Protein phosphorylation is an abundant posttranslational modification that functions as an important signaling cue. The balance between protein kinase and phosphatase activity mediates the dynamics of phosphorylation to transmit information flow, ultimately directing the activity, localization, and function of signaling effectors. Notably, disequilibrium in the phosphorylation–dephosphorylation balancing act can have severe repercussions, and is a major contributing factor to the manifestation of pathologies such as inflammatory diseases, diabetes, neurodegenerative diseases, metabolic disorders, and cancers. Assigning individual protein phosphorylation events to their cognate kinases and phosphatases is key to understanding the organization and dynamics of signaling networks. While advances in mass spectrometry have facilitated the identification of phosphorylation, the protein kinase(s) or phosphatase(s) responsible for regulating any particular phosphorylation event in most cases is unknown. Moreover, distinguishing functional from spurious phosphorylation remains a significant hurdle to mapping signaling pathways and information processing (Lienhard, 2008). The challenge now lies in trying to pinpoint critical phosphorylation events, their cognate kinase(s), the regulatory organization between kinases and phosphatases, and their relevance to organismal physiology.

In recent years, building on the knowledge of genome sequences and technological advances in molecular biology and proteomics, a number of approaches have been developed to systematically identify protein kinase substrates (Figure 1; Boyle and Koleske, 2007; Sopko and Andrews, 2008). For instance, protein and peptide chips have been used extensively to identify protein kinase substrates (Ptacek et al., 2005), protein kinase consensus motifs (Alexander et al., 2011; Hutti et al., 2004; Mok et al., 2010; Songyang et al., 1994; Turk et al., 2006), and those proteins with which protein kinases can associate (Fasolo et al., 2011; Kaushansky et al., 2008). Similarly, pools of radiolabeled proteins generated from coupled in vitro transcription–translation of verified cDNA libraries have served as substrates to screen for targets of recombinant protein kinases (Gao et al., 2000; Lee et al., 2005). Additionally, reverse phase protein microarrays, bearing spotted lysates rather than purified proteins, have been probed with phosphospecific antibodies to identify kinase targets (Gulmann et al., 2009). All these methods rely on interaction between kinase and substrate in an environment lacking cellular context, and the latter method additionally relies on antibody specificity and avidity. Genetically engineered analog-sensitive kinases capable of accepting bulky ATP analogs have enabled the labeling of direct protein kinase targets in cell lysates (Blethrow et al., 2008; Carlson et al., 2011; Dephoure et al., 2005; Holt et al., 2009; Ubersax et al., 2003); however, this approach relies on inconsequential movement of the kinase ATP-binding pocket. Approaches coupling affinity purification with mass spectrometry have proven fruitful in identifying physical protein kinase–substrate interactions (Breitkreutz et al., 2010; Rohila et al., 2009), although this approach, like yeast two hybrid experiments, is hindered by its inability to identify dynamic kinase–substrate interactions. The varying degree of protein kinase activity and inhibitor sensitivity across cancer cell lines and tumors has been revealed with kinase activity assay for kinome profiling (KAYAK) – a multiplexed kinase assay to simultaneously measure protein kinase activity in lysates (Kubota et al., 2009). This assay provides a survey of protein kinase activity and can be used to identify protein kinases responsible for phosphorylation of a given peptide designed a priori. Loss-of-function screens have revealed the influence of individual protein kinases on the phenotype of genetically altered cells, to predict epistasis and kinase-substrate relationships (Baldwin et al., 2008; Bommi-Reddy et al., 2008; Fiedler et al., 2009; van Wageningen et al., 2010). In the reciprocal direction, kinase substrates have been successfully

