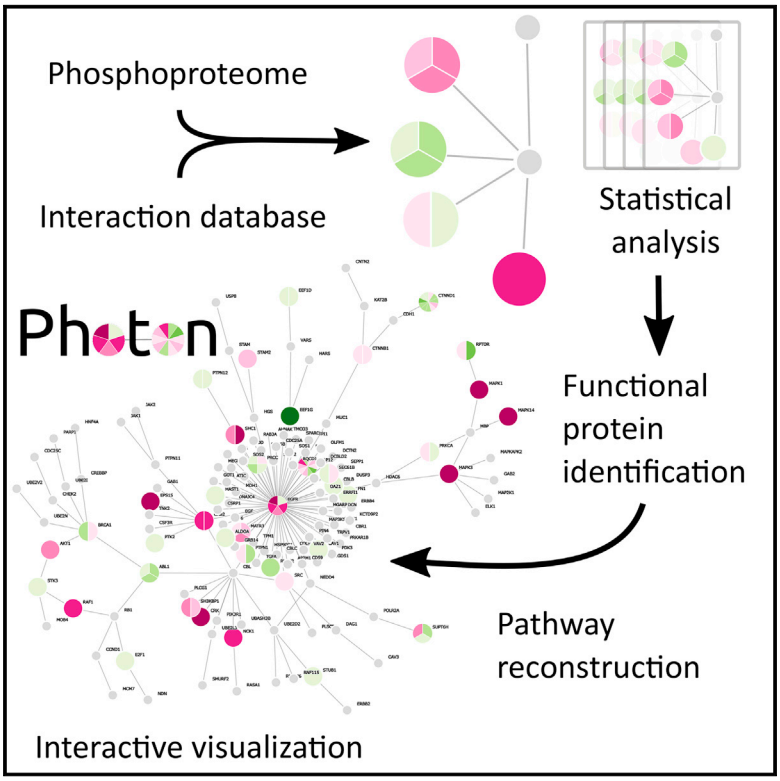
Cell Systems

Report

Elucidation of Signaling Pathways from Large-Scale Phosphoproteomic Data Using Protein Interaction Networks

Graphical Abstract



Authors

Jan Daniel Rudolph, Marjo de Graauw, Bob van de Water, Tamar Geiger, Roded Sharan

Correspondence

geiger@tauex.tau.ac.il (T.G.), roded@post.tau.ac.il (R.S.)

In Brief

Rudolph et al. present PHOTON, an interactive software for the identification of functional proteins and reconstruction of signaling pathways through integrated analysis of phosphoproteomic data and protein-protein interaction networks.

Highlights

- Analysis of phosphoproteomic data in the context of a protein interaction network
- Identification of functional proteins and derivation of signaling pathways
- Functional phosphosite prediction from site-specific and protein-based features
- The PHOTON interactive software for network-based analysis of phosphoproteomic data

Data Resources

PXD005032

Rudolph et al., 2016, Cell Systems 3, 585–593
December 21, 2016
http://dx.doi.org/10.1016/j.cels.2016.11.005

CellPress

Cell Systems
Report

Elucidation of Signaling Pathways from Large-Scale Phosphoproteomic Data Using Protein Interaction Networks

Jan Daniel Rudolph,[1,2] Marjo de Graauw,[3] Bob van de Water,[3] Tamar Geiger,[1,4,*] and Roded Sharan[2,*]
[1]Department of Human Molecular Genetics and Biochemistry, Sackler Faculty of Medicine, Tel Aviv University, 69978 Tel Aviv, Israel
[2]Blavatnik School of Computer Sciences, Tel Aviv University, 69978 Tel Aviv, Israel
[3]Division of Toxicology, Leiden Amsterdam Center for Drug Research, Leiden University, 2311 Leiden, the Netherlands
[4]Lead Contact
*Correspondence: geiger@tauex.tau.ac.il (T.G.), roded@post.tau.ac.il (R.S.)
http://dx.doi.org/10.1016/j.cels.2016.11.005

SUMMARY

Phosphoproteomic experiments typically identify sites within a protein that are differentially phosphorylated between two or more cell states. However, the interpretation of these data is hampered by the lack of methods that can translate site-specific information into global maps of active proteins and signaling networks, especially as the phosphoproteome is often undersampled. Here, we describe PHOTON, a method for interpreting phosphorylation data within their signaling context, as captured by protein-protein interaction networks, to identify active proteins and pathways and pinpoint functional phosphosites. We apply PHOTON to interpret existing and novel phosphoproteomic datasets related to epidermal growth factor and insulin responses. PHOTON substantially outperforms the widely used cutoff approach, providing highly reproducible predictions that are more in line with current biological knowledge. Altogether, PHOTON overcomes the fundamental challenge of delineating signaling pathways from large-scale phosphoproteomic data, thereby enabling translation of environmental cues to downstream cellular responses.

INTRODUCTION

Mass-spectrometry-based phosphoproteomic analysis is a powerful approach for the elucidation of signaling changes in cells upon perturbations. In this approach, phosphopeptides are typically enriched by immobilized metal ion affinity chromatography (IMAC), titanium dioxide or phosphotyrosine antibodies followed by mass spectrometric (MS) analysis. Downstream computational analysis identifies the phosphorylated peptides and localizes the modification site. However, despite the great depth of tens of thousands of phosphorylation sites in single experiments (Humphrey et al., 2013, 2015; Sharma et al., 2014; Yi et al., 2014), this technique still suffers from major limitations: (1) the functionality of the great majority of phosphorylation sites is unknown, and thus the phosphorylation change itself does not necessarily reflect a change in the signaling functionality of the protein; (2) the identification of the phosphosites by themselves does not lead to an integrated view of a larger signaling network; and (3) phosphoproteomic studies suffer under-sampling of the phosphoproteome, which is estimated to include hundreds of thousands of sites. As a result, the overlap of identified phosphosites between separate MS runs is always limited, leading to reduced analytical reproducibility.

The end result of most phosphoproteomic studies is a list of regulated phosphosites that differ between two or more cell states. Only some studies follow up with a functional examination of only a few sites. While some studies rely on statistical testing based on replicate analyses to extract regulated phosphosites from a large-scale dataset (Alcolea et al., 2012; Humphrey et al., 2015; Wilkes et al., 2015), many phosphoproteomic studies simply apply a fold-change cutoff, without proper statistical consideration (Cantin et al., 2006; Dephoure et al., 2008; Gruhler et al., 2005; Olsen et al., 2006, 2010; Pan et al., 2009; Yu et al., 2011), or determine a specific deviation from the mean (Humphrey et al., 2013; Rinschen et al., 2010; Wu et al., 2011) in a single phosphorylation measurement. Due to the under-sampling of the phosphoproteome, these site-centric approaches yield noisy and non-reproducible results. To alleviate these issues, an alternative approach was proposed by Terfve et al. (Terfve et al., 2015), where a Gaussian mixture model is fitted to the distribution of each phospho-peptide across a panel of kinase inhibitors to identify perturbed phosphorylation sites. However, this approach requires a large number of experimental conditions, making it inapplicable to standard proteomic studies.

Recently, various tools were developed to derive kinase activities from phosphoproteomic data and to put the phosphorylation data in a network context. Tools such as Inference of Kinase Activities from Phosphoproteomics (IKAP) (Mischnik et al., 2016) and Kinase Set Enrichment Analysis (KSEA) (Casado et al., 2013) exploit the kinase-substrate relations to derive the activity of a kinase from the phosphorylation state of its substrates. Aiming to put the regulated sites and proteins in a signaling context, networks in general and specifically kinase-substrate interactions can be explored using several computational tools such as KinomeXplorer and PHONEMeS (AlQuraishi et al., 2014; Casado et al., 2013; Horn et al., 2014; Mischnik et al., 2016; Schwartz

and Gygi, 2005; Terfve et al., 2015). All of these methods rely on a limited number of known site-specific kinase-substrate interactions. While a number of interactions is available in databases, the majority of the site-specific interactions are unknown. Mapping the phosphorylation data to these partial interactions greatly reduces the amount of data utilized in the analysis. For example, in the IKAP method, less than 5% (1,069 / 24,714) of the measured phosphorylation sites were used to predict the activity of 118 kinases during mitosis (Mischnik et al., 2016). In addition, such kinase-focused approaches ignore non-kinase proteins, which have an important part in mediating signal transduction. Similarly, placing phosphorylation events in a known signaling context, such as the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa et al., 2016) pathway under study, enables the interpretation and visualization of the data but excludes non-canonical interactions and cannot be applied to novel signaling pathways. A complete but cluttered picture of the signaling context can be obtained by extracting the potential interactions of regulated proteins from interaction databases.

