application for these various assays. The combination of several assays will be required to generate a comprehensive map of phosphotyrosine networks.

Tyrosine phosphorylation identification and quantification of temporal dynamics

Identification of tyrosine phosphorylation sites has traditionally been performed on a protein-by-protein basis, but advancements in mass spectrometry-based approaches have now enabled rapid, unbiased identification and quantification of hundreds of tyrosine phosphorylation sites. Specifically, enrichment based on peptide immunoprecipitation with anti-phosphotyrosine antibodies followed by immobilized metal affinity chromatography (IMAC) has now overcome the principal challenge in identification and quantification of tyrosine phosphorylation: the low level of these signals in a large background of serine/threonine phosphorylation [3,4]. Furthermore, accurate quantification of tyrosine phosphorylation across multiple biological samples is now available by encoding samples with stable isotope labels such as stable isotope labeling with amino acids in cell culture (SILAC) or isobaric tag for relative and absolute quantitation (iTRAQ). As the name implies, SILAC labeling occurs through culturing cells in media containing stable-isotope encoded amino acids which are incorporated into proteins during turnover and cell division. As demonstrated in Figure 1a, cell lysates from different conditions, cultured in different SILAC-labeled media, are mixed, and peptides resulting from proteolytic digestion of the labeled proteins can then be analyzed by liquid chromatography tandem mass spectrometry (LC-MS/MS). Since peptides containing different stable-isotope encoded amino acids differ in mass, quantification of SILAC samples occurs by comparing the peak height (or, more accurately, the chromatographic elution profile) for a given peptide found in each sample. In comparison to SILAC, iTRAQ labeling occurs later in the process, with labeling occurring through nucleophilic reaction of N-termini and lysine ϵ -amines with the iTRAQ reagent, an N-hydroxy succinimide ester. For iTRAQ, up to 8 different conditions may be labeled, mixed, and analyzed by LC-MS/MS. As the iTRAQ reagent is isobaric, a given peptide present in all samples and labeled with each of the 8 isoforms of iTRAQ will still appear as a single peak (with associated isotope envelope) in the full scan mass spectrum (Figure 1b). Quantification occurs from the MS/MS spectrum, by comparing peak areas of the iTRAQ marker ions. Since SILAC labeling occurs during cell culture, quantification errors associated with differential sample processing are minimized. At present the SILAC strategy is limited to simultaneous analysis of three samples, while iTRAQ is capable of analyzing 8 samples in a single analysis. iTRAQ labeling is inherently applicable to a broad range of samples, as labeling occurs post-cell lysis, while SILAC is primarily limited to cell culture, although stable-isotope labeling of a variety of model organisms has recently been reported. Both of these quantification strategies have additional advantages and disadvantages which have been extensively described [5]; in general either technique provides more accurate quantification compared to stable isotope label-free methods.

Even with these improvements in methodology, one of the fundamental limitations of all mass spectrometric analyses is the accuracy of peptide identification and post-translational modification site assignments. Despite advancements in searching algorithms, Chen and colleagues have demonstrated that there is no predetermined cut-off score or *p* value that can distinguish correct from incorrect peptide identification [6], leading to the conclusion that all tandem mass spectra should be manually validated to decrease false positive identifications [7].

To move from quantitative catalogs of phosphorylation sites to functional association, incorporation of phenotypic measurements has led to correlation of phosphorylation events with downstream cellular responses [8,9]. This integrative approach has been used to define differential tyrosine phosphorylation signaling downstream of ErbB receptors upon epidermal growth factor (EGF) or heregulin (HRG) stimulation in ErbB2 overexpressing human mammary epithelial cells (HMEC) [10]. In this study, computational analysis was used to correlate changes in tyrosine phosphorylation with quantitative cellular responses of cell proliferation and migration, providing a rank-ordered list of the most strongly correlated phosphorylation sites with each phenotype and thereby enabling follow-on validation studies for novel associations between phosphorylation and phenotype [11]. In a separate study, quantitative analysis of tyrosine phosphorylation in glioblastoma cells expressing titrated levels of the EGFRvIII mutant demonstrated altered utilization of the EGFR signaling pathway by the EGFRvIII mutant, leading to the identification of cross-talk between EGFRvIII and the c-Met receptor tyrosine kinase. Combinatorial inhibition of these receptors led to decreased cell viability, thus illustrating the utility of combining phosphoproteomic and phenotypic assays in the identification of novel cancer therapeutics [12,13].

