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

Julian Müller[1], Florian P. Bayer[1], Mathias Wilhelm[2], 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.

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

Abstract

Post-translational modifications (PTMs) play a governing role 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 on top of ~3000 well-curated pathway diagrams. 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 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 extending existing pathways by uncovering previously unknown relationships between PTMs and pathway components. By facilitating the discovery of novel PTM-pathway interactions and enhancing our understanding of cellular signaling dynamics, 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 PTMs 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], for example using online tools such as g:Profiler[14]. Multiple PEA algorithms exist, including Gene Set Enrichment Analysis (GSEA)[15], all of which consist of the same two steps: First, the list of perturbed genes is compared to the set of genes contained in each canonical pathway; Second, pathways for which the overlap is larger than expected by random chance are identified by some statistical test, using a background gene list to compare against. If possible, it is advised not to use a whole genome as a background, but only the subset of genes expressed in the studied tissue, which makes the comparison fairer. As a result, all PEAs output a list of pathway names together with their statistical significance[16]. While very useful for the analysis of proteome expression changes, this approach has several shortcomings in the context of PTM analysis. First, in a PEA, a pathway is regarded simply as a set of genes, and it is not considered that some of them might be more important to the activity of the pathway than others. Second, all genes that are part of the pathway are regarded the same no matter if they are negatively or positively correlated with the activity of the pathway. Third, investigators do not see whether the regulations are randomly distributed across the pathway (which might imply that the significance is over-estimated) or whether there is a discernible flow of signal. Finally, PEA requires a set of genes or proteins as input, therefore it cannot be applied to PTM sites directly. A simple workaround is to aggregate PTMs on the protein level. However, sites even on the same protein often do not agree in their response questioning the approach in principle. Currently, the only tool that partly addresses this problem is PTM-signature enrichment

analysis (PTM-SEA), which uses a specifically engineered database of associations between sites of PTMs and perturbations, pathways, or kinase activities (PTM signatures database; PTMsigDB)[17]. This constitutes a valuable orthogonal approach to conventional PEA. However, PTM-SEA so far only covers a fraction of the pathways deposited in KEGG and WikiPathways, and has the same shortcoming as other PEA tools in that users only obtain the names of the putatively regulated pathways, rather than how their data maps to the nodes in the pathways.

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.

To overcome the aforementioned problems of pathway enrichment on the PTM level and to empower researches to quickly examine pathway engagement via PTM-regulation, we introduce PTMNavigator, a web application that projects PTM perturbation datasets onto canonical pathways in the form of interactive graphs. 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. The core visualization component of our software, which we termed *biowc-pathwaygraph*, can be easily reused in web interfaces outside of ProteomicsDB. We demonstrate the utility of our application using two phosphoproteomic drug profiling studies[2, 26]. First, we show how PTMNavigator improves the interpretation of PEA by enabling users to explore why the top-

ranked pathways are reported as significant and whether the enrichment is systematic or random (i.e., the differentially expressed peptides are not connected within the pathway diagram). 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 Visualization Interface for Pathways

PTMNavigator is a web-based graphical interface to visualize PTMs from perturbation experiments within pathway diagrams. 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 (Fig. 1A). 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 whether or not it was up- or down-regulated between experimental conditions. The second type of input are the canonical pathways of 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 (Fig. 1B).

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. 1C) 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. The available pathways are sorted by their relevance derived from the pathway enrichment tool g:Profiler. The pathway diagram 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. 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.[27] into ProteomicsDB.

The final pathway diagrams can be downloaded as scalable vector graphics (SVG). In the following, we exemplify the capabilities and benefits of PTMNavigator using two recent PTM perturbation datasets.

Interpreting Pathway Enrichment Analysis Results with 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 were profiled for their effect on the phosphoproteome of a human retinal pigment epithelial cell line (RPE1)[26]. 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 g:Profiler resulted in enrichment scores (negative log-transformed p-values; ES) for each combination of drug and pathway (Fig. 2A). Some results confirmed prior knowledge on the compounds, e.g. AG-1478 and Lapatinib, both targeting EGFR, were strongly enriched for 'WP437: EGF/EGFR signaling pathway' (ES=8.2 and ES=3.8). Yet, for many drug perturbations, the results were not as easy to interpret. For example, the JNK inhibitor SP-600125 showed a significant enrichment for 'WP2380 - Brain derived Neurotrophic Factor (BDNF) signaling pathway' (ES=3.1, which corresponds to a p-value<0.001) and the same was observed for 9 other drugs. Another pathway

involving BDNF, 'WP3676: BDNF-TrkB signaling' also achieved a high enrichment score in the SP-600125 data (Fig. 2B). However, BDNF signaling is unlikely to be of relevance for RPE1 cells. 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 SP-600125 (Fig. 2B). The projection of the experimentally measured phosphopeptides onto 'WP2380 - Brain derived Neurotrophic Factor (BDNF) signaling pathway' (Supplementary Fig. 6) showed that there was, in fact, no regulation in signaling that could be attributed to altered activity of BDNF. Instead, the pathway scored high because the diagram includes many common signaling cascades that can not only be triggered by BDNF but also by other growth factors such as EGF. Most proteins with regulations, such as SRC and SHC1, are known to be part of the canonical EGFR signaling pathway or pathways downstream of it[28, 29]. The analysis also highlighted a group of regulations downstream of AKT1, notably on TSC2 (down), mTOR, EIF4EBP1, and FOXO3 (all up). Many other high-scoring pathways for SP-600125 (ranks 3-7 and 9) shared this signaling cascade or a variation of it (Fig. 2C; Supplementary Fig. 8-12, 14). The other three pathways (Supplementary Fig. 7, 13, 15) do not include this subnetwork, but two of them contain a shared set of regulations as well (Supplementary Fig. 13, 15). Other differentially phosphorylated proteins that were found in multiple of the high-scoring pathways included IRS1, GAB2, and AKT1S1, all of which are known actors in EGFR and AKT signaling[30-33]. It has been shown before that JNK, the target of SP-600125, phosphorylates IRS1 (insulin receptor substrate 1), which causes a negative feedback loop that inhibits AKT signaling[30, 31]. The inhibition of JNK by SP-600125 in RPE1 cells might thus have resulted in an activation of AKT1 signaling through this mechanism. The visualizations of PTMNavigator thus align well with prior knowledge on the drug under

consideration. While the set of enriched pathways seemed diverse at first, most regulations could be attributed to a single altered signaling pattern likely caused by JNK inhibition.

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 (Fig. 3A). Projection of their perturbed PTM peptides on 'hsa04150 - mTOR signaling pathway' (Fig. 3C), 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. 3C, 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[34-36]. 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. 3C, middle panel). Likewise, Temsirolimus regulated even fewer sites and all were downstream of mTOR (Fig. 3C, 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), which confirms 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. 3B). 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 had a wider impact than Sorafenib, which was visible in PTMNavigator's

visualization of 'hsa04010 – MAPK Signaling Pathway' (Fig. 3D). 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[37]. 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 successful BRAF-dimerization[38]. 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[39, 40]. 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] (Fig. 4A). 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 or CurveCurator[41] 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 (mechanism of action shown in Fig. 4B).

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 'hsa04150 – mTOR signaling pathway' (Fig. 4C) 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. 3C, 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. 4D), 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. 4E). 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[39], 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[42]. 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 (Supplementary Fig. 4).

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 significantly enriched for both datasets, the diagram 'WP51 – Regulation of actin cytoskeleton' depicted the signaling around the compounds' targets most comprehensively. Projection of the drug-regulated peptides onto this pathway using the 'potency' color scheme (Fig. 5A) showed that Mirdametinib regulates MEK signaling more potently than Refametinib. This is in agreement with the reported higher affinity observed in pulldown experiments[39].

By focusing on dose-dependent phosphorylation signals in common between both drugs, we obtain 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]. Less straightforward to interpret were several downregulations upstream of MEK at SOS1 sites as well as RAF1_S43 in both datasets. RAF1_S43 is a PKACA substrate[43], suggesting that the activity of this kinase (which is not visible in the pathway) was also perturbed by the treatment. 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). Among the regulated SOS1 sites are SOS1_S1134 and SOS1_S1161, which are substrates of p90RSK[37].

The activity of this protein is regulated by ERK, therefore this regulation likely is the result of an inhibited feedback loop. SOS1_S1064 was also observed less phosphorylated in both experiments. This site has no annotated kinase, but the other regulations suggest it could also be a p90RSK substrate. 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) and Vemurafenib (which has MEK5 as a potent off-target[39, 44]). 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 (Fig. 5B and Supplementary Fig. 5) suggests that this site could also be part of the general PTM signature of MEK inhibition.

The visualization also revealed another regulation in the 'WP51 – Regulation of actin cytoskeleton' pathway that was present in both datasets, namely GRLF_S589, which has no functional annotation or putative upstream kinase in PhosphoSitePlus. It showed similar dose-response behavior as the signature MEK inhibition sites on MAPK1, MAPK3, and also MYPT1_S507 (Fig. 5C and Supplementary Fig. 5). This suggests that the site is also part of the same phosphoproteomic signature. The pathway diagram shows no perturbation upstream of GRLF1, and, unlike for SOS1, there is no known feedback loop which could explain this observation. 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. Nevertheless, this shows that PTMNavigator can support researchers in the functionalization of uncharacterized PTM sites and thus in the extension of pathways by combining prior knowledge with experimental data (Fig. 5D).