Beyond the elucidation of signaling networks, MS-based phosphoproteomics has the potential to unravel novel phosphorylation sites and reveal new modes of protein regulation. To date, the site information was implemented into several databases, such as PhosphoSitePlus (Hornbeck et al., 2015), PHOSIDA (Gnad et al., 2007), and Phospho.ELM (Dinkel et al., 2011), but functional information is available only for a small number of well-studied sites. Even on a small scale, the experimental validation of a site's function is difficult (Lienhard, 2008), making computational approaches for function prediction much-needed. Toward this goal the PHOSIDA database (Gnad et al., 2007) provides functional annotations based on sequence, from which evolutionary (Tan et al., 2009) and structural evidence is derived.

In this work, we developed a phosphoproteomic analysis pipeline, which we termed PHOTON (phosphoproteomics dissection using networks), to tackle the above mentioned shortcomings of current approaches. First, we obtained robust estimates of protein signaling functionality from the raw phosphorylation data. After extracting a set of significantly functional proteins, we integrated those into a high-confidence protein-protein interaction (PPI) network and extracted the subnetwork linking the signaling source and the downstream effects. We benchmarked our approach using two published datasets related to epidermal growth factor (EGF) and insulin signaling. We further validated our approach against a novel dataset of EGF signaling, showing that it leads to results that better agree with current biological knowledge and exhibit higher reproducibility compared to the cutoff approach.

RESULTS

Signaling Functionality-Based Reconstruction of Phosphoproteomic Data

MS-based phosphoproteomic analysis has become the main experimental approach for the discovery of novel phosphorylation sites and elucidation of signaling pathways. The output of such experiments includes ratio values (between distinct samples) of thousands of phosphorylation sites; however, the functional significance of these sites and their integration into signaling networks is still a major obstacle in subsequent analyses. To transform the raw ratio data into functionality scores of proteins in the context of a signaling network, we first mapped the phosphorylation changes onto a weighted protein-protein interaction (PPI) network. The weights of the edges between the proteins provide a confidence measure for each protein-protein interaction, which is based on the experimental techniques by which each interaction was measured and the number of times it was observed (see STAR Methods). We postulate that the signaling functionality of a protein can be determined based on the phosphorylation changes of its interactors; for example, when the phosphorylation levels of multiple binding partners of a protein increase, we predict that it is functional. Since the PPI network includes multiple proteins with no catalytic activity (e.g., adaptor proteins), functionality does not necessarily refer to enzymatic activity but more broadly denotes the active mediation of signal transduction. To calculate the signaling functionality of each protein, we averaged the phosphorylation changes measured on the neighbors of each protein, weighted by the confidence of the corresponding interactions. By comparing the resulting functionality across proteins, we could estimate the p value for elevated or reduced functionality and infer a functionality score *a* of the protein (Figure 1A; and STAR Methods). The sign of the score corresponds to the direction in which the functionality changes (increases or decreases).

To examine the utility of the computed signaling functionality scores, we evaluated our approach on two publicly available studies (Humphrey et al., 2013; Olsen et al., 2006). The first study by Olsen et al. (Olsen et al., 2006) recorded phosphorylation changes after EGF treatment in the human HeLa cell line. The second study by Humphrey et al. (Humphrey et al., 2013) measured the insulin response of mouse adipocytes. After determining the functional proteins in these datasets, we examined the power of the signaling functionality scores by computing their enrichments with relevant Gene Ontology (GO) terms and KEGG pathways. In the Olsen dataset we identified 118 proteins that were significantly functional (upon 10 min EGF stimulation). We termed those the "functional group." For comparison, we used two approaches to define the significant changes based on the raw phosphorylation change, without signaling functionality calculation. The first approach was based on the fold change between the stimulated sample and the control. In order to obtain a group of proteins of the same size, we applied an absolute fold-change threshold of 2.31 and termed these sites the "phosphorylated group." The second approach based on significant outlier detection (STAR Methods) resulted in 139 changing proteins and was termed the 'outlier group'. Enrichment analysis to identify the relevance of these results showed much higher enrichment of EGF receptor (EGFR)-related pathways in the functional group. For example, the "regulation of epidermal growth factor receptor signaling pathway" category was found to be most highly enriched in the functional group ($q = 3.20 * 10^{-8}$), while the same category was ranked 58th in the phosphorylated group ($q = 3.74 * 10^{-4}$). In the outlier group, it ranked 70th ($q = 3.26 * 10^{-2}$) (Figure 1B; Table S1). Additionally, the functional group contained more than three times the number of proteins from this category compared to the other groups (15 vs. 4 and 4, respectively). Similarly, the category "ERBB signaling pathway," which underlies the data, was found

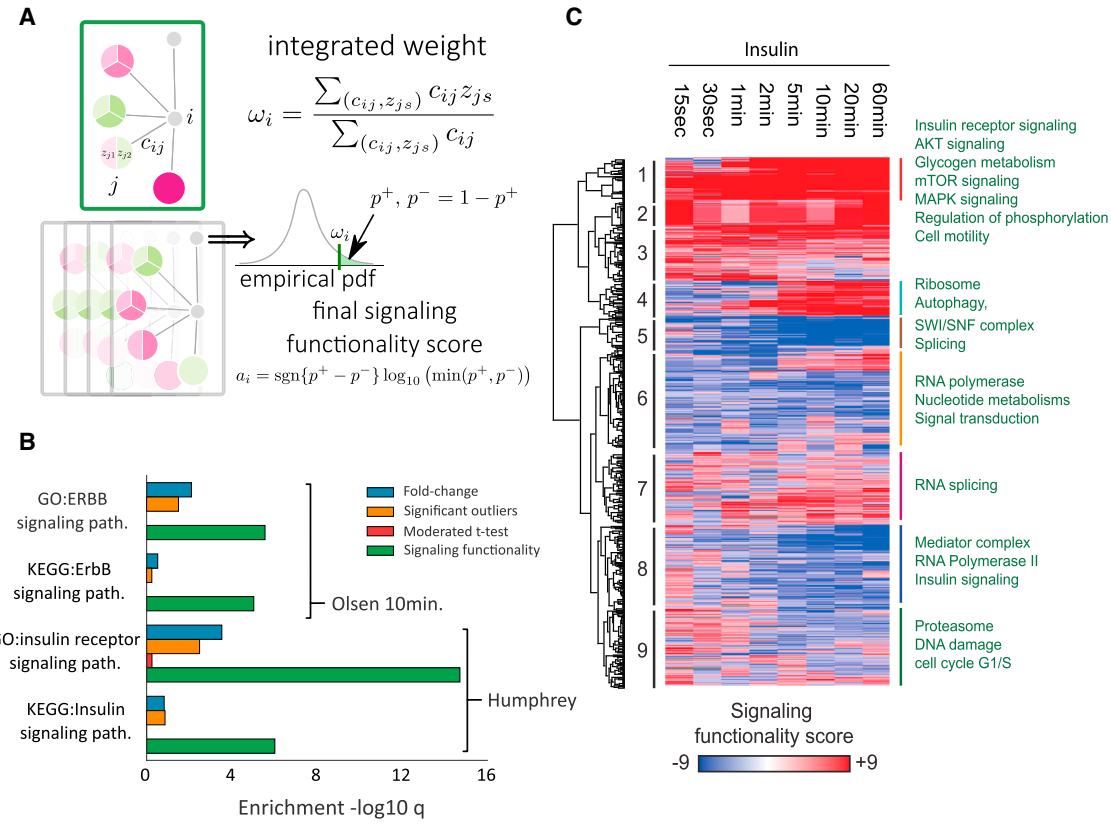


Figure 1. Signaling Functionality-Based Data Reconstruction

(A) For each protein *i*, phosphorylation measurements were aggregated across its interactors. The empirical distribution of the aggregated measurements, ω_i, was estimated using random permutations of the input data. The resulting p value was log-transformed to yield the final signaling functionality score.

