

# PhosProViz

A Web-based Platform for Interactive Visual Exploration of Phosphoproteomics Network Data



Mount  
Sinai

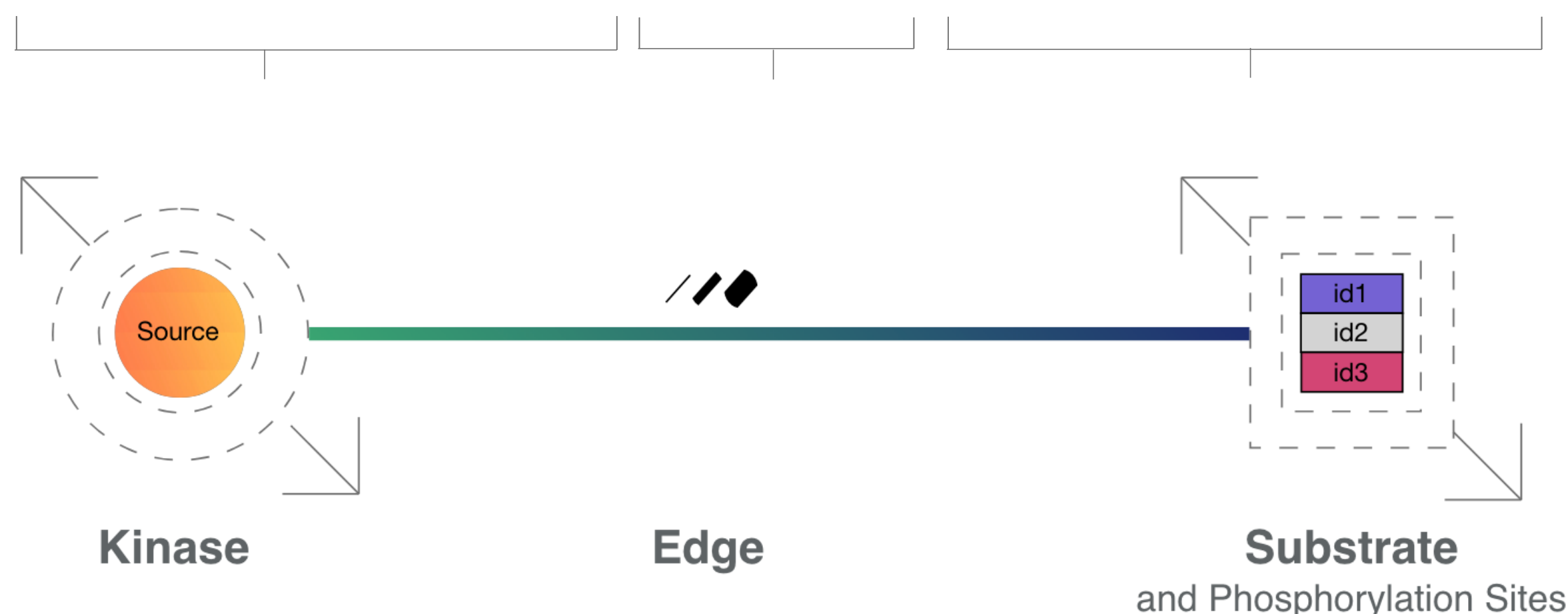
Shreya Chandrasekar<sup>1</sup>, Irene Font Peradejordi<sup>1</sup>, Berk Turhan<sup>2</sup>, Selim Kalayci<sup>3</sup>, Jeffrey Johnson<sup>3</sup>, Zeynep H. Gümüş<sup>3</sup>

CORNELL  
TECH

HOME OF THE JACOBS  
TECHNION-CORNELL  
INSTITUTE

<sup>1</sup>Jacobs Technion-Cornell Institute at Cornell Tech and Department of Information Science at Cornell University, United States, <sup>2</sup>Sabanci University, Turkey, <sup>3</sup>Icahn School of Medicine at Mount Sinai, United States

