Cytocopter Manual

CytoCopteR provides an intuitive and easy to learn graphical user interface (GUI) to CellNOptR methods through Cytoscape.

This results in a point and click interface where users can run the same steps as they would using an R script without having to actually write any code. Given that this is a front-end to the R algorithms, consistency is ensured between the results obtained through the GUI and those obtained through the corresponding scripts.

This tutorial is intended for Cytocopter v2.0 that requires Cytoscape v3.1 and R 3.x. The following materials are all available online on Cytocopter homepage: http://www.ebi.ac.uk/saezrodriguez/cno/cytocopter

Emanuel Gonçalves
Saez-Rodriguez Group
EMBL-EBI
emanuel@ebi.ac.uk

September 2, 2014

Table of Contents

Software Requirements

Please be sure that you have the following software installed and working:

1. R framework (version 3.x) - http://www.r-project.org/
2. Cytoscape (version 3.1.x) - http://www.cytoscape.org/

It is very important that Cytoscape version is equal or greater than 3.1.0

Linux users must run the following commands before using Cytocopter

- sudo apt-get install libcairo2-dev
- sudo apt-get install libxt-dev

Introduction

www.ebi.ac.uk/saezrodriguez/cno/cytocopter

Cytocopter[1] is a Cytoscape[2,3] application (App) that provides an intuitive graphical interface to CellNOptR R package.

This is developed under a GNU open-source license and the source code can be accessed from the respective GitHub project.

Installation

Cytocopter requires Cyrface[4], therefore Cyrface has **always** to be installed before Cytocopter.

Install from Cytoscape

Click *Apps* on Cytoscape top bar followed by *Apps Manager* menu then search for **Cyrface** and click install. This may take a few minutes. When it is done search now for **Cytocopter** and click install as before.

Install from file

Download the jar files of **Cyrface** and **Cytocopter** from the following webpages:

- www.ebi.ac.uk/saezrodriguez/cyrface
- www.ebi.ac.uk/saezrodriguez/cno/cytocopter

Click *Apps* on Cytoscape top bar followed by *Apps Manager* then click *Install from file…* button and search first for Cyrface jar, then repeat the last step for Cytocopter.

Install manually

Cytocopter can also be installed manually by copying the jar files mentioned before to CytoscapeConfiguration folder. CytoscapeConfiguration folder is kept in the user home folder. Drag the Apps jar files into the following folder:

- *~/ CytoscapeConfiguration / 3 / apps / installed /*

After moving the files start Cytoscape.

CellNOptR Tutorial

This tutorial assumes that Cytoscape as well as R is already installed. The necessary files for this tutorial is a network file in SBML-Qual format and the corresponding experimental data in MIDAS format.

Study case

To illustrate the use of Cytocopter we will use a biologically plausible prior knowledge network (PKN). This network includes a subset of intracellular signalling networks known to be activated downstream of EGF and TNFa stimulation[5]. The presented model is available for download in BioModels data-base.

The accompanying *in silico* data (MIDAS file format[6]) replicates biologically plausible behaviour that has been seen in such networks, such as the transient behaviour of ERK activation and the oscillatory dynamics of NFkB translocation from the cytoplasm to the nucleus.

SBML-Qual network format

SBML-Qual format[7] is an extension to the System Biology Markup Language (SBML) for Qualitative Models (Qual). In one sentence, SBML-Qual is designed to provide a standard mean for the exchange of logical models or regulatory and signalling networks.

For more details regarding the specifications please see [7].

MIDAS experimental data format

The MIDAS format (Minimum Information for Data Analysis in Systems Biology)[6] is a comma-separated file that specifies the layout of experimental data files.

Each row represents a single experimental sample; each column represents one sample attribute, such as treatment condition, or value obtained from an experimental assay.

A column header consists of two values

- Two-letter code defining the type of column, (e.g. TR for treatment, DV for data value),
- Short column name (e.g. a small molecule inhibitor added or a protein assayed).

