PTMNavigator: Interactive Visualization of Differentially Regulated Post-Translational Modifications in Cellular Signaling Pathways

Julian Müller[1], Florian P. Bayer[1], Mathias Wilhelm[2], Maximilian G. Schuh[1†], Bernhard Kuster[1,3], and Matthew The[1]*

[1]Proteomics and Bioanalytics, School of Life Sciences, Technical University of Munich; 85354 Freising, Germany.

[2]Computational Mass Spectrometry, School of Life Sciences, Technical University of Munich; 85354 Freising, Germany.

[3]German Cancer Consortium (DKTK), Partner Site Munich, 80336 Munich, Germany.

[†]Present Address: Organic Chemistry II, School of Natural Sciences, Technical University of Munich; 85748 Garching, Germany.

*Corresponding author. Email: matthew.the@tum.de

Abstract

Post-translational modifications (PTMs) play pivotal roles in regulating cellular signaling, fine-tuning protein function, and orchestrating complex biological processes. Despite their importance, the lack of comprehensive tools for studying PTMs from a pathway-centric perspective has limited our ability to understand how PTMs modulate cellular pathways on a molecular level. Here, we present PTMNavigator, a tool integrated into the ProteomicsDB platform, which offers an interactive interface for researchers to overlay experimental PTM data with pathway diagrams. PTMNavigator provides ~3000 canonical pathways from manually curated databases and further enables users to modify and create custom diagrams, tailored to their data. Additionally, PTMNavigator automatically runs multiple kinase and pathway enrichment algorithms whose results are directly integrated into the visualization. This offers a comprehensive view of the intricate relationship between PTMs and signaling pathways. To demonstrate the utility of PTMNavigator, we applied it to two phosphoproteomics perturbation datasets. First, PTMNavigator enhanced pathway enrichment analysis by showing how the regulated peptides and proteins are distributed in the pathways with high enrichment scores. Second, it visualized how drug treatments result in a discernable flow of PTM-driven signaling within pathways. Third, PTMNavigator aided in proposing extensions to an existing pathway by suggesting putative new links between both PTMs and pathway components. By enhancing our understanding of cellular signaling dynamics and facilitating the discovery of novel PTM-pathway interactions, PTMNavigator advances our knowledge of PTM biology and its implications in health and disease.

Introduction

Protein post-translational modifications (PTMs) are among the key regulators of cellular pathways. The presence of a certain PTM on a specific residue can, for instance, determine whether or not the protein can interact with other proteins, activate or inactivate an enzyme, alter a proteins' location in the cell, or mark it for degradation. Understanding the specific role of various types of PTMs within intracellular signaling cascades is essential to gain insights into the molecular mechanisms that govern biological processes. Perturbation studies that examine the effects of altered PTM patterns on pathways have been particularly useful in advancing our knowledge[1-7]. While proteome-wide measurements of PTMs have become a routine procedure for proteomics laboratories, its correct interpretation in the context of cellular signaling pathways continues to pose a daily challenge.

Multiple databases are available to the scientific community that are dedicated to the study of pathways[8-11], including the Kyoto Encyclopedia of Genes and Genomes (KEGG) and WikiPathways. KEGG[9] contains 563 manually curated pathway maps with annotations for almost 9,000 different organisms. WikiPathways[10] is an open-source initiative to collect, maintain, and disseminate data on biological pathways. Its March 2023 release consists of 3,130 pathways across 33 species. Both of these databases are designed to study relationships between genes, gene products, and compounds (such as metabolites or drugs). They offer no (KEGG) or only limited (WikiPathways) direct support for research on the role of specific PTM-sites within pathways. For the study of PTM data, several resources exist, the most notable being PhosphoSitePlus[12], which includes data from a large number of research articles not only on phosphorylation, but also acetylation, ubiquitinylation, and other PTMs. Where evidence exists, the sites in PhosphoSitePlus are annotated with their functions and putative upstream or downstream interactors or modifiers.

While PhosphoSitePlus has been a tremendously useful resource, it remains tedious to explore multiple or all sites of a signaling cascade together in order to uncover potential relationships between them and perform higher-level analysis of the data.

To place all regulated peptides into the context of signaling networks, a commonly employed method is pathway enrichment analysis (PEA)[13-15], implemented in tools such as g:Profiler[16]. While very useful for the analysis of proteome expression changes, this approach has several shortcomings in the context of PTM analysis, as has been pointed out by *Krug et al.*[17].

The authors introduced two algorithms that address the problems of traditional PEA: One is gene-centric redundant (GCR) single-sample gene set enrichment analysis (ssGSEA), in which genes are potentially counted multiple times during the calculation of enrichment scores, depending on how many regulated PTMs they harbor. The second is PTM-signature enrichment analysis (PTM-SEA), which uses a specifically assembled database of associations between sites of PTMs and perturbations, pathways, or kinase activities (PTM signatures database; PTMsigDB) and thus can calculate enrichments on PTM level. This constitutes a valuable complementary approach to conventional PEA and both methods have shown superior performance compared to non-redundant gene-centric PEA. A number of software tools to visualize PTM datasets and pathways have been developed[18-23]. For instance, Phosphomatics[23] is a web service that helps researchers to examine kinase-substrate relationships in their phosphoproteomic datasets, whereas KEGGViewer[19] can integrate gene expression profiles and KEGG pathways. A functionality which these tools have been lacking so far is the visualization of regulated PTMs directly within biological networks. The Cytoscape[18] app PhosphoPath[21] has addressed this issue for time-course PTM data, but the app is no longer supported by Cytoscape. PHONEMeS[22] is a software that creates data-driven pathways from phosphoproteomics datasets. Starting from a prior knowledge

network of protein-protein and kinase-substrate interactions, it uses mathematical optimization to extract the subset of interactions that best explain the observations. Running the software and visualizing the results is, however, not trivial for non-expert users.

To overcome the aforementioned problems of pathway enrichment on the PTM level and to empower researchers to visually examine pathway engagement via PTM-regulation, we introduce PTMNavigator, a web application that projects PTM perturbation datasets onto pathway diagrams in the form of interactive graphs. To this end, we created internal representations in ProteomicsDB for thousands of pathways from public databases. This enables users to trace signaling cascades across a cell and identify pivotal PTMs within a pathway. PTMNavigator is available as part of the 'Analytics Toolbox' of ProteomicsDB[24, 25] and can visualize data hosted on ProteomicsDB as well as user-uploaded data. In addition, we implemented the functionality to design custom pathway diagrams within PTMNavigator, in order to create bespoke representations of any PTM dataset. The core visualization component of our software can be easily reused in web interfaces outside of ProteomicsDB. Moreover, we implemented a service that provides programmatic access to several PTM analysis algorithms, including GCR-ssGSEA, PTM-SEA, and PHONEMeS[17, 22, 26-31]. This service, which can also be used independently, is automatically invoked when a dataset is uploaded, and the results are integrated into PTMNavigator's user interface. We demonstrate the utility of our application using two phosphoproteomic drug profiling studies[2, 32]. First, we show how PTMNavigator improves the interpretation of PEA by integrating the results of multiple enrichment algorithms with visualizations of all potentially relevant pathway diagrams. Second, we visualize target and pathway engagement of kinase inhibitors at different steps in a phosphorylation-dependent signaling cascade. Last, we show how PTMNavigator leverages the

information contained in dose-dependent drug perturbation studies to aid the functionalization of PTM sites.

Results

A PTM-Centric Interface for Pathway-Level Data Analysis

PTMNavigator is a web-based graphical interface to analyze the results of PTM perturbation experiments in the context of pathways (Fig. 1A). It combines interactive visualization and various enrichment analysis algorithms (Fig. 1B) into a single application. The interface is embedded into the ecosystem of ProteomicsDB, a multi-omics resource for life science data and data analysis. The software combines two types of inputs. The first input, which needs to be provided by the user, is a list of regulated modified peptides, e.g. the result of a differential or dose-response analysis. Each peptide must be annotated with a regulation type (i.e. not, up, or down). The second type of input are pathway diagrams. These can be of two different types: First, users can display canonical pathways from KEGG and WikiPathways, for which we created internal representations in ProteomicsDB. For *Homo sapiens*, we imported 1,121 pathways (779 from WikiPathways and 342 from KEGG) from the latest releases at the time of retrieval (see Methods). In addition, we imported 1,838 pathways for 9 model organisms, including other mammals, bacteria, and plants (Suppl. Fig. S1). Second, users can also create own diagrams by either modifying the canonical pathways or by drawing custom diagrams from scratch, using the 'Editing Mode' of PTMNavigator. The custom diagrams can be stored in ProteomicsDB and shared with other researchers.

PTMNavigator combines the experimental data and the pathway information into a projection of modified peptides onto their corresponding genes within a pathway diagram. The main interface of PTMNavigator (Fig. 2A) is accessible at www.proteomicsdb.org/analytics/ptmNavigator. Users can visualize datasets already contained in ProteomicsDB or upload their own (templates for input formats are available as Supplementary Table 1). It is possible to access (previously) uploaded

datasets by a private, user-specific session ID and to visualize one or multiple datasets at a time. When the user selects a pathway diagram, it is rendered together with additional nodes for each PTM peptide next to its corresponding protein. These additional PTM nodes are colored according to their regulation (up, down, not). The initial coordinates of the nodes are the same as in the reference databases so that it resembles the original pathway diagram. However, users can rearrange the layout by dragging nodes around, highlighting interesting subnetworks, and filtering their data by experiments or regulation categories. Where available, additional information on PTM sites is displayed as part of a tooltip and kinase-substrate relationships can be visualized in the form of additional arrows in the diagram. To this end, we imported a list of functional annotations and upstream regulators from PhosphoSitePlus and a list of predicted phosphosite functional scores calculated by *Ochoa et al.*[33] into ProteomicsDB. It is also possible to visualize protein-level data in combination with PTM-level data (see Suppl. Fig. S5 for an example, where this has been exemplified using a dataset from *Lee et al*.[34]).

In addition to the pathway visualization component, we implemented a service that hosts a number of established enrichment algorithms (Fig. 1B). After a dataset is uploaded, this service is automatically invoked and two pathway enrichment analysis algorithms are run over the dataset ('Gene-Centric' and 'Gene-Centric-Redundant' ssGSEA). For phosphoproteomics datasets, additional algorithms are run, including PTM-SEA, KSEA, and PHONEMeS. The results of the enrichments are displayed in tabular form in PTMNavigator (Fig. 2B). In the pathway selection menu in PTMNavigator, the pathways are sorted by 'Gene-Centric-Redundant' ssGSEA scores to show to the users which pathways are likely most relevant for the dataset. Some of the information is also integrated into the pathway view, e.g., for phosphoproteomics datasets, kinase activities calculated with RoKAI and KSEA can be shown as colored borders around the nodes in the graph.

The final pathway diagrams as well as the results of the enrichment analysis can be downloaded as scalable vector graphics (SVG) or in tabular format, respectively. In the following, we exemplify the capabilities and benefits of PTMNavigator using two recent PTM perturbation datasets.

Enrichment Analysis and Pathway Visualization Synergize in PTMNavigator

To demonstrate how PTMNavigator aids in interpreting the results of a pathway enrichment analysis (PEA), we re-analyzed data from a recent phosphoproteomics study, in which 30 different kinase inhibitors targeting members of the EGF signaling pathway (Fig. 3A) were profiled for their effect on the phosphoproteome of a human retinal pigment epithelial cell line (RPE1)[32]. Re-analysis of all 30 datasets yielded 4,989 regulations (2,563 up and 2,426 down) on 1,757 peptides (out of 7,871 peptides measured in each sample) at 10% FDR. PEA for each compound with gene-centric-redundant ssGSEA resulted in enrichment scores (negative log-transformed p-values; ES) for each combination of drug and pathway (Fig. 3B). Some results confirmed prior knowledge on the compounds, e.g., the six drugs inhibiting either EGFR or MEK, which are both located upstream of MAPK, were strongly enriched for 'WP: MAPK Cascade'. A PTM-SEA analysis of the same dataset (Suppl. Fig. S6) showed the datasets of the mTOR inhibitors Everolimus and Temsirolimus to be enriched for both the signatures 'PERT-PSP_Rapamycin' and 'PERT-P100-DIA2_Sirolimus'. Rapamycin and Sirolimus are different names for the same compound that also inhibits the kinase mTOR. Yet, for many drug perturbations, the results were not as easy to interpret. E.g., the KEGG pathway 'Progesterone-Mediated Oocyte Maturation' received a significant score in 3 of the datasets, however, since the cells studied were retinal cells, it is not obvious where this significance arose from.

To gain more insight into how the regulated peptides were distributed within the enriched pathways, we used PTMNavigator to investigate the high-scoring pathways of the pan-PKC inhibitor Go-6983, one of the datasets which was enriched in 'Progesterone-Mediated Oocyte Maturation'. Other pathways that received high Enrichment Scores in the ssGSEA analysis of this dataset (Fig. 3C) include signaling networks related to prostate cancer and the nervous system. Manual inspection of the pathways showed that genes with strong fold changes are often shared. Most of the high-scoring pathways contain 4E-BP1, RPS6KA, MAPK1, MAP3K7, mTOR, and/or AKT, all of which have peptides with strong fold changes in the Go-6983 data (Suppl. Fig. S7). Therefore, these proteins likely were the drivers of the high enrichment scores. In the PTM-SEA analysis of the dataset (Fig. 3D), the highest positively and negatively enriched signatures, respectively, were the mTOR and RPS6KA1 PhosphoSitePlus signatures (Scores=5.78 and -5.76, FDR=0.022 for both). mTOR and RPS6KA1 were also reported to be increased (decreased) in activity by KSEA. We investigated these scores in more depth, using the PhosphoSitePlus kinase-substrate annotations integrated into PTMNavigator (Suppl. Fig. S8). Highlighting the substrates of mTOR and RPS6KA1 within the diagram 'mTOR Signaling Pathway - hsa04150' showed that most of the peptides phosphorylated by RPS6KA1 had strong negative fold changes (e.g. EIF4B_S422 and SOS1_S1134), while most mTOR substrates were increased in phosphorylation (e.g. the autophosphorylation sites T2474 and S2478). This demonstrates how PTMNavigator can help in interpreting the results of enrichment analyses.

Visualizing Target and Pathway Engagement in PTMNavigator

We next employed PTMNavigator on the same dataset to compare the kinase inhibitors Lapatinib, Wortmannin, and Temsirolimus, which target the same signaling cascade, but at different steps.

