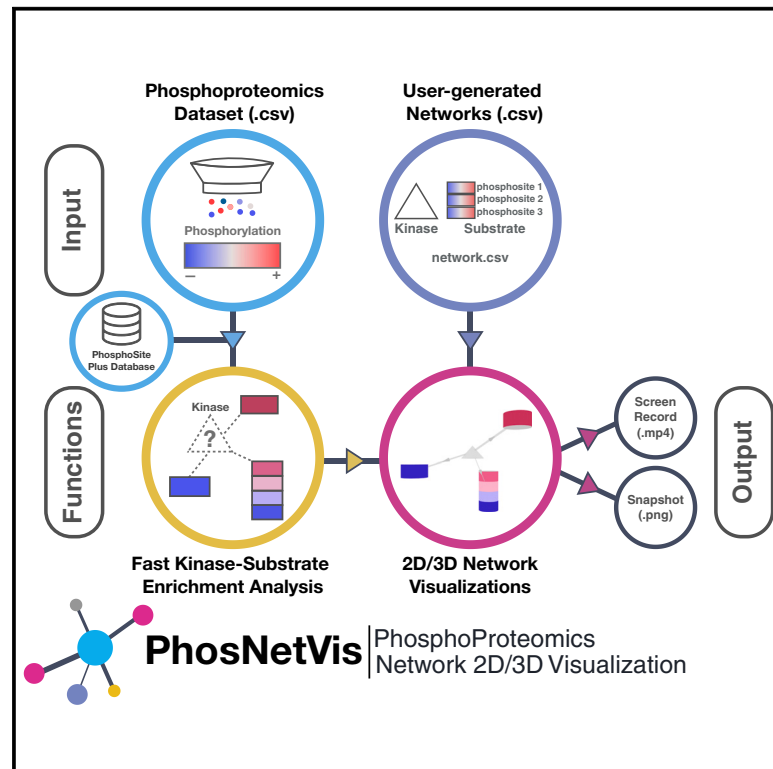


Patterns

PhosNetVis: A web-based tool for fast kinase-substrate enrichment analysis and interactive 2D/3D network visualizations of phosphoproteomics data

Graphical abstract



Authors

Osho Rawal, Berk Turhan, Irene Font Peradejordi, ..., Jeffrey Johnson, Mehdi Bouhaddou, Zeynep H. Gümüş

Correspondence

zeynep.gumus@mssm.edu

In brief

PhosNetVis is a web-based platform designed to infer kinases and visualize kinase-substrate interaction (KSI) networks from phosphoproteomics data. It enables users to generate, interactively explore, and compare multiple KSI networks in both 2D and 3D, complete with detailed phosphorylation site annotations. Additionally, users can explore pre-existing networks, including those from the Clinical Proteomic Tumor Analysis Consortium (CPTAC). While optimized for KSIs, its flexible features support visualizing other biomolecular networks. The platform is free, open to all, and requires no login.

Highlights

- PhosNetVis is a unified platform to interactively explore phosphoproteomics data
- Users infer kinases and visually explore the resulting kinase-substrate networks
- Users can visually compare multiple networks, altered phosphosites in 2D or 3D
- Secure, browser-based, open-source tool with no login required



Resource

PhosNetVis: A web-based tool for fast kinase-substrate enrichment analysis and interactive 2D/3D network visualizations of phosphoproteomics data

Osho Rawal,^{1,9} Berk Turhan,^{1,2,9} Irene Font Peradejordi,^{1,3} Shreya Chandrasekar,^{1,3} Selim Kalayci,¹ Sacha Gnjatic,^{4,5} Jeffrey Johnson,⁶ Mehdi Bouhaddou,^{7,8} and Zeynep H. Gümüş^{1,4,10,*}

¹Department of Genetics and Genomics, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA

²Faculty of Engineering and Natural Sciences, Sabanci University, 34956 Istanbul, Türkiye

³Cornell Tech, Cornell University, New York, NY 10044, USA

⁴Marc and Jennifer Lipschultz Precision Immunology Institute, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA

⁵Department of Immunology and Immunotherapy, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA

⁶Department of Microbiology, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA

⁷Department of Microbiology, Immunology, and Molecular Genetics, University of California, Los Angeles, Los Angeles, CA 90095, USA

⁸Institute for Quantitative and Computational Biosciences, University of California, Los Angeles, Los Angeles, CA 90095, USA

⁹These authors contributed equally

¹⁰Lead contact

*Correspondence: zeynep.gumus@mssm.edu

<https://doi.org/10.1016/j.patter.2024.101148>

SUMMARY

Protein phosphorylation involves the reversible modification of a protein (substrate) residue by another protein (kinase). Liquid chromatography-mass spectrometry studies are rapidly generating massive protein phosphorylation datasets across multiple conditions. Researchers then must infer kinases responsible for changes in phosphosites of each substrate. However, tools that infer kinase-substrate interactions (KSIs) are not optimized to interactively explore the resulting large and complex networks, significant phosphosites, and states. There is thus an unmet need for a tool that facilitates user-friendly analysis, interactive exploration, visualization, and communication of phosphoproteomics datasets. We present PhosNetVis, a web-based tool for researchers of all computational skill levels to easily infer, generate, and interactively explore KSI networks in 2D or 3D by streamlining phosphoproteomics data analysis steps within a single tool. PhosNetVis lowers barriers for researchers by rapidly generating high-quality visualizations to gain biological insights from their phosphoproteomics datasets. It is available at <https://gumuslab.github.io/PhosNetVis/>.

INTRODUCTION

Protein phosphorylation is a vital process in cellular signaling where a kinase protein modifies a residue on a substrate protein. This reversible modification can occur at multiple sites (phosphosites) on a single substrate, with different kinases targeting various substrates. Advances in liquid chromatography-mass spectrometry (LC-MS) technology have enabled the rapid generation of extensive protein phosphorylation datasets across different cellular states. To identify significant kinase-substrate interactions (KSIs) from these large datasets, bioinformatic analysis tools are essential. These tools infer which kinases are responsible for the observed changes in protein phosphorylation, using available KSI database resources, and facilitate the visual exploration of the resulting KSI networks.

Over the past decade, a wide array of computational tools has been developed to infer kinases.^{1,2} Many of these tools involve kinase-substrate enrichment analysis, which employs a gene

set enrichment analysis algorithm to determine whether a set of query proteins is enriched in substrates known to interact with specific kinases. Commonly employed kinase enrichment tools are KEA3,³ KSEAapp,⁴ PhosFate Profiler,⁵ RoKAI,⁶ KSEA,⁷ KSEAplus,⁸ and pCHIPS.⁹ These tools typically utilize publicly available KSI databases such as Phospho.ELM,¹⁰ PhosphoSitePlus,¹¹ the Human Protein Reference Database,¹² Swiss-Prot,¹³ and Kinase Library.¹⁴ PhosphoSitePlus¹¹ is often preferred, as it is updated regularly. As of the latest update, PhosphoSitePlus lists 10,995 KSI pairs (8,004 in human), and 23,800 phosphosites (14,663 in human) (version downloaded on Monday, October 28, 2024, from the database update of Thursday, October 17, 2024, 11:21:06 Eastern Daylight Time (EDT)).