Discussion

PTMNavigator is a tool for visualizing quantitative PTM data in the context of cellular pathways which supports user 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 the integration with PhosphoSitePlus adds another layer of functional annotation. 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.

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. In the future, we envision to incorporate more PTM annotations e.g. substrate motifs and predicted kinase-substrate using tools such as the Kinase Library[45] to further improve and streamline PTM data analysis. A combination with Kinase Activity Inference tools such as KSTAR[46], INKA[47], or RoKAI[48] also presents an opportunity for future enhancements.

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 'WP51 – Regulation of actin cytoskeleton' (Fig. 5A), would have worked just as well with the eponymous KEGG diagram 'hsa04810 – Regulation of actin cytoskeleton', as it contains the same subnetwork of SOS1, RAF1, MAPK1/3, GRLF1, and MYPT1.

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). 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. For example, 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. 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.

As alluded to in our final example, a desirable feature is the visualization of user-defined pathways that are either variations of KEGG/WikiPathways diagrams or entirely new pathways. The *biowc-*

pathwaygraph package already enables users to create their own signaling networks and project their data onto it but currently requires hosting one's own web server. We plan to make this feature more convenient in the future by allowing users to create, edit and load pathway diagrams within PTMNavigator. We expect that features like this will further enhance the versatility of our software and its use by the scientific community to interpret PTM data in the context of their respective pathways.

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 (Supplementary Fig. 1). 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 (Supplementary Fig. 2):

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 (Supplementary Fig. 1). 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 (Supplementary Fig. 3). 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 (Supplementary Fig. 3). 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 biowc-pathwaygraph

We developed a WebComponent (www.webcomponents.org) termed *biowc-pathwaygraph* for the dynamic rendering of pathway diagrams together with peptide-level experimental data. *biowc-pathwaygraph* was implemented using the Lit library (https://lit.dev/), Version 2.0.2. Our WebComponent 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.

biowc-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, *biowc-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. 1C; users can toggle between the two representations).

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.

Implementation of PTMNavigator

We implemented PTMNavigator as a Vue.js component (https://vuejs.org). It is essentially a wrapper around *biowc-pathwaygraph* that facilitates interaction of the WebComponent with both the user and the backend of ProteomicsDB. 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 *biowc-pathwaygraph*.

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.

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. During the upload, a pathway

enrichment analysis is performed automatically so that each pathway in our database is associated with an enrichment score (negative log-transformed p-value) for the uploaded dataset. The uploaded data is stored temporarily in ProteomicsDB and is deleted if not accessed for 14 days. 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.

Pathway Enrichment Analysis

Throughout this study, pathway enrichment analysis was performed as follows: All regulated peptides or proteins of a dataset were mapped to Uniprot accession numbers. The set of all unique Uniprot accession numbers was then sent as a POST request to the API of g:Profiler (https://biit.cs.ut.ee/gprofiler/api/gost/profile/), version e109_eg56_p17_773ec798 (accessed June 16, 2023). The 'organism' parameter was set to 'hsapiens' in all datasets discussed in the main text (in the case of the automatic PEA during data upload, it is set to the organism selected by the user). The 'sources' parameter was set to '["KEGG","WP"]' to only receive enrichment scores for KEGG and WikiPathways entries. The 'all_results' parameter was set to 'true' to retrieve also insignificant enrichment results. All others parameters were set to their default values.

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[27].

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.

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.). The following files are available on Zenodo (https://doi.org/10.5281/zenodo.10185633):

- For the Bekker-Jensen data:
  - Perseus output (.txt)
  - A Jupyter Notebook describing how to convert the Perseus output into PTMNavigator input (.ipynb)
  - PTMNavigator input (.xlsx)
  - Pathway enrichment analysis results (.xlsx)
- For the Zecha data:
  - CurveCurator parameter files (.toml)
  - CurveCurator output (.xlsx)
  - Pathway enrichment analysis results (.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. *biowc-pathwaygraph* is available as a JavaScript module here: www.npmjs.com/package/biowc-pathwaygraph, and 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 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

Data Analysis: JM, MT, FB

Writing – original draft: JM

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

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.

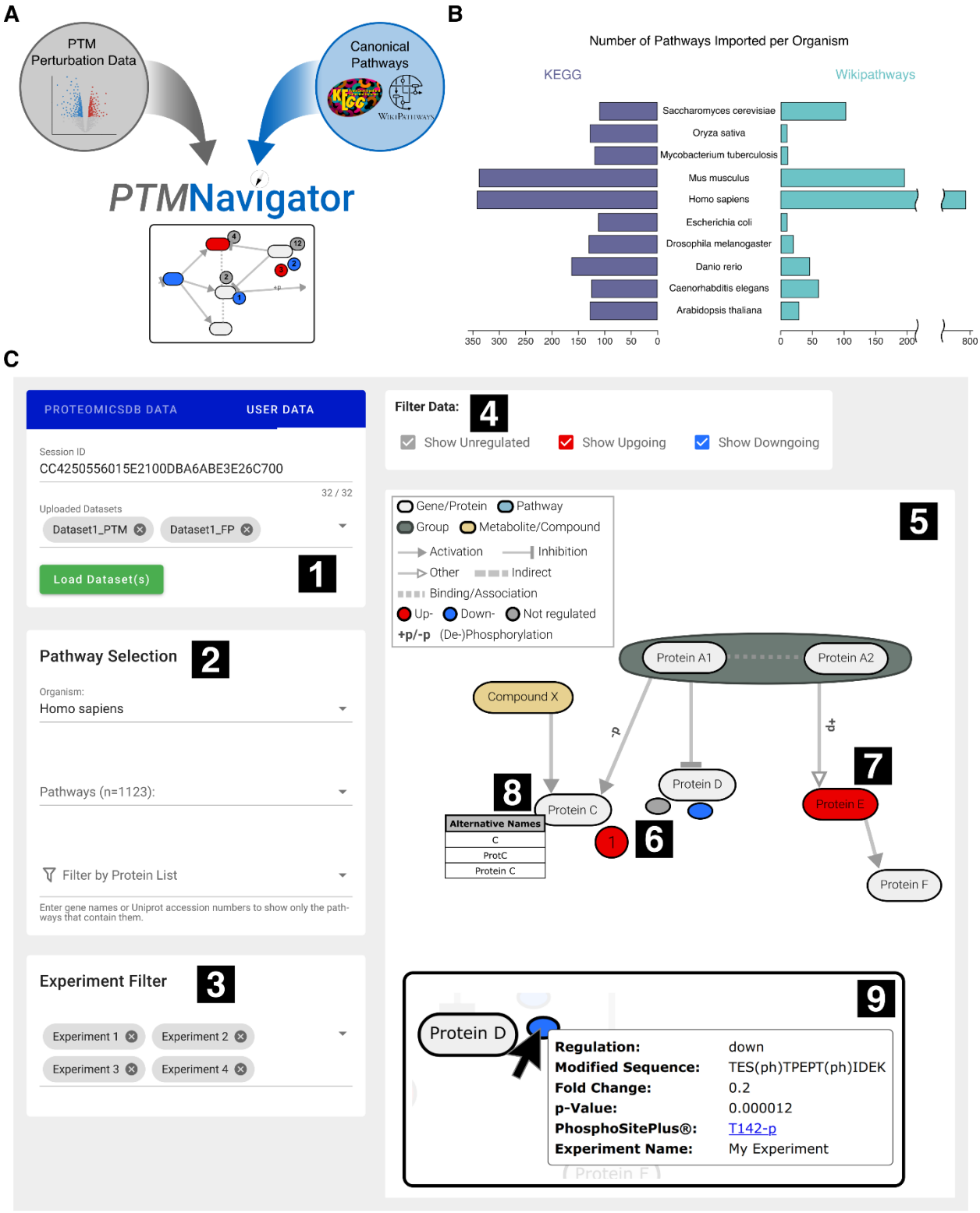
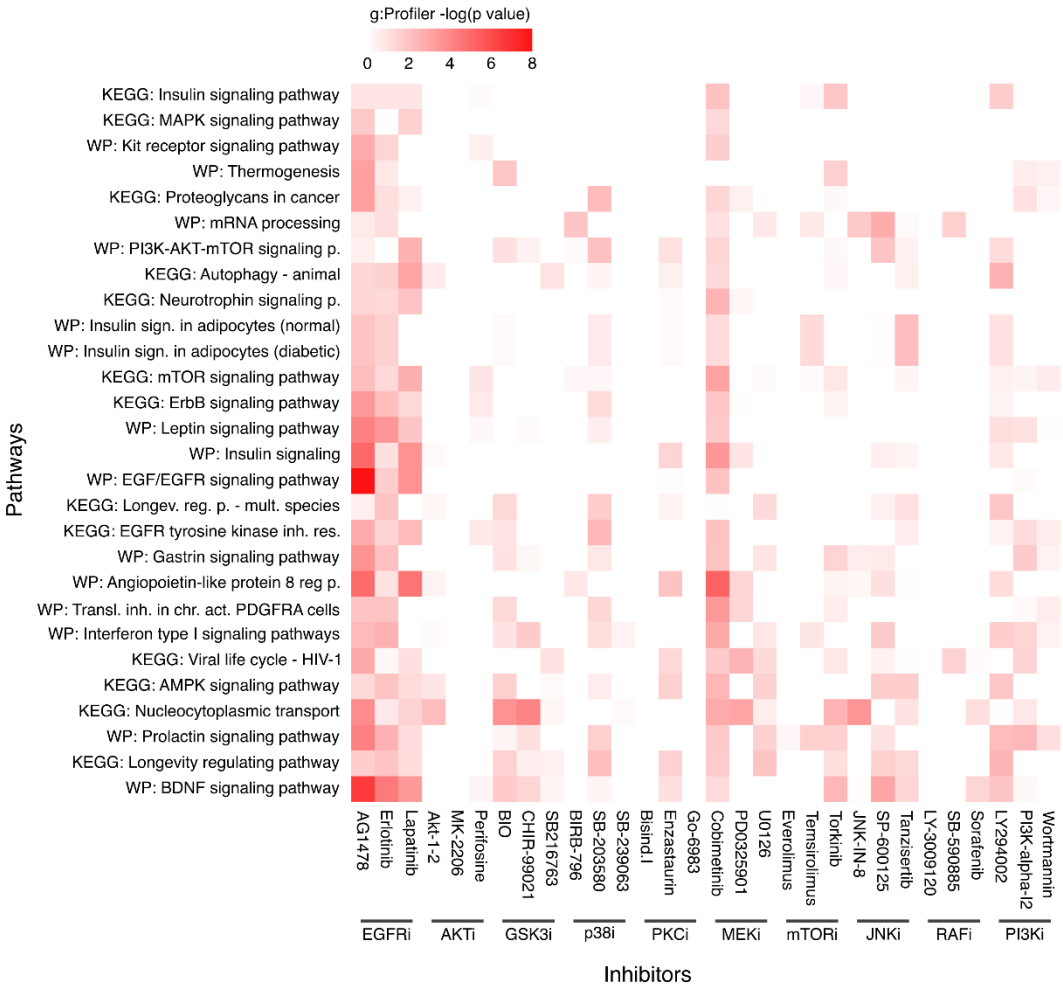