**Figure 1** Methods to link phosphorylation with culpable enzymes. Direct targets can be best predicted from the overlap of in vitro, in vivo, and in silico approaches to distinguish kinase and phosphatase substrates.
predicted by examining the viability of a sensitized kinase mutant in the context of target gain-of-function (Sopko et al., 2006). In addition to these experimental approaches, predictions of kinase substrates using computational methods are commonplace. A number of computational resources predicting kinase-specific phosphosites or phospho-binding motifs are available (Tan and Linding, 2009); however, their use is limited due to their over-predictive nature. Unquestionably, matching protein phosphorylation sites with culpable kinases is challenging, due to the limited specificity and number of kinase consensus motifs, and a number of factors including localization, expression, and association with regulatory subunits or scaffolds (Ubersax and Ferrell, 2007). As such, proving a direct link between any potential kinase-target pair still requires a considerable amount of verification. Mass spectrometry continues to develop rapidly in terms of its application to phosphoproteomics (Jünger and Aebersold, 2013). This technology is sensitive enough to capture and measure global phosphoproteome alterations in different genetic scenarios, and serves as a way to monitor altered information flow, using phosphorylation as a proxy. Notably, the increased sensitivity, speed, and dynamic range of peptide identification with current proteomics platforms is providing near comprehensive and reproducible description of simple proteomes and phosphoproteomes, making comparative analyses between genetic variants, for instance wild-type to mutants, accessible (Ahrens et al., 2010; Picotti et al., 2013). Additionally, the application of isobaric labeling strategies has facilitated multiplexing for relative peptide quantification in tissues and whole organisms. While less expansive scale in terms of the number of protein measurements, mass cytometry provides exquisite sensitivity in concurrently measuring dozens of proteins in individual cells (Bendall et al., 2012). This platform, combining inductively coupled plasma mass spectrometry with flow cytometry, currently enables the detection of trace levels of 37, although theoretically 70–100, proteins simultaneously in 500–1000 cells, per second. This eliminates the hassle of obtaining large quantities of starting material, as would be necessary for traditional phosphoproteomic approaches. The application of this technology for phosphoprotein detection however relies on the specificity of metal isotope-tagged phospho-antibodies in fixed cells. Current detection limits are 100 molecules per cell.

In essence, the identification of phosphoproteins altered specifically in genetic contexts where a protein kinase or phosphatase is mutated, removed, or depleted should drive substrate predictions, as well as help assign function to those corresponding kinases and phosphatases. Such types of experiments have been carried out in budding yeast, Arabidopsis, and cell culture using mass spectrometry to measure phosphopeptide levels (Bodenmiller et al., 2010; Chen et al., 2010; Gnad et al., 2013; Hilger et al., 2009; Smolka et al., 2007; Umezawa et al., 2013). While large-scale application of this type has been carried out in yeast with kinase and phosphatase knockouts (Bodenmiller et al., 2010), a lack of scalable genetic tools for similar systematic studies has stalled in vivo analyses in metazoan systems. Given the recently improved efficacy of RNAi during Drosophila melanogaster embryogenesis (Ni et al., 2011) and the relative ease with which one can collect large numbers of staged D. melanogaster embryos, we generated a library of shRNA-based RNAi reagents for analogous quantitative proteomic studies (Sopko et al., 2014). The library accounts for three-quarters of kinase and phosphatase encoding genes in the in D. melanogaster genome, 95% of which are conserved to human. The smaller genome and hence proteome of in D. melanogaster, as compared to human, allows for more extensive peptide detection and deeper protein coverage making in D. melanogaster more accessible for quantitative studies. Expression in the female in D. melanogaster germline of efficient protein kinase and phosphatase-targeting shRNAs permits the generation of eggs depleted of these single gene products. In order to perform a quantitative assessment of all phosphorylated proteins in eggs depleted for specific protein kinases (essentially a phosphorylation signature), we employed isobaric peptide labeling, which enables concurrent quantitation of multiple shRNA-expressing genotypes relative to one another and a common control. As a proof of principle, we carried out phosphoproteomic profiling on embryos depleted of 19 different D. melanogaster protein kinases for the purpose of systematically linking protein phosphosites with these maternally inherited kinases. To date, in D. melanogaster is the most advanced organism for which whole-animal studies of this kind have been performed. While most identified phosphosites were unchanged in terms of levels relative to a mock RNAi included in all experiments, a number of known and candidate kinase-targeted phosphosites, were altered in specific knockdown conditions. Examples of known substrates of shRNA-targeted kinases that exhibited a measurable change in phosphorylation in respective kinase-depleted embryos included Cdk1, Klp61F, and Hsp83 in ‘wee’ deficient embryos, and Histone H3, Med13 and Stat92E in Cdk8 deficient embryos. In fact, we found that phosphorylation of one-third of respective D. melanogaster orthologs of literature-curated yeast Cdk8 substrates (Sharifpoor et al., 2011) that we identified were downregulated in Cdk8 deficient D. melanogaster embryos. One phosphoproteome which exhibited a large number of candidate substrates based on pathway enrichment of downregulated phosphoproteins was that generated from embryos depleted of gilgamesh, the D. melanogaster casein kinase I gamma ortholog: the majority of phosphorylation we observed downregulated resided in proteins involved in Hedgehog (Hh) and Wnt/Wingless (Wg) pathways, consistent with a characterized role for Gilgamesh in mediating signaling of these pathways (Davidson et al., 2005; Hummel et al., 2002; Schertel et al., 2013). Specifically, we observed downregulated phosphorylation of the G protein-coupled receptor Smoothened; Combgap, a direct repressor of Cubitus interruptus (Ci) and Wg target gene expression (Campbell and Tomlinson, 2000; Song et al., 2000); Debra, a mediator of Ci polyubiquitination and degradation (Dai et al., 2003); Retinoid- and fatty acid-binding glycoprotein, a lipophorin that binds and regulates Hh and Wg diffusion (Panakova et al., 2005); Dyrk3, the human ortholog that directly phosphorylates Ci and induces its degradation (Varjosalo et al., 2008); Interference Hedgehog (Ihog) interacting proteins Modulo and Karst (Guruharsha et al., 2011); and other regulators of Hh signaling, namely Smg5, Belle, Adenomatous polyposis coli tumor suppressor homolog 2, Bx42, Chromator, Pitslre, and Rm62 (Nybakken et al., 2005). This list of altered phosphoproteins undoubtedly includes both direct and indirect
Gilgamesh targets. Obviously, validation of *bona fide* direct Gilgamesh substrates would require additional information such as *in vitro* kinase activity towards the potential substrate, a protein–protein interaction, and functional assays. Based on the observations described above, we are confident that identified changes in the *D. melanogaster* phosphoproteome can be used to screen for authentic substrates and alterations in signaling downstream of the targeted kinase.