After performing kinase enrichment, significant KSIs identified based on user-defined criteria are often visualized as networks, with kinases and substrates represented as nodes and their interactions as edges. Visualizing the altered phosphorylation



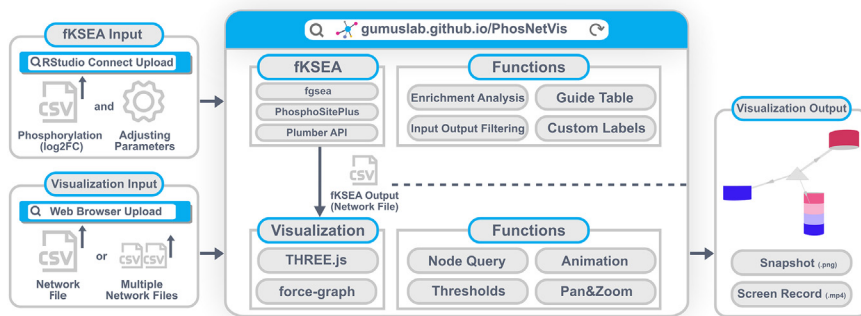


Figure 1. PhosNetVis overall architecture and main components

states of phosphosites in substrate proteins is also essential. Despite the proliferation of tools to infer KSIs from phosphoproteomics datasets, currently available tools are not optimized for visualizing and generating large, complex KSI networks along with their associated phosphorylation sites and states. Furthermore, there are no dedicated tools that integrate kinase enrichment with interactive visualizations of the resulting KSI networks and phosphosites across multiple conditions.

Current network data exploration workflows involve manual adjustments to maximize usefulness, minimize clutter, and improve visualization design. This process can be improved with scripting languages; however, proteomics researchers should not need to learn programming skills to create visualizations. Even then, tools for network visualizations typically do not include the functionalities to visually represent phosphorylation data of protein residue sites and states.¹⁵ At the same time, tools specifically designed for visual exploration of proteomics datasets do not allow users to upload and explore their own data. For instance, ProNetView,¹⁶ a web-based interactive 3D network visualization tool developed by our group, is tailored for exploring specific proteomics network data from the Clinical Proteomic Tumor Analysis Consortium (CPTAC).¹⁷ Furthermore, similar to network visualization tools, ProNetView does not visually represent the phosphorylation sites and states. One exception is the Cytoscape¹⁸ app, Omics Visualizer,¹⁹ which allows visual representations of phosphosite data but does not include a kinase enrichment component. At the same time, a popular web-based platform to investigate kinase activity data, PhosFate Profiler,⁵ only enables data exploration in tabular format without KSI network or phosphosite data visualization functionalities. Another web-based platform, RoKAI,⁶ for inferring kinase activity, also does not include capabilities for enhanced network visualizations and interactive explorations. There is thus a clear need for a tool that facilitates user-friendly generation, interactive exploration, visualization, and communication of phosphoproteomics datasets.

Here, we present PhosNetVis, a freely accessible web-based tool designed for users of all computational proficiency levels to explore kinase activity from phosphoproteomics studies. PhosNetVis integrates multiple analysis steps within a single platform, allowing users to effortlessly perform kinase-substrate enrichment analysis and/or build and visualize shareable interactive 2D or 3D KSI networks. The tool provides a versatile environment, offering visual representations of protein networks, phosphorylation sites, and states of interest along with associated differential phosphorylation and statistical signifi-

cance data. Users can interact with the data through KSI network visualizations and tabular formats, all within a single interface. The KSI networks can visually include phosphorylation sites and their respective states (increased or decreased). Input is straightforward, requiring users to upload their datasets as comma-separated files on their web browser. While PhosNetVis is tailored for the analysis of phosphoproteomics datasets, its adaptable architecture extends its utility to other biomolecular network applications even in the absence of phosphorylation data. This versatility enhances its potential impact across a wide range of applications.

The PhosNetVis web portal provides detailed tutorials, a frequently asked questions (FAQ) section, and interactive examples that highlight its user-friendly design and versatility. Additionally, as a resource for the research community, the portal hosts a complete catalog of KSI networks across 7 tumor immune subtypes derived from over 1,000 tumors across 10 different cancers, sourced from the CPTAC initiative.²⁰ As we describe under Illustrative examples, users can seamlessly query, visualize, interactively explore, compare, and download these KSI networks with phosphorylation state data, facilitating easy access to this extensive dataset.

RESULTS

User research and prototyping

Before designing and building PhosNetVis, we conducted preliminary user research with five domain practitioners. This research involved observing their current workflows, identifying pain points, and gathering requirements for ideal features. We analyzed the user input network data and how these data were typically mapped to 2D visual elements using existing tools. To establish the general workflow and convert these mappings from 2D to 3D, we first created low-fidelity (low-fi) prototypes on paper. This approach allowed us to quickly and efficiently present our assumptions on the functionalities and workflow. Following the low-fi prototypes, we developed high-fidelity (hi-fi) prototypes using Adobe XD (adobe.com/products/xd) and Figma (figma.com). We created multiple separate prototypes for both stages, which were eventually merged into a single final design. We presented the hi-fi prototype to prospective users within the Human Immunology Project Consortium and CPTAC. Based on their feedback, we made further improvements to both the prototype and the technical implementation.

PhosNetVis workflow design

Building KSI networks from phosphoproteomics datasets typically involves a two-step workflow, first using a kinase enrichment algorithm to establish KSI networks and then employing visualization tools to explore and analyze these networks. PhosNetVis streamlines this process by integrating both kinase

A Input File Format

Upload a .csv file in the format below for enrichment and then to build your interactive 3D phosphoproteomics network!

ProteinAccession (Required)	log2FC (Required)	PhosphoSiteID (PhosphositePlusID or Residue ID)(Optional)	pValue (Optional)
P08183	0.25	S671	0.66
P08183	1.24	S683	0.015
Q9UEY8	-2.12	S693	0.02

B Adjust Parameters

Input / Output Parameters

Input p-value cutoff:

Output p-value cutoff:

Output network node labels: gene name accession id

Enrichment Analysis Parameters

minSize:

maxSize:

eps:

[Click here for more information about fgsea library's analysis parameters](#)

C Upload Data & Run Analysis

Upload your .csv file

Choose File: kea-original_p...ial-dataset.csv

Upload Complete

Run

Download

Visualize Network ▶

Run on Sample Data

Figure 2. fKSEA page snapshot

(A) The Input File Format section describes the input .CSV file format for fKSEA input.

(B) The Adjust Parameters section enables users to adjust parameters based on their needs before running the FGSEA algorithm.

(C) The Upload Data & Run analysis section enables users to upload their input file from their local directory, run fKSEA, download the generated KSI network file, and visualize the generated network.

enrichment and interactive visual network exploration into a single, unified solution.