Figure 1. PTMNavigator puts PTM perturbation data in the context of signaling networks. (**A**) PTMNavigator combines two inputs: PTM Perturbation data, which is provided by the user,

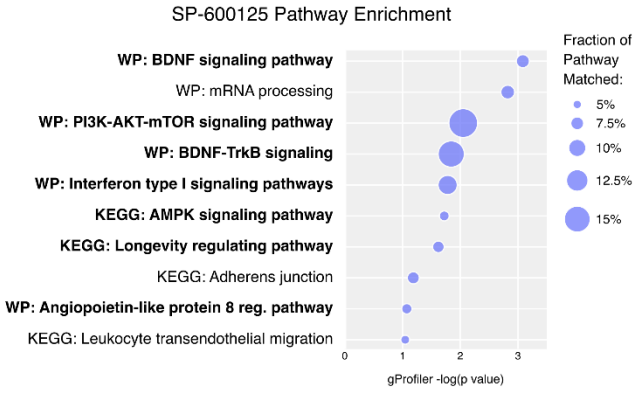
and pathway diagrams, which are provided by KEGG and WikiPathways. (**B**) Number of pathways per organism and reference database imported into ProteomicsDB and available in PTMNavigator. (**C**) The main interface of PTMNavigator. (1) The datasets can be accessed via a 32-digit Session ID that users receive during data upload. An arbitrary combination of datasets can be loaded at the same time. (2) The list of pathways to choose from is sorted by g:Profiler enrichment scores. 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. (3) A dataset may contain multiple experiments (specified in the uploaded CSV file). In this panel, they can dynamically be added to and removed from the diagram. (4) In this panel, users can show and hide regulation groups (up-/down-/unregulated). This feature is particularly helpful because many data sets contain mostly unregulated entries. (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 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). (7) If a user uploaded protein expression data, the proteins will also be colored according to their regulation group. (8) Many proteins may have multiple labels. By right-clicking on a node, an alternative name can be selected. (9) 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.

624

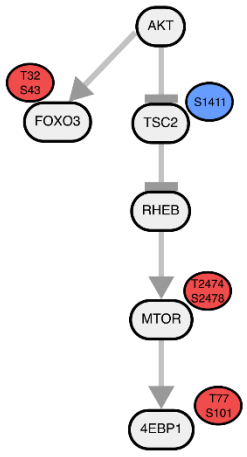
A


g:Profiler -log(p value)
0 2 4 6 8
Pathways
KEGG: Insulin signaling pathway
KEGG: MAPK signaling pathway
WP: Kit receptor signaling pathway
WP: Thermogenesis
KEGG: Proteoglycans in cancer
WP: mRNA processing
WP: PI3K-AKT-mTOR signaling p.
KEGG: Autophagy - animal
KEGG: Neurotrophin signaling p.
WP: Insulin sign. in adipocytes (normal)
WP: Insulin sign. in adipocytes (diabetic)
KEGG: mTOR signaling pathway
KEGG: ErbB signaling pathway
WP: Leptin signaling pathway
WP: Insulin signaling
WP: EGF/EGFR signaling pathway
KEGG: Longev. reg. p. - mult. species
KEGG: EGFR tyrosine kinase inh. res.
WP: Gastrin signaling pathway
WP: Angiopoietin-like protein 8 reg p.
WP: Transl. inh. in chr. act. PDGFRA cells
WP: Interferon type I signaling pathways
KEGG: Viral life cycle - HIV-1
KEGG: AMPK signaling pathway
KEGG: Nucleocytoplasmic transport
WP: Prolactin signaling pathway
KEGG: Longevity regulating pathway
WP: BDNF signaling pathway
AG1478
Erlotinib
Lapatinib
Akt-1-2
MK-2206
Perifosine
BIO
CHIR-99021
SB216763
BIRB-796
SB-203580
SB-239063
Bisind.I
Enzastaurin
Go-6983
Cobimetinib
PD0325901
U0126
Everolimus
Temsirolimus
Torkinib
JNK-IN-8
SP-600125
Tanziserfib
LY-3009120
SB-590885
Sorafenib
LY294002
PI3K-alpha-I2
Wortmannin
EGFRi
AKTi
GSK3i
p38i
PKCi
MEKi
mTORi
JNKi
RAFi
PI3Ki
Inhibitors