(B) Enrichments of the target categories (ERBB signaling pathway and insulin receptor pathway) were higher for functional over phosphorylated proteins in GO and KEGG.

(C) Hierarchical clustering of signaling functionality scores at different time points of insulin stimulation. Representative enriched pathways in selected clusters are indicated.

significantly enriched for the functional $(q = 2.75 * 10^{-6})$, phosphorylated $(q = 7.86 * 10^{-3})$, and outlier groups $(q = 2.5 * 10^{-2})$. Further comparing the two significant enrichments, the expected category was not only more highly ranked in the functional group compared to the phosphorylated or outlier group (13 vs. 26 and 43) but also contained more proteins (24 vs. 9 and 9; Table S1). For the KEGG database, we found an even larger difference between the methods; in the functional group, the "ErbB signaling pathway" ranked first $(q = 9.30 * 10^{-6})$, while being non-significant $(q > 0.1)$ in the phosphorylated and outlier groups.

The Humphrey et al. (Humphrey et al., 2013) dataset yielded 275 significantly functional signaling proteins. A set of proteins of the same size was extracted by applying a fold-change cutoff of 2.76, yielding the phosphorylated group. Significant outlier detection yielded 183 proteins. Additionally, since multiple replicates were available (as opposed to the Olsen dataset), we also applied a moderated Student's *t* test to the data, which yielded 365 significant proteins, termed "*t*-test group." The most enriched GO category for the functional group was the "insulin receptor signaling pathway" $(q = 1.94 * 10^{-15})$ with 37 protein hits. While the same category was also found significantly enriched in the phosphorylated $(q = 2.95 * 10^{-4})$ and outlier $(q = 2.15 * 10^{-2})$ groups, it was lower ranked at positions 11 and 7 and contained fewer protein hits (17 and 13). The insulin receptor signaling pathway group was not found to be significantly enriched in the *t*-test group, which in general produced only few significant enrichment hits, with the top-ranked one being "eukaryotic translation initiation factor 4F complex" $(q = 3.38 * 10^{-2})$. Similar to the Olsen dataset, the KEGG enrichment analysis showed a greater difference between the methods. For the functional group, the "PI3K-Akt signaling pathway" $(q = 5.60 * 10^{-10})$ ranked first, with "insulin signaling pathway" ranking fifth $(q = 9.52 * 10^{-7})$. In the phosphorylated, outlier, and *t*-test groups, none of these pathways were significantly enriched. Neither the significant outlier detection nor the moderated *t* test, proved advantageous over the fold-change cutoff. This was also true for a combined approach using both the phosphorylated proteins and significantly functional proteins (Figure S1A). Therefore, in the following analyses, we focused exclusively on comparing between the signaling functionality-based approach and the fold-change cutoff.

Notably, when we tried to restrict our analysis to the kinase-substrate interactions from PhosphoSitePlus (Hornbeck et al.,

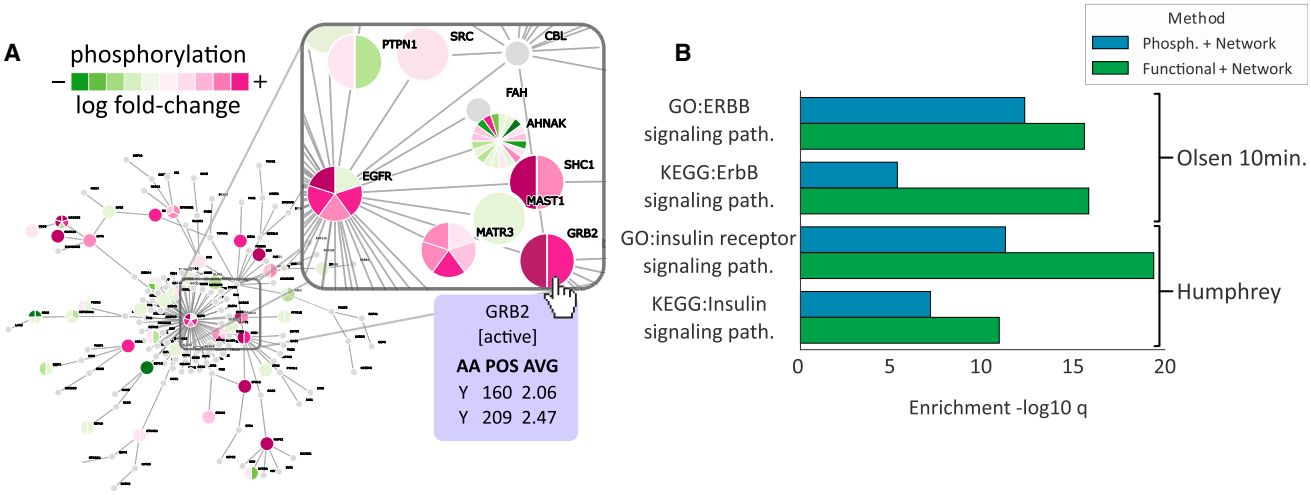


Figure 2. Signaling Network Reconstruction

(A) Reconstructed signaling pathway for the Olsen dataset. The measured phosphorylation state of each protein is displayed as a pie chart. The interactive visualization enables the exploration of the data, with additional information being displayed for selected proteins.

(B) Using the sets of functional and phosphorylated proteins as sources, we inferred signaling networks. The networks derived from functional signaling proteins showed higher enrichment than their phosphorylated counterparts.

2015), most of the data could not be mapped. While 40% of the 3652 phosphorylation sites in the Olsen dataset could be mapped to the PPI network, only 167 sites (4.5%) had a known kinase in the PhosphoSitePlus kinase-substrate network, and not a single protein was found to be significantly functional (Figure S2). Therefore, the broad PPI view is capable of capturing considerably larger areas of the network. Altogether, these results highlight the power of the signaling functionality determination in increasing the coverage and significance of the known, biologically relevant signaling pathways.

To examine the quantitative capabilities of the signaling functionality scores, we took advantage of the entire time course in the published Humphrey dataset, which included eight time points and unstimulated control cells. After calculation of the signaling functionalities of the proteins across the entire dataset, we extracted 2,340 significantly changing proteins (in any of the eight time points; $p < 0.05$). Hierarchical clustering of the significant proteins showed nine protein clusters representing the patterns of signaling functionality changes during 60 min of insulin stimulation. Enrichment analysis showed the pathways and processes that are regulated in each pattern. For example, the cluster of proteins that are stimulated within one minute of insulin treatment were enriched for "insulin receptor signaling," "AKT signaling," "glycogen metabolism," etc. (cluster 1) (Figure 1C; Table S2). A slower stimulation pattern, which initiates only after 5 min, showed enrichment for "ribosome" and "autophagy" (cluster 4). Four of the clusters showed patterns involving reduced functionality upon insulin stimulation. These were found to be enriched for epigenetic regulators and splicing factors (clusters 5), "mediator complex," and "RNA polymerase II" (cluster 8) and cell cycle and DNA damage regulation (cluster 9). Thus, the functionality view provides a means of evaluating the dynamic changes in biological processes in the course of a measured response.