The overall architecture and main functional components of PhosNetVis are represented in Figure 1. Briefly, to infer a network of significant KSIs from a phosphoproteomics study, users input their differential phosphorylation data (log2 fold change [log2FC]) into the fast kinase-substrate enrichment algorithm (fKSEA) interface of PhosNetVis. For fKSEA, PhosNetVis uses the fast gene set enrichment (FGSEA) algorithm,²¹ which allows having more permutations and thereby more fine-grained *p* values than using standard multiple hypothesis correction methods. The FGSEA process generates a list of KSIs with their associated Benjamini and Hochberg (BH) corrected *p*-values, with a KSI considered significant if its BH-adjusted *p* value is less than or equal to a user-defined cutoff. Users have the flexibility to adjust the input parameters and labels as needed.

The fKSEA page generates KSI network file(s) in .CSV format, which can then be explored in the network visualization page. On the network visualization page, users can visually explore the KSI network(s), examine phosphorylation sites and their states, query for specific proteins, adjust significance thresholds for differential phosphorylation, pan, zoom, and animate network changes across different states, such as multiple treatments or time points. Additionally, PhosNetVis allows for the direct upload of users' own phosphoproteomics network data, whether custom built or from other kinase enrichment tools. This feature enables users to leverage the tool's interactive visual exploration functionalities independent of its kinase enrichment capabilities. Alternatively, users can also download the results of their fKSEA runs to visually explore the resulting network data files using other tools.

fKSEA page

A snapshot of the PhosNetVis fKSEA interface is shown in Figure 2 (descriptive text not shown). This page has three sections: input file format (Figure 2A), adjust parameters (Figure 2B), and upload data and run analysis (Figure 2C).

fKSEA input file format

Prior to using the fKSEA interface, users should process their proteomics data for differential phosphorylation analysis be-

tween two states (e.g., baseline vs. perturbed) to obtain log2FC (perturbed/baseline) phosphorylation values and associated *p* values. Figure 2A illustrates the format of a sample input file, with mandatory fields marked by an asterisk.

A sample fKSEA input file is also provided in Table S1. An input file should be in .CSV format and include at least two columns:

- protein accession ID (identifies the protein) and
- log2FC differential phosphorylation value at a phosphosite on that protein.

Optionally, users can include additional columns:

- unique phosphosite ID for when multiple phosphosites are involved (e.g., phosphoSitePlus ID, residue ID, or any other custom ID) or
- *p* value-associated with the differential phosphorylation fold change.

fKSEA input parameters

Users can adjust the default parameters of the FGSEA algorithm²¹ through the "Adjust Parameter" section (Figure 2B). These parameters include

- differential phosphorylation *p* value cutoff (default = 0.05),
- output BH-adjusted *p* value cutoff (default = 0.05),
- minimum size of a KSI set to test (minSize) (default = 15),
- maximum size of a KSI set to test (maxSize) (default = 100), and
- boundary for calculating the *p* value (eps) (default = 0).

For protein labels, users can choose between UniProt accession IDs or Human Genome Organisation (HUGO) gene IDs. After uploading data and customizing parameters, the fKSEA pipeline maps the input data to KSIs within the PhosphoSitePlus database¹¹ (version downloaded on Monday, October 28, 2024, from the database update of Thursday, October 17, 2024, 11:21:06 EDT) to identify KSIs that pass user-defined significance thresholds. Additionally, each kinase is assigned an enrichment score by the FGSEA algorithm,²¹ which has been demonstrated in previous studies to effectively estimate kinase activities.²²

fKSEA output

After fKSEA runs, the resulting KSI network is output into a .CSV file. Users are then prompted with a success message that redirects them to download the KSI network connectivity data and/or to directly visualize the network on the network

FORMAT AND UPLOAD YOUR DATA

KinaseID (Required)	TargetID (Required)	KinaseSize (Optional)	KinaseActivity (Optional)	EdgeWeight (Optional)	EdgeHue (Optional)	TargetSize (Optional)	PhosphoSiteID (Optional)*	log2FC (Optional) (Phosphorylation)	pValue (Optional) (Phosphorylation)
CDK1	NUCKS1	1.2	0.16	1	1.44	0.9	S19	-0.1	0.05
CDK1	NUCKS1	1.2	0.16	1	1.44	0.9	T202	-4.32	0.006
CDK1	CDK1	1.2	0.16	1.5	4.28	1.2	T161	1.47	0.12
CDK1	NPM1	1.2	0.16	3.8	0.07	1.12	T78	2	0.0235
CSNK2A1	TOP2A	4.56	-2.27	1.67	2	2.96	S29	-3.059	0.03
MAPK3	TOP2A	1	1.2	2.28	1.5	2.98	K431	-2.51	0.01

*PhosphositePlusID or Residue ID

If you have more than one dataset, select the number of datasets and make sure to upload all of them!

Number of Datasets: No file chosen

UPLOAD

or

FAST KINASE-SUBSTRATE ENRICHMENT ANALYSIS

Figure 3. Snapshot of the network visualization input page (descriptive text not shown)

This page allows users to upload network files for visualization. Users can upload their custom network files or customize those generated by the fKSEA page. For comparative analyses across networks, users can upload multiple files by selecting the number of datasets they want to upload. This page also displays a table that provides the required network data format. To visualize network data, users need a .CSV file with at least two columns: (1) kinases (KinaseID) and (2) target nodes (TargetID) for substrates. Optionally, users can add columns to customize the node and edge attributes.

visualization page. A sample fKSEA output file is provided in [Table S2](#).

KSI network visualization

To visualize the KSI networks, users can either go directly from the fKSEA page to the network visualization page or skip the fKSEA step and input their datasets directly for network visualizations on the network data input page.

Network data input page

This page allows users to input one or more phosphoproteomics network dataset files corresponding to different cellular states (e.g., time, treatment, exposure) for KSI network visualizations. It also includes descriptive text and a formatting guide for optional customizations of the network input files. A snapshot of this page is provided in [Figure 3](#) (descriptive text on its website not shown).

Network data input file format

A network input file should include at least two columns: kinases (kinase ID) and substrates (target IDs). A sample input file format is also provided in [Table S3](#). Users can optionally include additional attributes to customize their network visualizations according to their needs. These optional attributes include the following:

- KinaseSize: node size associated with each kinase
- KinaseActivity: kinase node color representing positive or negative direction of kinase activity
- EdgeWeight: custom thickness for each edge
- EdgeHue: edge hue changes between user-defined minimum and maximum values.
- TargetSize: node size associated with each substrate
- PhosphoSiteID: unique ID of any phosphosite users deem important to visualize
- log2FC: phosphosite color representing the log fold change in phosphorylation at that site
- pValue: associated *p* value of phosphorylation

These attributes allow users to tailor their visualizations for better clarity and insight depending on their needs. However, please note that additional customizations are not required to visually explore KSI networks generated by the fKSEA page, as

these can be visualized directly within PhosNetVis. In addition, if users prefer to upload their own KSI networks, as in the CPTAC networks described under Illustrative examples, the minimum requirement of PhosNetVis is a simple .CSV file that includes two columns: KinaseID and TargetID. Furthermore, while the interactive network exploration page is optimized for KSI networks, by design it can accommodate visual explorations of any directed biomolecular network as long as edge directionality for the source and target nodes is provided by using the KinaseID and Target ID column names.