B

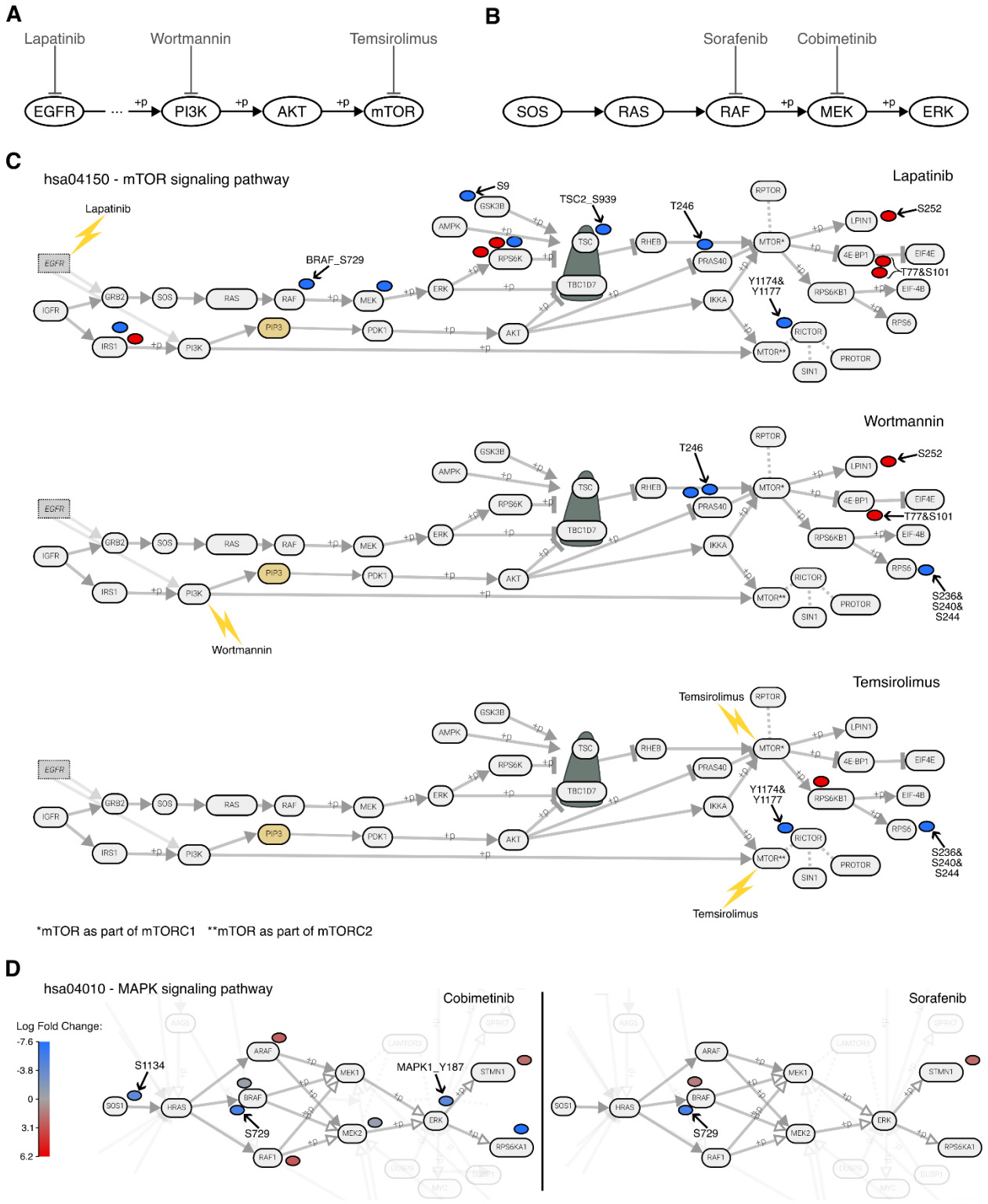

SP-600125 Pathway Enrichment
WP: BDNF signaling pathway
WP: mRNA processing
WP: PI3K-AKT-mTOR signaling pathway
WP: BDNF-TrkB signaling
WP: Interferon type I signaling pathways
KEGG: AMPK signaling pathway
KEGG: Longevity regulating pathway
KEGG: Adherens junction
WP: Angiopoietin-like protein 8 reg. pathway
KEGG: Leukocyte transendothelial migration
0 1 2 3
gProfiler -log(p value)
Fraction of Pathway Matched:
5%
7.5%
10%
12.5%
15%

C


AKT
T32
S43
FOXO3
TSC2
S1411
RHEB
T2474
S2478
MTOR
T77
S101
4EBP1

625

Figure 2. PTMNavigator aids in interpreting pathway enrichment analyses results. (**A**) Results of g:Profiler enrichments showing pathways with p-value<0.05. Pathways are sorted by the number of datasets in which they were reported as enriched. For illustrative purposes, only pathways that are significant for at least 3 datasets are shown. Inhibitors are grouped by their designated targets. (**B**) 10 highest-scoring pathways for the JNK inhibitor SP-600125. Circle size indicates the fraction of genes in the pathway that were contained in the set of regulations. The pathways that contain the signaling cascade from panel C are highlighted in bold. (**C**) The AKT-mTOR signaling axis that appears with slight variations in 7 of the 10 highest scoring pathways of the SP-600125 data. The small nodes next to the proteins indicate phosphorylation sites that were regulated up (red) or down (blue).



637

638 **Figure 3. Analysis of a phosphoproteomic perturbation dataset in PTMNavigator (A)**

639 Simplified schematic depiction of the PI3K-AKT-mTOR pathway and its interactions with

Lapatinib, Wortmannin, and Temsirolimus. (**B**) Simplified schematic depiction of the RAF-MEK-ERK pathway and its interactions with Cobimetinib and Sorafenib. (**C**) Edited screenshot from PTMNavigator showing an excerpt from the pathway diagram 'hsa04150 – mTOR Signaling Pathway' and the modified peptides 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. *Top:* Lapatinib. *Middle:* Wortmannin. *Bottom:* Temsirolimus. (**D**) Edited screenshot from PTMNavigator showing an excerpt from the pathway diagram 'hsa04010 – MAPK signaling pathway' 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.

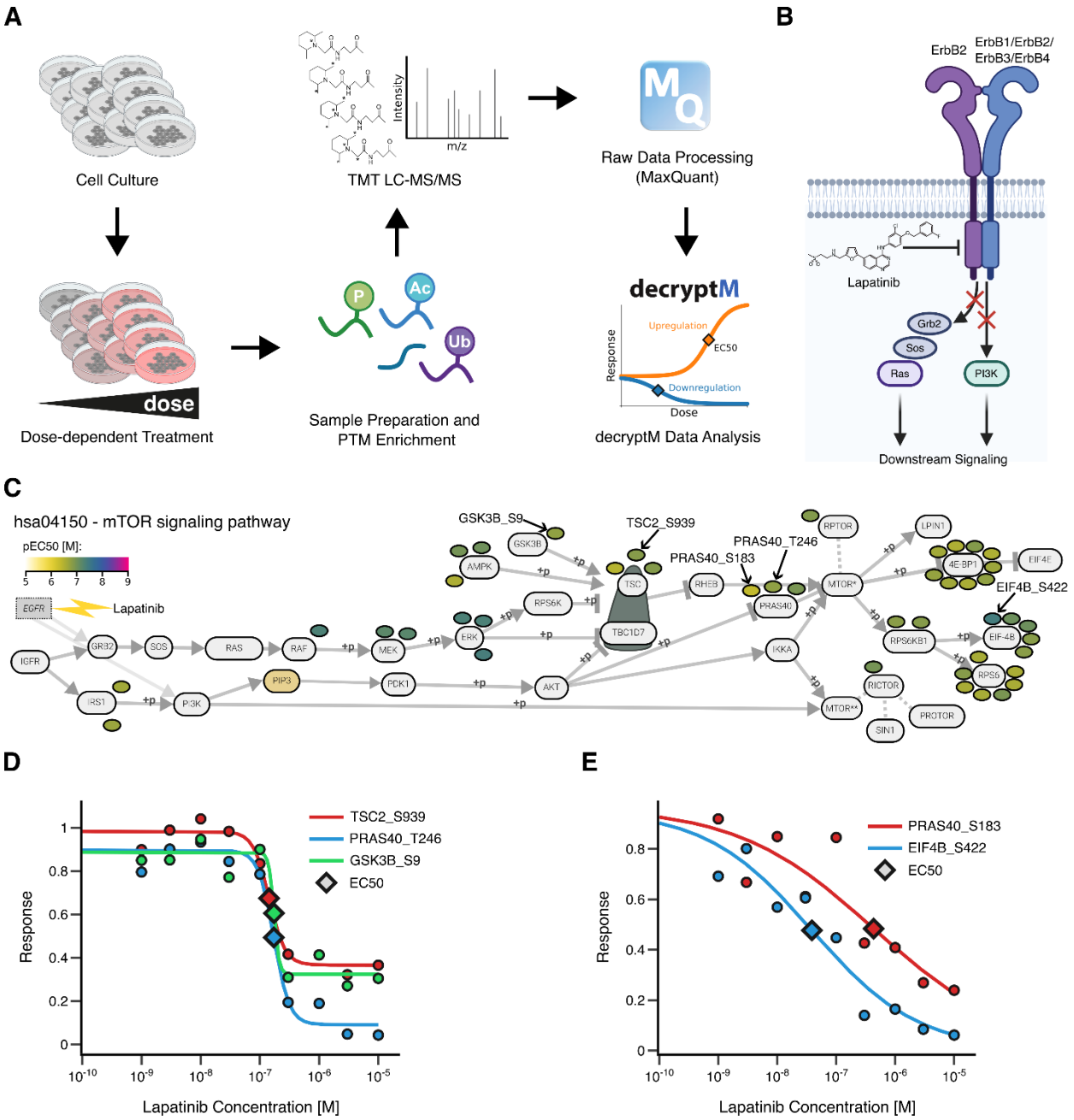


Figure 4. PTMNavigator supports analysis of drug dose-dependent perturbation data. (**A**) decryptM workflow. P, phosphorylation; Ac, acetylation; Ub, ubiquitinylation; LC-MS/MS, liquid chromatography tandem mass spectrometry; TMT, tandem mass tags. (**B**) Lapatinib inhibits ErbB2 by binding to its ATP-binding pocket, thereby inactivating its downstream signaling cascade. (**C**)

Edited screenshot from PTMNavigator showing an excerpt from the pathway diagram 'hsa04150 – mTOR Signaling Pathway' 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 corresponding pathway diagram of the Lapatinib experiment from Bekker-Jensen et al. is shown in Figure 3C, upper panel. (**D**) decryptM curves from the Lapatinib dataset for those peptides that were also significantly downregulated in the Bekker-Jensen data. (**E**) decryptM curves with the highest (pEC_{50}=7.41) and lowest (pEC_{50}=6.37) potency in the mTOR signaling pathway.