Projection of their perturbed PTM peptides on 'mTOR signaling pathway - hsa04150' (Fig. 4A), which includes all members of said cascade except for EGFR, confirmed that all three drugs perturbed the pathway. Lapatinib had the largest number of both up- and downregulated peptides and also the regulations most upstream in the pathway (Fig. 4A, upper panel). Among the sites exclusively regulated by Lapatinib were GSK3B_S9, TSC2_S939, and BRAF_S729, whose importance has previously been pointed out in the literature[35-37]. In contrast, the effects of Wortmannin in this pathway were limited to sites on PRAS40, LPIN, 4E-BP1 and RPS6, all of which are downstream of PI3K in the depicted network (Fig. 4A, middle panel). Likewise, Temsirolimus regulated even fewer sites and all were downstream of mTOR (Fig. 4A, lower panel). Most of the regulated peptides for Wortmannin and Temsirolimus were shared with Lapatinib (PRAS40_T246, 4EBP1_T77/S101, LPIN1_S252 for Wortmannin, RICTOR_Y1174&Y1177 for Temsirolimus), or one another (RPS6_S236&S240&S244). This was further substantiated by the results of kinase activity scoring on the three datasets: KSEA produced significant negative scores for AKT, RAF, RPS6KA, and RPS6KB in the Lapatinib and Wortmannin data. For Temsirolimus, only RPS6KA and RPS6KB were reported to have decreased activity. RPS6KB is the only of these kinases that appears downstream of mTOR, and the high score of RPS6KA probably stems from the large number of substrates that are shared with RPS6KB. Together, these observations confirm that all drugs perturb the same signaling network at different stages.

A more surprising example was presented when comparing the effects of Sorafenib and Cobimetinib. Both compounds target the RAF-MEK-ERK pathway, another phosphorylation axis downstream of EGFR (Fig. 3A). Since Sorafenib is perturbing the network at a more upstream position (RAF), one would expect its impact on the PTM level to be more pronounced. In reality,

Cobimetinib (which inhibits MEK) had a wider impact than Sorafenib, which was visible in PTMNavigator's visualization of 'MAPK Signaling Pathway - hsa04010' (Fig. 4B). As expected, Cobimetinib significantly inhibited the phosphorylation of the site MAPK1_Y187, which is essential to MAPK1 activity, with a Log2 Fold Change (logFC) of -4.39 between conditions (p-value=0.01). There also were effects upstream in the pathway, e.g. on SOS1_S1134. This is likely the result of an inhibited feedback loop, since this site is a substrate of RPS6KA1, which is itself regulated by ERK[38]. In contrast, Sorafenib only inhibited the phosphorylation of one site in this pathway (BRAF_S729), which was also strongly downregulated by Cobimetinib (logFC=-5.44, p=0.07 for Sorafenib and logFC=-5.41, p=0.09 for Cobimetinib) and is known to be important for BRAF-dimerization[39]. PTMNavigator's visualization suggests that Sorafenib does engage its target, but the perturbation in signaling does not progress beyond RAF. Previous research on the compound pointed out that Sorafenib, while being an effective agent against renal cell carcinoma and other cancer types, might not exert its efficacy via inhibition of RAF[40, 41]. In both examples presented here, PTMNavigator provided a visual insight into the target engagement and signaling cascades that one would not get as easily from other depictions such as heatmaps.

Exploring Dose-Dependent Perturbation Data in PTMNavigator

A very important characteristic of a drug's mode of action is how potently it perturbs different parts of cellular signaling, which can be captured by applying a series of doses and measuring the changes on the level of PTMs for each dose. Such a concentration-dependent approach, called decryptM, was recently published by *Zecha et al.*[2]. decryptM allows determination of dose-response characteristics for each peptide, including drug potency (negative logEC50 or pEC50 value) and drug effect size (curve fold change). Output files of the decryptM pipeline[2] or CurveCurator[42] can directly be imported by PTMNavigator. We implemented additional features

for dose-dependent data into PTMNavigator, which we demonstrate by using the tool to compare data from *Zecha et al.* to data from *Bekker-Jensen et al*. for the EGFR inhibitor Lapatinib.

The decryptM study included an experiment in which Lapatinib was applied to the breast cancer cell line MDA-MB-175-VII for 2 hours in 10 different concentrations (from 1 nM to 10 µM). The experiment yielded 180 upgoing and 441 downgoing dose-response curves classified as significant (see Methods). PTMNavigator's projection of the regulated peptides onto the pathway 'mTOR signaling pathway - hsa04150' (Fig. 5A; this was the highest-scoring pathway in the gene-centric and gene-centric-redundant ssGSEA of the dataset) showed that the overall perturbation of the network was very similar to that observed in the Lapatinib experiment by *Bekker-Jensen et al.* on RPE1 cells (Fig. 4A, upper panel). Many proteins had regulated peptides in both experiments (for example, RAF, MEK, and RICTOR), and some downregulated sites were shared by both biological systems (TSC2_S939, PRAS40_T246, and GSK3B_S9). However, the decryptM data adds the potency dimension for every regulation as an additional layer of information (Fig. 5B), represented in PTMNavigator by PTMs colored according to their pEC50 value. This visualization reveals that all peptides perturbed by Lapatinib in this pathway were regulated at a pEC50 between 6.4 and 7.4, which corresponds to a concentration range between 40 nM and 400 nM (Fig. 5C). The dose-response curves of each shown peptide can be displayed directly within PTMNavigator. Lapatinib is known to be a highly selective EGFR/ErbB2 inhibitor[40], and the fact that all potencies were within one order of magnitude suggests that this pathway is only regulated by a single initial perturbation. Here, this is likely the inhibition of ErbB2, as EGFR has been reported to be not active in MDA-MB-175-VII cells[43]. In other biological systems, the effect might differ. Such cell-type specific observations can also be highlighted by PTMNavigator, which we exemplified by comparing the effect of Dasatinib on 3 different cell lines (Suppl. Fig. S9).

PTMNavigator Aids in Refinement and Expansion of Pathway Diagrams

We applied PTMNavigator to two other decryptM datasets to put phosphorylation sites with and without functional annotation into context. Refametinib and Mirdametinib are two highly selective kinase inhibitors that both target MAP2K1 (MEK1) and MAP2K2 (MEK2). Both inhibitors were applied to A549 lung carcinoma epithelial cells in the decryptM study (in ten concentrations ranging from 1 nM to 10 µM). Both drugs elicited a selective response in the A549 phosphoproteome, with 148 regulated dose-response curves for Refametinib (38 up, 110 down) and 73 for Mirdametinib (15 up, 58 down). Among the pathways that were enriched for both datasets, the diagram 'Regulation of actin cytoskeleton – WP51' depicted the signaling around the compounds' targets most comprehensively. Projection of the drug-regulated peptides onto this pathway using the 'potency' color scheme (Fig. 6A) showed that Mirdametinib regulates MEK signaling more potently than Refametinib. This is in agreement with the reported higher affinity observed in pulldown experiments[40].

By focusing on dose-dependent phosphorylation signals in common between both drugs, we obtained a more complete understanding of the MEK signaling in this cellular system. Both drugs potently downregulated phosphorylation of MAPK1 and MAPK3 at their most important activating sites (T185 and Y187 for MAPK1, T202 and Y204 for MAPK3), which are all direct substrates of MEK1 and MEK2 and are part of the MEK1 kinase activity signature in PTMSigDB[17]. Downstream of MAPK, both drugs elicited downregulation of several peptides including MYPT1_S507, a site that has previously been associated with treatment by Selumetinib (a MEK1 and MEK2 inhibitor)[44] and Vemurafenib (which has MEK5 as a potent off-target)[40, 45]. The fact that it is also affected by the two inhibitors studied here and that its dose-response profiles are similar to other peptides associated with MEK inhibition (Suppl. Fig. S10) suggests that this

site could also be part of the general PTM signature of MEK inhibition. Less straightforward to interpret were several downregulations upstream of MEK at SOS1 sites as well as RAF1_S43 and GRLF1_S589 in both datasets. This is an example where the pathway diagram does not depict all relevant interactions for the cellular system, a frequently observed issue of the canonical pathway definitions (see Discussion). To obtain a more complete network of the effects of MEK inhibition, we investigated the pathway that PHONEMeS generated for the dataset (Fig. 6B) and consulted previous research on the matter. PHONEMeS reproduced the RAF-MAP2K-MAPK axis that is the essential part of the network. It also proposed a flow of signal from MAPK to SOS via RPS6KA and from MAPK to RAF1. Prior literature confirmed the first proposal: the regulated SOS1 residues S1134 and S1161 are RPS6KA substrates[38], and RPS6KA activity is controlled by MAPK[46]. Therefore, this regulation likely is the result of an inhibited feedback loop. SOS1_S1064 has no annotated kinase, but the other regulations suggest it could also be an RPS6KA substrate. RAF1_S43, however, is a substrate of PKACA, not MAPK[47]. The 'Editing Mode' of PTMNavigator allowed us to integrate this information into an extended pathway, using the diagram 'Regulation of actin cytoskeleton – WP51' as a starting point (Fig. 6C). GRLF_S589 has no functional annotation or putative upstream kinase in PhosphoSitePlus and PHONEMeS did not find a link, however, the similarities in regulation (Suppl. Fig. S10) suggest that it is part of the same phosphoproteomics signature. Therefore, it remains to be investigated which kinases are directly responsible for phosphorylating this site, how essential these phosphorylations are to the signaling pathway and whether this finding translates to other biological systems. Taken together, this shows that PTMNavigator can support researchers in the functionalization of uncharacterized PTM sites and can thus in aid in proposing putative new links in pathways by combining prior knowledge with experimental data.

Discussion

PTMNavigator is a tool for visualizing quantitative PTM data in the context of cellular pathways which supports users to analyze complex PTM proteomics data. In its current implementation in the ProteomicsDB platform, it allows exploring pathways of 10 organisms using pathway annotations from KEGG and WikiPathways and functional annotations from PhosphoSitePlus. The integration of automated enrichment analyses adds another dimension of information and the possibility of creating user-defined pathways from scratch and sharing them makes PTMNavigator versatile and unique. Other databases, such as BioCyc[48] and Reactome[8], may be utilized in future iterations of the tool to further extend the list of available pathways. We envision that it will become a platform for knowledge exchange in the systems biology field. We consciously decided to realize this software as a website instead of a desktop application. We believe that this makes it more accessible, as users will not have to install it or be concerned about the computational power of their desktop computers. The core visualization component of our software was built on the Web Component standards, making it easy to reuse in web interfaces outside of ProteomicsDB.

We demonstrated how PTMNavigator facilitates the interpretation of pathway enrichment analyses, makes perturbed signaling traceable, leverages the novel information contained in dose-dependent PTM data, and generates hypotheses for follow-up experiments to study unannotated modification sites. To our knowledge, PTMNavigator is the only available software that can graphically illustrate how regulated PTMs are distributed within pathways.

PTMNavigator gives a bird's-eye view of PTMs within a pathway, while at the same time integrating multiple sources of information on particular sites. This accelerates PTM perturbation analysis considerably, especially in large scale data sets. Furthermore, it fosters the identification

of new kinase-substrate relationships and signaling pathway crosstalk, as we exemplified for the phosphoproteomic response to MEK inhibitors in A549 cells.

There are usually multiple pathways that depict a signaling cascade of interest, and choosing a single one can seem arbitrary. We'd like to note that for the purpose of conciseness, in this manuscript we always selected a single KEGG or WikiPathways diagram for visualization of a dataset, while in practice we recommend that the user should look at all diagrams of highly enriched pathways to get an idea of how interchangeable the visualizations are and see if specific subnetworks are shared between them. For example, the visualization of MEK inhibition, for which we chose 'Regulation of actin cytoskeleton – WP51' (Fig. 6A), would for our purpose have worked just as well with the eponymous KEGG diagram 'Regulation of actin cytoskeleton – hsa04810', as it contains the same subnetwork of SOS1, RAF1, MAPK1/3, GRLF1, and MYPT1. Meanwhile, for other datasets or visualization purposes one diagram might be preferable over the other, and it is usually not possible to decide that without inspecting both. This is also the case if one of the pathways has a substantially higher relevance score – this can simply happen because the other pathway contains more additional proteins that 'water down' the score, however, that pathway can still be the better choice for visualization if the subnetwork of interest is more appropriately represented.

Current PEAs are far from ideal for analyzing PTM data sets, and their outcomes must, therefore, be taken with considerable caution. Nonetheless, PEAs provide currently the best starting point, which is why PTMNavigator uses PEA to pre-sort pathways by putative relevance. The initial ranking should help users to prioritize pathway engagement, but users should keep in mind that some high-ranked pathways are not biologically relevant and a PEA artifact. This is often rooted in the fact that pathway annotations are highly redundant (i.e., the same well studied proteins show

up in many well-studied pathways). On the other hand, also pathways with non-significant scores are often worth examining and additional resources, such as previous literature, should be considered when selecting diagrams for visualization. If better pathway ranking procedures are developed in the future, PTMNavigator can easily be updated to use those. We emphasize that PTMNavigator is a visualization tool, and that the interpretation of the results remains the responsibility of the user. While we believe the current work constitutes a substantial advance, we point out several caveats as follows: i) Differential or dose-response analysis is not part of PTMNavigator and users need to perform this prior to data upload. PTMNavigator only shows what its users already deemed significantly regulated. This is also a benefit, as it makes PTMNavigator flexible to many experimental settings. ii) The currently available canonical pathway diagrams are heterogenous in structure and, for PTM research, they may sometimes even be misleading. One example are the phosphorylation and dephosphorylation annotations in KEGG diagrams ('+/- p'). A phosphorylating connection from kinase *A* to protein *B* just means that at least one site on *B* is a substrate of *A*, not necessarily (actually not likely) all of them, but for a user the latter may seem implied. Furthermore, we could only retain those details of the pathway diagram that were also represented in the KGML or GPML files. Some of the pathways resemble drawings rather than formal diagrams and the relationships between pathway members are not encoded as edges in the graph (consider for example the KEGG pathway 'Nucleocytoplasmic transport - hsa03013' and the WikiPathways *D. melanogaster* pathway 'Homologous recombination - WP1205'). In PTMNavigator, these pathways contain few or no edges and can only serve as a collection of functionally related proteins. iii) Pathway definitions are typically generalized from a large amount of heterogeneous data. Consequently, they cannot be specific for a certain tissue or cell type, and some proteins depicted in a pathway might not be expressed or

active in the cells in which an experiment was performed. iv) A single node may represent a single gene or gene product, but can also summarize a subfamily of proteins (such as MAPK) or even an entire class of proteins (such as all receptor tyrosine kinases). Conversely, a ‘group’ node can represent a protein complex, but can also represent a set of alternatives for a step in the pathway. Users of the pathway diagram need to be wary of these disparate definitions. v) The current knowledge of pathways is limited and many relationships are not yet discovered. PTMNavigator can only show what is already known and its use for the discovery of novel signaling cascades or nodes is necessarily limited. As a consequence, PTMNavigator will become more powerful as pathway databases become more comprehensive in the future.

Materials and Methods

Creating a Pathway Database in ProteomicsDB

We assembled a collection of pathways from 10 different species within ProteomicsDB based on the reference pathway databases in KEGG (Release 107.0, July 1, 2023) and WikiPathways (Release 20230610, June 6, 2023). Both projects provide textual pathway definitions in variations of the extensible markup language (XML) format (KEGG: KEGG Markup Language (KGML); WikiPathways: Graphical Pathway Markup Language (GPML)), which we used as a starting point. There are differences in the detailed pathway specifications of KEGG and WikiPathways (see the next two subsections), but essentially, both define a pathway in the form of a mixed graph, consisting of a set of nodes and a set of edges (both directed and undirected) which connect the nodes. There are two classes of nodes: *regular nodes* have a description that includes 2-dimensional cartesian coordinates as well as information on the entity the node represents (for example a gene name), *group nodes* do not contain any additional description and are instead defined solely by a set of regular nodes contained in them. We wrote a Python package termed *pathway-importer* that converts both types of pathways into JavaScript Object Notation (JSON) format. The JSON format is better suited for use with a web application, since it can be parsed at runtime without further conversion (in contrast to XML). We did not import pathways that contained zero gene or gene product nodes, since they are not interesting for the study of cellular signaling (this includes, for example, metabolic pathways that are focused only on metabolites).