KSI network visualization interface

Once the KSI network files are generated in the fKSEA page or uploaded directly through the network data input page, the user is directed to the interactive network visualization interface. The PhosNetVis interactive interface is user-optional 2D or 3D. For a comparative understanding of the 2D versus 3D layouts, [Figure 4](#) provides two snapshots of the same KSI network generated with PhosNetVis: [Figure 4A](#) shows a rendering in 3D layout, and [Figure 4B](#) shows it in 2D layout. The 2D view also includes node labels. This KSI network visualization interface features three components.

The interactive control panel ([Figures 4A and 4B](#), left corner) allows users to seamlessly transition between 2D and 3D network layouts, query for specific nodes, dynamically adjust node color thresholds based on phosphorylation levels and fold change criteria, customize the background color, toggle the visibility of phosphorylation site partitions and node labels, enter or exit full-screen mode, capture screenshots, and restore the network visualization to default settings. Additionally, users can perform gene enrichment analyses by sending their list of genes in the network to the Enrichr tool^{23–25} with a simple click, enabling more in-depth biological insights.

Users can interact with the 2D or 3D visualized network ([Figures 4A and 4B](#), center) by rotating, zooming in/out, dragging nodes to different positions, or selecting nodes or edges. Substrates are represented as cylinders. The height of each cylinder depends on the number of phosphosites harbored by the substrate it represents, with each differentially phosphorylated phosphosite depicted as a slice of the cylinder. The color of each cylinder slice indicates its level of differential

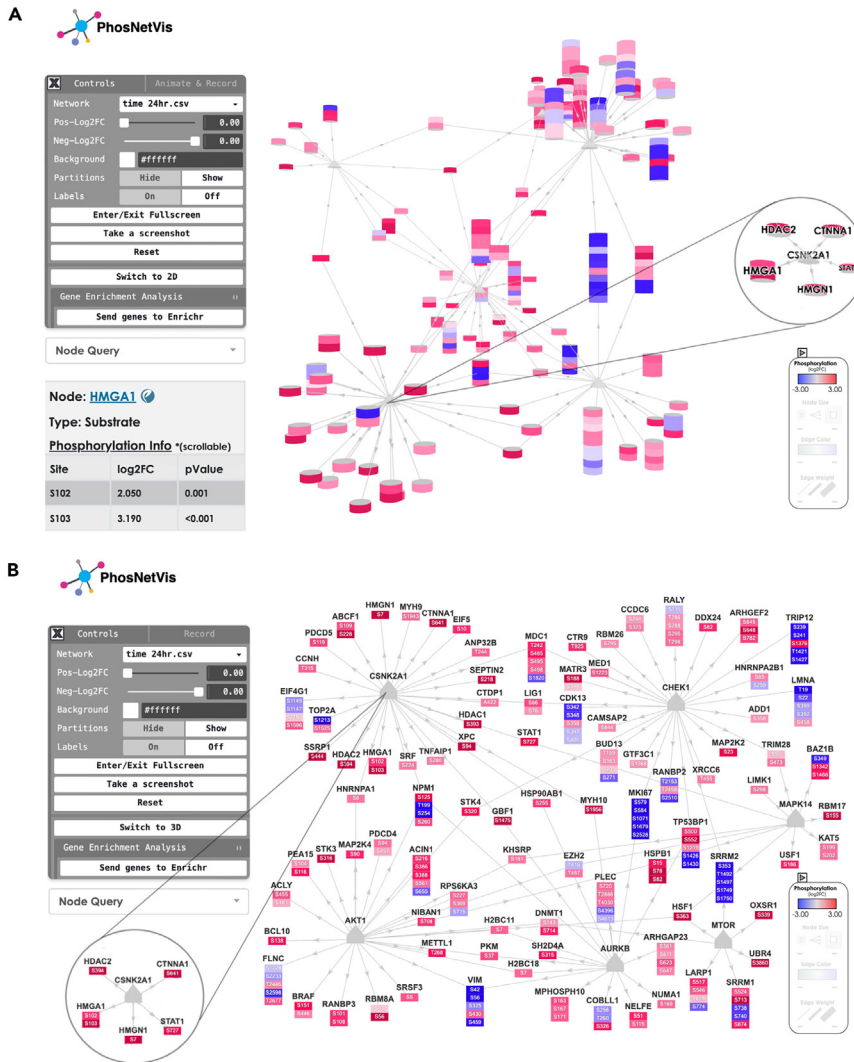


Figure 4. Snapshots of the network visualization interface featuring the phosphorylation landscape of SARS-CoV-2 infection

This page allows users to view and interact with the network curated from the global phosphorylation landscape of SARS-CoV-2 infection (Bouhaddou et al.). Users can rotate the network, zoom in/out, pan through, see node information by double-clicking on it, or reset the view by double-clicking on the background. Additionally, it provides options for the user to switch between 3D or 2D networks. Users can also send their gene list to the Enrichr tool^{23–25} for gene enrichment analyses and further exploration of enriched pathways and functional gene groups. Double-clicking on any node opens an interactive table that displays detailed information about the node, as shown in (A), bottom left, for node *HMGA1*.

(A) Snapshot of the 3D view of the network, including a magnified view of a *CSNK2A1*-central subnetwork in 3D. This subnetwork highlights phosphorylation sites known to be associated with cytoskeleton remodeling, filtered from the main network.

(B) Snapshot of the 2D view of the same network. The left corner shows the control panel, the center displays the network, and the bottom right corner contains the legend. The starred section indicates the *CSNK2A1* phosphorylation subnetwork, including a magnified view of the same *CSNK2A1* node with select interactors in a subnetwork in 2D. These features enable comprehensive exploration and analysis of the phosphorylation events during SARS-CoV-2 infection.

The legend (Figures 4A and 4B, bottom right corner) explains the user-selected range of phosphorylation log fold change values. Each cylinder slice

representing a phosphosite is colored based on the maximum negative and positive differential phosphorylation levels in the study. Lighter hues indicate smaller values, while darker hues correspond to larger values, with blue representing the most negative and red the most positive differentially phosphorylated phosphosite. If optional parameters were included in the input file, then the legend also shows the range for these attributes, such as node size, edge color, and edge weight (edge thickness). In the given example (Figures 4A and 4B), the input file includes only fold change information, so the legend displays only the range of differential fold change of phosphorylation.

