PTMapper: a Cytoscape plugin for PTM site-oriented protein network analysis

User Manual

Yuta Narushima, Hiroko Kozuka-Hata, Kouhei Tsumoto, Jun-ichiro Inoue, Masaaki Oyama

1. Overview 1
2. Example data 1
3. Usage 2
3.1 Installation of PTMapper 2
3.2 Import protein-protein interaction network 3
3.3 Import kinase-specific phosphorylation site information via PTMapper 4
3.4 Annotation with the quantification data on proteins/PTM sites. 5
3.5 Extraction of sub-networks 5
4. References 8

1. Overview

Post translational modification (PTM)-dependent cellular signaling is known to play a diverse role in regulating multiple cellular processes such as proliferation, differentiation and apoptosis. PTM mapper (PTMapper) is a Cytoscape plugin for PTM-dependent protein interaction network analysis. Users can add the information on PTM sites and their upstream regulators to protein-protein interaction (PPI) networks using PTMapper. This document shows how to use PTMapper for construction of the phosphorylation site-oriented networks in Cytoscape based on quantitative phosphoproteome data. Furthermore, it introduces extraction of significantly regulated sub-networks from PTMapper-based high-resolution PPI networks using jActiveModules.

PTMapper has been tested in Cytoscape version 3.2.1 and 3.3.0 with Java 8 on 64-bit version of Windows 7 and 8.

2. Example data

As the example data to construct phosphorylation site-oriented networks, we provided the following four files generated from our previous phosphoproteome data (Kozuka-Hata *et al.*, 2012).

Download these files from https://www.github.com/y-narushima/PTMapper.

1. The file named “ppi_network_SIF.txt” contains the information on the proteins (Uniprot ID) and their interactions regarding the phosphoproteome data based on the PPI public database, Pathway Commons (Cerami *et al.*, 2011). This file is used to construct PPI networks.

2. The file named “phosphosite_kinase_data.txt” contains the information on the substrates (Uniprot ID), phosphorylation sites and their upstream kinases (Uniprot ID) regarding the phosphoproteome data, which is used to create PTM site-oriented networks via PTMapper. The detailed information on kinase-phosphorylation site relationships is based on the integrated dataset generated from four public phosphorylation site-related databases; PhosphoSitePlus (Hornbeck *et al.*, 2012), Phospho.ELM (Dinkel *et al.*, 2011), PhosphoNetworks (Hu *et al.*, 2014) and UniprotKB (Magrane *et al*, 2011).

3. The file named “all_phosphosite_data.txt” contains the information on the phosphorylated proteins (Uniprot ID) and their PTM sites regarding the phosphoproteome data, which is used to add the nodes for all quantified phosphorylation sites via PTMapper.

4. The file named “quantification_data.txt” contains the information on the phosphorylated proteins and their PTM sites (Uniprot ID, description) along with quantification data (log2[ratio], significance A) based on the phosphoproteome data, which is used to extract significantly regulated sub-networks.