Preprocessing of KEGG Pathway Diagrams

We retrieved the list of pathways for each organism and downloaded the KGML files using the KEGG REST API (https://rest.kegg.jp). KGML files distinguish four types of nodes: *gene*, *compound*, *ortholog*, and *map*. To reduce the complexity of our internal representation, we

summarized the *gene* and *ortholog* nodes into a common *gene_protein* node type. *map* nodes were renamed as *pathway* nodes since they describe connections to other pathways, and *compound* nodes were not modified (Suppl. Fig. S2). In KGML files, nodes are labeled using KEGG identifiers, which we mapped to Uniprot accession numbers using the 'idmapping' service of the Uniprot REST API (as described here: https://www.uniprot.org/help/id_mapping). In order to maximize the number of nodes that had a human-readable name, we subsequently mapped the Uniprot accession numbers to gene names (also using the Uniprot API).

KEGG pathways have two classes of edges: *relations*, which connect exactly two nodes, and *reactions*, which can also represent many-to-many relationships by specifying a list of substrates and a list of products. We intended to only represent one-to-one relationships in ProteomicsDB, therefore we replaced each *reaction* by an equivalent set of *relations*, connecting each substrate-product pair by an individual edge. KEGG *relations* can have 19 different subtypes, which we retained in our JSON representation with two exceptions (Suppl. Fig. S3):

a) The types 'indirect' and 'indirect effect' were summarized to one type 'indirect'.

b) The types 'inhibition' and 'repression' were summarized to one type 'inhibition'.

Preprocessing of WikiPathways Pathway Diagrams

We downloaded GPML files from http://data.wikipathways.org/current/gpml. These pathways, which are maintained by a large, open community, are less standardized than in KEGG. The pathway collection we downloaded distinguishes 13 types of nodes, which we mapped to the same types as the nodes from the KEGG pathways. In addition to the *gene_protein*, *compound*, and *pathway* types, we introduced a fourth type *misc* to account for WikiPathways nodes that could not be assigned one of the other categories (Suppl. Fig. S2). In GPML files, each node is annotated

with a reference to an external database, such as Entrez or Ensembl. To unify these diverse annotations, we mapped them to Uniprot Accession Numbers where possible, using the Uniprot API (Suppl. Fig. S4). Edges in GPML files have 24 different types. 10 of these were converted into one of the KEGG edge types, the other 14 were retained as additional types (Suppl. Fig. S3). A specific feature of the GPML specification is that edges may also end on another edge instead of a node. This is represented by an additional tag in the XML schema, called *anchor*. To keep our representation simple, we replaced *anchor* tags by edges that had other edges as their start or end point.

Implementation of the *pathwaygraph* Web Component

We developed a Web Component (www.webcomponents.org), termed *pathwaygraph*, for the dynamic rendering of pathway diagrams together with peptide-level experimental data. *pathwaygraph* was implemented using the Lit library (https://lit.dev/), Version 2.0.2. Our Web Component expects two inputs: A pathway diagram in JSON format (we call this 'pathway skeleton'), and a list of PTMs, each annotated with a regulation type (possible values are 'up', 'down', and 'not'). Optionally, a list of proteins and their regulation types can be supplied in addition.

pathwaygraph maps each entry of the PTM list to nodes of the pathway skeleton by comparing the gene names and/or Uniprot accession numbers that are associated with the nodes. For each matched PTM, an additional node is added to the graph definition, as well as an edge from the new node to its associated node in the skeleton. A PTM can be matched to multiple nodes of the skeleton, resulting in multiple PTM nodes for the same entry of the PTM list. This can happen, for example, when a protein appears twice in the same pathway because the creators wanted to describe two different routes of a signaling cascade.

After the matching process has completed, *pathwaygraph* renders the graph as an SVG object using the software library D3.js[49] (Version 6.2.0). The cartesian coordinates of the pathway skeleton nodes are used for the initial layout. The edges are drawn in such a way that they connect their two adjacent nodes in a straight line. When a node changes its position (e.g. by a user dragging it around), the endpoints of the edge are automatically updated. The PTM nodes, which do not have cartesian coordinates assigned, are positioned using simulated physical forces (using the d3-force package (https://github.com/d3/d3-force), Version 2.1.1), which cause them to remain close to their reference skeleton node without colliding with one another. PTM nodes can either be represented individually, or in the form of a single summary node for each regulation type that only displays the number of regulations (step 6 in Fig. 2A; users can toggle between the two representations). Optionally, a label consisting of the modified residue and position within the protein can be displayed above the PTM node.

The different node types (*gene_protein*, *compound*, *pathway*) are distinguished by different colors. Group nodes are represented as a polygon around the convex hull of all members of the group. 4 types of edges are distinguished visually by different dash patterns or arrowheads: *activation*, *inhibition*, *binding/association*, and *indirect*. All remaining edge subtypes are rendered with the same style. Only edges from the skeleton are rendered, the edges between PTM nodes and skeleton nodes are invisible (but their presence is required for the force simulation).

There are three possible color schemes for the PTM nodes: the default and most simple option is to color them by their regulation type ('up' nodes are colored red, 'down' nodes blue, and 'not' nodes grey). Alternatively, nodes can be colored by fold change or (in the case of decryptM data) by potency (using EC50 values estimated from the dose-response curve fit). In both cases, the full range of values is determined after mapping the PTM nodes to the skeleton. A continuous color

gradient is then calculated between the minimum and the maximum value and applied to each node.

Each PTM in the input can be annotated with arbitrary details. In the case of phosphoproteomics data, these details may include upstream kinases. If such annotations are present, and the kinases happen to also be present in the currently displayed pathway skeleton, *pathwaygraph* will automatically add the kinase-substrate relationships as additional edges to the diagram that can optionally be shown during the analysis.

In addition to PTM-level and protein-level data, a list of perturbation targets (e.g. estimated by a KSEA) may be supplied. These targets can optionally be highlighted within the pathway diagram by blue and red circles around the respective nodes.

In addition to pathway viewing and projection of PTM datasets, we implemented a second application mode for *pathwaygraph*, in which pathway diagrams can be edited. This mode allows the addition, removal, and rearrangement of nodes, node groups, and edges. Details of nodes and edges (such as labels or gene identifiers) can be edited as well. Diagrams can be exported in the same JSON format as the canonical 'pathway skeletons' and they can also be stored in ProteomicsDB. It is also possible to upload JSON files from a user's local machine for editing and storing in ProteomicsDB.

Using the package @api-client/context-menu, v0.4.1, we developed a context menu for *pathwaygraph* which allows the user to interact with the application. The context menu is dynamic and its contents depend on the current state of the application (e.g., it shows different options in 'viewing' mode and in 'editing' mode).

Implementation of PTMNavigator

We implemented PTMNavigator as a Vue.js component (https://vuejs.org). It is essentially a wrapper around *pathwaygraph* that facilitates interaction of the Web Component with both the user and the backend of ProteomicsDB and can provide additional information such as the results of enrichment analyses (see below). PTMNavigator handles user interactions such as the selection of datasets, organisms, and pathways, or the application of filters on the loaded data, by calling the necessary endpoints of the ProteomicsDB API, processing the returned data, and passing it on to *pathwaygraph*. Custom pathway diagrams can also be saved to and restored from ProteomicsDB (they are associated with the user's Session ID), as well as downloaded and uploaded in JSON format.

PTMNavigator is embedded into ProteomicsDB, which is also Vue.js-based since its latest release[24]. The tool can be found in the 'Analytics' section of ProteomicsDB: https://www.proteomicsdb.org/analytics/ptmNavigator.

Implementation of a Service for Enrichment Analyses

Using the *flask* framework (version 2.0.3), we implemented a Python-based web server that provides a number of common enrichment analysis algorithms for proteomics data, with a focus on phosphoproteomics. An Application Programming Interface (API) allows applications as well as human users to send POST requests to the server. Since each of the method that the service offers requires a different way of structuring the input data, we implemented preprocessing scripts that convert JSON input into the required formats, so that the actual input that the user has to provide is as similar as possible for all algorithms. The result of each algorithm is then also converted into JSON and sent back to the requester. We mostly used existing Python or R packages, except where noted. The service is publicly available under this web address: https://enrichment.kusterlab.org/main_enrichment-server/. The code can be accessed here:

https://github.com/kusterlab/enrichment-server. In the following, we describe the methods that were offered by our service at the time of writing. For details on the input formats, please refer to the Github repository, where we provide examples for every endpoint.

PTM Signature Enrichment Analysis (PTM-SEA)

We use the R package ssGSEA2.0 (https://github.com/broadinstitute/ssGSEA2.0), commit hash 4b5198f (July 20, 2023). As database, we use PTMSigDB v2.0.0, "flanking" and "uniprot". Expected input is a list of PTM sites and a numeric value for each experiment (e.g. normalized measured expressions). Duplicates are eliminated during preprocessing by only retaining the entry with maximal absolute expression value for each site. The algorithm is invoked with the *export (-e)* parameter set to FALSE, and all other parameters set to their default values. From the resulting .gct file, the 'Signature ID' column and the 'Overlap', 'adj p-val' and 'Score' columns for every processed experiment are returned.

Gene-Centric and Gene-Centric Redundant single sample Gene Set Enrichment Analysis

Again, we use the R package ssGSEA2.0. As database, we use the subset of all KEGG and WikiPathways signatures from MSigDB. In order to match the set of pathways that are available in PTMNavigator, we excluded 'KEGG medicus' signatures. The expected input is a list of gene symbols and their expression in each experiment. In the case of Gene-Centric analysis, duplicates are eliminated as described for PTM-SEA. In the case of Gene-Centric Redundant analysis, duplicates are not eliminated. The same columns as for PTM-SEA are returned.

Kinase-Substrate Enrichment Analysis (KSEA)

We use the *kinact* Python package (https://github.com/saezlab/kinact, commit hash 572d3a8) which performs KSEA as described by *Casado et al.* by comparing the mean fold change among the set of substrates of a kinase to an expected value. As prior knowledge, we use the set of

kinase-substrate relationships from PhosphoSitePlus, retrieved using OmniPath (https://pypath.omnipathdb.org/) on February 2nd, 2024, and converted into an adjacency matrix by a custom Python script. Expected input is a list of phosphorylation sites and a numeric value for every experiment. Duplicates are eliminated as described for PTM-SEA. For each experiment, scores and adjusted p-values are returned.

Robust Kinase Activity Inference (RoKAI) + KSEA

We use the R package RokaiApp (https://github.com/serhan-yilmaz/RokaiApp), commit hash e2eb10a. As prior knowledge, we use the file 'rokai_network_data_uniprotkb_human.rds' provided by the RokaiApp repository. Expected user input is the same as for KSEA. In addition to RoKAI's default Kinase-Substrate network, all optional components of the RoKAI circuit system are used (PPI, SD, CoEv). The resulting refined phosphorylation profiles are subsequently processed by KSEA as described above.

PHONEMeS

We use the R package PHONEMeS (https://github.com/saezlab/PHONEMeS), v2.0.1, as well as Cytoscape v3.10.1 equipped with the yFiles plugin (https://www.yworks.com/products/yfiles-layout-algorithms-for-cytoscape), v1.1.4. As prior knowledge network, we use the concatenation of the 'phonemesPKN' and 'phonemesKSN' networks, which describe ground truth protein-protein interactions and kinase-substrate relationships. The expected user input is, for each experiment, a list of up- and down-regulated perturbation targets (in gene symbol format), and a list of sites with numeric values for each experiment (e.g., normalized measured expression, Fold Change, or t-statistic). The PHONEMeS algorithm is then invoked for each experiment with n_steps_pruning set to 2, and default values for all other parameters, and using the IBM cplex optimizer (https://www.ibm.com/products/ilog-cplex-optimization-studio/cplex-optimizer) as

solver for the Integer Linear Program (ILP). If no solution is found, the algorithm is rerun with n_steps_pruning set to 3. Then, the resulting network for each experiment is converted into a *pathwaygraph* pathway skeleton: First, all nodes representing a phosphorylation site are replaced by a node for the protein to which the phosphosite corresponds. Then, duplicated *node-edge-node* triplets and self-loops are removed. Next, the network is loaded into Cytoscape (we use the Python package py4cytoscape, v1.9.0., for programmatic interaction with Cytoscape), and the yFiles 'Hierarchic Layout' is applied and the graph is exported as a '.cx' file. Using custom Python code, this file is converted into a JSON file with the 'pathway skeleton' format that can be parsed by *pathwaygraph*. The JSON file is returned to the requester.

Motif Enrichment

We use inhouse Python code that essentially reimplements the service provided here: https://kinase-library.phosphosite.org/ea. In brief, from the kinase library data we created position-specific scoring matrices for each kinase that reflect the preference of every amino acid to appear at every position of the kinase's substrate motif. We scored every phosphopeptide in PhosphoSitePlus against the matrix of each kinase to obtain a background score distribution. The expected user input is then a list of phosphopeptide sequences, the Uniprot Accession Numbers of the corresponding proteins, and for each experiment the regulation category of this peptide ('up'/'down'/'not'). Every peptide from the user input is scored against every kinase and the result is compared to the background to obtain quantiles. The 15 highest scoring kinases for every site are determined. Then, each kinase that appears in the result is tested for significant overrepresentation by means of a Fisher's exact test (including multiple testing correction using the Benjamini Hochberg method). The test results for each kinase in each experiment are converted into JSON format and returned to the requester.

KEA3

We use the application programming interface (API) of KEA3. The expected user input is a list of proteins (gene symbols) for each experiment. The enrichment service sends a POST request to the URL https://amp.pharm.mssm.edu/kea3/api/enrich/. The tables 'Integrated--meanRank' and 'Integrated--topRank' are extracted from the response and returned to the requester.

KSTAR

We use the kstar Python package, v0.5.3. The expected user input is a list of phosphopeptide sequences, the associated Uniprot Accession Numbers and a numeric value for each peptide in each experiment. The peptides are converted into ‘flanking sequence’ format (7 adjacent amino acids in both directions) and annotated with their position within the protein using the psite_annotation package (v.0.5.0) and the reference proteome from PhosphoSitePlus (Version from November 2nd, 2020). KSTAR is then executed using ‘mean’ as aggregation method and 0 as threshold (which means that values are regarded as significant, therefore non-significant regulations need to be filtered out by the requester in advance). The algorithm is run twice, first with ‘greater’ set to FALSE, then to TRUE, to capture both over- and underactive kinases. This is executed for both modification types (‘ST’ and ‘Y’). The p-values of the hypergeometric test are log-transformed and, in the ‘greater==TRUE’ run, multiplied by -1. This ensures that activated kinases receive a score >0, while inactivated kinases get a score <0. These scores are returned to the requester. Note that the KSTAR protocol recommends to afterwards perform analyses on randomized datasets and conduct Mann Whitney U-tests to obtain more robust p-values. Since this is a time-intensive calculation that takes significantly bigger amounts of memory then the other enrichment algorithms, we decided against including this step in our enrichment service.