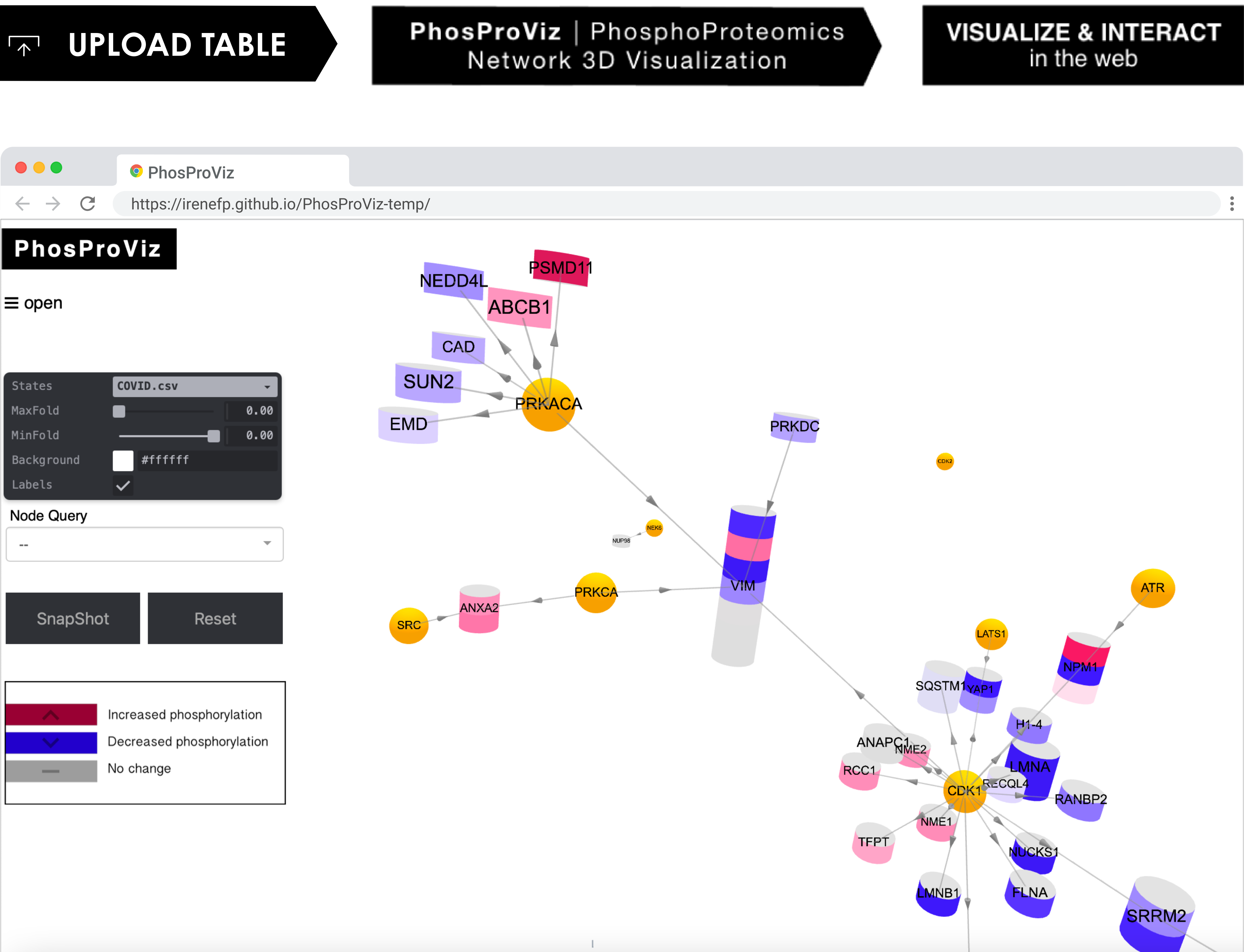
State	Source ID	Source Color	Source Size	Edge Color	Edge Weight	Target ID	Target Size	PhosphoSite ID	Target Color
TimeStamp0	A	-0.38	1.85	-2.44	1	Z	0.9	Z1	-0.1
TimeStamp0	A	-0.38	1.85	-2.44	1	Z	0.9	Z2	-4.32
TimeStamp0	A	-0.38	1.85	-2.44	1	Z	0.9	Z3	8.4
TimeStamp0	A	-0.38	1.85	0.07	3.8	Y	1.12	Y1	2
TimeStamp0	A	-0.38	1.85	0.07	3.8	Y	1.12	Y2	1.75
TimeStamp0	B	2.69	0.72	2	1.67	X	2.98	X1	0.01
TimeStamp0	B	2.69	0.72	2	1.67	X	2.98	X2	3.2
TimeStamp0	B	2.69	0.72	2	1.67	X	2.98	X3	-6.51
TimeStamp0	B	2.69	0.72	-5.1	0.03	W	1	W1	0.965