Reconstruction of Signaling Pathways Using ANAT

Next, we used the Advanced Network Analysis Tool (ANAT) algorithm (Yosef et al., 2011) to reconstruct the signaling pathways underlying the measured responses. Starting from a global PPI network, the reconstruction aims to connect in an efficient manner a set of affected proteins, which we call "terminals," to the source of the response. In our case, we selected EGF (Olsen) or insulin (Humphrey) as the response source and examined the pathway reconstruction for two choices of terminals: (1) a set derived from the raw phosphorylation data and (2) a set based on significantly functional signaling proteins (as described above). These networks provide a powerful visualization of the interaction between the signaling proteins and their measured phosphorylation states (Figures 2A, S3, and S4). The source is located at the center of the network, and for each protein between the source and the terminals, the phosphorylation state of each site is indicated. Each node (protein) is represented as a pie chart, divided into slices according to the number of measured phosphorylation sites and colored according to the ratio of that specific site. In general, the insulin network is much larger than the EGF one, due to the higher coverage of the phosphoproteome in that study, and also shows a higher fraction of downregulated sites than in the EGF data, presumably due to different signaling dynamics of these systems. A closer investigation of these networks shows MAPK1 and AKT as the major nodes downstream of both receptors. Specifically, the insulin network also shows the nuclear factor κB (NF-κB) pathway (e.g., IKBKG and IRAK) to be affected by insulin, and BCL2 and other family members are shown to be activated downstream of the insulin receptor, as well as MDM2 and its interactors.

To evaluate the two networks, we calculated their enrichment with the two relevant GO processes and respective KEGG pathways: ERBB signaling pathway for the Olsen data, and insulin receptor signaling pathway for the Humphrey study. The

enrichments of the target categories were dramatically improved when network reconstruction was based on the functional group (Figure 2B). To ensure that the observed enrichments were not merely an artifact of the PPI network and the signaling pathway reconstruction, we computed their empirical p value as follows. We compared the enrichment scores from n = 1,000 random network reconstructions, which were generated by choosing a random set of potentially functional proteins as terminals, to the original reconstruction that used the significant proteins. Reassuringly, we found the p values for the target pathways to be significant ($p \leq 0.001$) for all datasets. Overall, these networks reinforce the advantage of functionality-based analysis and provide a global visualization of the pertaining signaling network.

Analysis of EGF Stimulation of MCF7 Cells

After establishing the utility of our method, we applied it to analyze a novel phosphoproteomic dataset, which measured the response of human MCF7 cells to EGF stimulation. Overall, we identified 4,269 phosphorylation sites on 1,786 proteins. We filtered the data to retain only sites with localization probability higher than 0.75 (class I) to focus the analysis on confident phosphosites only. Following the same workflow of signaling functionality determination, using this in-house EGF stimulation dataset, we reproduced protein-functionality predictions of a curated Boolean model of the EGFR signaling pathway (Samaga et al., 2009). This Boolean model was built using antibody-based analysis and heavily depended on prior knowledge and small-scale experiments. In contrast, our method, which is derived from large-scale phosphoproteomics, enables prediction of functional proteins without any prior knowledge and therefore can be much more broadly applied. Notably, splitting the proteins according to their functionality, as predicted by the Boolean model, resulted in a significant ($p = 0.0059$) increase in the calculated functionality score of the functional group (Figure 3A), providing further support to our predictions.

Based on the computed signaling functionality scores, we extracted a group of 128 significantly functional proteins, and inferred the underlying signaling network (Figure S5), as described above. We found the resulting network to be highly enriched in relevant, signaling-related categories, such as "epidermal growth factor signaling pathway" ($q = 2.92 * 10^{-17}$). The superiority of our method over the fold-change approach was confirmed by its higher enrichments of the target GO and KEGG categories (Figure 3B; Table S1). A comparison between the two EGF datasets in HeLa and in MCF7 cells showed multiple similarities, including the three central hubs of the EGF receptor, MAPK1, and AKT1, and in addition both showed the E3 ubiquitin ligase CBL as a central node, signaling to CRK, NCK, and PLC-gamma. Beyond these parallels, we also found signaling hubs with cell specificity, including RPS6KB1 and its effectors, which were specifically activated in MCF7 cells, or signaling to CDK2, which was more prominent in HeLa cells. These analyses show the ability to extract valuable signaling information and perform functional comparisons between systems.

Inference of Functional Activity of Phosphorylation Sites

In addition to reconstructing the signaling network for each dataset, we aimed to generalize this knowledge by linking a phosphorylation site to its function, providing basic biochemical information on protein regulation based on the specific phosphorylation site. To this end, we trained a logistic-regression model to predict functional phosphorylation sites. For each site, we assembled four features: (1) the observed phosphorylation fold change, (2) the derived signaling functionality score of the protein, (3) the evolutionary conservation of the phosphorylation site, and (4) the number of phosphorylation sites on the protein. For the in-house dataset, complete information was available for 984 sites on 353 proteins. As a training set, we obtained a list of known functional sites from PhosphoSitePlus (Hornbeck et al., 2015). We assigned a positive label to 106 sites that were listed in the reference, while the remaining 878 sites were labeled as negatives.

We measured the performance of our model in a 10-fold cross validation procedure using the area under the curve (AUC) of the receiver-operator characteristic (ROC). The functional phosphorylation site predictor achieved an average AUC of 0.71. In comparison to baseline methods, such as ranking the sites by the phosphorylation change or the raw protein signaling functionality, the logistic-regression classifier showed superior performance in functional site prediction (Figure 3C). Additionally, the coefficients of the logistic regression model can be used to understand the contribution of each of the features. We found that high protein signaling functionality, increased phosphorylation, and a small number of evolutionary conserved sites on the same protein characterize functional phosphorylation sites (Figure 3D). The 20 highest-scoring predictions include phosphorylations on 17 proteins, among them the EGF receptor itself, Shc1, Gab2, and Sos1, which are components of the signaling complex of the EGFR (see Table S3 for the entire list). While most of these sites are known, their activities are not annotated in PhosphositePlus and were therefore not included in the training set. In addition to those, we found several components of the S6 kinase signaling pathway, including RPS6KA1 and URI1. Our predictions included phosphorylation of RPS6KA1 on Ser389 (residue 380–381 in other isoforms), which was shown to be induced upon mitogen-activated protein kinase (MAPK) activation and autophosphorylation of the enzyme (Dalby et al., 1998). We also identified two core components of the MAPK family, namely MAPK14 (p38) and MAP2K2 (Mek2). These results demonstrate our ability to derive reliable functional information. Similar results were obtained for the Olsen dataset. Here, the classifier was evaluated on 63 positive and 813 negative samples. We again found the logistic regression model to outperform the baseline methods with a mean AUC of 0.70 (Figure S6). The same analysis could not be performed on the Humphrey dataset due to phosphorylation site mapping incompatibility of the mouse proteins.