Innovations in microarray technology have also enabled quantitative screening of tyrosine phosphorylation [14]. In antibody microarrays, an array is imprinted with antibodies, incubated with cell lysate, and bound proteins are detected and quantified via a secondary antibody, mimicking a sandwich ELISA. Antibody microarrays have been used to analyze protein expression and tyrosine phosphorylation of EGFR and ErbB2 signaling networks in breast cancer cell lines following EGF stimulation [15]. Although the throughput is high and the sample requirements are minimal, assays are performed with antibodies recognizing a given phosphorylation site and are therefore limited to analysis of known tyrosine phosphorylation sites. Reverse-phase or lysate microarrays utilize a different format in which cell lysates are arrayed on a microarray and probed with pan- or phosphorylation-specific antibodies. This approach has been applied to multiple biological samples ranging from cultured cell lysates to tumor tissue lysates [16]. For both of these microarray formats, the applicability to a wide range of signaling networks and the quality of the data is limited by antibody specificity. However, the accuracy of the antibody microarray tends to be greater than the reverse-phase microarray and is less reliant on highly specific phosphorylation-specific antibodies due to the combination of two antibodies recognizing different epitopes on a given protein. Lack of specificity of phosphorylation-specific antibodies is a major concern, as highlighted in a recent study of the ErbB signaling network in which phosphorylation-specific antibodies were tested for their ability to provide consistent data when used in the reverse-phase microarray or in a 1-dimensional immunoblot against the same sample. Of the 61 antibodies tested, only 4 were consistent between the two formats, and only 12 antibodies gave similar trends when microarrays were compared to immunoblots [17]. This poor specificity can severely limit the accurate interpretation of reverse-phase phosphorylation microarrays: although the general trends tend to be correct, in many cases non-specific binding generates incorrect data, both qualitatively and quantitatively. As with gene expression microarrays, alternate assays (e.g. ELISA, mass spectrometry) are recommended to validate the results from reverse-phase

phosphorylation microarrays. On a positive note, with improved antibodies targeting more phosphorylation sites in the network, microarray-based analysis of signaling networks may eventually provide a higher-throughput, lower-sample consumption alternative to MS-based phosphoproteomics for evaluating phosphorylation changes caused by oncogenic mutations.

Regulation of tyrosine phosphorylation by kinases and phosphatases – substrate identification

The equilibrium of tyrosine phosphorylation is tightly regulated by the antagonistic actions of kinases and phosphatases, and many oncogenic mutations disrupt this critical balance [18,19]. In many cases, it is not obvious how increased signaling at one node in the network affects the other nodes, or how this translates to oncogenesis. Identification of substrates and enzymatic activities of kinases and phosphatases under different biological conditions would provide insight into this process, but this identification has proven to be challenging due to several reasons, including low-abundance substrates, complex feed-back loops in the network, and overlapping substrate specificities [20].

Conventional *in vitro* kinase assays using recombinant proteins to identify kinase substrates are not only arduous, as each kinase-substrate pair must be evaluated individually, but also may be physiologically irrelevant. To bypass this process and identify endogenous substrates in an unbiased manner, Shokat and colleagues engineered analog-specific (AS) kinases containing ‘gate-keeper’ mutations to expand the ATP-binding pocket, thereby enabling the utilization of bio-orthogonal ATP analogs (Figure 2a). When combined with analog ATP γ S, AS-kinases enable direct tagging of substrates with thiophosphate. Purification of the substrates via a catch-and-release strategy followed by mass spectrometry allows for substrate and phosphorylation site identification in an unbiased, highly sensitive, and high-throughput manner [21,22]. Several AS kinase-substrates screens have been performed with yeast and murine cells, and a recent screen using AS-Cdk1, a serine/threonine kinase involved in mitotic regulation, has been performed in human cells [21,23–25]. In this latter study, nearly one third of the phosphorylation sites identified did not occur in a Cdk consensus sequence, indicating either that Cdk1 can phosphorylate non-consensus motif substrates, or that non-AS-kinases in the cell lysate can use the ATP γ S analog. We have recently shown that wild-type kinases can utilize the ATP γ S analog; by coupling stable isotope-tagging to AS-kinase substrate determination, it is possible to quantify off-target utilization of the ATP γ S analog and determine true substrates for the selected kinase [Carlson *et al.*, unpublished]. In the future, this combination of stable-isotope labeling, AS-kinases, and ATP γ S analogs will be used to discover unknown substrates of tyrosine kinases that are overexpressed or have increased activity in cancer, such as c-Src or Abl. These results will be critical in developing a better understanding of the connectivity and topology of these highly complex signaling networks.