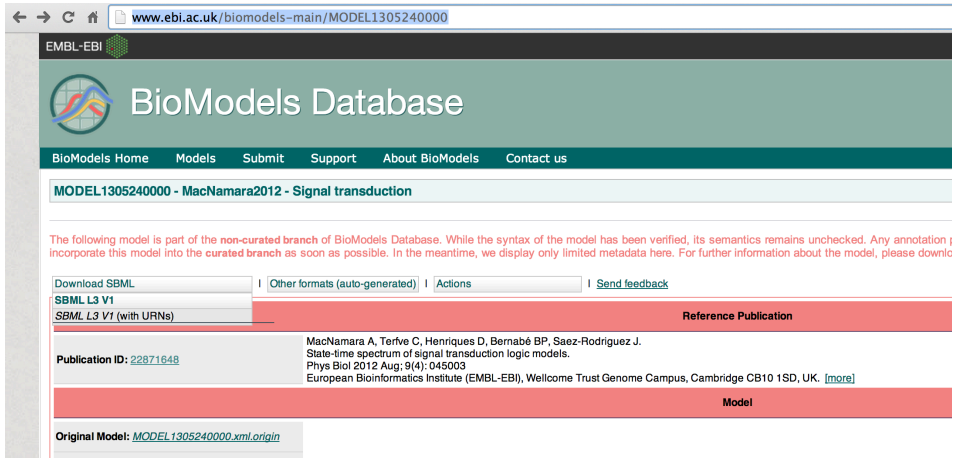
The body of each column stores the corresponding value for each row (sample) such as a plate/well name, reagent concentration, time point, or data value. For more details about the MIDAS format please see [6].

Download model file

Network model

1. Visit the URL below to download the MODEL1305240000 from BioModels data-base
    a. www.ebi.ac.uk/biomodels-main/MODEL1305240000

2. Click *Download SBML* menu and choose any of the SBML levels available for instance *SBML L3 V1*



NOTE: The network model can be stored in any of the formats supported by Cytoscape, for instance we also provide the same network in SIF format.

www.ebi.ac.uk/saezrodriguez/cno/cytocopter/resources/ToyModelPB.sif

Download experimental data

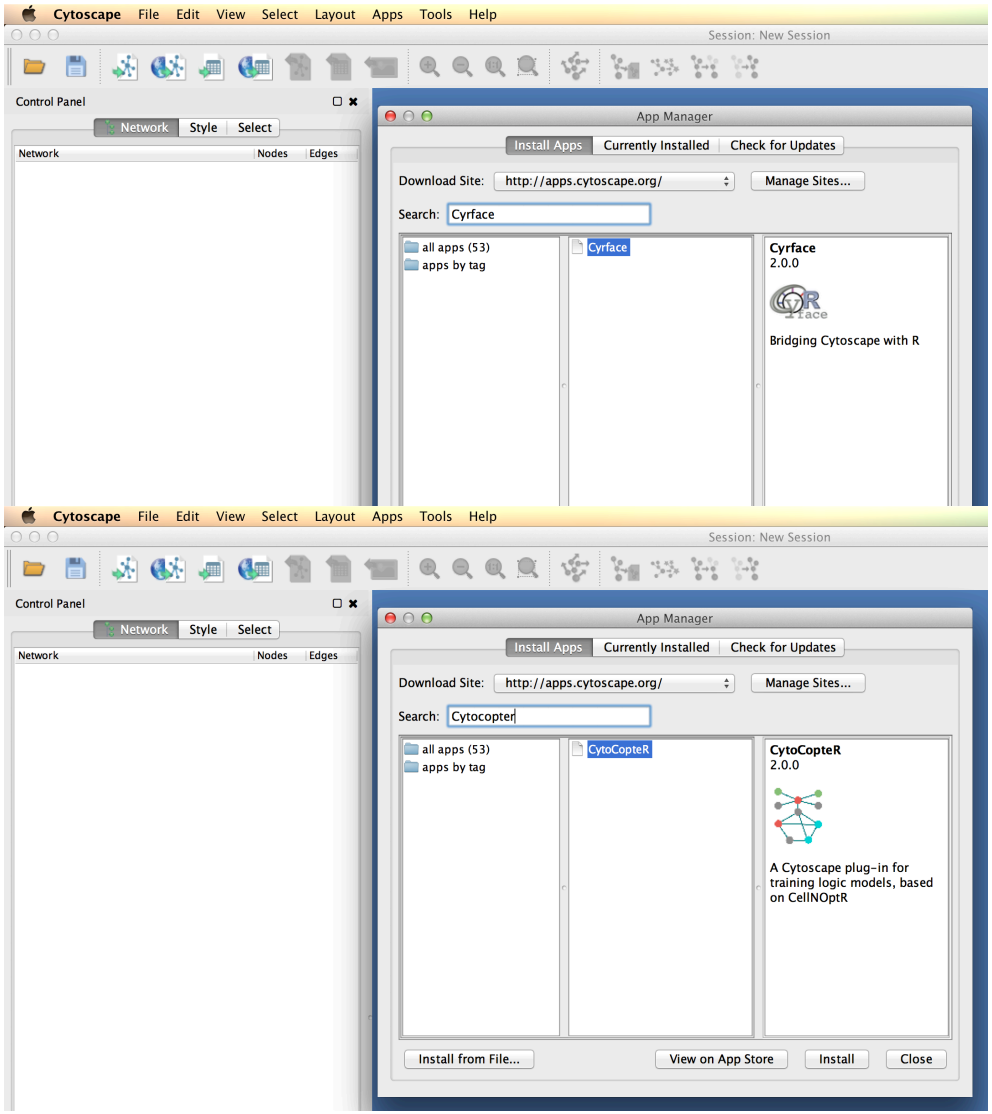
1. Visit the URL below and download the linked MIDAS file
    a. www.ebi.ac.uk/saezrodriguez/cno/cytocopter/resources/ToyModelPB.csv