Processing of User Data Upload

We extended the upload page of ProteomicsDB (www.proteomicsdb.org/analytics/customDataUpload) to allow the upload of data for PTMNavigator in csv format. It is possible to upload conventional quantitative proteomics ('Fold Change') data as well as decryptM data on both PTM and protein level. In the case of phosphorylation data, users can upload data on both peptide and p-site level; for other PTMs, only peptide-level upload is possible at this time. Example files for all input types are provided for download on the website (see also Supplementary Table 1). In the case of decryptM data, the 'curves.txt' file output by the decryptM pipeline or CurveCurator can directly be uploaded together with the TOML file containing the experimental parameters. The uploaded data is stored temporarily in ProteomicsDB and is deleted if not accessed for 14 days. After the completion of the upload, requests are sent to the endpoints of the enrichment service (see above). For phosphoproteomics data, all endpoints are queried, for other PTM and protein-level data, only the gene-centric and gene-centric redundant GSEA are performed. In the case of the PHONEMeS+Cytoscape endpoint, perturbation targets are required as input in addition to the phosphoproteomic data. Here, the results of the RoKAI+KSEA analysis are used. For each experiment, the 3 highest scoring kinases are regarded as perturbation targets. The pathways returned by the endpoint are added to the user's custom pathway diagram list. All other enrichment results are inserted into a dedicated table in ProteomicsDB. To limit the access to the data, the user receives a personalized 32-digit alphanumeric session identifier during the upload that needs to be entered in PTMNavigator to retrieve the uploaded data.

When a user visits https://proteomicsdb.org/analytics/ptmNavigator, the custom pathways associated with their Session ID are automatically retrieved. When the user requests a dataset they

had previously uploaded, all enrichment analysis results for this dataset are also retrieved. The canonical KEGG and WikiPathways pathways are sorted by gene-centric redundant GSEA score. All kinases that received an absolute score >0.5 in the RoKAI+KSEA analysis are considered perturbed and are passed as a property to the *pathwaygraph* instance. If those kinases appear in a pathway diagram, they can optionally be highlighted by red (up) and blue (down) circles.

Import of PhosphoSitePlus annotations and Functional Scores

We downloaded the files 'Regulatory_sites.gz' and 'Kinase_Substrate_Dataset.gz' from PhosphoSitePlus (https://www.phosphosite.org/staticDownloads; accessed October 18, 2023). Using a custom Python script, we mapped all annotations to the *in silico* digested human reference proteome present in ProteomicsDB and created API endpoints to make the information accessible from the PTMNavigator frontend. The phosphosite functional scores were imported analogously, using Supplementary Table 3 from the paper by Ochoa and colleagues[33]. During retrieval of a dataset, the associated annotations are collected from ProteomicsDB and forwarded to PTMNavigator alongside the data. This allows, for example, to display kinase-substrate annotations from PhosphoSitePlus as additional edges within the diagram.

Preprocessing of data from Bekker-Jensen et al.

For the detailed experimental procedures of the kinase inhibitor screen, please refer to the original publication. Briefly, RPE1 cells were treated with either 0.1 µM or 1 µM of inhibitor for 30 minutes, followed by stimulation with EGF for 10 minutes. All experiments were performed in triplicates, except for an EGF-stimulation-only control, which was performed in duplicates. After extraction of proteins, digestion into peptides, and Ti-IMAC-based phosphopeptide-enrichment, the samples were measured using LC-MS/MS in Data-Independent Acquisition (DIA) mode. The data was then processed using Spectronaut to identify and quantify phosphorylated peptides.

Our analysis of the data differed from that of *Bekker-Jensen et al.* in that they performed multiple sample significance testing (ANOVA), comparing all conditions against the control together, whereas we intended to compare each condition to the control separately using *t*-tests. We downloaded the Spectronaut output file from ProteomeXChange using the Identifier PXD014525 ('20190416_214744_Kinase-inhibitor-screen-3reps-STY-version13-LocProb-075-Score-1_Report.txt') and imported it into Perseus (version 1.6.2.3). We transformed the normalized intensities into phosphorylation site tables using the Peptide Collapse plugin (https://github.com/AlexHgO/Perseus_Plugin_Peptide_Collapse), specifying EG.PTMAssayProbability as grouping column. We filtered for three valid values in at least one treatment group. We log-transformed and replaced missing values by sampling from a normal distribution (setting the parameters 'width' to 0.3, 'downshift' to 1.8, and 'mode' to 'Total matrix'). We then performed two-sided *t*-tests between the 60 treatment conditions (30 drugs, 2 doses) and the control condition for each quantified peptide. *p*-values were corrected for multiple testing using the Benjamini-Hochberg procedure (FDR=10%). Finally, we summarized the data into 30 conditions (one per inhibitor) and eliminated duplicated peptides using the following rules:

I. If the peptide appeared multiple times in the same experiment (which can happen for multiply phosphorylated peptides), retain only the most significant entry (lowest adjusted *p*-value).

II. If a peptide was neither significantly regulated in the low-dose experiment nor in the high-dose experiment of a compound, retain only the entry from the high-dose experiment.

III. If a peptide was significantly regulated in the high-dose experiment, retain only this entry.

IV. If a peptide was significantly regulated *only* in the low-dose experiment, retain only this entry.

The resulting tables was formatted to fulfil the input criteria of PTMNavigator.

Pathway enrichment was performed as described above.

The code of this analysis (starting from the Perseus output and ending at the creation of PTMNavigator input files) is available as a Jupyter Notebook (see 'Data Availability').

Preprocessing of data from Zecha et al.

For details on the decryptM experimental protocol, please refer to the supplementary methods in the original publication. In brief: In the Lapatinib experiment, MDA-MB-175-VII cells (human breast cancer) were treated with a series of Lapatinib for 2h (concentrations: 1, 3, 10, 30, 100, 300, 1000, 3000, 10000 nM) or a DMSO-only control. In the Refametinib/Mirdametinib experiments, human lung adenocarcinoma A549 cells were treated with a series of Mirdametinib (PD325901) or Refametinib concentrations for 60 minutes (concentrations: 1, 3, 10, 30, 100, 300, 1000, 3000, 10000 nM), or a DMSO-only control. In the Dasatinib experiments, A431 epidermoid carcinoma epithelial cells, K562 lymphoblast cells and A549 cells were treated with Dasatinib (concentrations: 0.03, 0.3, 1, 3, 10, 30, 100, 300, 1000, 10000 nM for A431 cells; 1, 3, 10, 30, 100, 300, 1000, 3000, 10000 nM for A549 cells; 0.01, 0.03, 0.1, 0.3, 1, 3, 10, 30, 100 nM for K562 cells) or a DMSO-only control for 30 minutes (60 minutes for A549 cells). Proteins were extracted from the cells and digested using Trypsin. The samples were multiplexed using tandem mass tags (TMT). Phosphopeptide enrichment was performed using Fe-IMAC, after which the samples were separated into six fractions using bRP-fractionation to increase phosphoproteomic coverage. The fractions were then measured separately using LC-MS/MS in data-dependent acquisition (DDA) mode. TMT phosphopeptides were identified and quantified using MaxQuant (version 1.6.12.0).

We downloaded the following archives containing MaxQuant output from ProteomeXChange using the identifier:

- Lapatinib experiment: HER2_Inhibitors_Phosphoproteome_dd.zip
- Refametinib experiment: 10_Kinase_Inhibitors_Rep_Phosphoproteome.zip
- Mirdametinib experiment: 10_Kinase_Inhibitors_2KI_Phosphoproteome.zip
- Dasatinib A431 experiment: EGFR_Inhibitors_Phosphoproteome.zip
- Dasatinib A431 experiment: 10_Kinase_Inhibitors_6KI_Phosphoproteome.zip
- Dasatinib A431 experiment: Dasatinib_Triplicates_Phosphoproteome_MS3.zip

We filtered the evidence.txt files for rows containing the experiments of interest. The filtered files, together with TOML parameter files which we constructed manually, were used as input for CurveCurator (https://github.com/kusterlab/curve_curator). Using CurveCurator, we fitted 4-parameter log-logistic dose-response curves for every quantified peptide and obtained significance estimates for the regulations. Then we imported the curves file and the TOML into PTMNavigator. For the Refametinib and Mirdametinib datasets, pathway enrichment was performed as described above.

Preprocessing of Data from Lee et al.

We downloaded the Datasets EV4 (Full Proteome Data) and EV5 (Phosphoproteome Data) from here: https://doi.org/10.1038/s44320-023-00004-7 and used the sheets containing log-transformed and median-centric normalized iBAQ Values with imputations (‘normalized_FP_log10_imput_hmin’ and ‘normalized_PP_log10_imput_hmin’, respectively). We calculated the difference in iBAQ log fold changes between the A204 and RD-ES cell lines. As a threshold for regulation, we set an iBAQ log fold change difference of 3 for the full proteome data and 4 for the phosphoproteomic data (i.e., peptides/proteins higher in A204 are annotated as

729 ‘up’, whereas peptides/proteins higher in RD-ES are annotated as ‘down’). The resulting files were

730 loaded into PTMNavigator.

Data availability

Experimental data was downloaded from ProteomeXChange as described in the Methods, using the identifiers PXD014525 (*Bekker-Jensen et al.*) and PXD037285 (*Zecha et al.*), as well as from the website of Molecular Systems Biology: https://doi.org/10.1038/s44320-023-00004-7 (*Lee et al.*). The following files are available on Zenodo (https://doi.org/10.5281/zenodo.12720528):

- For the Bekker-Jensen data:
    - Perseus output (log-transformed phosphorylation site intensities) (.txt)
    - A Jupyter Notebook describing how to convert the Perseus output into PTMNavigator input (.ipynb)
    - PTMNavigator input (.xlsx)
    - Enrichment analysis results: GCR, PTM-SEA, KSEA, RoKAI+KSEA (all .xlsx)
- For the Zecha data:
    - CurveCurator parameter files (.toml)
    - CurveCurator output (.xlsx)
    - Enrichment analysis results: GCR, PTM-SEA, KSEA, RoKAI+KSEA (all .xlsx)
- For the Lee data:
    - PTMNavigator input (.xlsx)

All other data are available in the main text or the supplementary materials.

Code availability

The code to import and preprocess the pathway diagrams is available at https://github.com/kusterlab/pathway-importer. *pathwaygraph* is part of the series of Bio-Web Components developed by our lab, *@biowc,* and is available as a JavaScript module here: www.npmjs.com/package/@biowc/pathwaygraph. The code can be accessed here:

www.github.com/kusterlab/biowc-pathwaygraph. The source code of PTMNavigator can be accessed here:
https://github.com/wilhelm-lab/proteomicsdb-vue/blob/main/src/views/analytics/PTMNavigator.vue.

Acknowledgements

We thank Chien-Yun Lee and Stefanie Höfer for testing the software and providing valuable feedback. We thank all other members of the Kuster, Lee, and Wilhelm Labs as well as the Saez Lab for fruitful discussions. Parts of the figures were created with BioRender.com. This work was partly funded by the European Research Council (ERC Advanced Grant No. 833710), the German Federal Ministry of Education and Research (DIAS, FKZ 031L0168), and the European Union's Horizon 2020 Program (Grant Agreement ID: 823839).

Author contributions

Conceptualization: MT, FB, JM

Software: JM, MS

Data Analysis: JM, MT, FB

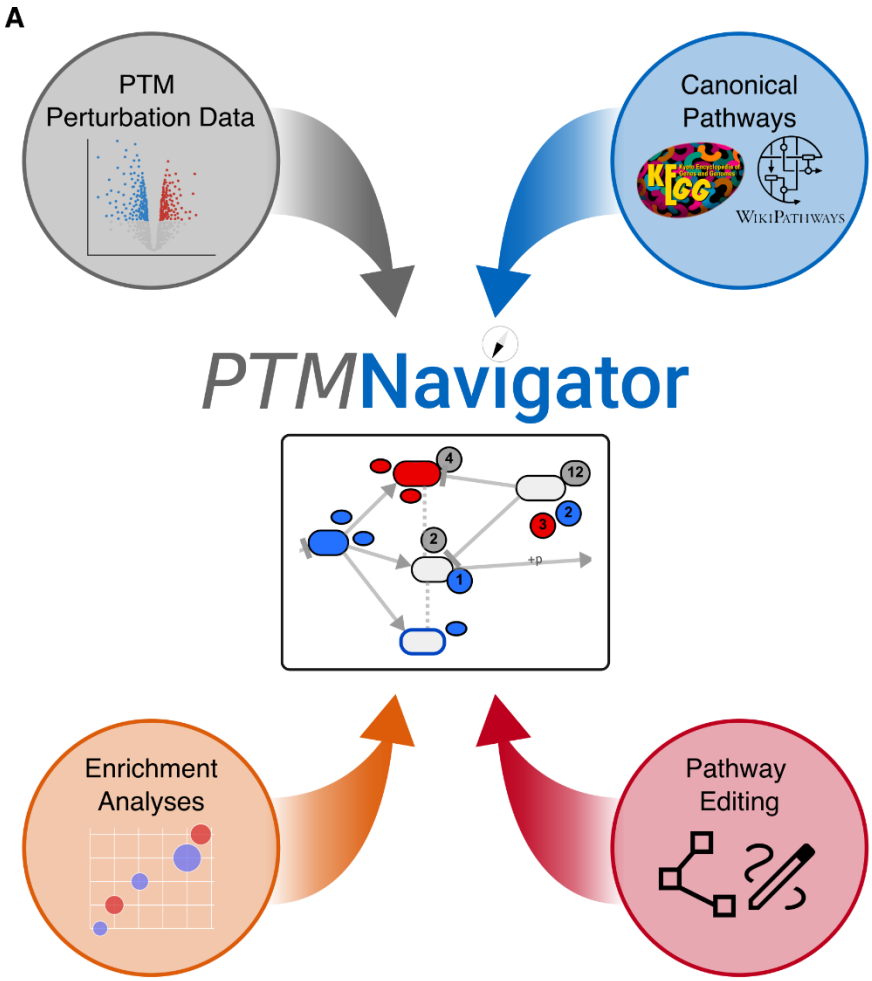
Writing – original draft: JM

Writing – review & editing: MT, BK, FB, MW, JM, MS

Supervision & Project Administration: MT, MW, BK

Conflict of Interest

BK and MW are founders and shareholders of OmicScouts and MSAID. They have no operational role in either company. All other authors declare that they have no competing interests.



B

Type	Algorithm Name	Publication
Gene Set Enrichment	ssGSEA - Gene-Centric	Krug et al.
Gene Set Enrichment	ssGSEA - Gene-Centric-Redundant	Krug et al.
PTM Set Enrichment	PTM-SEA	Krug et al.
Kinase Activity Inference	KSEA	Casado et al.
Kinase Activity Inference	RoKAI + KSEA	Yılmaz et al.
Kinase Activity Inference	KSTAR	Crowl et al.
Kinase Activity Inference	KEA	Kuleshov et al.
Motif Analysis	Motif Enrichment	Johnson et al. / Inhouse
Pathway Inference	PHONEMeS*	Gjerga et al.