Consistency and Reproducibility

A major challenge in phosphoproteomic analysis is the lack of reproducibility, which results from undersampling of the phosphoproteome and identification of distinct sites in independent experiments. To test whether PHOTON improves on standard phosphoproteomic analyses in terms of consistency and reproducibility, we compared its results obtained from the two EGF datasets to the standard (site and protein-centric phosphorylation). We considered the overall set of proteins and sites under study, and the set of proteins of interest, as chosen by the

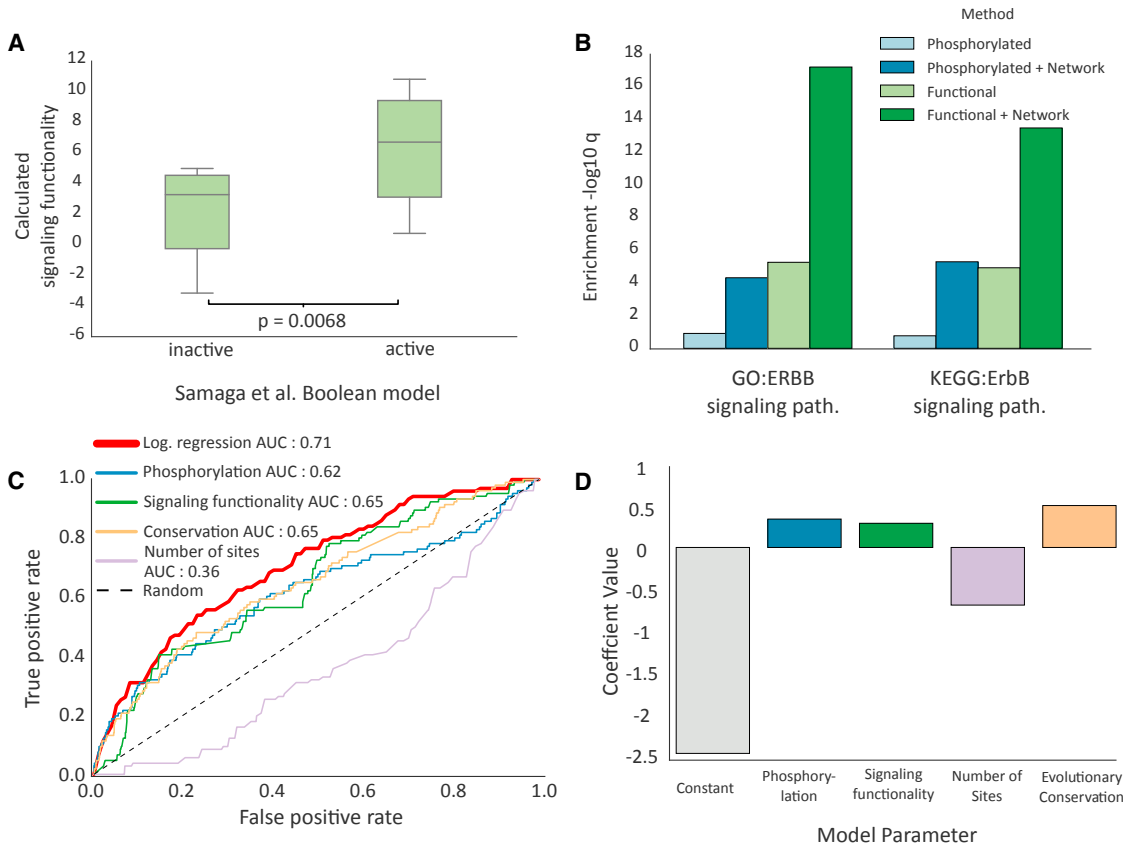


Figure 3. Analysis of the In-House EGF Dataset

(A) Calculated signaling functionality scores were compared to model predictions by the Boolean model from Samaga et al. (2009). Functional proteins showed a significantly increased score.

(B) Comparison of the enrichment in the target GO and KEGG categories for the different methods. "Phosphorylated" proteins were identified using the cutoff approach; "Functional" proteins were found to have significant signaling functionality scores. Additionally, both protein sets were extended using the network reconstruction approach, which is denoted by "+ Network."

(C) Receiver-operator characteristic curves for functional phosphorylation site prediction. Four baseline predictors were used: (1) ranking phosphorylation sites by their fold-change, (2) ranking by the signaling functionality score of the protein, (3) ranking by the evolutionary conservation of the site, and (4) ranking by the number of known sites on the protein. The logistic-regression model was found superior to any single approach.

(D) The logistic regression coefficients for each feature. The value of the coefficient is an estimate of the influence of each feature on the log-odds of the functionality of a phosphorylation site. The features were *z* -scored in order to make the coefficients comparable. The large negative constant coefficient implies that most phosphorylation sites are not functional. The odds of being functional are further decreased if the number of sites on the protein is large. The odds are improved, however, if the site is phosphorylated, resides on a functional signaling protein, or is evolutionary conserved.

respective method (Figure 4). Looking at the overlap between these protein sets, we found the signaling-functionality-based approach to be more consistent. 865 (59%) of the proteins were tested in both datasets, and within the group of proteins found to be significantly functional, 46 (23%) were found in both datasets. In contrast to the signaling-functionality-based approach, from a site-centric perspective, 627 of the 7,196 phosphorylation sites (9%) were experimentally observed in both datasets. The intersection of the 239 differentially phosphorylated sites in both datasets was found to contain only a single phosphorylation site. In conclusion, our approach leads to higher reproducibility across multiple samples.

PHOTON Software

The PHOTON pipeline was implemented as an easy-to-use graphical application running in the browser (Figure S7A). The user interface allows for the adjustment of parameters and the application of PHOTON analyses to novel phosphoproteomic datasets (Figure S7B). Additionally, PHOTON includes a command-line tool suitable for running batch jobs. PHOTON can be installed on all major operating systems using Docker. The installation instructions, alongside with the source code, are available at the PHOTON website (https://www.github.com/jdrudolph/photon/). The PHOTON source code is published under an open-source license and is accessible to the scientific community.

DISCUSSION

We presented an integrative approach for the analysis of phosphoproteomic data in the context of signaling pathways. Using this integrated approach, we were able to extract proteins of

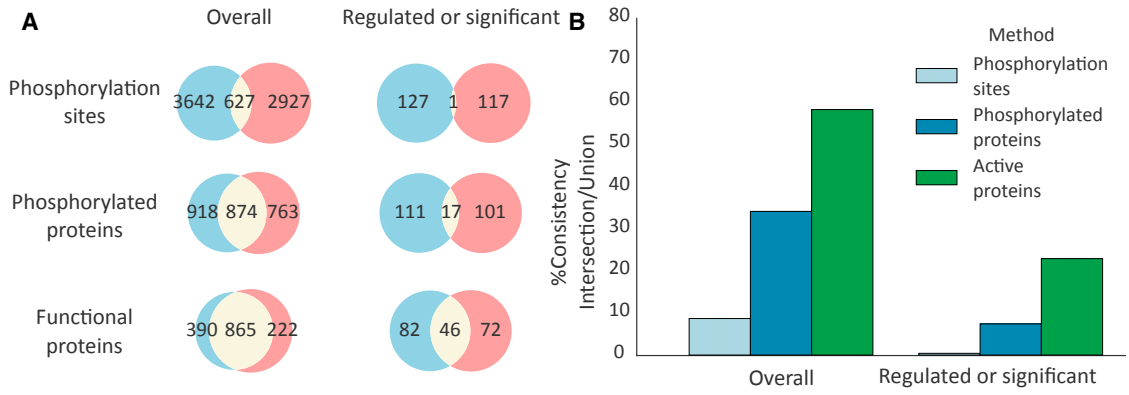


Figure 4. Reproducibility Analysis

(A) Overlap between the Olsen et al. and in-house EGF datasets for different analyses. Top row: experimentally observed phosphorylation sites. The second and third rows compare phosphorylated proteins with their functional counterparts.

(B) Consistency of the two datasets for the settings shown in (A). The overlap was quantified by the ratio of the sizes of the intersection and the union. The signaling-functionality-based approach was found to be most consistent. 59% of the proteins were tested in both datasets, and within the group of proteins found significantly functional, 24% were found in both datasets.

higher relevance, and with higher reproducibility, compared to other commonly used approaches. The derived protein signaling functionality scores allow for a more natural interpretation of protein action (promotion or repression of signal transduction) than changes in the phosphorylation of a specific site. While on a biochemical level only kinases phosphorylate specific residues on specific proteins, we did not limit our network only to kinases and substrates but rather considered all proteins of possible kinase-substrate complexes, including their adaptor proteins, as captured by the PPI network, to participate in the phosphorylation and signal transduction process. Regardless of the choice of network, the empirical scoring scheme used for signaling functionality determination is designed to correct for any biases introduced by the network. Due to the great complexity of the phosphoproteome, current high-throughput data are still considered sparse, therefore, gaps in the knowledge that one can derive from such data are expected. The network reconstruction with ANAT narrows those gaps and additionally provides a coherent interpretation of the data. The visualization of the resulting subnetwork allows for the exploration of the data and the generation of specific hypotheses for further testing.