Note: It is important that the file is stored with the existing extension *.csv*. Not doing so can impact incorrectly the application.

Install Cyrface and Cytocopter from Cytoscape

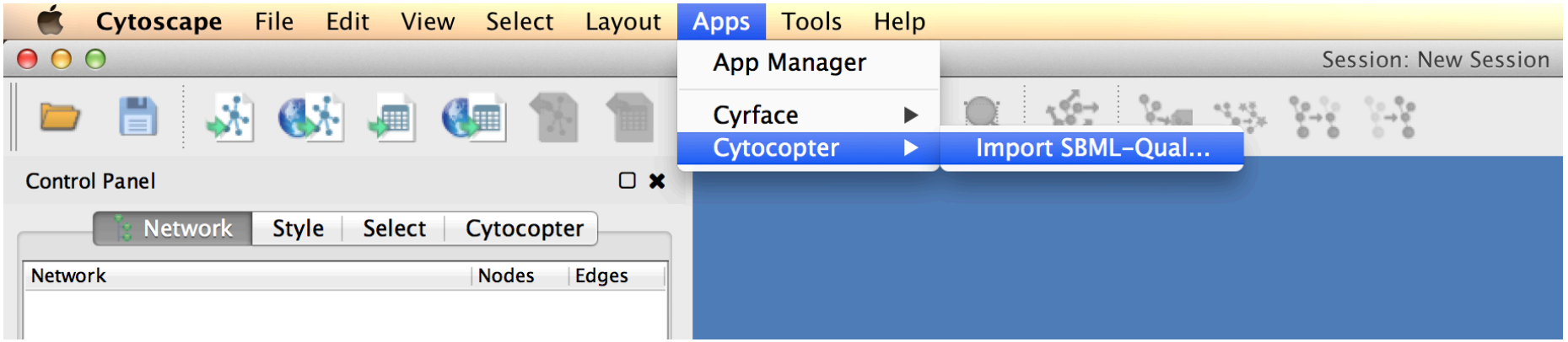
1. Click *Apps* then *Apps Manager*. On the opened window search for ***Cyrface*** and click *Install* button.