As expected the numbers of direct and indirect candidate kinase targets, which we considered as those phosphoproteins whose levels were downregulated relative to mock knockdown more than 1.5 fold, varied depending on the identity of the kinase depleted (Sopko et al., 2014). This was expected since the number of phosphoproteins expected to have an interaction with an RNAi-depleted kinase will differ depending on the kinase, due to multiple factors including the function of the kinase, its expression level, localization, and connectivity (physical interactions with other proteins) (Ubersax and Ferrall, 2007). The number of downregulated phosphosites for the 19 kinase-deficient contexts we surveyed ranged from 752 to 22 and correlated with the degree of phenotype resulting from kinase depletion, indicating that extensive modulation of the phosphoproteome often translates to major morphological changes. For example, embryos depleted of the cyclin-dependent kinase CaDK8 display extreme morphological defects and fail to hatch, consistent with the major effects on the phosphoproteome that we observed from knockdown of this general transcriptional regulator. On the other end of the spectrum, Bub1-deficient embryos look normal, consistent with the minimal effects on the phosphoproteome observed for near complete depletion of Bub1, hinting that this mitotic kinase functions redundantly with its paralog, BubR1 (also known as Mad3). Redundant kinase targets could be uncovered from phosphoproteomic profiling of embryos deficient for two redundant kinases, for instance Bub1 and BubR1, by co-expressing shRNAs targeting both, or one kinase-targeting shRNA in a background mutant for the other kinase. Such a scenario would reveal shared substrates, since the compensating kinase activity would be absent or hampered. Early *D. melanogaster* embryos serve up a superb experimental advantage in this respect since one can restrict RNAi specifically to the germline and hence negate any effects on viability, since an intact germline is dispensable for organismal development. In this way one can examine genetic scenarios where two mutations would be synthetic lethal and otherwise unattainable, for instance in yeast where the combination would result in arrested growth. Additionally, one can fine-tune the activity of the RNAi driver with the Gal4/UAS system in *D. melanogaster* (Brand and Perrimon, 1993) and therefore modulate the extent of knockdown, by altering temperature to look at graded effects (Staller et al., 2013). Though single kinase knockdown alone may be sufficient to hint at redundancy, if kinase activity is hindered sufficiently to warrant detectable phosphosite alteration. In this case, individual phosphoproteins for redundant kinases would exhibit a similar trend in change for compromised phosphorylation at shared sites.

Overlapping phosphorylation alteration patterns for any two kinases could additionally indicate participation of the two kinases in a common pathway, since one would expect shared changes in phosphorylation for targets downstream of both kinases. One could decipher where they function by the extent to which they share similarity with phosphoprofiles generated from knockdown of known components of characterized signaling pathways. For example, to identify new components and targets of MAPK signaling, one could first define the individual phosphoproteomes associated with knockdown of Raf, MEK, and ERK, three kinases directly acting on one another in a signaling cascade. The phosphorylation of peptides representing true output of the pathway should trend in a similar direction for all three kinase knockdowns – either all up or all down – relative to knockdown controls, while phosphorylation events unique to each kinase will not. Phosphoproteomes generated from knockdown of genes in the same pathway would be qualitatively similar and share this pattern or partial pattern (Figure 2). This concept is similar to the approach that has been used extremely successfully in developmental genetics to identify components of signaling pathways, whereby genes with similar mutant phenotypes usually encode molecules that are part of the same biochemical pathways (Green et al., 2011; Nusslein-Volhard and Wieschaus, 1980). This approach has been extremely powerful in particular using the *D. melanogaster* embryonic cuticle as a phenotypic readout. In relation to phosphoproteomic data, one could infer signaling hierarchy by substituting comparisons of cuticle phenotypes for similarity in patterns of thousands of phosphorylation measurements, generated from mass spectrometry analysis.