## Motivation and Background

To better understand the human immune system and how diseases influence cells and protein expression, researchers are profiling human cells before and after infection, vaccination, or treatment, collecting **massive \*omics data sets**. Some of these studies involve protein phosphorylation interactions and perturbations. **Protein phosphorylation** is a reversible post-translational modification and is a vital modification for cellular processes and signaling networks. Recent advances in mass spectrometry based quantitative proteomics have led to the rapid generation of **massive protein phosphorylation datasets at multiple states**. These are typically visualized as networks, with **kinases and substrates represented as nodes**, and their **interactions as edges** (see Figure 1). It is also important to visualize the states of all phosphorylation sites altered within each substrate.

Currently available visualization tools are not optimized for large and complex phosphoproteomics networks with their associated phosphorylation data. Furthermore, there are **currently no dedicated tools to generate, explore and share interactive visualizations** of these systems.



Link to Tool: [irenefp.github.io/PhosProViz-temp/](https://irenefp.github.io/PhosProViz-temp/)

## User Interface & Workflow

1. On the landing page of PhosProViz, users can either **directly upload their datasets** themselves; or they can **get help for formatting their data by answering several questions** on edge directionality; source color and size attributes; target color and size attributes; phosphorylation sites; and edge color and weight attributes.
2. After answering the questions, users can **inspect their data in PhosProViz input format**. A dialog window also includes guidelines to properly format and upload datasets.
3. **After successful completion of data upload**, users can simultaneously **explore multiple network datasets** that belong to **different states** (time point, treatment, other)

## Features & Technology

### Features

- **User-friendly**, intuitive web interface
- **Visualize** all network & phosphorylation data simultaneously
- **3D interactivity**, including pan, zoom, rotate, and drag nodes
- **Query** for certain kinases or substates of choice
- **Threshold** selection, where users can filter their network by applying different thresholds (e.g., log fold change values)
- **Annotations**, where each node contains a link that redirects to PhosphoSite
- **Modularity** which enables users to build more specialized interfaces
- **Easy accessibility** through a standard web browser

### Technology

- Users input network data as a **CSV file**
- PhosProViz utilizes multiple client-side **Javascript libraries** (e.g. **Three.js** and **3d-force-graph**)
- Code is **open-source** & the tool is publicly hosted on **GitHub Pages**

## Research in Progress

- **Uniprot to HUGO** annotation conversion feature for node IDs
- **Pre-populating kinases** from a public interactions dataset and creating the network only from user-uploaded substrate information
- **Animated transitions** for network graphs between different states
- **Color and size legends** that map color & size to numerical values

## References

We gratefully acknowledge funding from NIH/NIAID IOF Pilot under U19 grant # AI118610-01.

- [1] G. Su, J. H. Morris, B. Demchak and G. D. Bader, "Biological network exploration with Cytoscape 3," Current protocols in bioinformatics, vol. 47, 2014.
- [2] S. Kalayci, F. Petralia, P. Wang and Z. H. Gümüş, "ProNetView- ccRCC: A Web-Based Portal to Interactively Explore Clear Cell Renal Cell Carcinoma Proteogenomics Networks," Computational Proteomics: Focus on Deep Learning, vol. 20, no. 21-22, 2020.
- [3] F. Petralia, N. Tignor and B. Reva, "Integrated Proteogenomic Characterization across Major Histological Types of Pediatric Brain Cancer," Cell, vol. 183, no. 7, pp. 1962-1985, 2020.
- [4] D. E. Gordon, J. Hiatt and M. Bouhaddou, "Comparative host- coronavirus protein interaction networks reveal pan-viral disease mechanisms," Science, vol. 370, no. 6521, 2020.
- [5] D. Ochoa, M. Jonikas, R. T. Lawrence, B. El Debs, J. Selkrig, A. Typas, J. Villen, S. DM Santos and P. Beltrao, "An Atlas of Human Kinase Regulation," Molecular Systems Biology, vol. 12, p. 888, 2016.
- [6] P. V. Hornbeck, B. Zhang, B. Murray, J. M. Kornhauser, V. Latham and E. Skrzypek, "PhosphoSitePlus, 2014: mutations, PTMs and recalibrations," Nucleic Acids Res, vol. 43, no. D512-20, 2015.