Zooming in from the pathway view to a specific site, we proposed a data-driven approach for functional phosphorylation-site prediction based on a logistic-regression model. Previous computational approaches focused exclusively on protein structure or evolutionary conservation (Gnad et al., 2007). The integration of quantitative information outperformed predictions made on the basis of sequence-based information only. While our method is currently limited by training-data availability, it implements a promising data-driven approach and can be considered as a first step toward more involved classifiers. Additional data, such as kinase motifs, derived from the already used static sequence, protein structure data, or aggregates from multiple quantitative phosphoproteomics experiments, could be used to further improve future predictions.

PHOTON is applicable to the analysis of diverse datasets. If the process under study is not triggered by a single stimulus, the network reconstruction can be adjusted to account for multiple stimuli, or unstimulated scenarios, as described by Yosef et al. (Yosef et al., 2011). When changes in the protein level are expected for the process under study, site-occupancy data could be used as an input to PHOTON, though currently, their low coverage is limiting. Other post-translational modifications can be studied in the same manner as phosphorylation using PHOTON. Label-free approaches to peptide quantification could be accommodated for by adjusting the scoring scheme to model the difference between the case and control groups. In general, the empirical approach used to derive protein signaling functionality provides a flexible framework for any scoring function.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Experimental Data
- METHOD DETAILS
  - Protein-Protein Interaction Network
  - Signaling Functionality Scoring
  - Evolutionary Conservation of Phosphorylation Sites
  - Predicting Functional Phosphorylation Sites
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Differentially Phosphorylated Proteins
  - Enrichment Analysis
- DATA AND SOFTWARE AVAILABILITY
  - Software
  - Data Resources

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and four tables and can be found with this article online at http://dx.doi.org/10.1016/j.cels.2016.11.005.

AUTHOR CONTRIBUTIONS

J.D.R. performed the computational work under the supervision of R.S. and T.G. B.v.d.W. and M.D.G. performed the MCF7 experiment, and T.G. performed the phosphoproteomic analysis. J.D.R., R.S., and T.G. wrote the manuscript.

ACKNOWLEDGMENTS

We thank the members of the Geiger and Sharan groups for fruitful discussions. J.D.R.'s work was supported by the Edmond J. Safra Center for Bioinformatics at Tel Aviv University. T.G. was supported by I-CORE Centers of Excellence in Gene Regulation in Complex Human Disease (grant 41/11). R.S. was supported by research grants from the I-CORE Program (grant 757/12) and the Israel Science Foundation (grant 241/11).

Received: May 11, 2016
Revised: August 24, 2016
Accepted: November 9, 2016
Published: December 21, 2016

REFERENCES

Alcolea, M.P., Casado, P., Rodríguez-Prados, J.-C., Vanhaesebroeck, B., and Cutillas, P.R. (2012). Phosphoproteomic analysis of leukemia cells under basal and drug-treated conditions identifies markers of kinase pathway activation and mechanisms of resistance. Mol. Cell. Proteomics *11*, 453–466.

AlQuraishi, M., Koytiger, G., Jenney, A., MacBeath, G., and Sorger, P.K. (2014). A multiscale statistical mechanical framework integrates biophysical and genomic data to assemble cancer networks. Nat. Genet. *46*, 1363–1371.

Cantin, G.T., Venable, J.D., Cociorva, D., and Yates, J.R., 3rd (2006). Quantitative phosphoproteomic analysis of the tumor necrosis factor pathway. J. Proteome Res. *5*, 127–134.

Capra, J.A., and Singh, M. (2007). Predicting functionally important residues from sequence conservation. Bioinformatics *23*, 1875–1882.

Casado, P., Rodriguez-Prados, J.-C., Cosulich, S.C., Guichard, S., Vanhaesebroeck, B., Joel, S., and Cutillas, P.R. (2013). Kinase-substrate enrichment analysis provides insights into the heterogeneity of signaling pathway activation in leukemia cells. Sci. Signal. *6*, rs6–rs6.

Chatr-Aryamontri, A., Breitkreutz, B.-J., Oughtred, R., Boucher, L., Heinicke, S., Chen, D., Stark, C., Breitkreutz, A., Kolas, N., O'Donnell, L., et al. (2015). The BioGRID interaction database: 2015 update. Nucleic Acids Res. *43*, D470–D478.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. *26*, 1367–1372.

Dalby, K.N., Morrice, N., Caudwell, F.B., Avruch, J., and Cohen, P. (1998). Identification of regulatory phosphorylation sites in mitogen-activated protein kinase (MAPK)-activated protein kinase-1a/p90rsk that are inducible by MAPK. J. Biol. Chem. *273*, 1496–1505.

Dephoure, N., Zhou, C., Villén, J., Beausoleil, S.A., Bakalarski, C.E., Elledge, S.J., and Gygi, S.P. (2008). A quantitative atlas of mitotic phosphorylation. Proc. Natl. Acad. Sci. USA *105*, 10762–10767.

Dinkel, H., Chica, C., Via, A., Gould, C.M., Jensen, L.J., Gibson, T.J., and Diella, F. (2011). Phospho.ELM: a database of phosphorylation sites—update 2011. Nucleic Acids Res. *39*, D261–D267.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. *32*, 1792–1797.

Eppig, J.T., Blake, J.A., Bult, C.J., Kadin, J.A., and Richardson, J.E.; Mouse Genome Database Group (2015). The Mouse Genome Database (MGD): facilitating mouse as a model for human biology and disease. Nucleic Acids Res. *43*, D726–D736.

Gnad, F., Ren, S., Cox, J., Olsen, J.V., Macek, B., Oroshi, M., and Mann, M. (2007). PHOSIDA (phosphorylation site database): management, structural and evolutionary investigation, and prediction of phosphosites. Genome Biol. *8*, R250.

Gruhler, A., Olsen, J.V., Mohammed, S., Mortensen, P., Faergeman, N.J., Mann, M., and Jensen, O.N. (2005). Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway. Mol. Cell. Proteomics *4*, 310–327.

Horn, H., Schoof, E.M., Kim, J., Robin, X., Miller, M.L., Diella, F., Palma, A., Cesareni, G., Jensen, L.J., and Linding, R. (2014). KinomeXplorer: an integrated platform for kinome biology studies. Nat. Methods *11*, 603–604.

Hornbeck, P.V., Zhang, B., Murray, B., Kornhauser, J.M., Latham, V., and Skrzypek, E. (2015). PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. Nucleic Acids Res. *43*, D512–D520.

Humphrey, S.J., Yang, G., Yang, P., Fazakerley, D.J., Stöckli, J., Yang, J.Y., and James, D.E. (2013). Dynamic adipocyte phosphoproteome reveals that Akt directly regulates mTORC2. Cell Metab. *17*, 1009–1020.

Humphrey, S.J., Azimifar, S.B., and Mann, M. (2015). High-throughput phosphoproteomics reveals in vivo insulin signaling dynamics. Nat. Biotechnol. *33*, 990–995.

Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2016). KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res. *44*, D457–D462.

Lienhard, G.E. (2008). Non-functional phosphorylations? Trends Biochem. Sci. *33*, 351–352.

Mischnik, M., Sacco, F., Cox, J., Schneider, H.-C., Schäfer, M., Hendlich, M., Crowther, D., Mann, M., and Klabunde, T. (2016). IKAP: A heuristic framework for inference of kinase activities from Phosphoproteomics data. Bioinformatics *32*, 424–431.

Olsen, J.V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., and Mann, M. (2006). Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. Cell *127*, 635–648.

Olsen, J.V., Vermeulen, M., Santamaria, A., Kumar, C., Miller, M.L., Jensen, L.J., Gnad, F., Cox, J., Jensen, T.S., Nigg, E.A., et al. (2010). Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis. Sci. Signal. *3*, ra3.