Numerous computational techniques have been developed for predicting kinase-substrate relationships. Among these are methods, such as Scansite, which primarily rely on consensus sequence motifs recognized by the active site of the kinase, and NetworKIN, which also incorporates co-localization, co-expression, and scaffolding by using the probabilistic protein association network, STRING [26,27]. By integrating contextual information and consensus motif sequence, the prediction accuracy of NetworKIN was double that of consensus motif sequence alone. NetworKIN has also been shown to correctly predict known and novel phosphorylation sites involved in the DNA-damage response network [28]. Although this example is very promising, results from purely computational techniques such as NetworKIN are often incorrect due to inaccurate data sources: many of the automated curation algorithms incorrectly identify protein-protein interactions, and consensus motifs have been typically generated from *in vitro* kinase assays with peptide libraries, conditions which are likely not replicated in cells. Therefore, it is important to note

that although these computational methods can make reliable predictions, these predictions need to be validated using conventional techniques such as biochemical assays and *in vivo* kinase-substrate studies.

Most phosphatase substrates have been identified using substrate-trapping approaches with catalytically inactive protein tyrosine phosphatase (PTP) mutants that form stable interactions with their substrates. Similar to *in vitro* kinase assays, the traditional phosphatase substrate-trapping assay is limited to one phosphatase-substrate pair at a time, as detection is typically based on immunoblotting with commercially available antibodies [29]. Unfortunately, phosphatases have not been as extensively studied as kinases, and the knowledge of phosphatase substrates lags behind that of kinases. Nevertheless, techniques to analyze phosphatase substrates in a comprehensive and systematic manner are emerging [18,30]. Recently, mass spectrometry has been applied to the substrate-trapping technique, allowing for an unbiased analysis of PTP substrates (Figure 2b) [31,32]. Microarrays have also been used to characterize PTP substrates. In a recent study, phosphatase substrate preference was analyzed by comparing two PTP substrate-trapping mutants simultaneously: each PTP mutant was labeled with distinctive fluorescent tags and used in a microarray imprinted with known phosphotyrosine peptides [33]. Although dissociation constants were measured and used to evaluate substrate affinity, the *in vitro* nature of this approach may not be reflective of *in vivo* selectivity and enzyme substrate kinetics.

Phosphotyrosine dependent protein-protein interactions

Tyrosine phosphorylation can mediate cellular signaling by affecting protein activity, stability, localization, and by modulating protein-protein interactions through Src homology 2 (SH2) and phosphotyrosine binding (PTB) domains. Although many protein-protein interactions have been uncovered by co-immunoprecipitation experiments, this method does not work well for transient or weak interactions, and does not provide quantitative information regarding affinities of these interactions [34]. Protein microarrays provide a high-throughput solution to identification of phosphotyrosine-mediated protein-protein interactions, but can be severely limited by the non-physiological context of the assay. In one of the few quantitative applications of this technique, protein microarrays were used to identify SH2 and PTB binding partners of the phosphorylated tyrosines of the ErbB family. In this study, protein microarrays comprised of nearly every SH2 and PTB domain were probed with fluorescently labeled, tyrosine phosphorylated peptides from the ErbB family. By varying the concentration of each peptide and quantifying the fraction bound to the microarray, saturation binding curves and dissociation constants were generated for each interaction. This analysis led to the intriguing finding that overexpression increases the promiscuity of protein-protein interactions for EGFR and ErbB2, but not ErbB3, and that this increase in binding may be related to the oncogenic potential of these overexpressed proteins [35]. Thus far SH2/PTB protein microarrays have been used to evaluate tyrosine phosphorylation sites on receptor tyrosine kinases [35–38], but this technique can easily be applied to identify and evaluate binding partners of other tyrosine phosphorylation sites.