Additional information in terms of mapping network connectivity and shared enzymatic targeting could be gleaned from combining perturbations of protein kinases and phosphatases. For example, a shared protein kinase and phosphatase substrate should exhibit an exacerbated reduction in phosphorylation in a context where protein kinase perturbation, knockdown or mutation, is combined with overexpression of a counteractive phosphatase, as compared to kinase knockdown or phosphatase overexpression alone. In this scenario, the addition of phosphate onto the shared target would be hindered and the removal of phosphate enhanced. Conversely, overexpression of the culpable kinase and a reduction in opposing phosphatase activity could lead to elevated phosphorylation. Technically, an increase in phosphorylation may in fact be easier to detect by mass spectrometry than a decrease, though these types of combinatorial experiments have not yet been reported. These experiments should be relatively easy in *D. melanogaster* and budding yeast given the large number of available P-element insertion lines bearing UAS transgenes for Gal4-mediated overexpression and plasmid-based libraries for Gal4- and endogenous promoter-driven expression, respectively. Analogous genetic manipulations as those described above, to generate scenarios of target hyperactivity, have been demonstrated to result in altered cell growth and a means to successfully predict kinases substrates (Fiedler et al., 2009; Sopko et al., 2006). In these cases, growth was the phenotypic readout, however gene expression has also been used as a measure to predict kinase and phosphatase targets and redundancy between these enzymes (van Wageningen et al., 2010). We expect that global protein phosphorylation measurements by mass spectrometry should serve as a better predictor of these relationships as compared to global gene expression, since these enzymes...
directly impact phosphorylation. In this way not only would global proteomic analyses map pathways, but could also be used to reveal critical nodes in signaling that may partially or completely overcome mutations resulting in pathway hyperactivity.

Because the number of protein phosphatases is underrepresented in the genome relative to protein kinases, one might speculate their signatures would be less specific than protein kinases and their knockdown would affect larger proportions of the phosphoproteome. In fact, there appears to be no correlation between phosphatase ablation and impact on the phosphoproteome, relative to that of protein kinases – at least from the examination of yeast null mutants (Bodenmiller et al., 2010). However, this observation could in fact be related to compensation for missing activity in deletion mutants over generations, as has been persuasively reported in yeast (Teng et al., 2013). Further support for the adaptation of null mutants is observations of different phenotypes depending on whether a kinase is targeted by a small molecule inhibitor or by complete deletion (Carroll et al., 2001; Knight et al., 2006). Consistent with this, Bodenmiller et al. (2010) observed a higher proportion of changes reflective of indirect effects with permanent deletion mutants as compared to short temporal inhibition of analog-sensitive kinases with small molecules. Conceivably, by comparing deletion mutants over generations or developmental time, one could identify compensatory rewiring events by identifying altered phosphorylation-dependent signaling by mass spectrometry. Further, by comparing complete gene knockout to incomplete depletion by RNAi-mediated knockdown with the same approach, one could identify off-target effects of RNAi reagents and compensatory effects resulting from gene knockdown, for example,
the stabilization of alternative isoforms (Figure 3). Such types of analyses have recently become more accessible given current advances to efficiently make deleterious mutations in D. melanogaster using CRISPR (Kondo, 2014).

In our recent study we demonstrated how kinase-substrate relationships may be illuminated from examining correlations in phosphosite changes measured by mass spectrometry (Sopko et al., 2014). Specifically, we revealed how two phosphosites, correlating in their directionality of change among multiple genetic contexts, multiple kinase-deficient conditions in our study, have a tendency to comprise a kinase-substrate pair. For instance, correlation of kinase autophosphorylation and phosphorylation of another protein can indicate activation of the phosphoprotein by the kinase, either directly or indirectly (Figure 4, top panel: positive correlation). On the other hand, inhibitory phosphorylation of a kinase will always be out of phase with phosphorylation of that kinase’s targets (Figure 4, bottom panel: negative correlation). We found enrichment for

![Figure 3](image_url)

**Figure 3** A workflow for integrating systems genetics and proteomics to identify similar responses to genetic manipulation. A phosphoprofile is generated from phosphosites (PS) identified by multiplexed mass spectrometry methods analyzing genetically modified cells or organisms, manipulated by conditions (c) of RNAi, CRISPR-mediated knockout, overexpression, or combinations thereof. Clustering of phosphoprofiles generated from individual conditions are dictated by PS alterations, either down (cyan) or up (yellow) in levels relative to control. Neighboring phosphoprofiles post-clustering share similarity in their cellular response to the specific genetic manipulation(s).