Double-clicking on a node either in 2D or 3D configuration opens an interactive pop-up node details table (Figure 4A, bottom left corner) displaying detailed information about the node. This includes the node type (kinase or substrate) and, for each differentially phosphorylated site on that node, its position, differential phosphorylation fold change value (perturbation/baseline), and differential phosphorylation *p* value. Additionally, it includes hyperlinks to the PhosphoSitePlus database¹¹ for more detailed protein information.

representing a phosphosite is colored based on the maximum negative and positive differential phosphorylation levels in the study. Lighter hues indicate smaller values, while darker hues correspond to larger values, with blue representing the most negative and red the most positive differentially phosphorylated phosphosite. If optional parameters were included in the input file, then the legend also shows the range for these attributes, such as node size, edge color, and edge weight (edge thickness). In the given example (Figures 4A and 4B), the input file includes only fold change information, so the legend displays only the range of differential fold change of phosphorylation.

This setup ensures that users can thoroughly explore and analyze their KSI networks with a high degree of interactivity and customization.

Network visualization output file format

Users can take snapshots of their network visualizations or screen-record animations of differential phosphorylation changes across multiple states (from multiple input files). Snapshots are saved in .PNG file format. Screen recordings are saved in .MP4 file format.

Illustrative examples

We illustrate the functionalities of the PhosNetVis KSI network visualization interface with two case studies. Both are available for interactive exploration within the tool web portal. The portal also includes several toy datasets customized to help users get familiar with different KSI network attributes. By exploring these examples, users can visually interact with the networks, understand tool features, and experiment with further customizations of their datasets.

SARS-CoV-2 KSI network

This example is curated from a study that examined the global phosphorylation landscape of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection.²⁶ In this study, Vero-E6 cells, an African green monkey cell line, were infected with SARS-CoV-2, the virus responsible for COVID-19. Cells were harvested at various time points post infection (2, 4, 8, 12, and 24 h). Phosphoproteomics analysis using LC-MS quantified 4,624 human-orthologous phosphorylation sites across 3,036 human-orthologous proteins. A screenshot of this KSI network at the 24 h time point is shown in Figure 4A in 3D and Figure 4B in 2D.

One significant finding in this study was the increased activity of casein kinase II (*CSNK2A1*) following infection. This was evidenced by a marked increase in the abundance of known *CSNK2A1* phosphorylation sites post infection. *CSNK2A1* was localized within filopodium protrusions that emerged from the cell surface during infection, suggesting a role in filopodium formation. Further analysis of phosphorylation sites affecting cytoskeletal changes revealed increased phosphorylation of *CTNNA1*, *HDAC2*, *HMG1A1*, *HMG1N1*, and *STAT1* proteins at sites known to be associated with cytoskeleton remodeling. Specifically, phosphorylation was observed at *CTNNA1*:S641, *HDAC2*:S394, *HMG1A1*:S102-103, *HMG1N1*:S7, and *STAT1*:S727. Zooming in on *CSNK2A1* in this network and inspecting its substrates and known phosphosites visually confirms this finding. These interactions were identified by running the fKSEA algorithm on the phosphoproteomics data and are depicted in the magnified subnetwork of *CSNK2A1* (Figures 4A and 4B, magnified circles). Clicking on one of its substrates, *HMG1A1*, brings up a pop-up table of its differentially phosphorylated phosphosites, S102 and S103, along with their log2FC (infection/mock) and *p* values (Figure 4A, bottom left). For further details on *HMG1A1*, the table links to PhosphoSitePlus.¹¹

This KSI network is available for interactive exploration in PhosNetVis at <https://gumuslab.github.io/PhosNetVis/existing-networks.html>. Users can inspect the phosphorylation status of the network at a time point of interest by simply using the drop-down menu in the control panel (Figures 4A and 4B, top left), toggle across different time points, or inspect an animation of the changes in phosphorylation within the KSI network over time by selecting the Animate & Record tab from the control panel. These functionalities help users pinpoint potential mechanisms of interest for further studies, providing a dynamic and detailed view of the KSI network and the phosphorylation events during SARS-CoV-2 infection.

CPTAC pan-cancer immune subtype KSI networks

Recently, through CPTAC, we investigated the immune landscape of over 1,000 tumors from 10 different cancer types.²⁰ This effort aimed to enhance the understanding of immune cell

surveillance mechanisms and the various strategies tumors use to evade immune responses. Using CPTAC's comprehensive pan-cancer proteogenomic data,²⁷ this study identified and characterized seven unique immune subtypes, and by analyzing kinase activities within these subtypes, it uncovered potential therapeutic targets specific to each subtype. Here, we provide the full catalog of PhosNetVis interactive visualizations of KSI networks of each immune subtype both as a resource for the research community and as use case examples. This PhosNetVis application skipped the fKSEA page. Instead, the networks were derived from simply mapping the list of proteins (both kinases and substrates) in each immune subtype to the PhosphoSitePlus database¹¹ to derive their connecting edges between the kinases and the phosphosites in the CPTAC immune study.²⁰ Then, the adjacency matrices of all immune subtype KSI networks were directly uploaded into PhosNetVis using its Network Data Input page to generate their interactive visualizations.

The PhosNetVis catalog of these networks allows easy queries, interactive visualizations, exploration, and download and is available through the PhosNetVis Existing Networks link at <https://gumuslab.github.io/PhosNetVis/cptac-vis.html>. Using PhosNetVis, users can comparatively analyze KSIs and differential phosphorylation across these 7 immune subtypes on the fly. For example, Figure 5 illustrates two of the largest connected network snapshots in a 2D layout. Figure 5A shows the most active pan-cancer immune subtype (CD8+/Interferon Gamma [IFNG]+), which leads to the upregulation of Protein Kinase cAMP-activated Catalytic subunit Alpha (PRKACA) and the phosphorylation of its substrates. Figure 5B depicts the least active cluster (CD8-/IFNG-), resulting in the upregulation of cell-cycle kinases such as CDK1 and the phosphorylation of its substrates.

These CPTAC networks provide good case studies on how users can interact with the networks in 2D or 3D layout and then directly embed their network snapshots in 2D into publications, as we illustrate in Figure 5. The drag function allows users to easily modify the network layout, emphasizing specific parts of the network as needed. In Figures 5A and 5B, quick manual alterations in the network connectivities and layouts highlight the main kinases more clearly. These visualizations enable researchers to pinpoint potential mechanisms of interest for follow-up studies. Furthermore, the catalog of these pan-cancer immune subtype KSI networks enables the broader community of researchers to explore complex KSI network datasets from the CPTAC initiative.