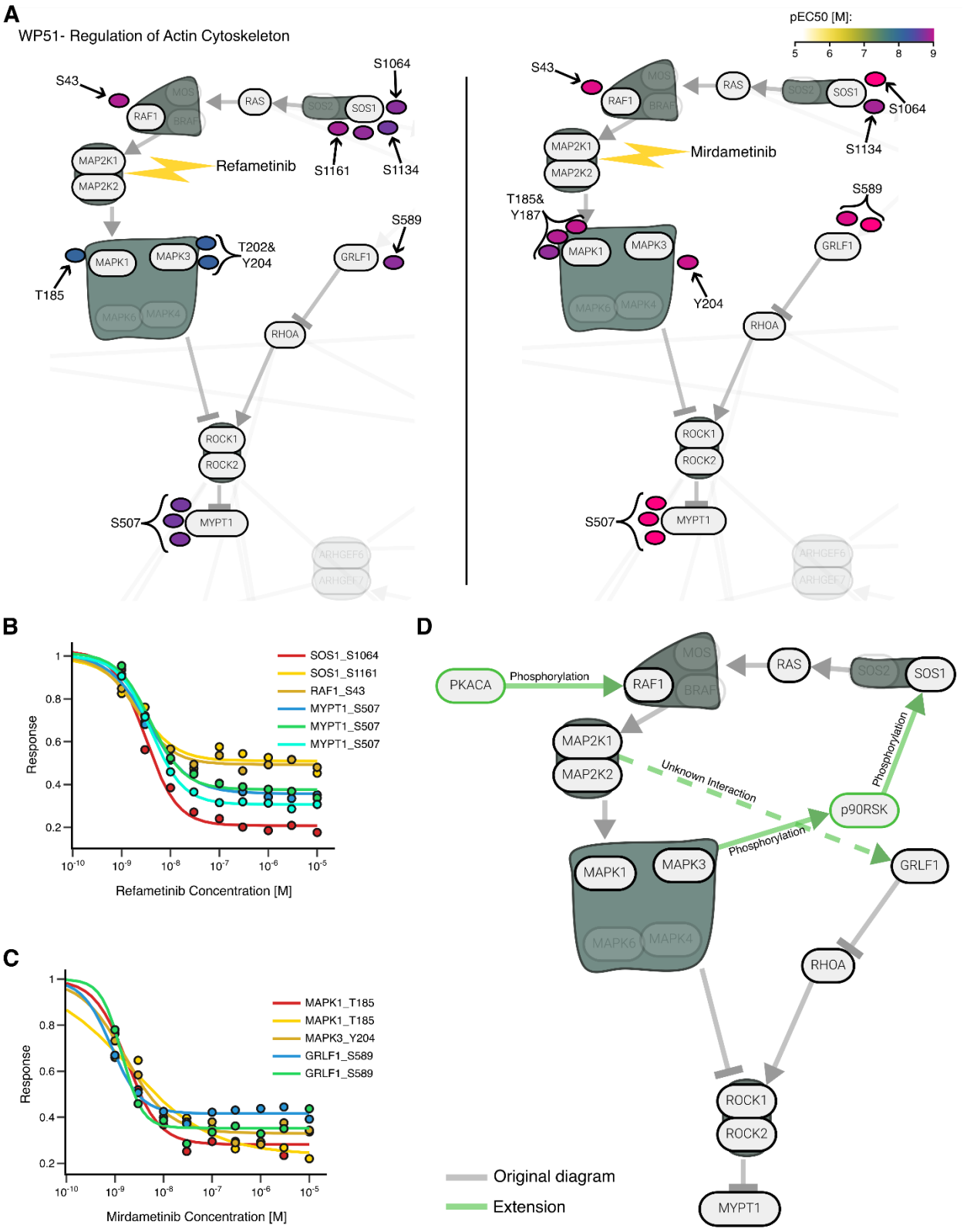


Figure 5. PTMNavigator helps to understand the pathway-level effects of MEK inhibition.

(**A**) Edited screenshot from PTMNavigator showing an excerpt from the pathway diagram 'WP51 – Regulation of Actin Cytoskeleton' and the modified peptides regulated in the Refametinib (*left*) and Mirdametinib (*right*) decryptM experiments. All regulations are downregulations. (**B**) decryptM curves for selected peptides on SOS1, RAF1, and MYPT1 in the Refametinib experiment. (**C**) decryptM curves for selected peptides on MAPK1, MAPK3, and GRLF1 in the Mirdametinib experiment. (**D**) Proposed extension to the MEK inhibition network based on the diagram 'WP51 – Regulation of Actin Cytoskeleton'. Additional nodes and edges are shown in green.

Supplementary Materials

Supplementary Table 1. Input examples for PTMNavigator.

Supplementary Figure 1. Mapping of Node Types from KEGG and WikiPathways to PTMNavigator.

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

Supplementary Figure 3. Gene Identifier Conversion for WikiPathways Diagrams.

Supplementary Figure 4. Comparison of dose-dependent Dactolisib treatment in three different cell lines.

Supplementary Figure 5. decryptM curves from the Refametinib and Mirdametinib datasets.

Supplementary Figure 6. 'WP2380 - Brain-derived neurotrophic factor signaling pathway' perturbed by SP-600125.

Supplementary Figure 7. 'WP411 - mRNA processing' perturbed by SP-600125.

Supplementary Figure 8. 'WP3844 - PI3K-AKT-mTOR signaling pathway and therapeutic opportunities' perturbed by SP-600125.

Supplementary Figure 9. 'WP3676 - BDNF-TrkB signaling' perturbed by SP-600125.

Supplementary Figure 10. 'WP585 - Interferon type I signaling pathways' perturbed by SP-600125.

Supplementary Figure 11. 'hsa04152 - AMPK signaling pathway' perturbed by SP-600125.

Supplementary Figure 12. 'hsa04211 - Longevity regulating pathway' perturbed by SP-600125.

Supplementary Figure 13. 'hsa04520 - Adherens junction' perturbed by SP-600125.

Supplementary Figure 14. 'WP3915 - Angiopoietin-like protein 8 regulatory pathway' perturbed by SP-600125.

Supplementary Figure 15. 'hsa04670 - Leukocyte transendothelial migration' perturbed by SP-600125.

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. 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).

15. 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).

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

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. Bekker-Jensen DB, *et al.* Rapid and site-specific deep phosphoproteome profiling by data-independent acquisition without the need for spectral libraries. *Nature Communications* **11**, 787 (2020).

27. Ochoa D, *et al.* The functional landscape of the human phosphoproteome. *Nature Biotechnology* **38**, 365-373 (2020).

28. Roberts PJ, Der CJ. Targeting the Raf-MEK-ERK mitogen-activated protein kinase cascade for the treatment of cancer. *Oncogene* **26**, 3291-3310 (2007).

29. Zheng Y, *et al.* Temporal regulation of EGF signalling networks by the scaffold protein Shc1. *Nature* **499**, 166-171 (2013).

30. Lee YH, Giraud J, Davis RJ, White MF. c-Jun N-terminal kinase (JNK) mediates feedback inhibition of the insulin signaling cascade. *The Journal of Biological Chemistry* **278**, 2896-2902 (2003).

31. Cui X, Kim H-J, Kuiatse I, Kim H, Brown PH, Lee AV. Epidermal growth factor induces insulin receptor substrate-2 in breast cancer cells via c-Jun NH(2)-terminal kinase/activator protein-1 signaling to regulate cell migration. *Cancer Research* **66**, 5304-5313 (2006).

32. Adams SJ, Aydin IT, Celebi JT. GAB2 - A Scaffolding Protein in Cancer. *Molecular cancer research : MCR* **10**, 10.1158/1541-7786.MCR-1112-0352 (2012).

33. Feng J, *et al.* The Role of JNk Signaling Pathway in Obesity-Driven Insulin Resistance. *Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy* **13**, 1399-1406 (2020).

34. Lohi H, *et al.* Novel glycogen synthase kinase 3 and ubiquitination pathways in progressive myoclonus epilepsy. *Hum Mol Genet* **14**, 2727-2736 (2005).

35. Zhan J, Chitta RK, Harwood FC, Grosveld GC. Phosphorylation of TSC2 by PKC-delta reveals a novel signaling pathway that couples protein synthesis to mTORC1 activity. *Mol Cell Biochem* **456**, 123-134 (2019).

36. MacNicol MC, Muslin AJ, MacNicol AM. Disruption of the 14-3-3 binding site within the B-Raf kinase domain uncouples catalytic activity from PC12 cell differentiation. *J Biol Chem* **275**, 3803-3809 (2000).

37. Saha M, *et al.* RSK phosphorylates SOS1 creating 14-3-3 docking sites and negatively regulating MAPK activation. *The Biochemical journal* **447**, 159-166 (2012).

38. Sun Q, Wang W. Structures of BRAF-MEK1-14-3-3 sheds light on drug discovery. *Signal Transduct Target Ther* **4**, 59 (2019).

39. Klaeger S, *et al.* The target landscape of clinical kinase drugs. *Science*, 44 (2017).

40. Cohen P, Cross D, Jänne PA. Kinase drug discovery 20 years after imatinib: progress and future directions. *Nature Reviews Drug Discovery*, 1-19 (2021).

41. Bayer FP, Gander M, Kuster B, The M. CurveCurator: A recalibrated F-statistic to assess, classify, and explore significance of dose-response curves. *bioRxiv* *https://doiorg/101101/20230807552260* *[PREPRINT]*, (2023).

42. Castoldi R, *et al.* TetraMabs: simultaneous targeting of four oncogenic receptor tyrosine kinases for tumor growth inhibition in heterogeneous tumor cell populations. *Protein Eng Des Sel* **29**, 467-475 (2016).

43. 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).

44. Reinecke M, *et al.* Chemoproteomic Selectivity Profiling of PIKK and PI3K Kinase Inhibitors. *ACS Chem Biol* **14**, 655-664 (2019).