SH2 domain binding assays provide another platform to profile the global tyrosine phosphorylation state and establish network topology. In one format, SH2-domains have been used to probe far-western blots, identifying protein molecular weights capable of binding the probe, and providing signatures of changes in tyrosine phosphorylation across biological samples [39]. Although this assay is not capable of directly determining the site of tyrosine phosphorylation bound to the SH2 domain, this information may be accessible by including data obtained by quantitative mass spectrometry-based tyrosine phosphoproteomics of the same samples. A high-throughput, quantitative, reverse-phase version of the SH2 binding assay, the Rosette assay, features immobilized protein samples on nitrocellulose in a 96-well plate. In this assay, each well can be tested with a recombinant

SH2 domain probe; detection occurs with chemiluminescence and densitometric quantification. Rosette-assay analysis of whole cell lysates from v-Abl, v-Src, and v-Fps transformed NIH-3T3 cells generated distinctive quantitative patterns of SH2 domain binding. These patterns might correlate with therapeutic sensitivity, but additional validation needs to be performed [40].

CONCLUSIONS

As oncogenic mutations frequently result in alternations in tyrosine phosphorylation, defining and understanding tyrosine phosphorylation networks is paramount. To this end, recent methodological advancements in mass spectrometry enable the unbiased identification and quantification of tyrosine phosphorylation. Future directions in the field will apply multiple analytical techniques, several of which have been described here, to generate complimentary data describing additional aspects of signaling network, including topology and connectivity (Figure 3). One crucial aspect to understanding the biological impact of network activation, heretofore largely disregarded, is to combine the data from these analyses with phenotypic data describing the biological system under a variety of conditions to enable the extraction of key regulatory nodes in the network controlling cellular biological activity. The more comprehensive understanding of tyrosine phosphorylation networks that emerges from these combined analyses will help to elucidate how oncogenic mutations perturb the network and lead to aberrant biological response.

Application of additional molecular and cellular quantitative analytical techniques to a given biological system, while providing additional insight into the network, incurs significant data validation and data integration challenges. Many of the above techniques are plagued by false positive (and false negative) results due to the *in vitro* nature of the technique and/or non-specificity of reagents and detection. Perhaps the most effective solution to these data quality issues involves confirmation of results from one technique through analysis with a second technique. For instance, data gathered on reverse-phase microarrays could be confirmed by phospho-MS, and phospho-MS results could be confirmed by quantitative ELISA. Integration of disparate data types at the cellular and molecular level requires computational algorithms such as partial least-squares regression (PLSR) or machine learning to identify molecular signatures that are most highly correlated with selected cellular measurements. Several examples of PLSR-based integration of signaling and phenotypic data have been published over the past five years. It is anticipated that multiple other computational techniques will be applied to this task in the near future, especially given the complexity of the proteomic, transcriptional, and genomic data that is currently being generated to characterize the molecular characteristics of given systems.

The analytical and computational techniques described here should be applicable to a wide variety of biological systems and can be used in the future to elucidate the framework of other, non-phosphorylation-mediated, signaling networks including acetylation and ubiquitination. Consequently, a combined systems approach to understanding multiple levels of signal transduction can lead to the much broader understanding of the interaction of regulatory networks. Eventually, this data will lead to the identification of novel cancer therapeutic targets and therapeutics that maximize efficacy while minimizing deleterious side effects.

Acknowledgments

We would like to thank members of the Forest White lab for helpful discussions regarding this review. This work was supported by National Cancer Institute Grant CA118705.