DISCUSSION

Large-scale phosphoproteomics experiments are characterizing protein phosphorylation sites and states across conditions to better understand cellular signaling in health and disease. While numerous tools are dedicated to inferring kinases from these datasets, there is a need for integrated visualization tools that allow users to explore the resulting KSI networks across multiple conditions. Here, we introduced PhosNetVis, an interactive web-based tool that enables users to perform fKSEA and then to visually explore, interpret, and communicate the inferred KSI networks and their associated phosphorylation data across multiple states. PhosNetVis provides a private

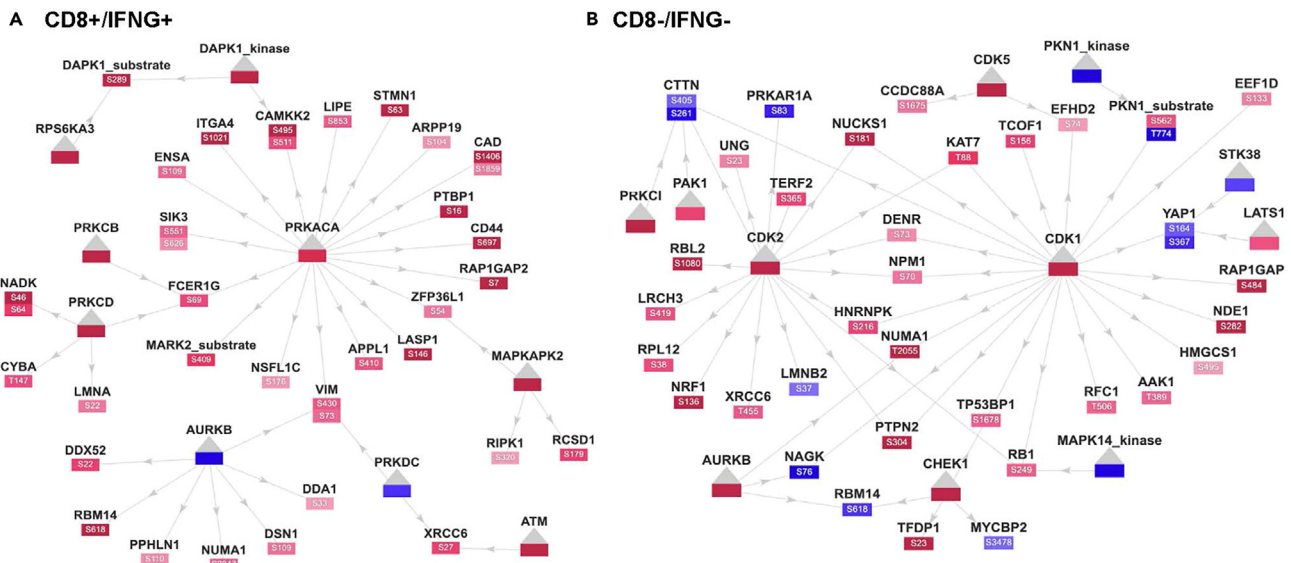


Figure 5. Snapshots of interactive 2D network visualizations for CPTAC pan-cancer immune subtypes

(A) Kinases and their specific altered phosphosites, phosphorylation levels, and positions in each altered substrate are clearly shown in the CD8+/IFNG+ network. The visualization clearly depicts upregulation of the global abundance of the kinase PRKACA (center) and the phospho-abundance of its substrates in red. (B) The CD8-/IFNG- subnetwork. The global proteomic expression of cell-cycle kinases such as CDK1 and the phospho-abundance of its substrates are upregulated and shown in red.

analysis environment, as all operations during KSI network visualization take place exclusively within the user’s local browser without relying on an external server or third-party site. In the case of fKSEA, datasets undergo processing on the protected RStudio Connect server at Mount Sinai (<https://rstudio-connect.hpc.mssm.edu>) to guarantee data privacy. By integrating fKSEA and network visualization in a single platform, PhosNetVis simplifies the workflow for analyzing phosphoproteomics data, making it accessible to users of varying computational proficiency levels.

PhosNetVis features an easy-to-navigate graphical user interface for KSI networks that visually represents phosphorylation sites and their respective states across multiple experiments. Users can explore KSI network visualizations, switch between different network states (i.e., time points or treatments), toggle between 2D or 3D visualizations, and adjust network parameters. The tool allows users to query nodes of interest for detailed phosphosite information, change differential phosphorylation parameters, rotate, pan, zoom in/out, drag nodes, download the network snapshots, or record animations of network changes over time or conditions. While PhosNetVis is specifically designed for phosphoproteomics analysis, its adaptable architecture extends its utility to various biomolecular network applications even when phosphorylation data are unavailable. This adaptability broadens its potential impact across diverse research problems, enabling researchers to visualize and interpret complex biological data in a dynamic and interactive manner. One limitation is that while fKSEA analysis employs the most recent version of the most popularly used kinase-substrate database, PhosPhoSitePlus,¹¹ the upstream kinases of a majority of phosphosites are still unknown. Furthermore, 20% of kinases that are known are upstream of the majority of known

phosphosites.²⁸ This, however, is a limitation within the field and an active area of research with ongoing studies to reveal the full extent of human KSIs.

Overall, PhosNetVis lowers the barriers between complex phosphoproteomics data and researchers who want rapid, intuitive, and high-quality tools to infer kinase activity and thereby visually explore KSI networks at multiple phosphorylation sites and states. This will empower investigators to translate rich datasets into biological insights and clinical applications. PhosNetVis is freely accessible at <https://gumuslab.github.io/PhosNetVis>. Its website hosts detailed tutorials, an FAQ page, and a variety of use case examples, including the full catalog of different KSI networks within 7 different tumor immune subtypes derived from a pan-cancer analysis of more than 1,000 CPTAC tumors across 10 different cancers.

METHODS

Website implementation

fKSEA page

We developed a user-friendly interface for fKSEA using hyper-text markup language (HTML), cascading style sheets (CSS), and Bootstrap 4. FGSEA is performed using a Plumber application programming interface (API) (<https://www.rplumber.io/>) deployed on the RStudio Connect server at Mount Sinai (<https://rstudio-connect.hpc.mssm.edu>). Users upload their .CSV files and initiate the analysis via a simple interface, which sends an API request to the server to perform FGSEA. This setup ensures efficient and seamless processing of kinase enrichment analysis, making it accessible and convenient for researchers. Once the fKSEA API call is finished, users are directed to download the network data files or to directly visualize the files.

Network data input page

Users can upload one or more network data files through this page. Network data files can be the outputs from PhosNetVis fKSEA page analysis or custom generated. The page provides detailed guidelines for network data input formats for custom-generated files. For .CSV file parsing and data processing, we integrated the PapaParse library (<https://www.papaparse.com/>) to enable users to concurrently upload and process multiple .CSV files. Once network data .CSV files are uploaded, PhosNetVis transforms each input .CSV file into the JavaScript object notation format using the Danfo.JS package ([danfo.jsdata.org](https://github.com/danfocorp/danfo.js)).