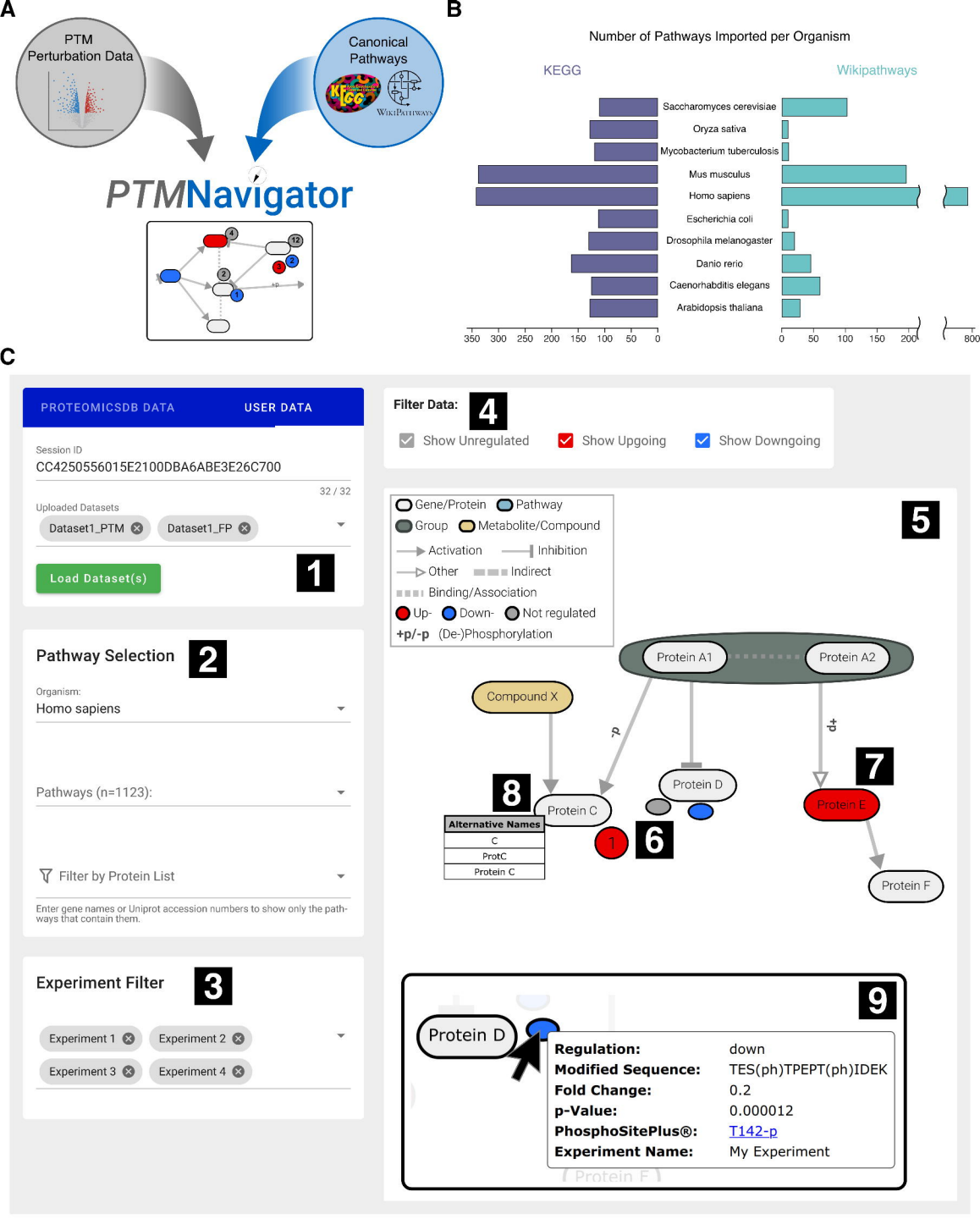
45. Johnson JL, *et al.* An atlas of substrate specificities for the human serine/threonine kinome. *Nature* **613**, 759-766 (2023).

46. Crowl S, Jordan B, Ahmed H, Ma C, Naegle KM. KSTAR: An algorithm to predict patient-specific kinase activities from phosphoproteomic data.). Systems Biology (2021).

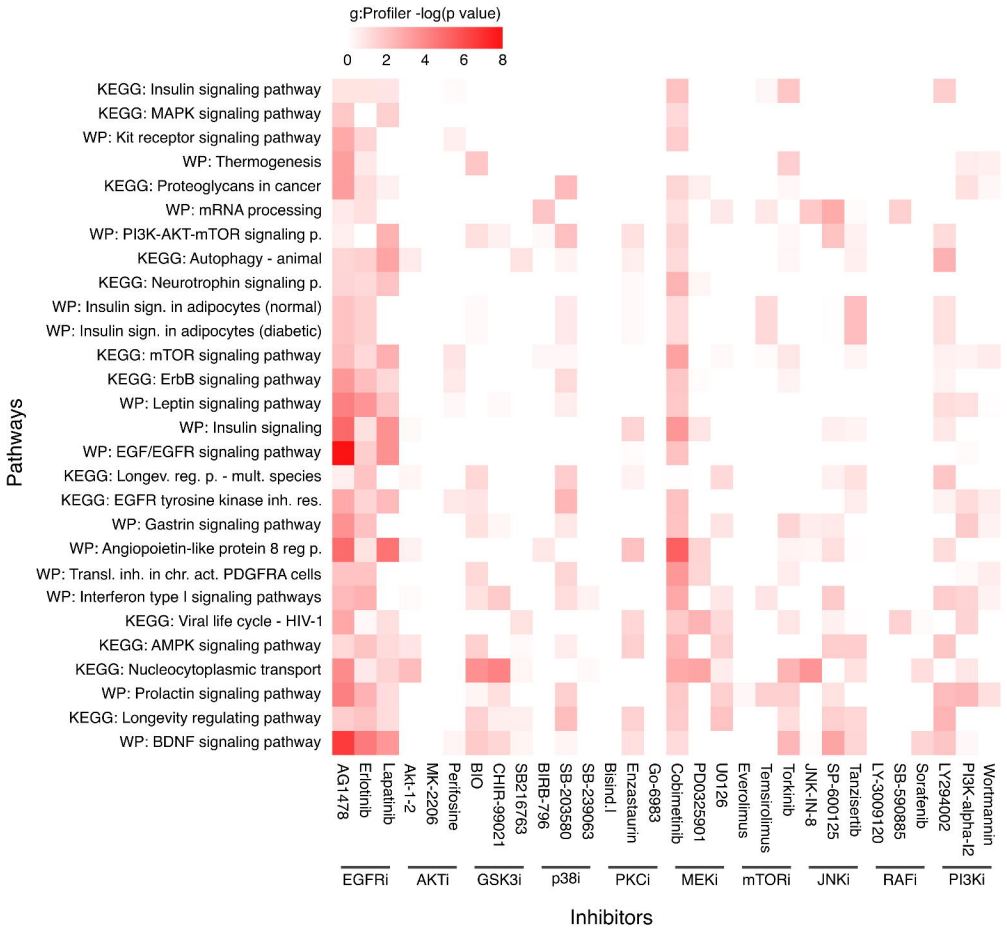
47. Beekhof R, *et al.* INKA, an integrative data analysis pipeline for phosphoproteomic inference of active kinases. *Molecular Systems Biology* **15**, e8250 (2019).

48. 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).

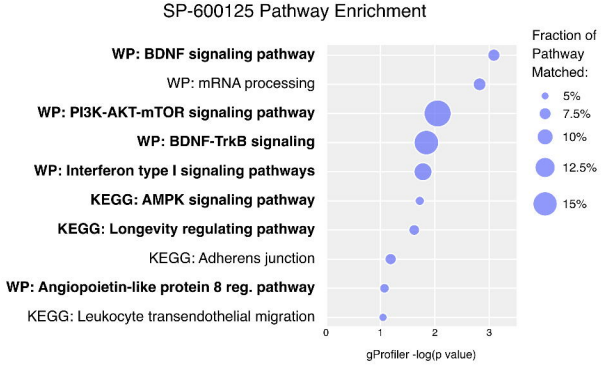
49. Bostock M, Ogievetsky V, Heer J. D³ Data-Driven Documents. *IEEE Transactions on Visualization and Computer Graphics* **17**, 2301-2309 (2011).