References

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest

- of outstanding interest

1. Ptacek J, Snyder M. Charging it up: global analysis of protein phosphorylation. *Trends Genet.* 2006; 22:545–554. [PubMed: 16908088]
2. Blume-Jensen P, Hunter T. Oncogenic kinase signalling. *Nature.* 2001; 411:355–365. [PubMed: 11357143]
3. Rush J, Moritz A, Lee K, Guo A, Goss V, Spek E, Zhang H, Zha X, Polakiewicz R, Comb M. Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. *Nat Biotechnol.* 2005; 23:94–101. [PubMed: 15592455]
4. Zhang Y, Wolf-Yadlin A, Ross P, Pappin D, Rush J, Lauffenburger D, White F. Time-resolved mass spectrometry of tyrosine phosphorylation sites in the epidermal growth factor receptor signaling network reveals dynamic modules. *Mol Cell Proteomics.* 2005; 4:1240–1250. [PubMed: 15951569]
- 5. Tedford N, Hall A, Graham J, Murphy C, Gordon N, Radding J. Quantitative analysis of cell signaling and drug action via mass spectrometry-based systems level phosphoproteomics. *Proteomics.* 2009; 9:1469–1487. This review discusses the most recent technological advances in mass spectrometry and how these techniques are being applied to quantitative phosphoproteomics. [PubMed: 19294625]
- 6. Chen Y, Zhang J, Xing G, Zhao Y. Mascot-derived false positive peptide identifications revealed by manual analysis of tandem mass spectra. *J Proteome Res.* 2009; 8:3141–3147. Thorough analysis of the five most common reasons for false-positives and details why manual validation of mass spectra is necessary for correct identification of peptides and posttranslational modifications. [PubMed: 19368407]
- 7. Nichols A, White F. Manual validation of peptide sequence and sites of tyrosine phosphorylation from MS/MS spectra. *Methods Mol Biol.* 2009; 492:143–160. A guide as to how to manually validate mass spectra; peptide sequence and phosphorylation of tyrosine, serine, and threonine validation. [PubMed: 19241031]
8. Olsen J, Blagoev B, Gnäd F, Macek B, Kumar C, Mortensen P, Mann M. Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell.* 2006; 127:635–648. [PubMed: 17081983]
- 9. Huang P, White F. Phosphoproteomics: unraveling the signaling web. *Mol Cell.* 2008; 31:777–781. Detailed review outlining how quantitative mass spectrometry data can generate a better understanding of cellular signaling networks; focusing on the DNA damage response and ErbB family signaling networks. [PubMed: 18922462]
10. Wolf-Yadlin A, Kumar N, Zhang Y, Hautaniemi S, Zaman M, Kim H, Grantcharova V, Lauffenburger D, White F. Effects of HER2 overexpression on cell signaling networks governing proliferation and migration. *Mol Syst Biol.* 2006; 2:54. [PubMed: 17016520]
11. Kumar N, Wolf-Yadlin A, White F, Lauffenburger D. Modeling HER2 effects on cell behavior from mass spectrometry phosphotyrosine data. *PLoS Comput Biol.* 2007; 3:e4. [PubMed: 17206861]
- 12. Huang P, Mukasa A, Bonavia R, Flynn R, Brewer Z, Cavenee W, Furnari F, White F. Quantitative analysis of EGFRvIII cellular signaling networks reveals a combinatorial therapeutic strategy for glioblastoma. *Proc Natl Acad Sci U S A.* 2007; 104:12867–12872. By using quantitative mass spectrometry, demonstrated that the mutant EGFRvIII alters the EGFR signaling pathway, and mediates cross-talk with the receptor tyrosine kinase c-Met. [PubMed: 17646646]
13. Huang P, Cavenee W, Furnari F, White F. Uncovering therapeutic targets for glioblastoma: a systems biology approach. *Cell Cycle.* 2007; 6:2750–2754. [PubMed: 17986864]