Figure 1. PTMNavigator combines PTM perturbation data, pathway diagrams, and enrichment analyses within a single application. (**A**) The four building blocks of PTMNavigator. Data from PTM perturbation experiments is provided by the user *(top left)*. This

data can then be projected onto canonical pathway diagrams from KEGG and WikiPathways *(top right)*. Uploaded datasets are automatically processed using several enrichment analysis algorithms and the results are integrated into the visualization as well as displayed in tabular form *(bottom left)*. The user also has the option to design custom pathways, either by using the canonical diagrams as templates or by starting from scratch *(bottom right)*. (**B)** Overview of the enrichment algorithms available in PTMNavigator. *PHONEMeS is not an enrichment algorithm, instead it is a method to reconstruct pathways for phosphoproteomics datasets using prior knowledge.

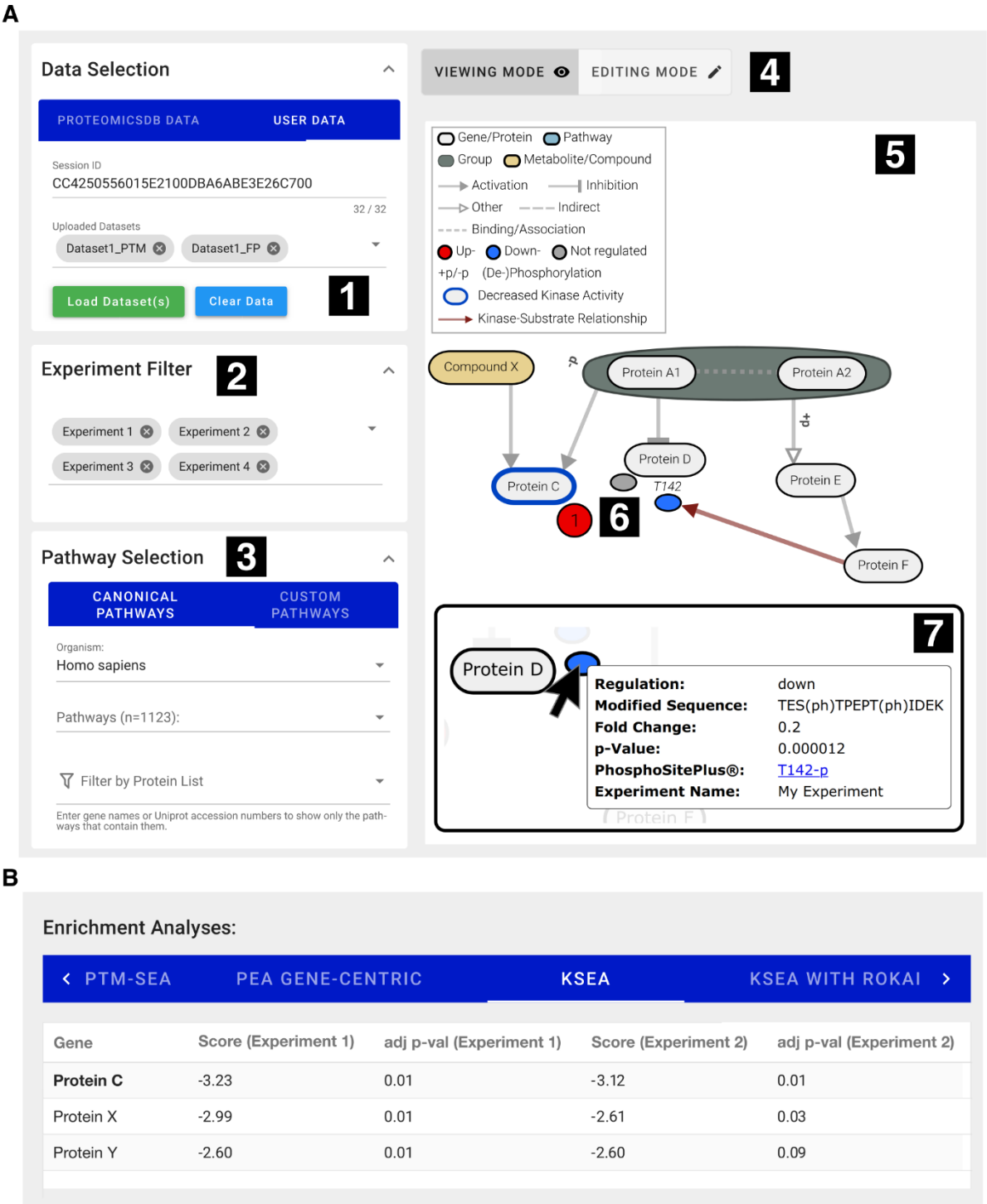


Figure 2. The User Interface of PTMNavigator. (**A**) The main interface of PTMNavigator. (1) Datasets can be accessed via a 32-digit Session ID that users receive during data upload. An arbitrary combination of datasets can be loaded at the same time. (2) A dataset may contain

multiple experiments (specified in the uploaded CSV file). In this panel, experiments can dynamically be added to and removed from the diagram. (3) The list of pathways to choose from is sorted by scores from a gene-centric redundant enrichment analysis. Users can also search this list for the name or identifier of a pathway, or filter for pathways that contain a certain set of genes or proteins. (4) Users can switch between a viewing mode, in which they can study the projection of their data onto pathway diagrams, and an editing mode, in which they can modify existing and create custom pathways. (5) The main canvas. After selecting a pathway, its diagram is rendered here. Additional smaller nodes are added for the peptides in the user data. (6) Initially, all peptides of a regulation group (up-/down-/unregulated) that belong to the same protein are summarized in one circular node that is labeled with the number of peptides it contains. By double-clicking on it, a summary node can be expanded into individual nodes for each peptide (elliptical shapes). The position of the modified residue can optionally be displayed above the ellipsis. (7) When hovering over a node, all information supplied during the upload is displayed as a tooltip. If a site is annotated in PhosphoSitePlus, a link to the annotation is retrieved. (**B**) The Enrichment Analyses View, showing the three most significant results from a Kinase-Substrate Enrichment Analysis (KSEA).

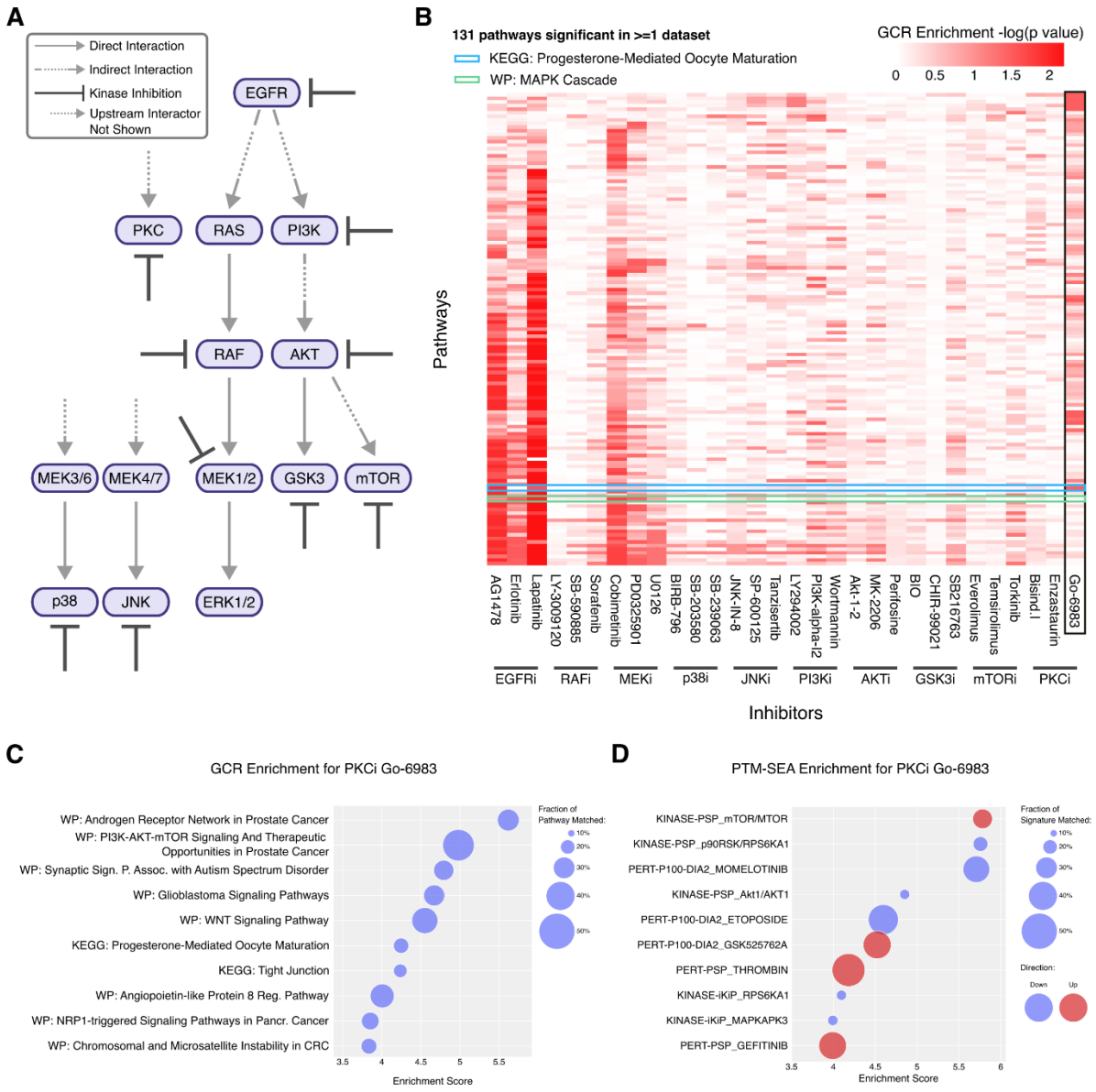


Figure 3. PTMNavigator aids in interpreting pathway enrichment analysis results. (**A**) Overview of the kinases which were targeted with inhibitors in the study by *Bekker-Jensen et al.* (**B**) Results of gene-centric-redundant ssGSEA showing pathways with p-value<0.05 for the 30 kinase inhibitor datasets. Pathways are sorted by the number of datasets in which they were reported as enriched (larger numbers on the bottom). Inhibitors are grouped by their designated targets. GCR: gene-centric redundant. (**C**) 10 highest-scoring pathways for the PKC inhibitor Go-6983 by gene-centric-redundant ssGSEA score. Circle size indicates the fraction of genes in the

pathway that were contained in the set of regulations. (**D**) 10 highest-scoring pathways for the PKC inhibitor Go-6983 by absolute PTM-SEA score. Circle size indicates the fraction of genes in the signature that were contained in the set of regulations.

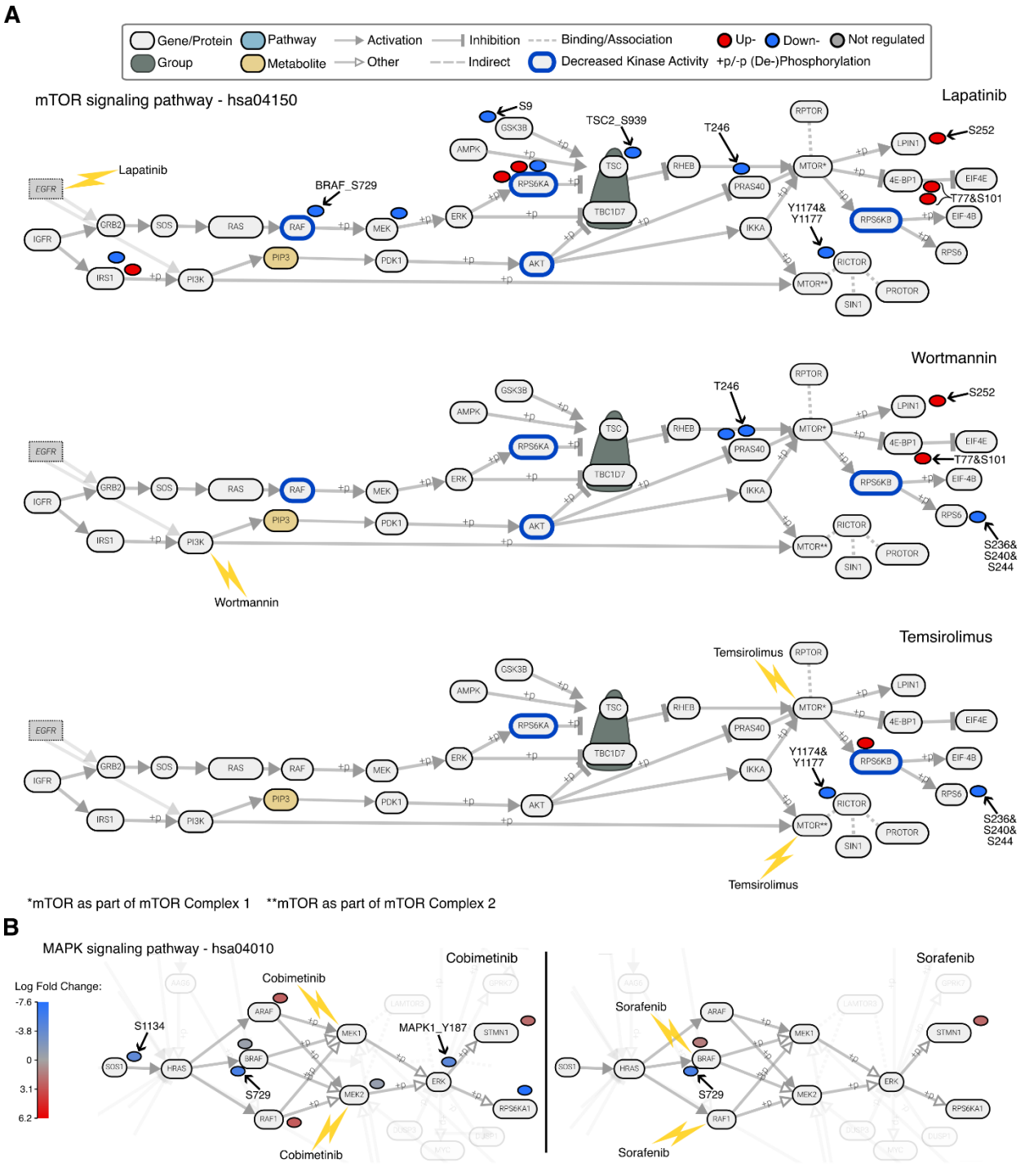


Figure 4. Analysis of a phosphoproteomic perturbation dataset in PTMNavigator. (**A**) Edited screenshot from PTMNavigator showing an excerpt from the pathway diagram 'mTOR Signaling Pathway - hsa04150' and the modified peptides significantly regulated by three of the drugs in

Bekker-Jensen et al. Note that the node for EGFR was added by the authors. It is not part of the original pathway diagram, but is known to interact with GRB2 and PI3K. The blue borders indicate kinases whose activity was significantly decreased (Score <-0.5, FDR < 0.1) according to KSEA analysis. *Top:* Lapatinib. *Middle:* Wortmannin. *Bottom:* Temsirolimus. (**B**) Edited screenshot from PTMNavigator showing an excerpt from the pathway diagram 'MAPK signaling pathway - hsa04010' and the modified peptides regulated by two of the drugs in *Bekker-Jensen et al.* Regulated PTMs are colored by the size of the effect (Log_2 Fold Change). *Left:* Cobimetinib. *Right:* Sorafenib.