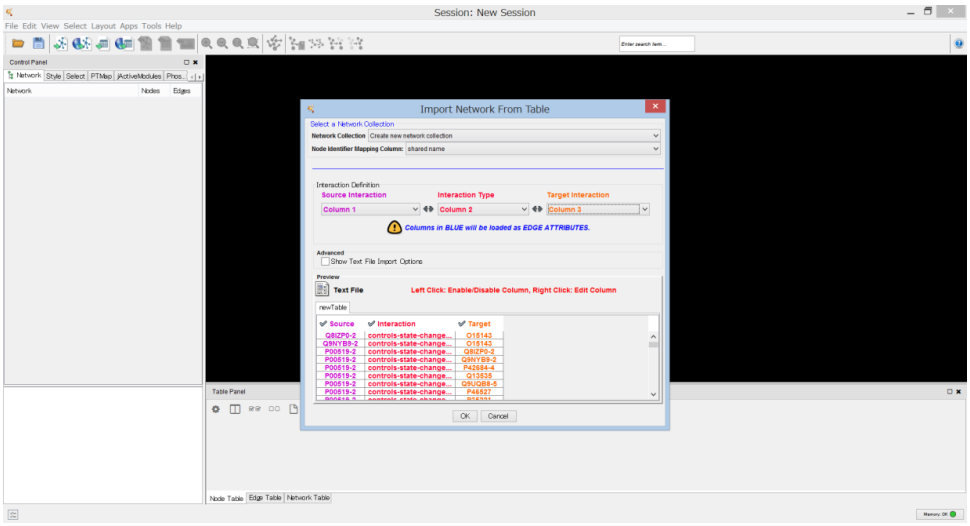
3. Usage

3.1 Installation of PTMapper

1. Download the PTMapper Jar file from Github (https://www.github.com/y-narushima/PTMapper/).
2. Start Cytoscape and install the PTMapper via “Apps” -> “App Manager”. Select “Install from File” and browse for the jar file.

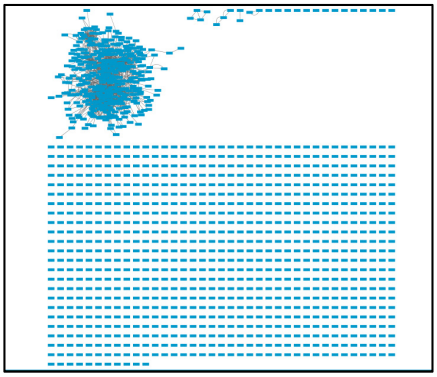
3.2 Import protein-protein interaction network

1. Import interaction data file via “File” -> “Import” -> “Network” -> “File…”. Choose “ppi_network_SIF.txt”.
2. Choose “Column 1” (Source), “Column 2” (Interaction) and “Column 3” (Target) for “Source Interaction”, “Interaction Type” and “Target Interaction”, respectively.
3. Click “OK”.



4. Click “Apply Preferred Layout”  to visualize the entire PPI networks.

Using the example data, the PPI network consisted of 1,276 nodes will be shown.



3.3 Import kinase-specific phosphorylation site information via PTMapper

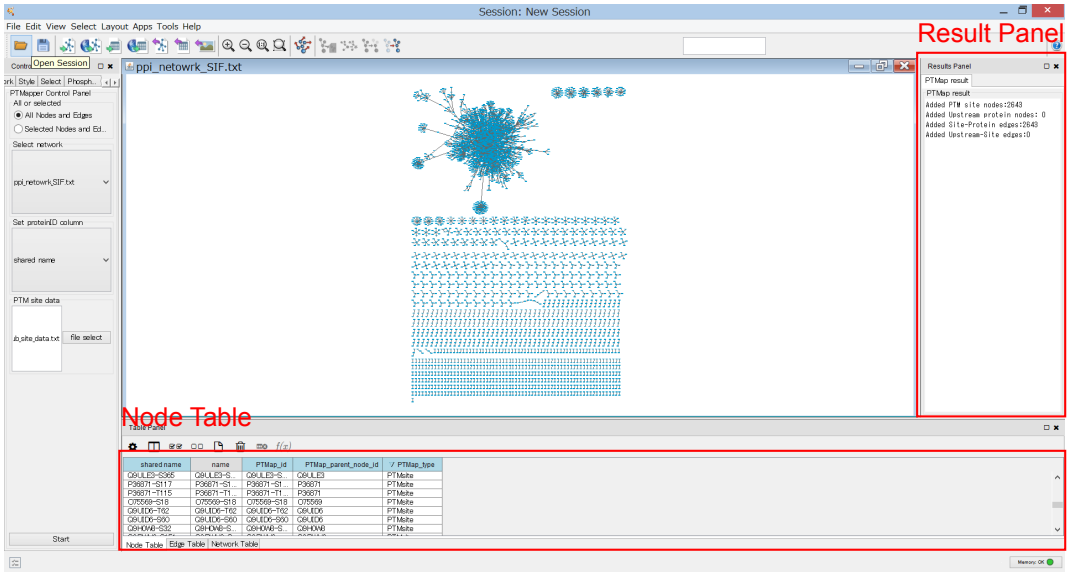
1. Select the “PTMapper” tab in the Control Panel.