- 14. Wolf-Yadlin A, Sevecka M, Macbeath G. Dissecting protein function and signaling using protein microarrays. *Curr Opin Chem Biol.* 2009; 13:1–8. Comprehensive review discussing advancements in protein microarray technology and biological applications. [PubMed: 19272831]
- 15. Nielsen U, Cardone M, Sinskey A, MacBeath G, Sorger P. Profiling receptor tyrosine kinase activation by using Ab microarrays. *Proc Natl Acad Sci U S A.* 2003; 100:9330–9335. [PubMed: 12876202]
- 16. Pierobon M, Calvert V, Belluco C, Garaci E, Deng J, Lise M, Nitti D, Mammano E, De Marchi F, Liotta L, et al. Multiplexed cell signaling analysis of metastatic and nonmetastatic colorectal cancer reveals COX2-EGFR signaling activation as a potential prognostic pathway biomarker. *Clin Colorectal Cancer.* 2009; 8:110–117.
- 17. Sevecka M, MacBeath G. State-based discovery: a multidimensional screen for small-molecule modulators of EGF signaling. *Nat Methods.* 2006; 3:825–831. [PubMed: 16990815]
- 18. Tiganis T, Bennett A. Protein tyrosine phosphatase function: the substrate perspective. *Biochem J.* 2007; 402:1–15. Reviews the challenges of phosphatase substrate identification, and proposes criteria that should be met to determine *bona fide* PTP substrates. [PubMed: 17238862]
- 19. Hunter T. Tyrosine phosphorylation: thirty years and counting. *Curr Opin Cell Biol.* 2009; 21:140–146. Fundamental review of tyrosine phosphorylation in cellular signaling. [PubMed: 19269802]
- 20. Manning B, Cantley L. Hitting the target: emerging technologies in the search for kinase substrates. *Sci STKE.* 2002; 2002:PE49. [PubMed: 12475999]
- 21. Allen J, Li M, Brinkworth C, Paulson J, Wang D, Hübner A, Chou W, Davis R, Burlingame A, Messing R, et al. A semisynthetic epitope for kinase substrates. *Nat Methods.* 2007; 4:511–516. [PubMed: 17486086]
- 22. Shah K, Liu Y, Deirmengian C, Shokat K. Engineering unnatural nucleotide specificity for Rous sarcoma virus tyrosine kinase to uniquely label its direct substrates. *Proc Natl Acad Sci U S A.* 1997; 94:3565–3570. [PubMed: 9108016]
- 23. Blethrow J, Glavy J, Morgan D, Shokat K. Covalent capture of kinase-specific phosphopeptides reveals Cdk1-cyclin B substrates. *Proc Natl Acad Sci U S A.* 2008; 105:1442–1447. [PubMed: 18234856]
- 24. Dephoure N, Howson R, Blethrow J, Shokat K, O'Shea E. Combining chemical genetics and proteomics to identify protein kinase substrates. *Proc Natl Acad Sci U S A.* 2005; 102:17940–17945. [PubMed: 16330754]
- 25. Ptacek J, Devgan G, Michaud G, Zhu H, Zhu X, Fasolo J, Guo H, Jona G, Breitkreutz A, Sopko R, et al. Global analysis of protein phosphorylation in yeast. *Nature.* 2005; 438:679–684. [PubMed: 16319894]
- 26. Obenauer J, Cantley L, Yaffe M. Scansite 2.0: Proteome-wide prediction of cell signaling interactions using short sequence motifs. *Nucleic Acids Res.* 2003; 31:3635–3641. [PubMed: 12824383]
- 27. Linding R, Jensen L, Pasculescu A, Olhovsky M, Colwill K, Bork P, Yaffe M, Pawson T. NetworKIN: a resource for exploring cellular phosphorylation networks. *Nucleic Acids Res.* 2008; 36:D695–699. [PubMed: 17981841]
- 28. Linding R, Jensen L, Ostheimer G, van Vugt M, Jørgensen C, Miron I, Diella F, Colwill K, Taylor L, Elder K, et al. Systematic discovery of in vivo phosphorylation networks. *Cell.* 2007; 129:1415–1426. [PubMed: 17570479]
- 29. Alonso A, Sasin J, Bottini N, Friedberg I, Osterman A, Godzik A, Hunter T, Dixon J, Mustelin T. Protein tyrosine phosphatases in the human genome. *Cell.* 2004; 117:699–711. [PubMed: 15186772]
- 30. Mertins P, Eberl H, Renkawitz J, Olsen J, Tremblay M, Mann M, Ullrich A, Daub H. Investigation of protein-tyrosine phosphatase 1B function by quantitative proteomics. *Mol Cell Proteomics.* 2008; 7:1763–1777. [PubMed: 18515860]
- 31. Kolli S, Zito C, Mossink M, Wiemer E, Bennett A. The major vault protein is a novel substrate for the tyrosine phosphatase SHP-2 and scaffold protein in epidermal growth factor signaling. *J Biol Chem.* 2004; 279:29374–29385. [PubMed: 15133037]