A
PTM Perturbation Data
Canonical Pathways
PTMNavigator
B
Number of Pathways Imported per Organism
KEGG
Wikipathways
Saccharomyces cerevisiae
Oryza sativa
Mycobacterium tuberculosis
Mus musculus
Homo sapiens
Escherichia coli
Drosophila melanogaster
Danio rerio
Caenorhabditis elegans
Arabidopsis thaliana
350 300 250 200 150 100 50 0
0 50 100 150 200 800
C
PROTEOMICSDB DATA
USER DATA
Session ID
CC4250556015E2100DBA6ABE3E26C700
32 / 32
Uploaded Datasets
Dataset1_PTM
Dataset1_FP
Load Dataset(s)
1
Pathway Selection
2
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.
Experiment Filter
3
Experiment 1
Experiment 2
Experiment 3
Experiment 4
Filter Data:
4
Show Unregulated
Show Upgoing
Show Downgoing
Gene/Protein
Pathway
Group
Metabolite/Compound
Activation
Inhibition
Other
Indirect
Binding/Association
Up-
Down-
Not regulated
+p/-p (De-)Phosphorylation
5
Protein A1
Protein A2
Compound X
-p
+p
Protein D
Protein C
Protein E
Protein F
8
7
6
1
Alternative Names
C
ProtC
Protein C
9
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

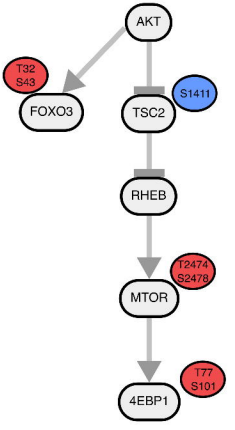
A

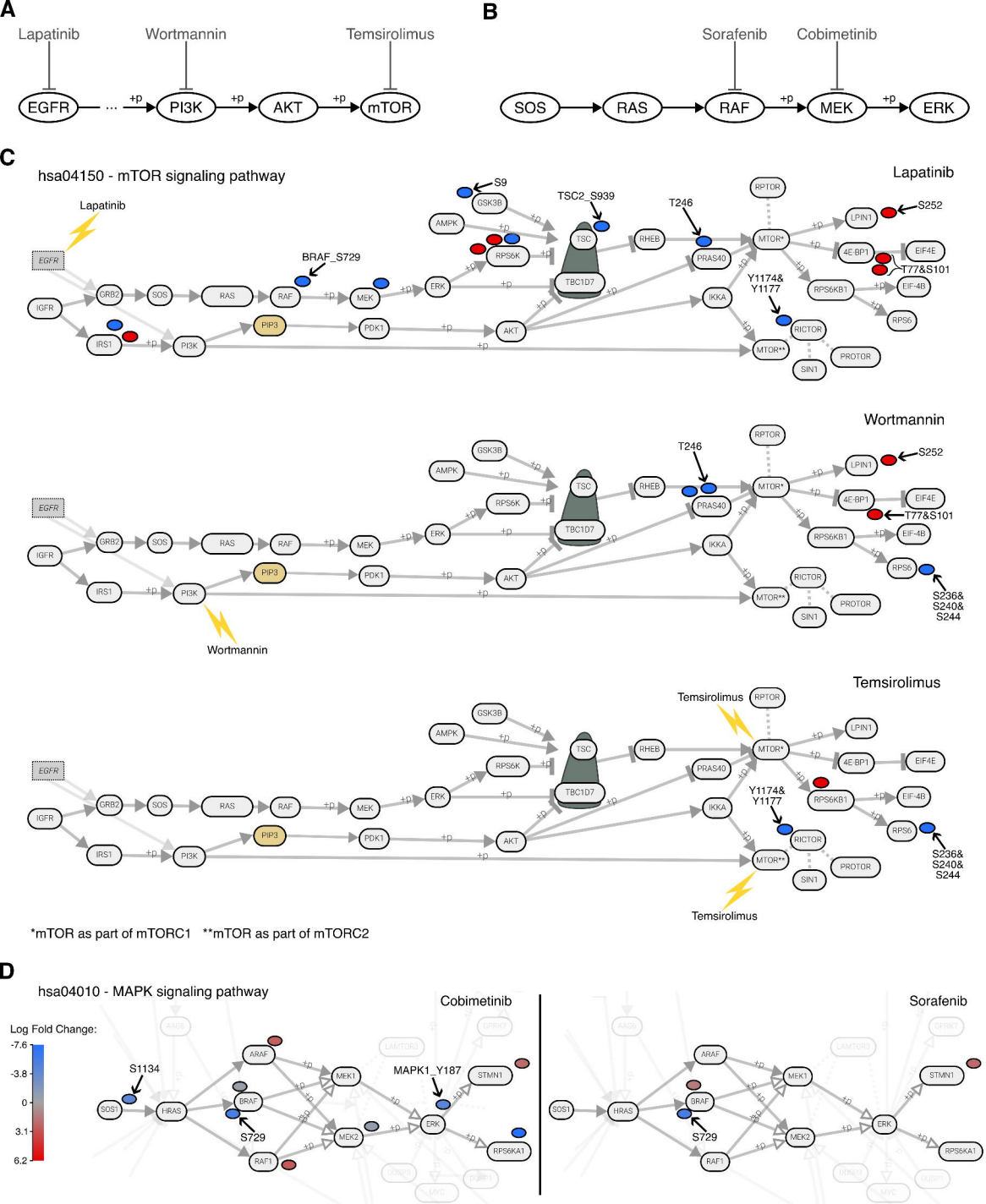

g:Profiler -log(p value)
0 2 4 6 8
Pathways
KEGG: Insulin signaling pathway
KEGG: MAPK signaling pathway
WP: Kit receptor signaling pathway
WP: Thermogenesis
KEGG: Proteoglycans in cancer
WP: mRNA processing
WP: PI3K-AKT-mTOR signaling p.
KEGG: Autophagy - animal
KEGG: Neurotrophin signaling p.
WP: Insulin sign. in adipocytes (normal)
WP: Insulin sign. in adipocytes (diabetic)
KEGG: mTOR signaling pathway
KEGG: ErbB signaling pathway
WP: Leptin signaling pathway
WP: Insulin signaling
WP: EGF/EGFR signaling pathway
KEGG: Longev. reg. p. - mult. species
KEGG: EGFR tyrosine kinase inh. res.
WP: Gastrin signaling pathway
WP: Angiopoietin-like protein 8 reg p.
WP: Transl. inh. in chr. act. PDGFRA cells
WP: Interferon type I signaling pathways
KEGG: Viral life cycle - HIV-1
KEGG: AMPK signaling pathway
KEGG: Nucleocytoplasmic transport
WP: Prolactin signaling pathway
KEGG: Longevity regulating pathway
WP: BDNF signaling pathway
AG1478
Erlotinib
Lapatinib
Akt-1-2
MK-2206
Perifosine
BIO
CHIR-99021
SB216763
BIRB-796
SB-203580
SB-239063
Bisind.I
Enzastaurin
Go-6983
Cobimetinib
PD0325901
U0126
Everolimus
Temsirolimus
Torkinib
JNK-IN-8
SP-600125
Tanzisertib
LY-3009120
SB-590885
Sorafenib
LY294002
PI3K-alpha-I2
Wortmannin
EGFRi
AKTi
GSK3i
p38i
PKCi
MEKi
mTORi
JNKi
RAFi
PI3Ki
Inhibitors