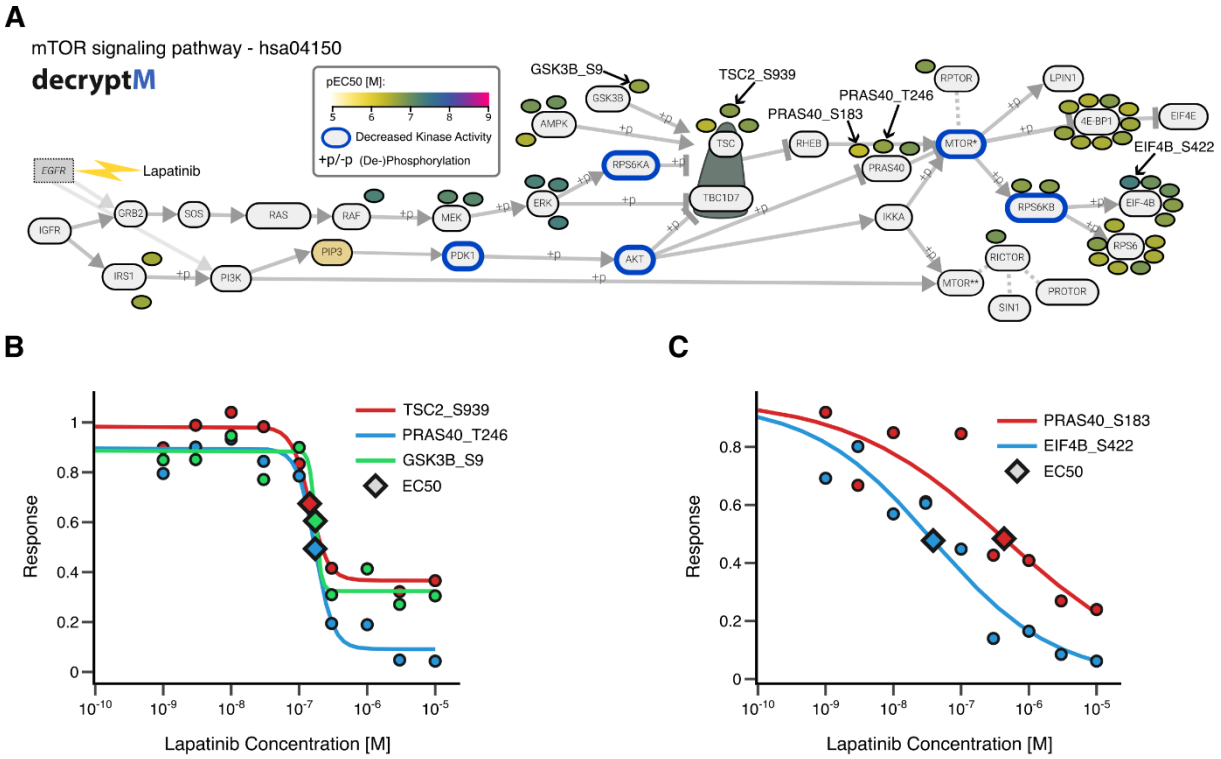
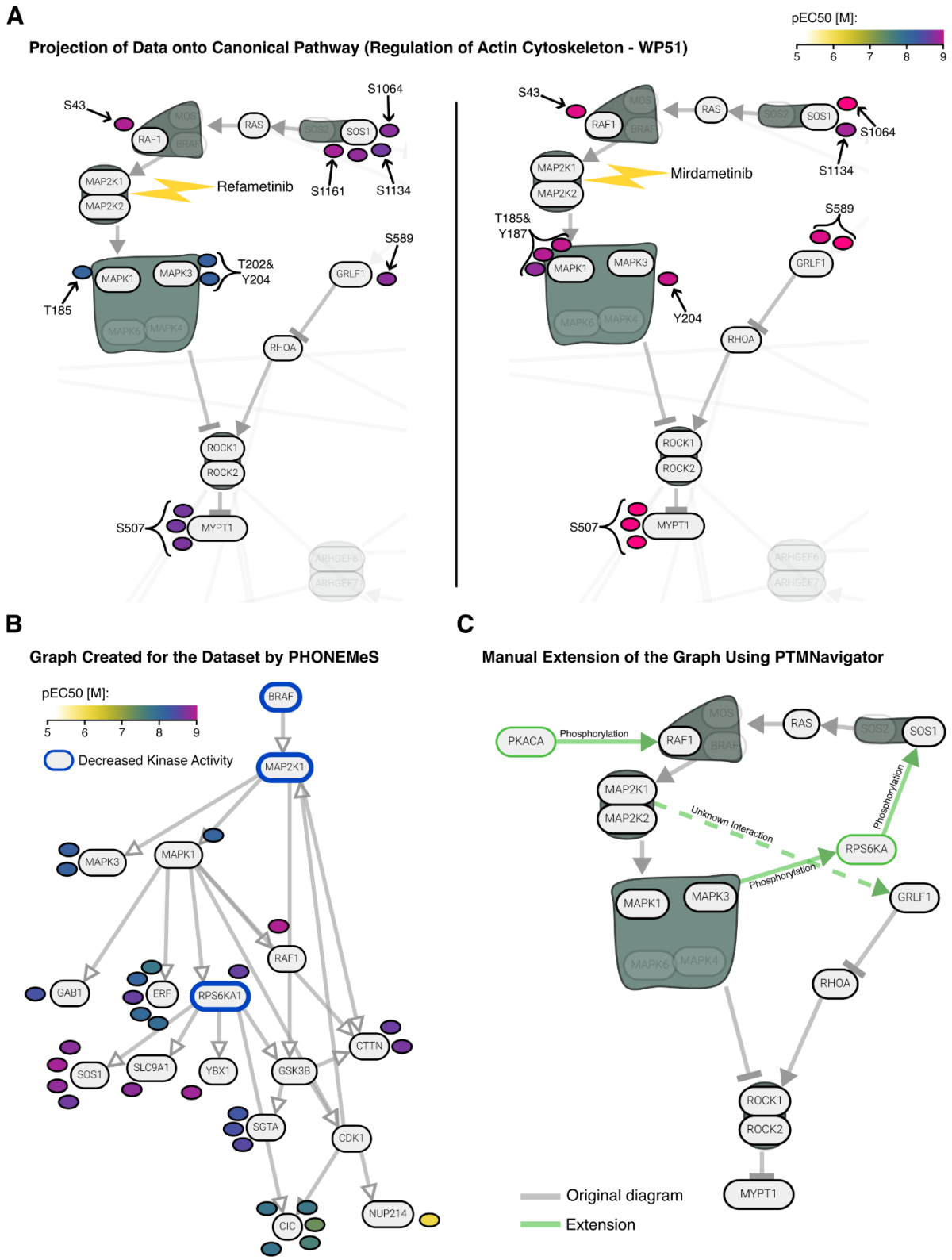


Figure 5. PTMNavigator supports analysis of drug dose-dependent perturbation data. (**A**) Edited screenshot from PTMNavigator showing an excerpt from the pathway diagram 'mTOR Signaling Pathway - hsa04150' and the modified peptides regulated in the Lapatinib experiment of *Zecha et al.* The node for EGFR was added by the authors, as it is not part of the original pathway diagram, but is known to interact with GRB2 and PI3K. The blue borders indicate kinases whose activity was significantly decreased (Score <-0.5, FDR < 0.1) according to RoKAI refinement followed by KSEA analysis. (**B**) decryptM curves from the Lapatinib dataset for those peptides that were also significantly downregulated in the *Bekker-Jensen* data. (**C**) decryptM curves with the highest (pEC_{50}=7.41) and lowest (pEC_{50}=6.37) potency in the mTOR signaling pathway.



842

843 **Figure 6. Extending a Canonical Pathway with PTMNavigator.** (**A**) Edited screenshot from

PTMNavigator showing an excerpt from the pathway diagram ‘Regulation of Actin Cytoskeleton – WP51’ and the modified peptides regulated in the Refametinib (*left*) and Mirdametinib (*right*) decryptM experiments. All regulations are downregulations. (**B**) Pathway graph of the Refametinib dataset generated by PHONEMeS. The kinases with decreased activity (determined by RoKAI + KSEA) were used as perturbation targets in the PHONEMeS input. (**C**) Proposed manual extension to the diagram ‘WP51’, created using the ‘Editing Mode’ within PTMNavigator. For illustrative purposes, nodes and edges added by hand are shown in green.

Supplementary Materials

Supplementary Table 1. Input examples for PTMNavigator.

Supplementary Figure 1. Number of pathways per organism and reference database that were imported into ProteomicsDB and are available in PTMNavigator.

Supplementary Figure 2. Mapping of Node Types from KEGG and WikiPathways to PTMNavigator.

Supplementary Figure 3. Mapping of Edge Types from KEGG and WikiPathways to PTMNavigator.

Supplementary Figure 4. Gene Identifier conversion for WikiPathways diagrams.

Supplementary Figure 5. Visualization of protein-level and peptide-level data in PTMNavigator.

Supplementary Figure 6. Results of PTM-SEA analysis of the Bekker-Jensen dataset.

Supplementary Figure 7. Common subnetworks of the high-scoring pathways in the Go-6983 dataset.

Supplementary Figure 8. PTMNavigator makes PTM-SEA and kinase activity scores transparent.

Supplementary Figure 9. Comparison of dose-dependent Dactolisib treatment in three different cell lines.

Supplementary Figure 10. decryptM curves from the Refametinib and Mirdametinib datasets.

References

1. Viéitez C*, et al.* High-throughput functional characterization of protein phosphorylation sites in yeast. *Nature Biotechnology*, 1-9 (2021).

2. Zecha J*, et al.* Decrypting drug actions and protein modifications by dose- and time-resolved proteomics. *Science*, eade3925 (2023).

3. Wiechmann S*, et al.* Chemical Phosphoproteomics Sheds New Light on the Targets and Modes of Action of AKT Inhibitors. *ACS Chem Biol* **16**, 631-641 (2021).

4. Ochoa D*, et al.* An atlas of human kinase regulation. *Molecular Systems Biology* **12**, 888 (2016).

5. Bodenmiller B*, et al.* Phosphoproteomic Analysis Reveals Interconnected System-Wide Responses to Perturbations of Kinases and Phosphatases in Yeast. *Science signaling* **3**, rs4 (2010).

6. Sharma K*, et al.* Ultradeep Human Phosphoproteome Reveals a Distinct Regulatory Nature of Tyr and Ser/Thr-Based Signaling. *Cell Reports* **8**, 1583-1594 (2014).

7. Humphrey SJ, Azimifar SB, Mann M. High-throughput phosphoproteomics reveals in vivo insulin signaling dynamics. *Nat Biotechnol* **33**, 990-995 (2015).

8. Gillespie M*, et al.* The reactome pathway knowledgebase 2022. *Nucleic Acids Research* **50**, D687-D692 (2022).

9. Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Research* **51**, D587-D592 (2023).

10. Martens M*, et al.* WikiPathways: connecting communities. *Nucleic Acids Research* **49**, D613-D621 (2021).

11. Rodchenkov I*, et al.* Pathway Commons 2019 Update: integration, analysis and exploration of pathway data. *Nucleic Acids Research* **48**, D489-D497 (2020).

12. Hornbeck PV, Zhang B, Murray B, Kornhauser JM, Latham V, Skrzypek E. PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Research* **43**, D512-D520 (2015).

13. Reimand J*, et al.* Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. *Nature Protocols* **14**, 482-517 (2019).

14. Subramanian A*, et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550 (2005).

15. Chicco D, Agapito G. Nine quick tips for pathway enrichment analysis. *PLOS Computational Biology* **18**, e1010348 (2022).

16. Kolberg L, Raudvere U, Kuzmin I, Adler P, Vilo J, Peterson H. g:Profiler-interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). *Nucleic Acids Res* **51**, W207-W212 (2023).

17. Krug K*, et al.* A Curated Resource for Phosphosite-specific Signature Analysis. *Molecular & Cellular Proteomics* **18**, 576-593 (2019).

18. Shannon P*, et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research* **13**, 2498-2504 (2003).

19. Villaveces JM, Jimenez RC, Habermann BH. KEGGViewer, a BioJS component to visualize KEGG Pathways. *F1000Research* **3**, (2014).

20. Kutmon M*, et al.* PathVisio 3: An Extendable Pathway Analysis Toolbox. *PLOS Computational Biology* **11**, e1004085 (2015).

21. Raaijmakers LM*, et al.* PhosphoPath: Visualization of Phosphosite-centric Dynamics in Temporal Molecular Networks. *Journal of Proteome Research* **14**, 4332-4341 (2015).

22. Gjerga E, Dugourd A, Tobalina L, Sousa A, Saez-Rodriguez J. PHONEMeS: Efficient Modeling of Signaling Networks Derived from Large-Scale Mass Spectrometry Data. *Journal of Proteome Research* **20**, 2138-2144 (2021).

23. Leeming MG*, et al.* Phosphomatics: interactive interrogation of substrate–kinase networks in global phosphoproteomics datasets. *Bioinformatics* **37**, 1635-1636 (2021).

24. Lautenbacher L*, et al.* ProteomicsDB: toward a FAIR open-source resource for life-science research. *Nucleic Acids Research* **50**, D1541-D1552 (2022).

25. Wilhelm M*, et al.* Mass-spectrometry-based draft of the human proteome. *Nature* **509**, 582-587 (2014).

26. Casado P*, et al.* Kinase-Substrate Enrichment Analysis Provides Insights into the Heterogeneity of Signaling Pathway Activation in Leukemia Cells. *Science Signaling* **6**, rs6-rs6 (2013).

27. Yılmaz S, Ayati M, Schlatzer D, Çiçek AE, Chance MR, Koyutürk M. Robust inference of kinase activity using functional networks. *Nature Communications* **12**, 1177 (2021).

28. Johnson JL*, et al.* An atlas of substrate specificities for the human serine/threonine kinome. *Nature* **613**, 759-766 (2023).

29. Yaron-Barir TM*, et al.* The intrinsic substrate specificity of the human tyrosine kinome. *Nature* **629**, 1174-1181 (2024).

30. Kuleshov MV*, et al.* KEA3: improved kinase enrichment analysis via data integration. *Nucleic Acids Res* **49**, W304-W316 (2021).