Interactive network visualization page

Once network data files are uploaded, all networks are visualized on the network visualization page. We built this interface mainly with HTML and CSS. We used JavaScript for the document object model element manipulations and the Bootstrap library to implement responsive layouts and to customize page architectures (getbootstrap.com). To create a screen-resizable responsive canvas where the 3D network visualization is rendered, we integrated the element-resize-detector library (github.com/wnr/element-resize-detector).²⁹

To prioritize user interactions and control, we included several graphical user interface (GUI) elements in the network visualization interface (Figures 4A and 4B). Briefly, to enable users to interactively modify their visualizations, adjust parameters, and conduct node queries, we implemented a GUI control panel by using the Tweakpane.js library ([cocopon.github.io/tweakpane/](https://github.com/cocopon/tweakpane)) (Figures 4A and 4B, top left corner). To enhance user interaction through specific node queries, we used the Fstdropdown.js library (github.com/VirtusX/fstdropdown). To offer users a 3D network visualization experience (Figure 4A), we used the 3d-force-graph library (github.com/vasturiano/3d-force-graph) to render the network layouts in 3D. The library leverages the WebGL-based Three.js (threejs.org) library to create interactive force-directed graphs in 3D space. Users can easily customize node colors and labels in 3D, for which we used RainbowVis-JS (github.com/anomal/RainbowVis-JS) and three-spritetext (github.com/vasturiano/three-spritetext) libraries, respectively. In addition to 3D, PhosNetVis also offers a 2D network visualization option (Figure 4B). To develop the 2D network visualization option, we implemented the force-graph library (github.com/vasturiano/force-graph), which uses a force-directed layout algorithm to effectively present network layouts on an HTML5 canvas. Finally, for users to export visualizations, we implemented jsScreenRecorder (github.com/manan657/jsScreenRecorder), a custom JavaScript script designed to capture and record screen interactions, which is particularly beneficial after rendering the networks in 3D. By implementing these tools and libraries, PhosNetVis ensures a comprehensive and user-friendly network visualization experience, supporting researchers in their exploration of KSI networks and associated phosphorylation data.

Tutorials, interactive examples, and toy datasets

To guide users, the documentation platform of PhosNetVis (<https://gumuslab.github.io/phosnetvis-docs/>) features tutorials that cover fKSEA and network visualization as well as instructions for input data formatting. In addition, we provide several

interactive examples and toy datasets for network visualization at <https://gumuslab.github.io/PhosNetVis/existing-networks.html>. All input files, including those for kinase substrate enrichment analysis are provided at: <https://github.com/GumusLab/PhosNetVis-DataExamples>.

RESOURCE AVAILABILITY

Lead contact

Zeynep H. Gümüş can be reached by e-mail (zeynep.gumus@mssm.edu).

Materials availability

PhosNetVis is hosted on GitHub pages and is freely available with no login requirements at <https://gumuslab.github.io/PhosNetVis>. For additional materials, please contact the [lead contact](#).

Data and code availability

PhosNetVis source code is available on the GitHub repository (github.com/GumusLab/PhosNetVis; <https://doi.org/10.5281/zenodo.14215570>),³⁰ is released under GNU's Not Unix Affero General Public License version 3.0 (GNU AGPL-3.0) (<https://www.gnu.org/licenses/agpl-3.0.en.html>), and is also available under a commercial license for enterprises seeking additional features or avoiding AGPL obligations.

ACKNOWLEDGMENTS

Z.H.G. gratefully acknowledges support from NIH R33 CA263705-01; S.G. from NIH U24 CA224319, U01 DK124165, and P01 CA196521; and M.B. from NIH K99 AI163868. The authors gratefully acknowledge valuable feedback from investigators within the NIH-funded Human Immunology Project Consortium (HIPC) and Clinical Proteomic Tumor Analysis Consortium (CPTAC), with special thanks to Francesca Petralia for kindly sharing the CPTAC pan-cancer immune subtype networks with the study team.

AUTHOR CONTRIBUTIONS

O.R., writing – review & editing, visualization, software, and investigation; B.T., writing – original draft, writing – review & editing, visualization, software, and investigation; I.F.P., writing – original draft, visualization, software, and investigation; S.C., writing – original draft, visualization, software, and investigation; S.G., writing – review and visualization; S.K., visualization and software; J.J., writing – original draft, writing – review & editing, visualization, software, and conceptualization; M.B., writing – review and visualization; Z.H.G., writing – original draft, writing – review & editing, conceptualization, visualization, and supervision.

DECLARATION OF INTERESTS

S.G. reports other research funding from Boehringer-Ingelheim, Bristol-Myers Squibb, Celgene, Genentech, Regeneron, and Takeda and consulting for Taiho Pharmaceuticals not related to this study.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.patter.2024.101148>.

Received: June 20, 2024

Revised: November 12, 2024

Accepted: December 11, 2024

Published: January 10, 2025

REFERENCES

- Hallal, M., Braga-Lagache, S., Jankovic, J., Simillion, C., Bruggmann, R., Uldry, A.-C., Allam, R., Heller, M., and Bonadies, N. (2021). Inference of kinase-signaling networks in human myeloid cell line