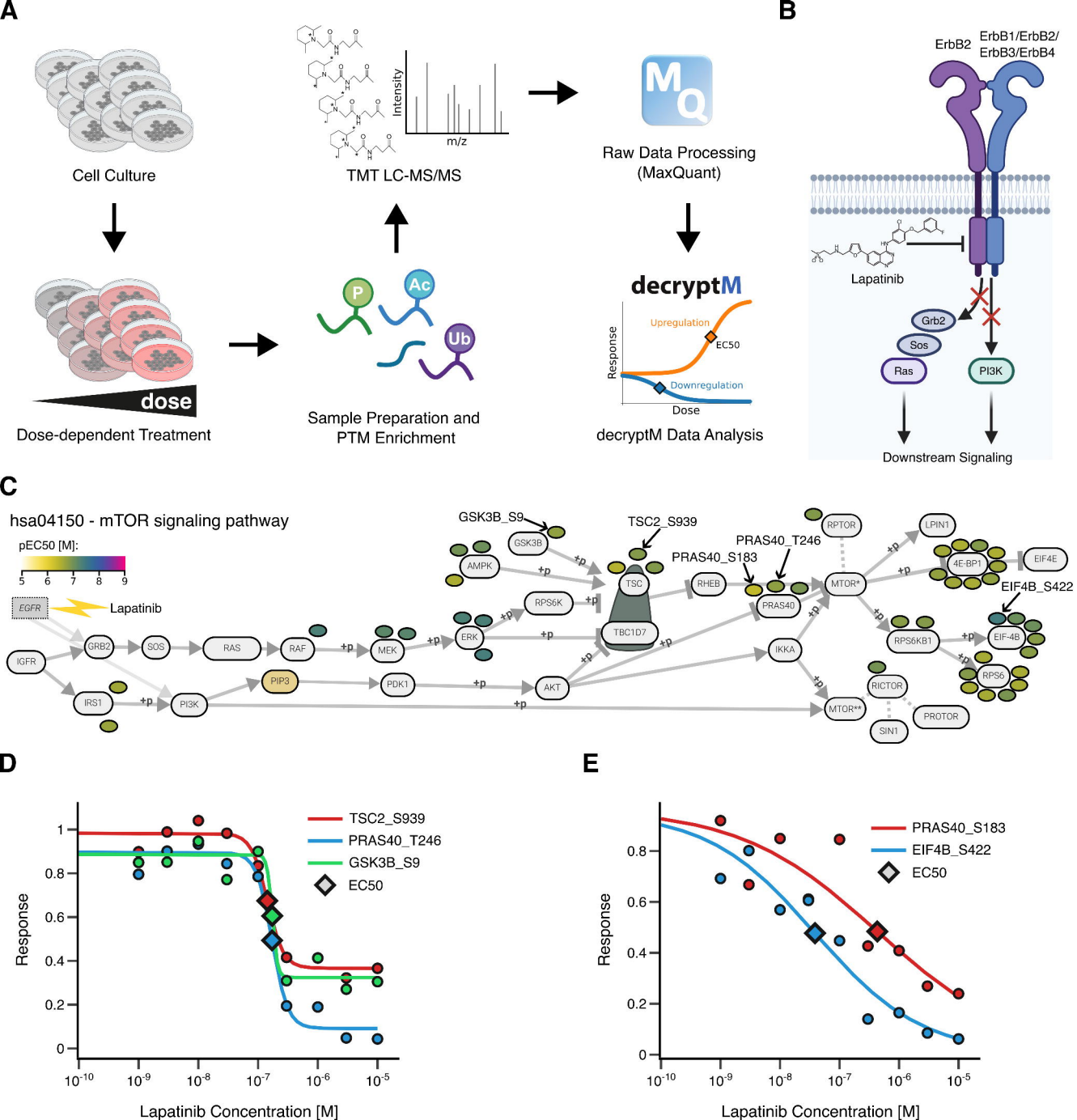
B

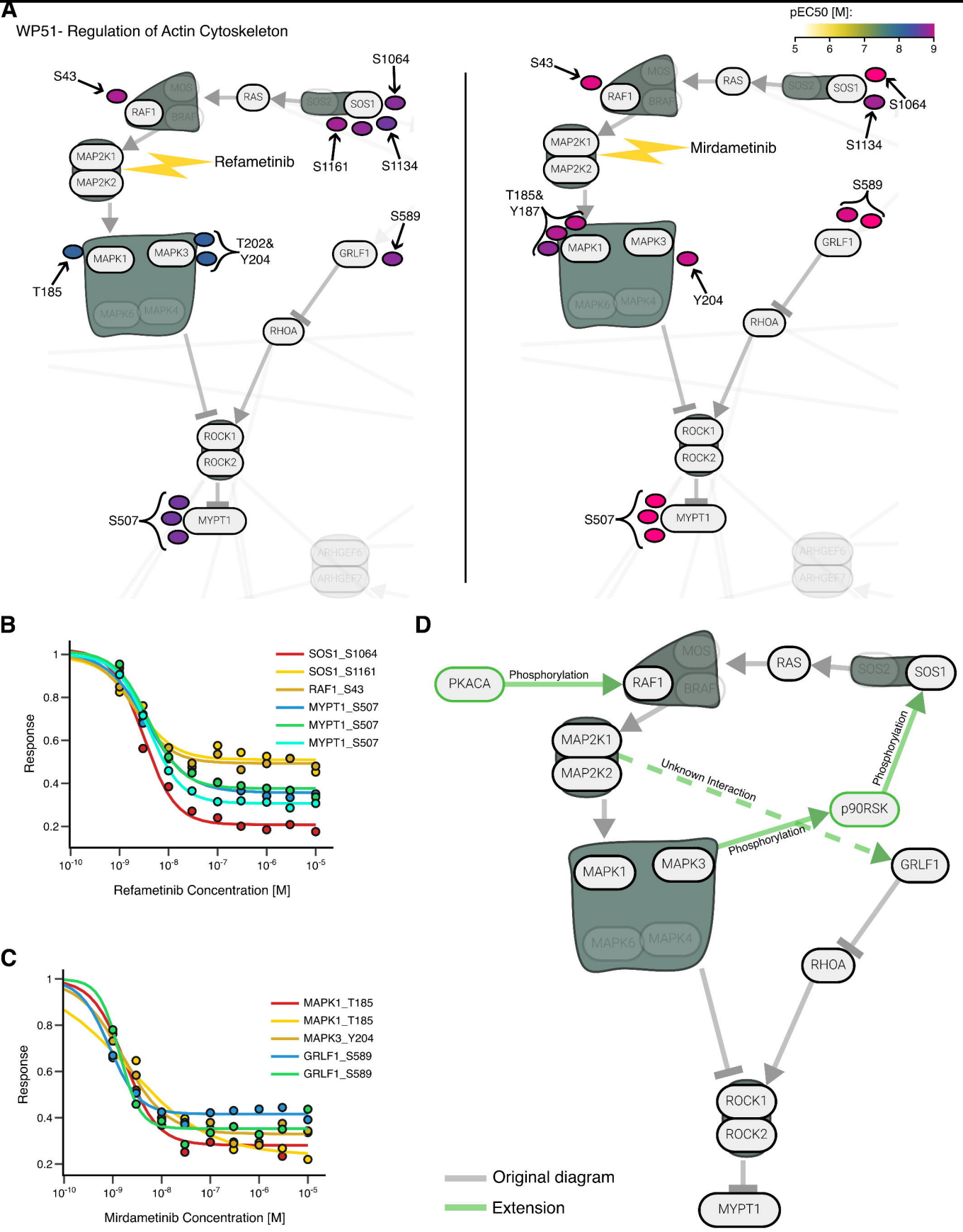

SP-600125 Pathway Enrichment
WP: BDNF signaling pathway
WP: mRNA processing
WP: PI3K-AKT-mTOR signaling pathway
WP: BDNF-TrkB signaling
WP: Interferon type I signaling pathways
KEGG: AMPK signaling pathway
KEGG: Longevity regulating pathway
KEGG: Adherens junction
WP: Angiopoietin-like protein 8 reg. pathway
KEGG: Leukocyte transendothelial migration
0 1 2 3
gProfiler -log(p value)
Fraction of Pathway Matched:
5%
7.5%
10%
12.5%
15%

C


AKT
T32
S43
FOXO3
S1411
TSC2
RHEB
T2474
S2478
MTOR
T77
S101
4EBP1


A
Lapatinib
Wortmannin
Temsirolimus
EGFR
PI3K
AKT
mTOR
+p
B
Sorafenib
Cobimetinib
SOS
RAS
RAF
MEK
ERK
C
hsa04150 - mTOR signaling pathway
Lapatinib
Wortmannin
Temsirolimus
S9
TSC2_S939
T246
BRAF_S729
Y1174&
Y1177
S252
T77&S101
S236&
S240&
S244
EGFR
IGFR
GRB2
SOS
RAS
RAF
MEK
ERK
IRS1
PI3K
PIP3
PDK1
AKT
GSK3B
AMPK
RPS6K
TSC
TBC1D7
RHEB
PRAS40
IKKA
RPTOR
MTOR*
MTOR**
RICTOR
SIN1
PROTOR
LPIN1
4E-BP1
EIF4E
RPS6KB1
EIF-4B
RPS6
*mTOR as part of mTORC1 **mTOR as part of mTORC2
D
hsa04010 - MAPK signaling pathway
Cobimetinib
Sorafenib
Log Fold Change:
-7.6
-3.8
0
3.1
6.2
S1134
MAPK1_Y187
S729
SOS1
HRAS
ARAF
BRAF
RAF1
MEK1
MEK2
ERK
STMN1
RPS6KA1


A
Cell Culture
TMT LC-MS/MS
Intensity
m/z
Raw Data Processing
(MaxQuant)
dose
Dose-dependent Treatment
P
Ac
Ub
Sample Preparation and
PTM Enrichment
decryptM
Upregulation
EC50
Downregulation
Response
Dose
decryptM Data Analysis
B
ErbB2
ErbB1/ErbB2/
ErbB3/ErbB4
Lapatinib
Grb2
Sos
Ras
PI3K
Downstream Signaling
C
hsa04150 - mTOR signaling pathway
pEC50 [M]:
5 6 7 8 9
EGFR
Lapatinib
IGFR
GRB2
SOS
RAS
RAF
MEK
ERK
IRS1
PI3K
PIP3
PDK1
AKT
GSK3B_S9
GSK3B
AMPK
RPS6K
TSC2_S939
TSC
TBC1D7
RHEB
PRAS40_S183
PRAS40_T246
PRAS40
RPTOR
MTOR*
IKKA
MTOR**
RICTOR
SIN1
PROTOR
LPIN1
4E-BP1
EIF4E
EIF4B_S422
RPS6KB1
EIF-4B
RPS6
+p
D
TSC2_S939
PRAS40_T246
GSK3B_S9
EC50
Response
Lapatinib Concentration [M]
E
PRAS40_S183
EIF4B_S422
EC50
Response
Lapatinib Concentration [M]


A
WP51- Regulation of Actin Cytoskeleton
pEC50 [M]:
5
6
7
8
9
S43
S1064
S1161
S1134
S589
T185
T202&
Y204
T185&
Y187
Y204
S507
RAF1
MOS
BRAF
RAS
SOS2
SOS1
MAP2K1
MAP2K2
MAPK1
MAPK3
MAPK6
MAPK4
GRLF1
RHOA
ROCK1
ROCK2
MYPT1
ARHGEF6
ARHGEF7
Refametinib
Mirdametinib
B
Response
Refametinib Concentration [M]
SOS1_S1064
SOS1_S1161
RAF1_S43
MYPT1_S507
MYPT1_S507
MYPT1_S507
C
Response
Mirdametinib Concentration [M]
MAPK1_T185
MAPK1_T185
MAPK3_Y204
GRLF1_S589
GRLF1_S589
D
PKACA
Phosphorylation
Unknown Interaction
Phosphorylation
Phosphorylation
p90RSK
Original diagram
Extension