* (
2. Select “All nodes and edges” in the “Select or Not” section.
3. Select “ppi_network_SIF.txt” in the “Select network” section.
4. Select “shared name” in the “Set proteinID column” section.
5. Click “file select” in the “PTM site data” section and choose data file containing kinase-specific phosphorylation information (“phosphosite_kinase_data.txt”).
6. Click “Start” at the bottom in the “PTMapper” panel.
)

7. Repeat the above steps (*) using the phosphorylation site data (“all_phosphosite_data.txt”). PTMapper enables us to add the nodes for phosphorylation sites to the target networks, optionally.

8. Click “Apply preferred layout”  to visualize the entire PTM site-oriented networks.

PTM site nodes, PTM regulator nodes and their relationships will be added to the network. In the “Node Table”, “PTMap_id”, “PTMap_parent_node_id” and “PTMap_type” (PTMsite, UpstRegulator and default) will be added. The number of nodes and edges added through PTMapper will be shown in the Result Panel.
Using the example data, the phosphorylation site-oriented network consisted of 4,171 nodes will be shown.



Added columns

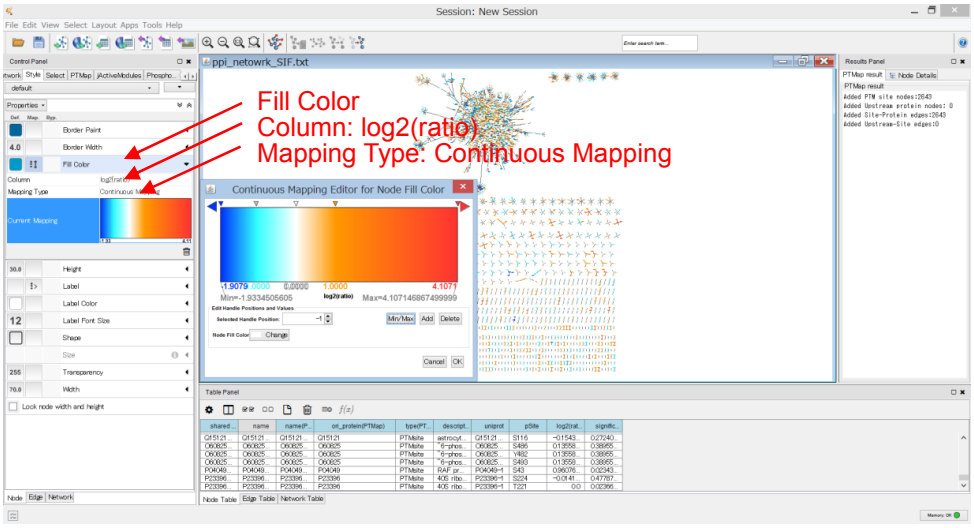
shared name	name	PTMap_id	PTMap_parent_node_id	PTMap_type	description	uniprot_id	pSite	log2(ratio)	significance_A
Q14498-2-S117	Q14498-2-S117	Q14498-2-S117	Q14498-2	PTMsite	RNA-binding protein 39 isoform...	Q14498-2	S117	0.038196324	0.468510913
Q14498-2-Y95	Q14498-2-Y95	Q14498-2-Y95	Q14498-2	PTMsite	RNA-binding protein 39 isoform...	Q14498-2	Y95	-0.117237675	0.322777127
Q14498-2-S136	Q14498-2-S136	Q14498-2-S136	Q14498-2	PTMsite	RNA-binding protein 39 isoform...	Q14498-2	S136	-0.058883689	0.408637279
Q8NEY1-S648	Q8NEY1-S648	Q8NEY1-S648	Q8NEY1	PTMsite	neuron navigator 1 isoform 1 [...	Q8NEY1-1	S648	0.149606344	0.378477297
Q8NEY1-S377	Q8NEY1-S377	Q8NEY1-S377	Q8NEY1	PTMsite	neuron navigator 1 isoform 1 [...	Q8NEY1-1	S377	0.19088462	0.34647044
Q8NEY1-S142	Q8NEY1-S142	Q8NEY1-S142	Q8NEY1	PTMsite	neuron navigator 1 isoform 1 [...	Q8NEY1-1	S142	0.199490411	0.339924457
Q8NEY1-S90	Q8NEY1-S90	Q8NEY1-S90	Q8NEY1	PTMsite	neuron navigator 1 isoform 1 [...	Q8NEY1-1	S90	-0.109562738	0.342262018

3.4 Annotation with the quantification data on proteins/PTM sites.

1. Import quantification data file via “File” -> “Import” -> “Table” -> “File…”.
2. Choose “quantification_data.txt”.
3. In the “Import Columns From Table” window, keep default setting and click “OK”.