- models by Phosphoproteomics using kinase activity enrichment analysis (KAEA). *BMC Cancer* 21, 789. <https://doi.org/10.1186/s12885-021-08479-z>.
2. Piersma, S.R., Valles-Marti, A., Rolfs, F., Pham, T.V., Henneman, A.A., and Jiménez, C.R. (2022). Inferring kinase activity from phosphoproteomic data: Tool comparison and recent applications. *Mass Spectrom. Rev.* 43, 725–751. <https://doi.org/10.1002/mas.21808>.
 3. Kuleshov, M.V., Xie, Z., London, A.B.K., Yang, J., Evangelista, J.E., Lachmann, A., Shu, I., Torre, D., and Ma'ayan, A. (2021). KEA3: improved kinase enrichment analysis via data integration. *Nucleic Acids Res.* 49, W304–W316. <https://doi.org/10.1093/nar/gkab359>.
 4. Wiredja, D.D., Koyutürk, M., and Chance, M.R. (2017). The KSEA App: a web-based tool for kinase activity inference from quantitative phosphoproteomics. *Bioinformatics* 33, 3489–3491. <https://doi.org/10.1093/bioinformatics/btx415>.
 5. Ochoa, D., Jonikas, M., Lawrence, R.T., El Debs, B., Selkrig, J., Typas, A., Villén, J., Santos, S.D., and Beltrao, P. (2016). An atlas of human kinase regulation. *Mol. Syst. Biol.* 12, 888. <https://doi.org/10.15252/msb.20167295>.
 6. Yılmaz, S., Ayati, M., Schlatzer, D., Çiçek, A.E., Chance, M.R., and Koyutürk, M. (2021). Robust inference of kinase activity using functional networks. *Nat. Commun.* 12, 1177. <https://doi.org/10.1038/s41467-021-21211-6>.
 7. Casado, P., Rodríguez-Prados, J.-C., Cosulich, S.C., Guichard, S., Vanhaesebroeck, B., Joel, S., and Cutillas, P.R. (2013). Kinase-substrate enrichment analysis provides insights into the heterogeneity of signaling pathway activation in leukemia cells. *Sci. Signal.* 6, rs6. <https://doi.org/10.1126/scisignal.2003573>.
 8. Casado, P., Hijazi, M., Gerdes, H., and Cutillas, P.R. (2022). Implementation of Clinical Phosphoproteomics and Proteomics for Personalized Medicine. *Methods Mol. Biol.* 2420, 87–106. https://doi.org/10.1007/978-1-0716-1936-0_8.
 9. Drake, J.M., Paull, E.O., Graham, N.A., Lee, J.K., Smith, B.A., Titz, B., Stoyanova, T., Faltermeier, C.M., Uzunangelov, V., Carlin, D.E., et al. (2016). Phosphoproteome Integration Reveals Patient-Specific Networks in Prostate Cancer. *Cell* 166, 1041–1054. <https://doi.org/10.1016/j.cell.2016.07.007>.
 10. Dinkel, H., Chica, C., Via, A., Gould, C.M., Jensen, L.J., Gibson, T.J., and Diella, F. (2011). Phospho.ELM: a database of phosphorylation sites—update 2011. *Nucleic Acids Res.* 39, D261–D267. <https://doi.org/10.1093/nar/gkq1104>.
 11. Hornbeck, P.V., Zhang, B., Murray, B., Kornhauser, J.M., Latham, V., and Skrzypek, E. (2015). PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res.* 43, D512–D520. <https://doi.org/10.1093/nar/gku1267>.
 12. Keshava Prasad, T.S., Goel, R., Kandasamy, K., Keerthikumar, S., Kumar, S., Mathivanan, S., Telikicherla, D., Raju, R., Shafreen, B., Venugopal, A., et al. (2009). Human Protein Reference Database—2009 update. *Nucleic Acids Res.* 37, D767–D772. <https://doi.org/10.1093/nar/gkn892>.
 13. Bateman, A., Martin, M.-J., Orchard, S., Magrane, M., Agivetova, R., Ahmad, S., Alpi, E., Bowler-Barnett, E.H., Britto, R., Bursteinas, B., et al. (2021). UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.* 49, D480–D489. <https://doi.org/10.1093/nar/gkaa1100>.
 14. Johnson, J.L., Yaron, T.M., Huntsman, E.M., Kerelsky, A., Song, J., Regev, A., Lin, T.-Y., Liberatore, K., Cizin, D.M., Cohen, B.M., et al. (2023). An atlas of substrate specificities for the human serine/threonine kinome. *Nature* 613, 759–766. <https://doi.org/10.1038/s41586-022-05575-3>.
 15. Liliushvili, V., Kalayci, S., Fluder, E., Wilson, M., Gabow, A., and Gümüş, Z.H. (2017). iCAVE: an open source tool for visualizing biomolecular networks in 3D, stereoscopic 3D and immersive 3D. *GigaScience* 6, 1–13. <https://doi.org/10.1093/gigascience/gix054>.
 16. Kalayci, S., Petralia, F., Wang, P., and Gümüş, Z.H. (2020). ProNetView-ccRCC: A Web-Based Portal to Interactively Explore Clear Cell Renal Cell Carcinoma Proteogenomics Networks. *Proteomics* 20, e2000043. <https://doi.org/10.1002/pmic.202000043>.
 17. Petralia, F., Tignor, N., Reva, B., Koptyra, M., Chowdhury, S., Rykunov, D., Krek, A., Ma, W., Zhu, Y., Ji, J., et al. (2020). Integrated Proteogenomic Characterization across Major Histological Types of Pediatric Brain Cancer. *Cell* 183, 1962–1985.e31. <https://doi.org/10.1016/j.cell.2020.10.044>.
 18. Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* 13, 2498–2504. <https://doi.org/10.1101/gr.1239303>.
 19. Legeay, M., Doncheva, N.T., Morris, J.H., and Jensen, L.J. (2020). Visualize omics data on networks with Omics Visualizer, a Cytoscape App. *F1000Res.* 9, 157. <https://doi.org/10.12688/f1000research.22280.2>.
 20. Petralia, F., Ma, W., Yaron, T.M., Caruso, F.P., Tignor, N., Wang, J.M., Charytonowicz, D., Johnson, J.L., Huntsman, E.M., Marino, G.B., et al. (2024). Pan-cancer proteogenomics characterization of tumor immunity. *Cell* 187, 1255–1277.e27. <https://doi.org/10.1016/j.cell.2024.01.027>.
 21. Korotkevich, G., Sukhov, V., Budin, N., Shpak, B., Artyomov, M.N., and Sergushichev, A. (2021). Fast gene set enrichment analysis. Preprint at bioRxiv. <https://doi.org/10.1101/060012>.
 22. Hernandez-Armenta, C., Ochoa, D., Gonçalves, E., Saez-Rodriguez, J., and Beltrao, P. (2017). Benchmarking substrate-based kinase activity inference using phosphoproteomic data. *Bioinformatics* 33, 1845–1851. <https://doi.org/10.1093/bioinformatics/btx082>.
 23. Xie, Z., Bailey, A., Kuleshov, M.V., Clarke, D.J.B., Evangelista, J.E., Jenkins, S.L., Lachmann, A., Wojciechowicz, M.L., Kropiwnicki, E., Jagodnik, K.M., et al. (2021). Gene Set Knowledge Discovery with Enrich. *Curr. Protoc.* 1, e90. <https://doi.org/10.1002/cpz1.90>.
 24. Kuleshov, M.V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S.L., Jagodnik, K.M., Lachmann, A., et al. (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* 44, W90–W97. <https://doi.org/10.1093/nar/gkw377>.
 25. Chen, E.Y., Tan, C.M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G.V., Clark, N.R., and Ma'ayan, A. (2013). Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinf.* 14, 128. <https://doi.org/10.1186/1471-2105-14-128>.
 26. Bouhaddou, M., Memon, D., Meyer, B., White, K.M., Rezelj, V.V., Correa Marrero, M., Polacco, B.J., Melnyk, J.E., Ulferts, S., Kaake, R.M., et al. (2020). The Global Phosphorylation Landscape of SARS-CoV-2 Infection. *Cell* 182, 685–712.e19. <https://doi.org/10.1016/j.cell.2020.06.034>.
 27. Li, Y., Dou, Y., Da Veiga Leprevost, F., Geffen, Y., Calinawan, A.P., Aguet, F., Akiyama, Y., Anand, S., Birger, C., Cao, S., et al. (2023). Proteogenomic data and resources for pan-cancer analysis. *Cancer Cell* 41, 1397–1406. <https://doi.org/10.1016/j.ccell.2023.06.009>.
 28. Needham, E.J., Parker, B.L., Burykin, T., James, D.E., and Humphrey, S.J. (2019). Illuminating the dark phosphoproteome. *Sci. Signal.* 12, eaau8645. <https://doi.org/10.1126/scisignal.aau8645>.
 29. Wiener, L., Ekholm, T., and Haller, P. (2015). Modular Responsive Web Design using Element Queries. Preprint at arXiv. <https://doi.org/10.48550/arXiv.1511.01223>.
 30. GumusLab (2024). Code for the Article “PhosNetVis: A Web-Based Tool for Fast Kinase-Substrate Enrichment Analysis and Interactive 2D/3D Network Visualizations of Phosphoproteomics Data (Zenodo). <https://doi.org/10.5281/zenodo.14215570>.