965 31. Crowl S, Jordan B, Ahmed H, Ma C, Naegle KM. KSTAR: An algorithm to predict patient-specific
966 kinase activities from phosphoproteomic data.). Systems Biology (2021).
967
968 32. Bekker-Jensen DB*, et al.* Rapid and site-specific deep phosphoproteome profiling by data-
969 independent acquisition without the need for spectral libraries. *Nature Communications* **11**, 787
970 (2020).
971
972 33. Ochoa D*, et al.* The functional landscape of the human phosphoproteome. *Nature Biotechnology*
973 **38**, 365-373 (2020).
974
975 34. Lee CY*, et al.* Illuminating phenotypic drug responses of sarcoma cells to kinase inhibitors by
976 phosphoproteomics. *Mol Syst Biol* **20**, 28-55 (2024).
977
978 35. Lohi H*, et al.* Novel glycogen synthase kinase 3 and ubiquitination pathways in progressive
979 myoclonus epilepsy. *Hum Mol Genet* **14**, 2727-2736 (2005).
980
981 36. Zhan J, Chitta RK, Harwood FC, Grosveld GC. Phosphorylation of TSC2 by PKC-delta reveals a
982 novel signaling pathway that couples protein synthesis to mTORC1 activity. *Mol Cell Biochem*
983 **456**, 123-134 (2019).
984
985 37. MacNicol MC, Muslin AJ, MacNicol AM. Disruption of the 14-3-3 binding site within the B-Raf
986 kinase domain uncouples catalytic activity from PC12 cell differentiation. *J Biol Chem* **275**, 3803-
987 3809 (2000).
988
989 38. Saha M*, et al.* RSK phosphorylates SOS1 creating 14-3-3 docking sites and negatively regulating
990 MAPK activation. *The Biochemical journal* **447**, 159-166 (2012).
991
992 39. Sun Q, Wang W. Structures of BRAF–MEK1–14-3-3 sheds light on drug discovery. *Signal*
993 *Transduction and Targeted Therapy* **4**, 1-2 (2019).
994
995 40. Klaeger S*, et al.* The target landscape of clinical kinase drugs. *Science*, 44 (2017).
996
997 41. Cohen P, Cross D, Jänne PA. Kinase drug discovery 20 years after imatinib: progress and future
998 directions. *Nature Reviews Drug Discovery*, 1-19 (2021).
999
1000 42. Bayer FP, Gander M, Kuster B, The M. CurveCurator: a recalibrated F-statistic to assess, classify,
1001 and explore significance of dose-response curves. *Nat Commun* **14**, 7902 (2023).
1002
1003 43. Castoldi R*, et al.* TetraMabs: simultaneous targeting of four oncogenic receptor tyrosine kinases
1004 for tumor growth inhibition in heterogeneous tumor cell populations. *Protein Eng Des Sel* **29**,
1005 467-475 (2016).
1006
1007 44. Stuart SA, Houel S, Lee T, Wang N, Old WM, Ahn NG. A Phosphoproteomic Comparison of B-
1008 RAFV600E and MKK1/2 Inhibitors in Melanoma Cells. *Molecular & Cellular Proteomics : MCP* **14**,
1009 1599-1615 (2015).
1010
1011 45. Reinecke M*, et al.* Chemoproteomic Selectivity Profiling of PIKK and PI3K Kinase Inhibitors. *ACS*
1012 *Chem Biol* **14**, 655-664 (2019).

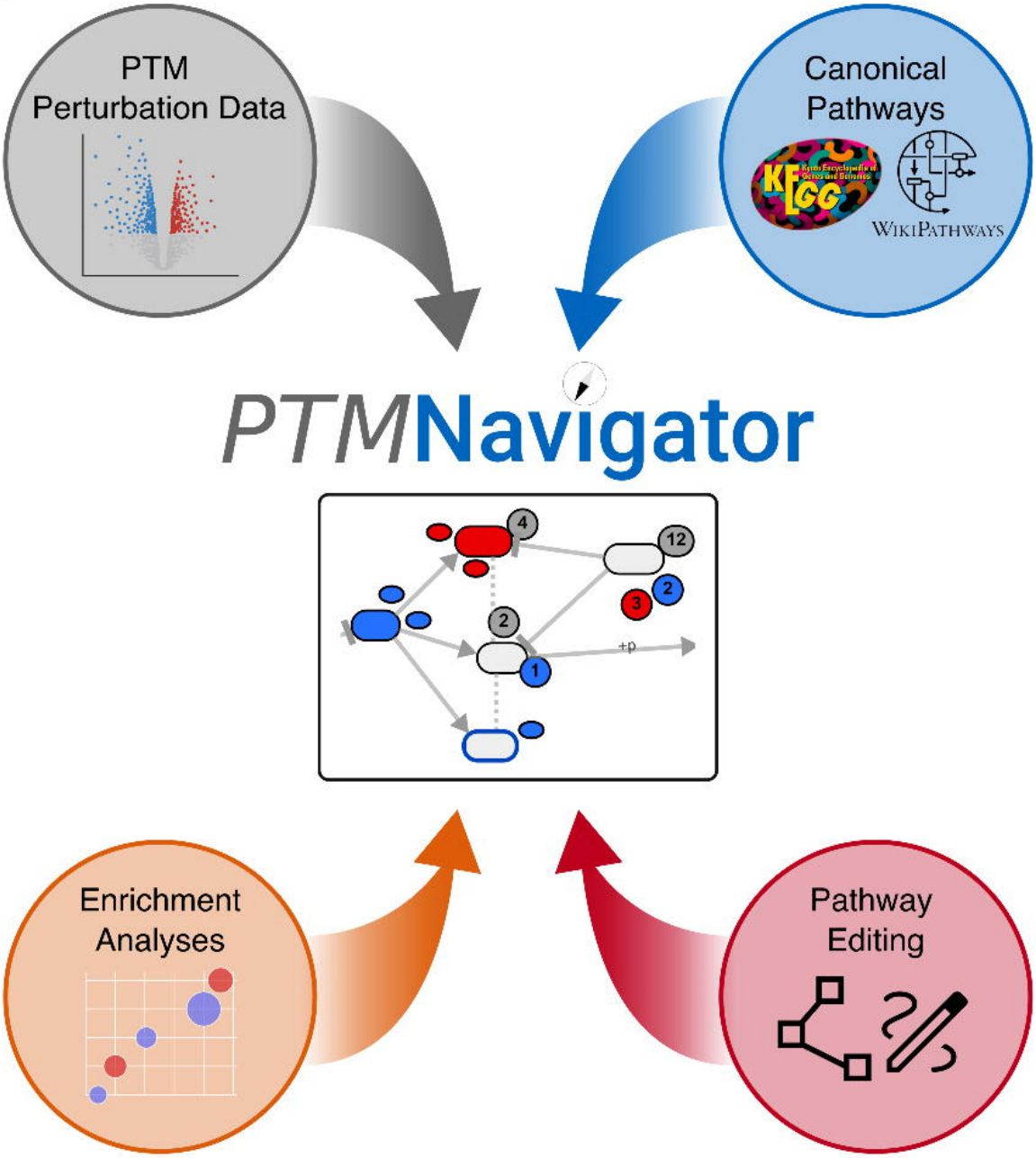
46. Cargnello M, Roux PP. Activation and function of the MAPKs and their substrates, the MAPK-activated protein kinases. *Microbiol Mol Biol Rev* **75**, 50-83 (2011).

47. Dhillon AS, Pollock C, Steen H, Shaw PE, Mischak H, Kolch W. Cyclic AMP-dependent kinase regulates Raf-1 kinase mainly by phosphorylation of serine 259. *Mol Cell Biol* **22**, 3237-3246 (2002).

48. Karp PD*, et al.* The BioCyc collection of microbial genomes and metabolic pathways. *Brief Bioinform* **20**, 1085-1093 (2019).

49. Bostock M, Ogievetsky V, Heer J. D^3 Data-Driven Documents. *IEEE Transactions on Visualization and Computer Graphics* **17**, 2301-2309 (2011).

A



B

Type	Algorithm Name	Publication
Gene Set Enrichment	ssGSEA - Gene-Centric	Krug et al.
Gene Set Enrichment	ssGSEA - Gene-Centric-Redundant	Krug et al.
PTM Set Enrichment	PTM-SEA	Krug et al.
Kinase Activity Inference	KSEA	Casado et al.
Kinase Activity Inference	RoKAI + KSEA	Yılmaz et al.
Kinase Activity Inference	KSTAR	Crowl et al.
Kinase Activity Inference	KEA	Kuleshov et al.
Motif Analysis	Motif Enrichment	Johnson et al. / Inhouse
Pathway Inference	PHONEMeS*	Gjerga et al.

A

Data Selection

PROTEOMICSDB DATA | USER DATA

Session ID
CC4250556015E2100DBA6ABE3E26C700

32 / 32

Uploaded Datasets
Dataset1_PTM Dataset1_FP

Load Dataset(s) Clear Data

1

Experiment Filter 2

Experiment 1 Experiment 2
Experiment 3 Experiment 4

Pathway Selection 3

CANONICAL PATHWAYS | CUSTOM PATHWAYS

Organism:
Homo sapiens

Pathways (n=1123):

Filter by Protein List

Enter gene names or Uniprot accession numbers to show only the pathways that contain them.

VIEWING MODE | EDITING MODE 4

5

Gene/Protein, Pathway, Group, Metabolite/Compound, Activation, Inhibition, Other, Indirect, Binding/Association, Up-, Down-, Not regulated, +p/-p (De-)Phosphorylation, Decreased Kinase Activity, Kinase-Substrate Relationship

Compound X, Protein A1, Protein A2, Protein C, Protein D, T142, Protein E, Protein F, 6

7

Protein D

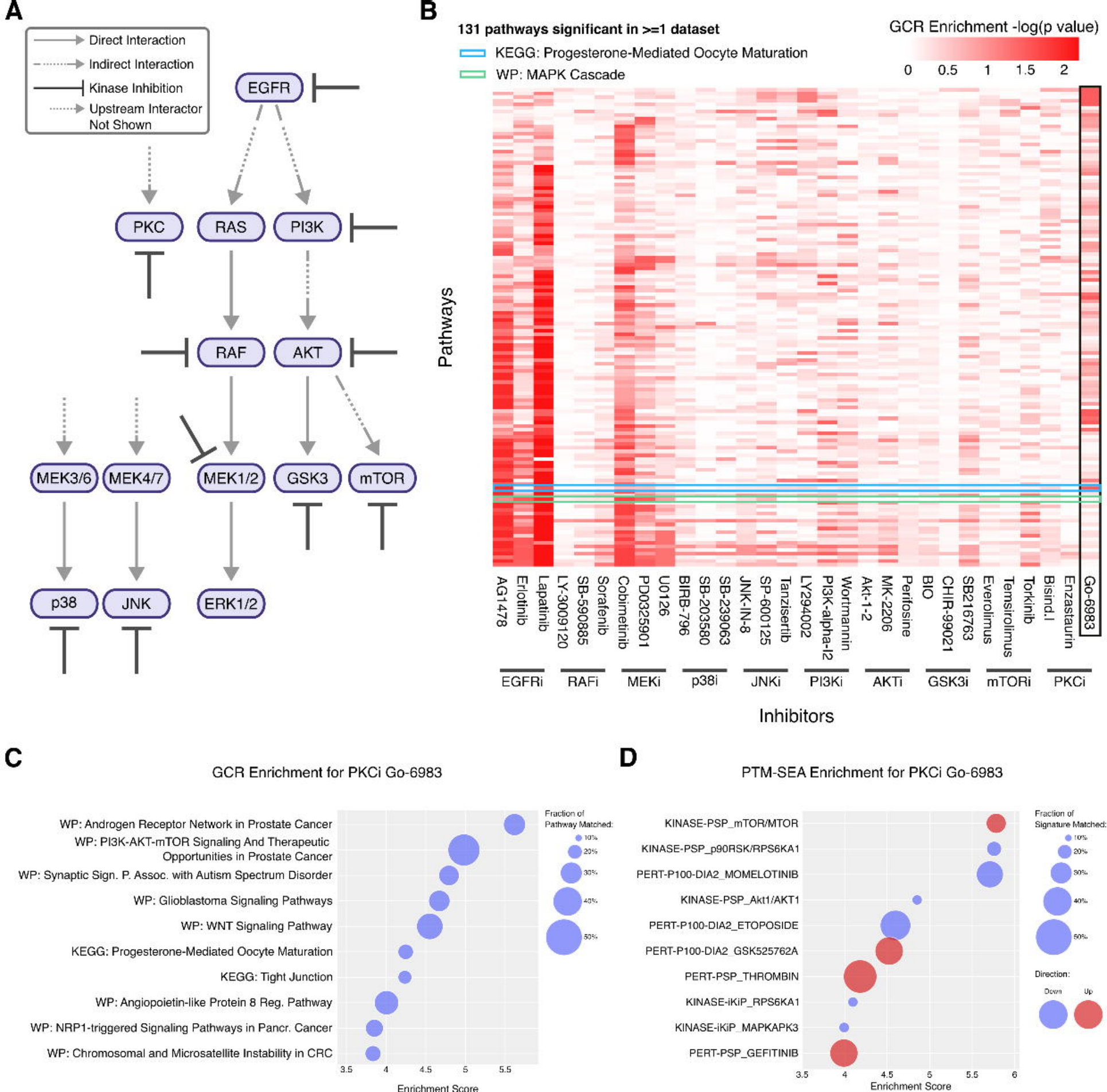
Regulation:	down
Modified Sequence:	TES(ph)TPEPT(ph)IDEK
Fold Change:	0.2
p-Value:	0.000012
PhosphoSitePlus®:	T142-p
Experiment Name:	My Experiment