In the “Node Table”, “description”, “uniprot_id”, “pSite”, “log2(ratio)” and “significance_A” will be added.

4. Select the “Style” tab in the Control Panel.
5. Select “Fill Color” and then choose the column “log2(ratio)”.
6. Select “Continuous Mapping” as mapping type and set the node color as shown below.

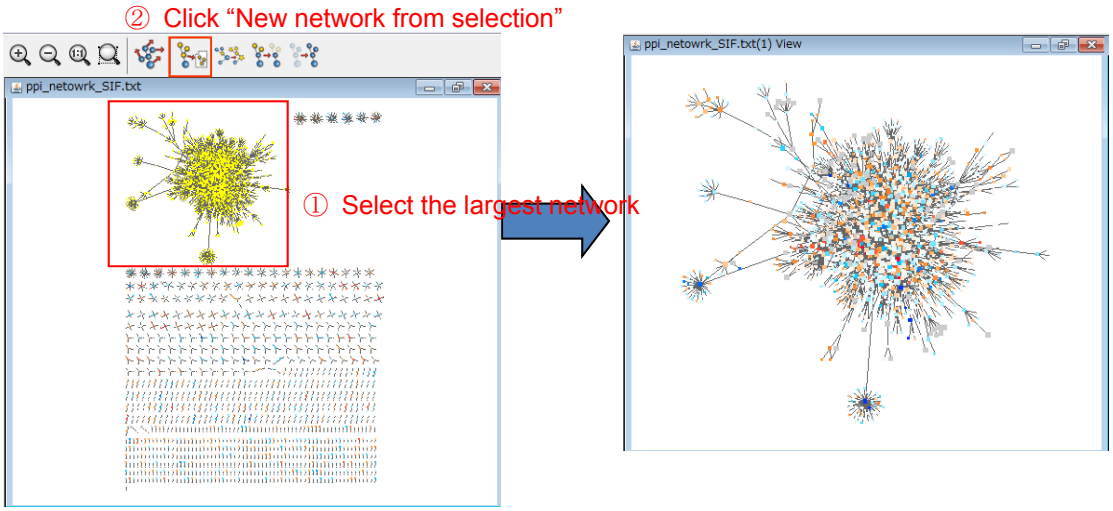


3.5 Extraction of sub-networks

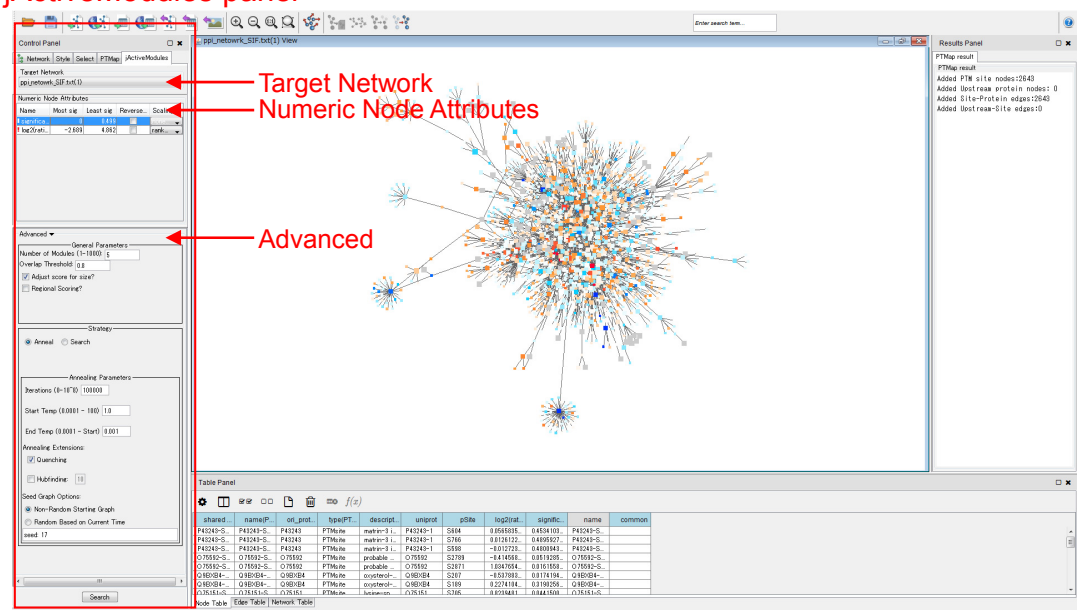
After the annotation of the nodes with the significance values, you can extract regulated sub-networks using the other Cytoscape application “jActiveModules” (Ideker *et al.*,