32. Wu J, Katrekar A, Honigberg L, Smith A, Conn M, Tang J, Jeffery D, Mortara K, Sampang J, Williams S, et al. Identification of substrates of human protein-tyrosine phosphatase PTPN22. *J Biol Chem*. 2006; 281:11002–11010. [PubMed: 16461343]
33. Sun H, Tan L, Gao L, Yao S. High-throughput screening of catalytically inactive mutants of protein tyrosine phosphatases (PTPs) in a phosphopeptide microarray. *Chem Commun (Camb)*. 2009:677–679. [PubMed: 19322419]
34. Pawson T, Gish G, Nash P. SH2 domains, interaction modules and cellular wiring. *Trends Cell Biol*. 2001; 11:504–511. [PubMed: 11719057]
35. Jones R, Gordus A, Krall J, MacBeath G. A quantitative protein interaction network for the ErbB receptors using protein microarrays. *Nature*. 2006; 439:168–174. [PubMed: 16273093]
36. Gordus A, Krall J, Beyer E, Kaushansky A, Wolf-Yadlin A, Sevecka M, Chang B, Rush J, MacBeath G. Linear combinations of docking affinities explain quantitative differences in RTK signaling. *Mol Syst Biol*. 2009; 5:235. [PubMed: 19156127]
37. Kaushansky A, Gordus A, Chang B, Rush J, MacBeath G. A quantitative study of the recruitment potential of all intracellular tyrosine residues on EGFR, FGFR1 and IGF1R. *Mol Biosyst*. 2008; 4:643–653. [PubMed: 18493663]
38. Kaushansky A, Gordus A, Budnik B, Lane W, Rush J, MacBeath G. System-wide investigation of ErbB4 reveals 19 sites of Tyr phosphorylation that are unusually selective in their recruitment properties. *Chem Biol*. 2008; 15:808–817. [PubMed: 18721752]
39. Machida K, Mayer B, Nollau P. Profiling the global tyrosine phosphorylation state. *Mol Cell Proteomics*. 2003; 2:215–233. [PubMed: 12754303]
40. Machida K, Thompson C, Dierck K, Jablonowski K, Kärkkäinen S, Liu B, Zhang H, Nash P, Newman D, Nollau P, et al. High-throughput phosphotyrosine profiling using SH2 domains. *Mol Cell*. 2007; 26:899–915. [PubMed: 17588523]

Figure 1. Peptide quantification via mass spectrometry

(a) Cells from 3 different conditions (e.g. unstimulated, stimulated for 5 minutes, or stimulated for 10 minutes) are cultured in different SILAC-labeled media (containing $^{12}\text{C}^{14}\text{N}$ -Arg, $^{13}\text{C}^{14}\text{N}$ -Arg, or $^{13}\text{C}^{15}\text{N}$ -Arg). Cell lysates are combined and proteolyzed to peptides, which are then analyzed via LC-MS/MS. While the peptide identification is determined by the MS/MS spectrum, the quantification data is acquired from the MS spectrum by comparing the peak height or chromatographic peak area. (b) For iTRAQ labeling, up to 8 biological samples (e.g. cells in culture or tumor samples) can be analyzed simultaneously. Following tryptic digestion, lysates are labeled with the isobaric mass tags, then combined and analyzed via LC-MS/MS. Peptide identification and quantification is determined from the MS/MS spectrum, with quantification determined by comparing peak areas of the iTRAQ marker ions.

Figure 2. Identification of kinase and phosphatase substrates

(a) Kinases engineered to accept an ATP analog (AS-kinase) can be used to identify direct substrates in an unbiased manner. By using bulky ATP γ S, an AS-kinase is able to directly thiophosphorylate its substrates, providing a chemical handle for solid-phase capture via iodoacetyl-agarose. Following proteolytic digestion, thiophosphorylated peptides are isolated on iodoacetyl-agarose beads, then released with a mild oxidizing reagent and analyzed by LC-MS/MS. (b) Phosphatase substrates can be identified in a substrate-trapping approach by using catalytically inactive protein tyrosine phosphatase (PTP) mutants that form stable interactions with their substrates. Subsequent to isolation of the PTP-substrate

pair, the captured substrates undergo proteolytic digestion and are then identified by mass spectrometry.

Figure 3. Incorporation of systems-level techniques to define tyrosine phosphorylation networks Each colored box depicts complimentary information about the EGFR network that can be gained from alternative approaches. Advances in mass spectrometry capabilities enable site specific identification and quantification of tyrosine phosphorylation (yellow box). The tyrosine phosphorylation data obtained from mass spectrometry analysis can be followed up by more targeted approaches to determine connectivity of the tyrosine phosphorylation sites. For instance, tyrosine phosphorylation dependent protein-protein interactions, *e.g.* between Shc and phosphorylated Y1173 of the EGFR, can be identified by using an SH2 domain protein microarray (red box). Kinase substrates of c-Src, can be identified through the use of an analog-specific (AS)-Src designed to utilize a bio-orthogonal ATP γ S analog (blue box). The knowledge gained from these techniques can then be combined to generate network topology, potentially leading to the development of novel cancer therapeutics.