Pan, C., Olsen, J.V., Daub, H., and Mann, M. (2009). Global effects of kinase inhibitors on signaling networks revealed by quantitative phosphoproteomics. Mol. Cell. Proteomics *8*, 2796–2808.

Rinschen, M.M., Yu, M.-J., Wang, G., Boja, E.S., Hoffert, J.D., Pisitkun, T., and Knepper, M.A. (2010). Quantitative phosphoproteomic analysis reveals vasopressin V2-receptor-dependent signaling pathways in renal collecting duct cells. Proc. Natl. Acad. Sci. USA *107*, 3882–3887.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. *43*, e47.

Samaga, R., Saez-Rodriguez, J., Alexopoulos, L.G., Sorger, P.K., and Klamt, S. (2009). The logic of EGFR/ErbB signaling: theoretical properties and analysis of high-throughput data. PLoS Comput. Biol. *5*, e1000438.

Schwartz, D., and Gygi, S.P. (2005). An iterative statistical approach to the identification of protein phosphorylation motifs from large-scale data sets. Nat. Biotechnol. *23*, 1391–1398.

Sharma, K., D'Souza, R.C.J., Tyanova, S., Schaab, C., Wiśniewski, J.R., Cox, J., and Mann, M. (2014). Ultradeep human phosphoproteome reveals a distinct regulatory nature of Tyr and Ser/Thr-based signaling. Cell Rep. *8*, 1583–1594.

Sonnhammer, E.L.L., and Östlund, G. (2015). InParanoid 8: orthology analysis between 273 proteomes, mostly eukaryotic. Nucleic Acids Res. *43*, D234–D239.

Tan, C.S.H., Bodenmiller, B., Pasculescu, A., Jovanovic, M., Hengartner, M.O., Jørgensen, C., Bader, G.D., Aebersold, R., Pawson, T., and Linding, R. (2009). Comparative analysis reveals conserved protein phosphorylation networks implicated in multiple diseases. Sci. Signal. *2*, ra39–ra39.

Terfve, C.D.A., Wilkes, E.H., Casado, P., Cutillas, P.R., and Saez-Rodriguez, J. (2015). Large-scale models of signal propagation in human cells derived from discovery phosphoproteomic data. Nat. Commun. *6*, 8033.

Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., and Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat. Methods *13*, 731–740.

Wilkes, E.H., Terfve, C., Gribben, J.G., Saez-Rodriguez, J., and Cutillas, P.R. (2015). Empirical inference of circuitry and plasticity in a kinase signaling network. Proc. Natl. Acad. Sci. USA *112*, 7719–7724.

Wu, C.-J., Chen, Y.-W., Tai, J.-H., and Chen, S.-H. (2011). Quantitative phosphoproteomics studies using stable isotope dimethyl labeling coupled with IMAC-HILIC-nanoLC-MS/MS for estrogen-induced transcriptional regulation. J. Proteome Res. *10*, 1088–1097.

Yi, T., Zhai, B., Yu, Y., Kiyotsugu, Y., Raschle, T., Etzkorn, M., Seo, H.-C., Nagiec, M., Luna, R.E., Reinherz, E.L., et al. (2014). Quantitative phosphoproteomic analysis reveals system-wide signaling pathways downstream of SDF-1/CXCR4 in breast cancer stem cells. Proc. Natl. Acad. Sci. USA *111*, E2182–E2190.

Yosef, N., Ungar, L., Zalckvar, E., Kimchi, A., Kupiec, M., Ruppin, E., and Sharan, R. (2009). Toward accurate reconstruction of functional protein networks. Mol. Syst. Biol. *5*, 248.

Yosef, N., Zalckvar, E., Rubinstein, A.D., Homilius, M., Atias, N., Vardi, L., Berman, I., Zur, H., Kimchi, A., Ruppin, E., and Sharan, R. (2011). ANAT: a tool for constructing and analyzing functional protein networks. Sci. Signal. *4*, pl1–pl1.

Yu, Y., Yoon, S.-O., Poulogiannis, G., Yang, Q., Ma, X.M., Villén, J., Kubica, N., Hoffman, G.R., Cantley, L.C., Gygi, S.P., and Blenis, J. (2011). Phosphoproteomic analysis identifies Grb10 as an mTORC1 substrate that negatively regulates insulin signaling. Science *332*, 1322–1326.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, Peptides, and Recombinant Proteins		
Titanium dioxide	GL Sciences	5020-75010
SCX column- Resource S	GE Healthcare	
Epidermal Growth Factor human	Sigma/Aldrich	E-9644
Deposited Data		
Raw MS files are deposited in PRIDE- ProteomExchange		PRIDE: PXD005032
Experimental Models: Cell Lines		
MCF7 cells	American Type Culture Collection (ATCC, Manassas, VA, USA)	ATCC HTB-22
Software and Algorithms		
ANAT	Yosef et al. (2011)	http://www.cs.tau.ac.il/~bnet/ANAT/
MUSCLE	Edgar (2004)	http://drive5.com/muscle/
PHOTON	This paper	https://www.github.com/jdrudolph/photon
Protein Residue Conservation Prediction	Capra and Singh (2007)	http://compbio.cs.princeton.edu/conservation/
Other		

CONTACT FOR REAGENT AND RESOURCE SHARING

As Lead Contact, Tamar Geiger is responsible for all reagent and resource requests. Please contact Tamar Geiger at geiger@tauex.tau.ac.il with requests and inquiries.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Experimental Data

The in-house EGF data was generated by stimulating MCF7 cells for 10 min with EGF. MCF7 cells were obtained from the American Type Culture Collection (ATCC, Manassas, VA, USA), where they were authenticated. Cell were routinely verified as having no mycoplasma contamination. For SILAC labeling cells were cultured in RPMI devoid of lysine and arginine and supplemented with light or heavy versions of these amino acids. After EGF stimulation cells were lysed with SDS-based buffer (4% SDS, 100 mM DTT in Tris-HCl pH 7.5). Proteins were digested using the FASP protocol followed by peptide separation into ten fractions on an SCX Resource S column. Phosphopeptide enrichment was performed with titanium dioxide. Phosphopeptides were analyzed on the LTQ-Orbitrap Velos mass spectrometer, with higher energy collisional dissociation fragmentation. Raw MS data were analyzed by MaxQuant (Cox and Mann, 2008), and included phospho(STY) sites as variable modifications. All protein identifiers were mapped to Entrez Gene IDs with duplications removed. Experiments were performed using two biological replicates. Replicate measurements were averaged and transformed into a z -scored $\log_2$ fold change. In the cases that the phosphorylation site was identified only in a single replicate the site was excluded from the analysis.

After pre-processing, which included the deletion of reverse hits, common contaminants, and sites with localization probability below 0.75, the in-house dataset contained 4269 sites on 1786 proteins measured in duplicates (Table S4). From the Olsen et al. (Olsen et al., 2006) time-series dataset we extracted the 10min / 0min ratio and obtained a single dataset of 3650 sites on 1637 proteins. Both datasets measured the response of human cell lines to stimulation with epidermal growth factor (EGF). A third dataset by Humphrey et al. (Humphrey et al., 2013) measured the response of mouse adipocytes to 20 min of insulin stimulation. In order to be able to compare the results directly to the human datasets, all mouse proteins were mapped to their human counterparts. To this end, a table of mouse ortholog genes was downloaded from MGI (Eppig et al., 2015) and used for the mapping, resulting in a dataset of 17177 sites on 3841 proteins. For the quantitative analysis, the entire time course was used (eight time points and control). For differential phosphorylation analysis the medians were used as provided by the authors.