2002), which can identify the highest scoring sub-network by simulated annealing algorithm.

1. Select the largest network (1,581 nodes) and click “New Network From Selection”  as shown below.



2. Select the “jActiveModules” tab in Control Panel.



3. Set the below parameters in the “jActiveModules” panel.

Target Network: ppi_network_SIF.txt(1)

Numeric Node Attributes: significance_A

Advanced:

Number of Modules: 5

Overlap Threshold: 0.8

Adjust score for size?: on

Regional Scoring: off

Strategy: Anneal

Iterations: 100000

Start Temp: 1.0

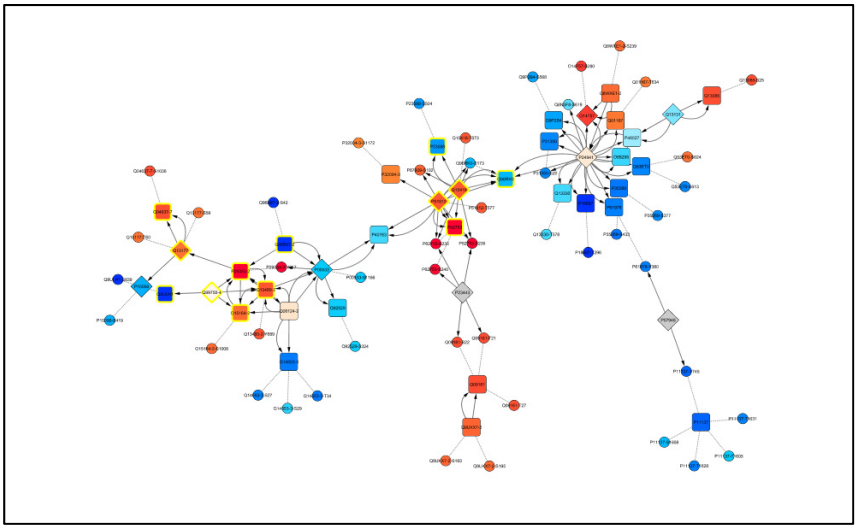
End Temp: 0.001

Annealing Extensions: Quenching

Seed Graph Options: Non-random Starting Graph (seed: 17)

4. Click the “Search” button.
5. Select the “Network” tab in the Control panel
6. Select “Module_0_1” in the “Network” panel.
7. Select the “Style” tab in the Control panel.
8. Change the style from “jActiveModules Module Style” to “default” at the top of the Style panel.

The significantly regulated sub-networks will be visualized as indicated below.



4. References

Cerami,E.G. *et al.* (2011) Pathway Commons, a web resource for biological pathway data. *Nucleic Acids Res.*, **39**, D685–D690.

Dinkel,H. *et al.* (2011) Phospho.ELM: a database of phosphorylation sites--update 2011. *Nucleic Acids Res.*, **39**, D261–D267.

Hornbeck,P. V. *et al.* (2012) PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res.*, **40**, D261–D270.

Hu,J. *et al.* (2014) PhosphoNetworks: a database for human phosphorylation networks. *Bioinformatics*, **30**, 141–142.

Ideker,T. *et al.* (2002) Discovering regulatory and signalling circuits in molecular interaction networks. *Bioinformatics*, **18 Suppl 1**, S233–S240.

Kozuka-Hata,H. *et al.* (2012) Phosphoproteome of human glioblastoma initiating cells reveals novel signaling regulators encoded by the transcriptome. *PLoS One*, **7**, e43398.

Magrane,M. and Consortium,U. (2011) UniProt Knowledgebase: a hub of integrated protein data. *Database (Oxford)*, **2011**, bar009.