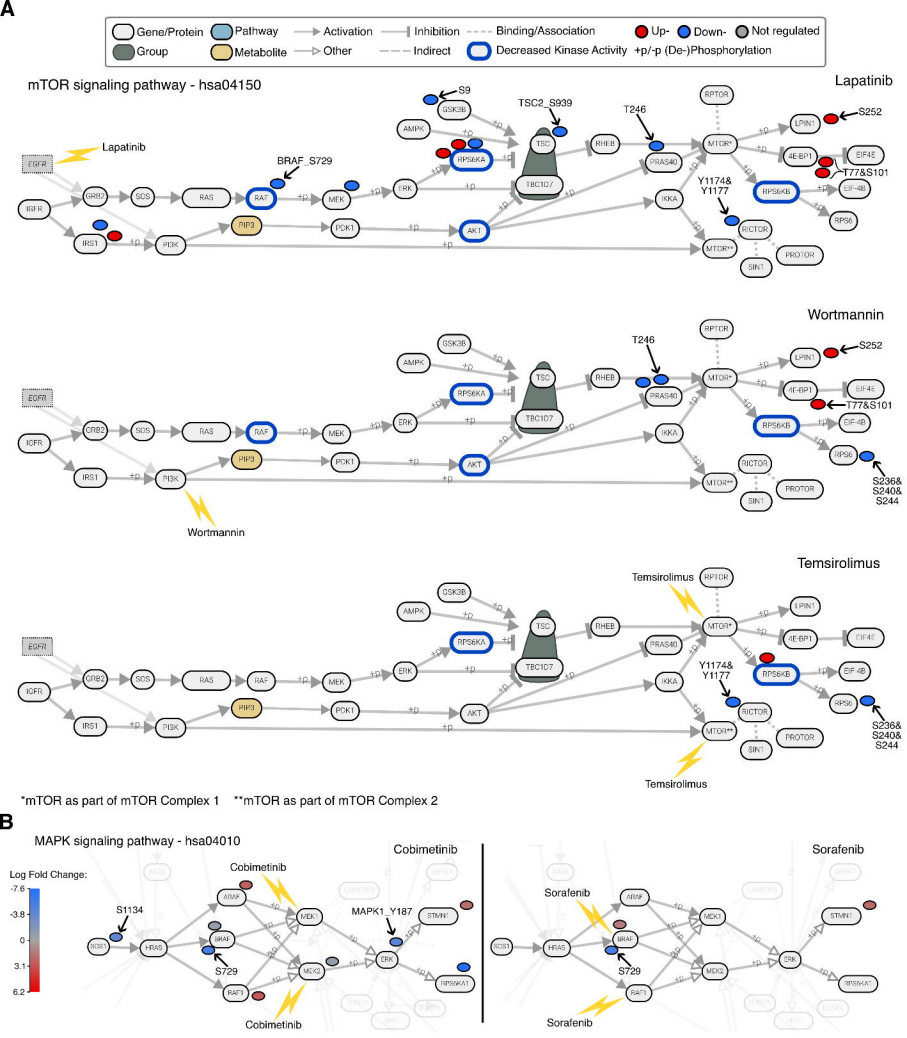
B

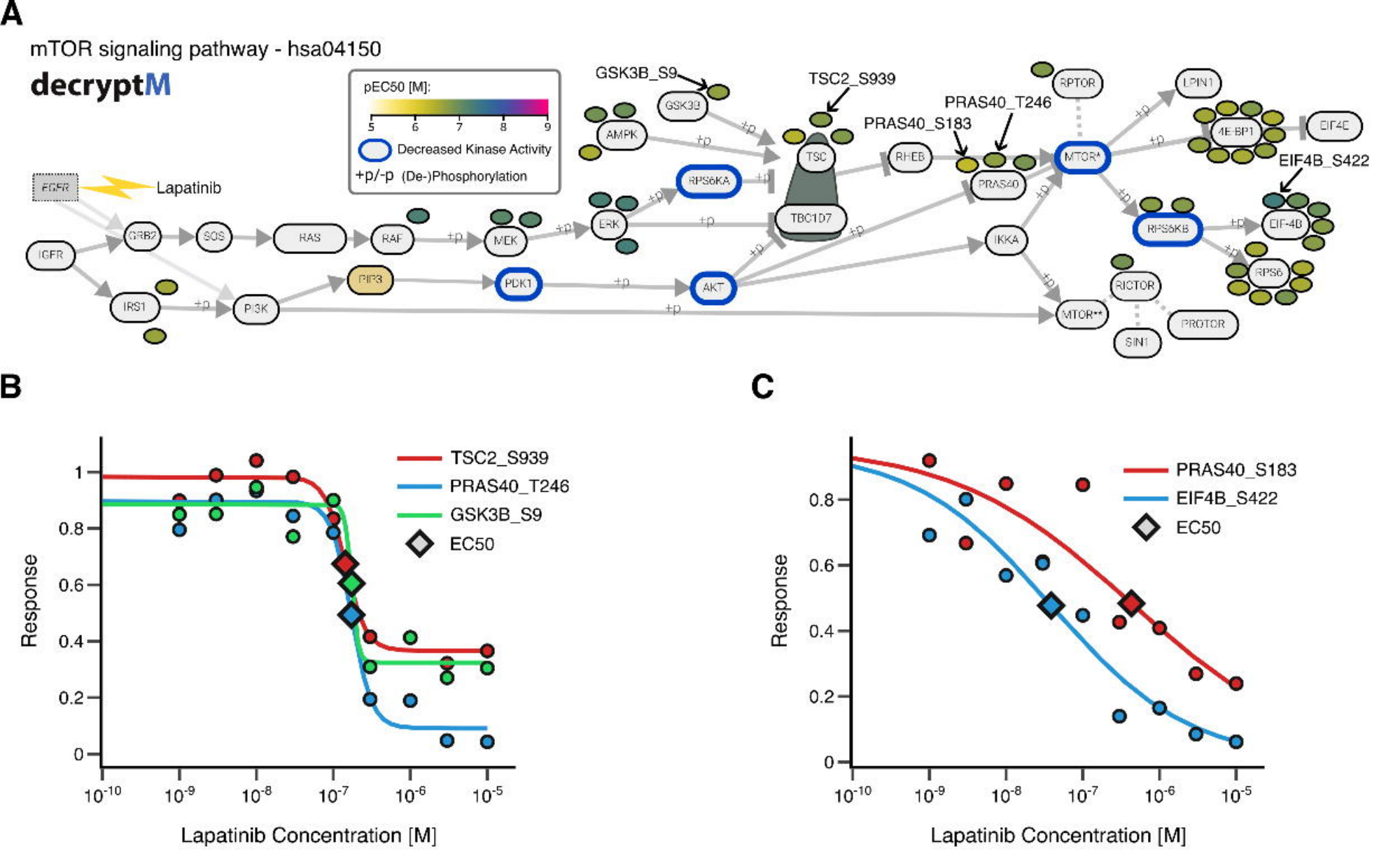
Enrichment Analyses:

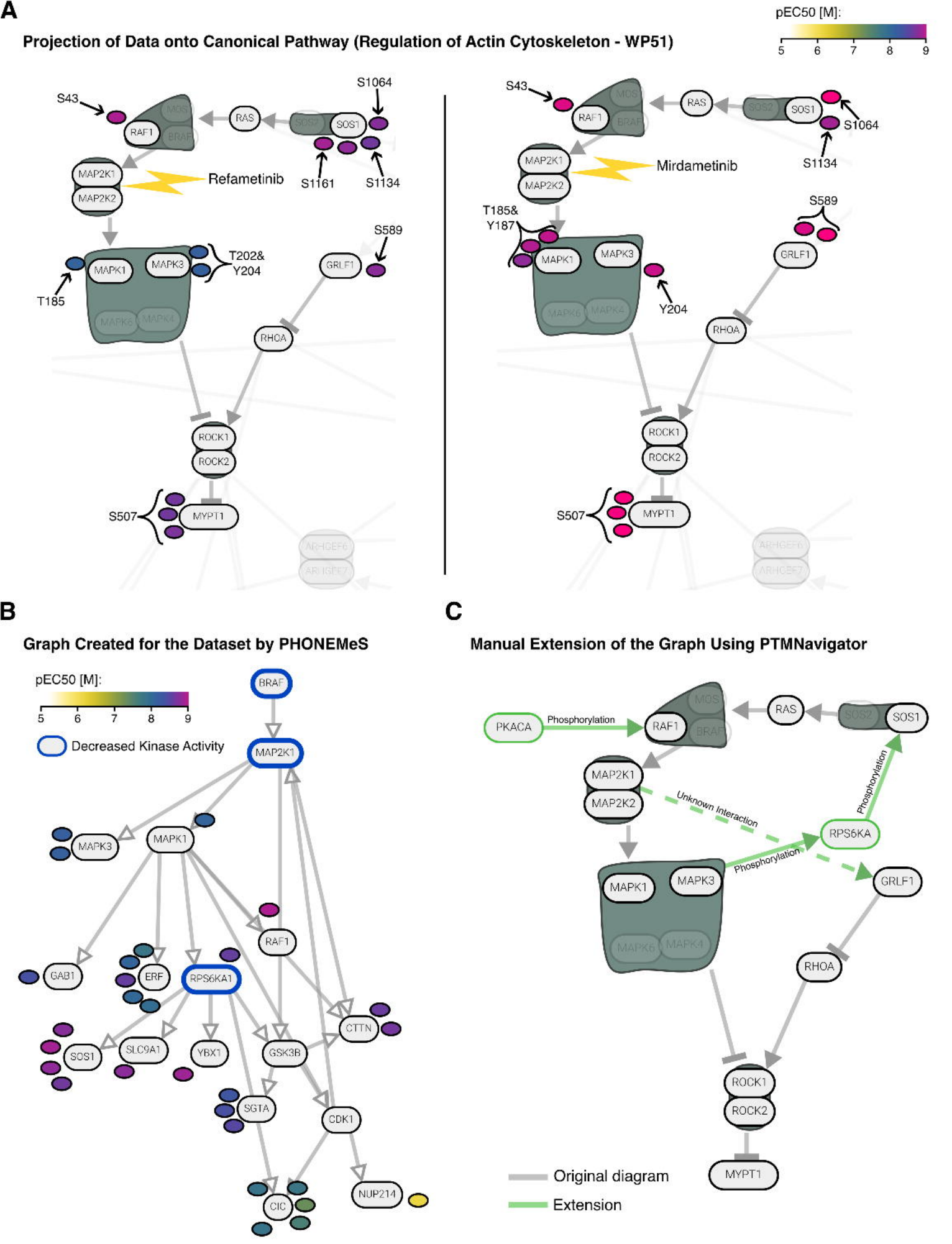
PTM-SEA | PEA GENE-CENTRIC | KSEA | KSEA WITH ROKAI

Gene	Score (Experiment 1)	adj p-val (Experiment 1)	Score (Experiment 2)	adj p-val (Experiment 2)
Protein C	-3.23	0.01	-3.12	0.01
Protein X	-2.99	0.01	-2.61	0.03
Protein Y	-2.60	0.01	-2.60	0.09


A
Direct Interaction
Indirect Interaction
Kinase Inhibition
Upstream Interactor Not Shown
EGFR
PKC
RAS
PI3K
RAF
AKT
MEK3/6
MEK4/7
MEK1/2
GSK3
mTOR
p38
JNK
ERK1/2
B
131 pathways significant in >=1 dataset
KEGG: Progesterone-Mediated Oocyte Maturation
WP: MAPK Cascade
GCR Enrichment -log(p value)
0
0.5
1
1.5
2
Pathways
AG1478
Erlotinib
Lapatinib
LY-3009120
SB-590885
Sorafenib
Cobimetinib
PD0325901
U0126
BIRB-796
SB-203580
SB-239063
JNK-IN-8
SP-600125
Tanzisertib
LY294002
PI3K-alpha-I2
Wortmannin
Akt-1-2
MK-2206
Perifosine
BIO
CHIR-99021
SB216763
Everolimus
Temsirolimus
Torkinib
Bisind.I
Enzastaurin
Go-6983
EGFRi
RAFi
MEKi
p38i
JNKi
PI3Ki
AKTi
GSK3i
mTORi
PKCi
Inhibitors
C
GCR Enrichment for PKCi Go-6983
WP: Androgen Receptor Network in Prostate Cancer
WP: PI3K-AKT-mTOR Signaling And Therapeutic Opportunities in Prostate Cancer
WP: Synaptic Sign. P. Assoc. with Autism Spectrum Disorder
WP: Glioblastoma Signaling Pathways
WP: WNT Signaling Pathway
KEGG: Progesterone-Mediated Oocyte Maturation
KEGG: Tight Junction
WP: Angiopoietin-like Protein 8 Reg. Pathway
WP: NRP1-triggered Signaling Pathways in Pancr. Cancer
WP: Chromosomal and Microsatellite Instability in CRC
3.5
4
4.5
5
5.5
Enrichment Score
Fraction of Pathway Matched:
10%
20%
30%
40%
50%
D
PTM-SEA Enrichment for PKCi Go-6983
KINASE-PSP_mTOR/MTOR
KINASE-PSP_p90RSK/RPS6KA1
PERT-P100-DIA2_MOMELOTINIB
KINASE-PSP_Akt1/AKT1
PERT-P100-DIA2_ETOPOSIDE
PERT-P100-DIA2_GSK525762A
PERT-PSP_THROMBIN
KINASE-iKiP_RPS6KA1
KINASE-iKiP_MAPKAPK3
PERT-PSP_GEFITINIB
3.5
4
4.5
5
5.5
6
Enrichment Score
Fraction of Signature Matched:
10%
20%
30%
40%
50%
Direction:
Down
Up


A
Gene/Protein
Pathway
Activation
Inhibition
Binding/Association
Up-
Down-
Not regulated
Group
Metabolite
Other
Indirect
Decreased Kinase Activity
+p/-p (De-)Phosphorylation
mTOR signaling pathway - hsa04150
Lapatinib
Wortmannin
Temsirolimus
*mTOR as part of mTOR Complex 1
**mTOR as part of mTOR Complex 2
B
MAPK signaling pathway - hsa04010
Cobimetinib
Sorafenib
Log Fold Change:
-7.6
-3.8
0
3.1
6.2
S1134
S729
MAPK1_Y187
S9
TSC2_S939
T246
S252
T77&S101
Y1174&
Y1177
S236&
S240&
S244
EGFR
IGFR
GRB2
SOS
RAS
RAF
MEK
ERK
IRS1
PI3K
PIP3
PDK1
AKT
AMPK
GSK3B
RPS6KA
TSC
TBC1D7
RHEB
PRAS40
MTOR*
RPTOR
LPIN1
4E-BP1
EIF4E
EIF-4B
RPS6KB
RPS6
IKKA
RICTOR
MTOR**
SIN1
PROTOR
HRAS
BRAF
MEK1
MEK2
RAF1
STMN1
RPS6KA1


A
mTOR signaling pathway - hsa04150
decryptM
pEC50 [M]:
5
6
7
8
9
Decreased Kinase Activity
+p/-p (De-)Phosphorylation
Lapatinib
EGFR
IGFR
GRB2
SOS
RAS
RAF
MEK
ERK
IRS1
PI3K
PIP3
PDK1
AKT
AMPK
GSK3B_S9
GSK3B
RPS6KA
TSC
TBC1D7
TSC2_S939
RHEB
PRAS40_S183
PRAS40_T246
PRAS40
IKKA
MTOR*
MTOR**
RPTOR
RICTOR
SIN1
PROTOR
LPIN1
4E-BP1
EIF4E
EIF4B_S422
RPS6KB
EIF-4B
RPS6
B
TSC2_S939
PRAS40_T246
GSK3B_S9
EC50
Response
Lapatinib Concentration [M]
C
PRAS40_S183
EIF4B_S422
EC50
Response
Lapatinib Concentration [M]


A
Projection of Data onto Canonical Pathway (Regulation of Actin Cytoskeleton - WP51)
pEC50 [M]:
5
6
7
8
9
S43
RAF1
MOS
BRAF
RAS
SOS2
SOS1
S1064
S1161
S1134
MAP2K1
MAP2K2
Refametinib
MAPK1
MAPK3
T202&
Y204
T185
MAPK6
MAPK4
S589
GRLF1
RHOA
ROCK1
ROCK2
S507
MYPT1
ARHGEF6
ARHGEF7
S43
RAF1
MOS
BRAF
RAS
SOS2
SOS1
S1064
S1134
MAP2K1
MAP2K2
Mirdametinib
T185&
Y187
MAPK1
MAPK3
Y204
MAPK6
MAPK4
S589
GRLF1
RHOA
ROCK1
ROCK2
S507
MYPT1
ARHGEF6
ARHGEF7
B
Graph Created for the Dataset by PHONEMeS
pEC50 [M]:
5
6
7
8
9
Decreased Kinase Activity
BRAF
MAP2K1
MAPK3
MAPK1
RAF1
GAB1
ERF
RPS6KA1
CTTN
SOS1
SLC9A1
YBX1
GSK3B
SGTA
CDK1
CIC
NUP214
C
Manual Extension of the Graph Using PTMNavigator
PKACA
Phosphorylation
RAF1
MOS
BRAF
RAS
SOS2
SOS1
MAP2K1
MAP2K2
Unknown Interaction
Phosphorylation
RPS6KA
Phosphorylation
MAPK1
MAPK3
MAPK6
MAPK4
GRLF1
RHOA
ROCK1
ROCK2
MYPT1
Original diagram
Extension