![Figure 4](image_url)

**Figure 4** Molecular scenarios explaining correlative and anti-correlative phosphorylation. Two phosphosites (PS) consistently responding in the same direction (positive correlation; +) or the opposite direction (negative correlation; −) in different genetic contexts can illuminate phosphosite functionality. Correlations, however, are not enough to distinguish between direct and indirect kinase targets.
genuine kinase-substrate pairs among 447 585 correlative pairs involving 2058 altered phosphosites (>1.5-fold relative to control RNAi) we identified from profiling 19 kinase-deficient genotypes. Significantly, the probability of identifying kinase-substrate pairs increased with the number of kinase-deficient conditions for which the directionality of change for the phosphosite pair correlated, positively or negatively. Application of this type of correlative analysis for the mining of existing large-scale phosphoproteomic data from different organisms and tissues has enormous implications in terms of mapping kinase-phosphatase networks and deciphering functional phosphorylation. Particularly, because uncovering functionality for a particular phosphosite does not rely on direct manipulation of the respective phosphoprotein, fewer experimental conditions than genetic manipulations are required to interpret the consequences of individual phosphorylation events. Moreover, having a priori knowledge of the function of a specific phosphosite – serving in an activating or inhibitory capacity – can facilitate interpretation of the other participating phosphosite of the correlative pair. We exemplified this in our study with a pair of phosphorylations that correlated positively in six kinase-depleted conditions: activating phosphorylation of the transcription factor STAT and a previously uncharacterized Sterile20-like kinase (Slik) phosphosite. Since this activating STAT phosphorylation is reduced in ‘slik’ deficient embryos, we predicted that Slik acts upstream of STAT and that phosphorylation of Slik kinase is required to promote STAT phosphorylation and DNA binding. We verified this prediction, by demonstrating that transcriptional targets of STAT are reduced in expression following slik knockdown. Our novel approach examining correlative phosphorylation to map kinase signaling uncovered a previously unrecognized role for Slik and pinpointed functional interaction of an individual Slik phosphosite. We anticipate that a similar strategy surveying the correlation of posttranslational modifications other than phosphorylation, derived *en masse* from mass spectrometry measurements, will shed light on the identity and mechanisms of relevant regulatory enzymes.

Given the large number of human kinase mutations associated with disease, understanding the global consequences of specific phosphorylation-impacted perturbations could prompt treatment tailored to aberrant signaling of specific pathways (Lahiry et al., 2010). While current methods such as mass spectrometry can easily identify altered phosphorylation between samples, understanding the mechanisms by which these changes arise and deciphering impactful versus inconsequential changes on cellular and organismal phenotype remains a major challenge. We have recently reported how global phosphorylation surveyed under various genetic contexts can be used to generate a matrix of phosphophosites (Sokpo et al., 2014), permitting profile comparisons and correlative phosphosite analysis to predict kinase redundancy and map signaling networks. We expect that matrix expansion, incorporating additional genetic (for instance, combinatorial gene overexpression, or gene ablation or gene product depletion) and environmental contexts, will strengthen kinase-substrate and other types of network signaling mechanism predictions. This generic strategy is attractive given it can be applied to any organism or cell type. In conclusion, while the ability to catalog cellular measurements – DNA mutations, transcript levels, splice isoforms, protein levels and modifications – is becoming cheaper, easier, and more comprehensive, efforts to understand the consequences of changes in cellular machinery are lacking and require the inception of new analysis methods, such as those described above.

**Acknowledgments**

The National Institutes of Health supported this work (5R01DK088718, 5P01CA120964). R.S. is a Special Fellow of the Leukemia and Lymphoma Society. N.P. is a Howard Hughes Medical Institute investigator.

See also: Dynamic Integration: Dynamics: Dynamics of Protein Kinase Cascades. Horizontal Integration: Omics: Phosphoproteomics: Approaches, Developments, and Challenges

**References**