METHOD DETAILS

Protein-Protein Interaction Network

A human PPI network was obtained from BIOGRID (Chatr-Aryamontri et al., 2015). Interactions between two proteins (i, j) were assigned a confidence score c_{ij} reflecting their reliabilities according to the logistic-regression framework described in Yosef et al. (Yosef et al., 2009) In brief, each interaction is scored according to the experimental techniques by which it was measured and the number of times it was observed. A high confidence subnetwork was obtained by applying a confidence cutoff of 0.5 and removing proteins with more than 150 interactions.

Signaling Functionality Scoring

We calculated a signaling functionality score a_i for each protein i in the network as follows: Phosphorylation measurements z_{js} were collected from all proteins $j \in N(i)$ and sites $s \in S_j$ in the neighborhood of protein i. Each measurement z_{js} was assigned a weight c_{ij} reflecting the confidence in the interaction between the proteins i and j. Aiming to derive elevated, or reduced, functionality from the measurements $z_i = \{(c_{ij}, z_{js}), j \in N(i), s \in S_j\}$, we calculated the weighted mean of the phosphorylation measurements:

$$\omega_i = \frac{\sum_{(c_{ij}, z_{js})} c_{ij} z_{js}}{\sum_{(c_{ij}, z_{js})} c_{ij}}.$$

In order to ensure robust estimation, proteins, which had fewer than $|z_i| < n_o = 4$ observations on their interaction partners, were not considered. By exploring the empirical distribution of ω for all proteins, we estimated two p values for elevated (p^+), and reduced (p^-) functionality. The p value estimates for a score of ω were then given by $p^+ = \sum_r I\{\omega < \omega_r\}/r$ and $p^- = 1 - p^+$, where ω_r was calculated for $r = 10,000$ random permutations of the phosphorylation data (Figure 1A).

The final signaling functionality score was the log transform of the lower of both p values. The sign of the score was chosen to correspond to increased or decreased functionality. In order to ensure that the functionality score is always finite, a continuity correction $\epsilon = 1/r$ was added.

$$a_i = -1^{I\{p^+ < p^-\}} \log_{10}\left[\min\left(p^+, p^-\right) + \epsilon\right].$$

After correcting for multiple testing using the Benjamini-Hochberg method at significance level $\alpha = 0.05$, we obtained a set of significantly functional signaling proteins.

Evolutionary Conservation of Phosphorylation Sites

To calculate the evolutionary conservation of the phosphorylation sites, we used the InParanoid (Sonnhammer and Östlund, 2015) database, which provides ortholog protein information. From the evolutionary tree provided by the database we extracted a subtree (of depth 6) including humans and 18 other closely related species (*M.putorius, F.catus, M.musculus, R.norvegicus, C.porcellus, O.garnettii, S.tridecemlineatus, N.leucogenys, C.jacchus, G.gorilla, P.troglodytes, P.abelii, C.familiaris, A.melanoleuca, L.africana, E.caballus, B.taurus, O.cuniculus*). We obtained orthologs for the proteins observed in the phosphoproteomic datasets across all chosen species. MUSCLE (Edgar, 2004) was used with default settings to obtain multiple sequence alignments for all proteins. Evolutionary conservation scores were then derived from the multiple sequence alignment using the method of Capra and Singh (Capra and Singh, 2007). Phosphorylation sites which could not be assigned a conservation score due to e.g., missing orthologs were assigned a median conservation score.

Predicting Functional Phosphorylation Sites

We constructed a logistic-regression model to classify sites as functional or non-functional. As a positive set, a list of regulatory sites (5202), and another list of disease-associated sites (443), were retrieved from the PhosphoSitePlus database (Hornbeck et al., 2014). After removing overlaps, 5308 unique sites on 1815 proteins with known functionality were collected. The negative set was chosen to contain all phosphorylation sites not listed in the positive set. Each experimentally observed site (j, s) on protein j was labeled as a positive sample if it had known function and as a negative sample otherwise. Assuming that the signaling functionality of each protein is modulated by the differential regulation of few of its evolutionary conserved phosphorylation sites the following four features were selected: (i) the log-fold change in phosphorylation of the site z_{js}, (ii) the signaling functionality score of the protein a_j, (iii) the evolutionary conservation of the site e_{js}, and (iv) the number of sites on the protein n_j. The logistic-regression model assumes independence between the features and no additional interaction terms were added.

$$logit\left(p_{js}\right) = \beta_0 + \beta_1 z_{js} + \beta_2 a_j + \beta_3 e_{js} + \beta_4 n_j.$$

The area under the receiver-operator curve was used to measure the predictive performance of the classifier across a 10-fold cross-validation. The β coefficients were extracted after fitting the model on all the available data.

We observed that proteins with a great number of detected phosphorylation sites, such as SRRM2 artificially bloated the ROC AUC of the logistic-regression model. These proteins contributed a large number of uncharacterized sites to the analysis, therefore giving the trivial prediction 'non-functional' for all sites a high performance score. We therefore decided to remove proteins with a large number of sites (> 100) from the analysis.

QUANTIFICATION AND STATISTICAL ANALYSIS

Differentially Phosphorylated Proteins

Three approaches were tested for extracting proteins of interest directly from their phosphorylation changes, using either (i) a fold-change cutoff, (ii) significant outlier detection or (iii) a moderated *t* test.

After ranking all experimentally observed proteins according to their phosphorylation change, a cutoff was used to extract a set of proteins with differentially regulated phosphorylation sites. We chose the $\log_2$ fold-change cutoff such that the resulting regulated group of proteins is of the same size as the functional group. The resulting cutoff on the mean absolute z-scored $\log_2$ fold change was therefore different for each dataset (2.38 for in-house, 2.31 for Olsen, and 2.76 for Humphrey).

Using a statistically more informed approach, we applied two-sided significant outlier detection SignificanceA (Cox and Mann, 2008) as implemented in the Perseus (Tyanova et al., 2016) (version 1.5.4.0) software. Significance A identifies significant outliers from a potentially asymmetrical distribution with normal tails. Prior to analysis, replicates were log-transformed, averaged and z-scored. The resulting p values were adjusted for multiple testing using Benjamini-Hochberg FDR correction at $\alpha = 0.01$

For the two-sided moderated *t* test, we used the limma (Ritchie et al., 2015) R package to analyze the in-house (2 replicates) and Humphrey (3 replicates) datasets. The Olsen dataset could not be analyzed using this approach, since only a single replicate was available. All analyses were done on z-scored log fold-changes values and p values were adjusted for multiple testing using Benjamini-Hochberg FDR correction at $\alpha = 0.01$.

Applying cutoff of $\alpha = 0.01$, as compared to $\alpha = 0.05$, for outlier detection and moderated *t* test, yielded significant groups more similar in size to the regulated sets of proteins. Using a less stringent cutoff of $\alpha = 0.05$ yielded qualitatively similar results (Figure S1).

Enrichment Analysis

GO term and KEGG pathway enrichments were computed via a hypergeometric test. The background set was chosen according to four different settings. For (i) differentially phosphorylated proteins, as determined by fold-change cutoff, outlier detection, or moderated *t* test, it was chosen to contain all experimentally observed proteins. Since the signaling functionality scores are derived from the network, the background for (ii) the analysis of the functional proteins was chosen to contain all network proteins, for which signaling functionality scores could be computed. For (iii) the proteins within reconstructed networks, the background contained all network proteins. For (iv) the combined approach, the backgrounds in (i) and (ii) were combined. An additional analysis was performed to probe the effect of the different background sets on the results. The combined background, containing experimental, and network proteins, was chosen as a common background for all methods. In this setting, too, PHOTON outperforms all other approaches (Figure S1B). All enrichment analyses were performed using the 'goenrich' (https://www.github.com/jdrudolph/goenrich) python package.

DATA AND SOFTWARE AVAILABILITY

Software

The PHOTON software, including installation instructions, and source code, are available at the PHOTON website (https://www.github.com/jdrudolph/photon/). The PHOTON source code is published under an open-source license, and accessible to the scientific community.

Data Resources

The in-house data is available in Table S4. The accession number for the mass spectrometry proteomics data reported in this paper is PRIDE: PXD005032.