2. When ***Cyrface*** is installed repeat the previous step but this time search for ***Cytocopter***.



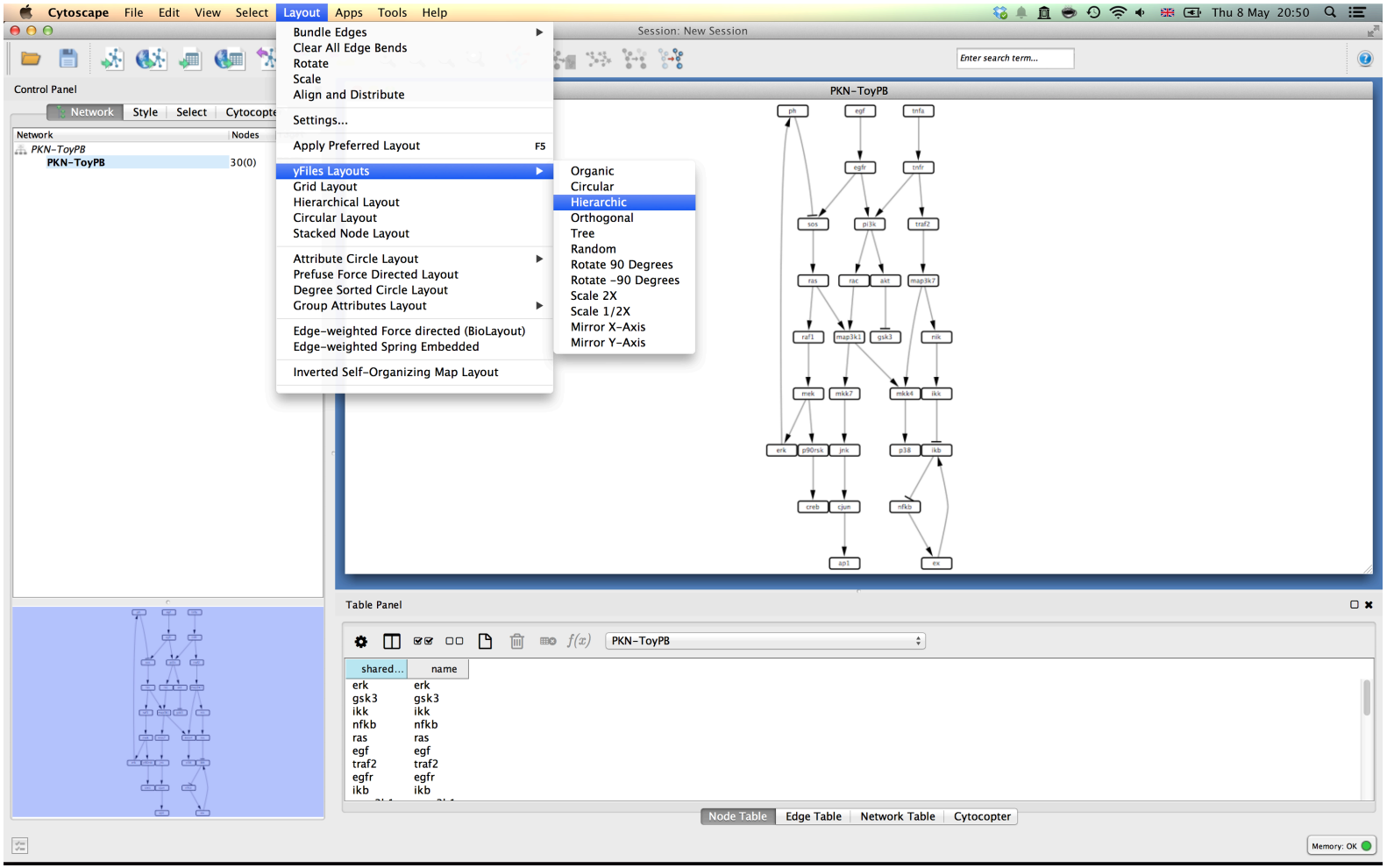
Load Model file

1. After Cytocopter is installed we can now import SBML-Qual formatted files.



2. Click *Apps*, *Cytocopter* then *Import SBML-Qual...* then look for the network model file downloaded before, **MODEL1305240000.xml**

3. An optional step is to apply one of the available network layouts in Cytoscape. In this case the yFile hierarchical layout is arguably the one that displays better the network. Click *Layout*, *yFiles Layouts* and then *Hierarchic*.



Preprocess experimental data and network

At this point we already have the network imported now we can use our experimental data and see how it maps into the network. In other words visualise which nodes are measured, inhibited and stimulated. For more details please consult *Terfve et al.* [1].

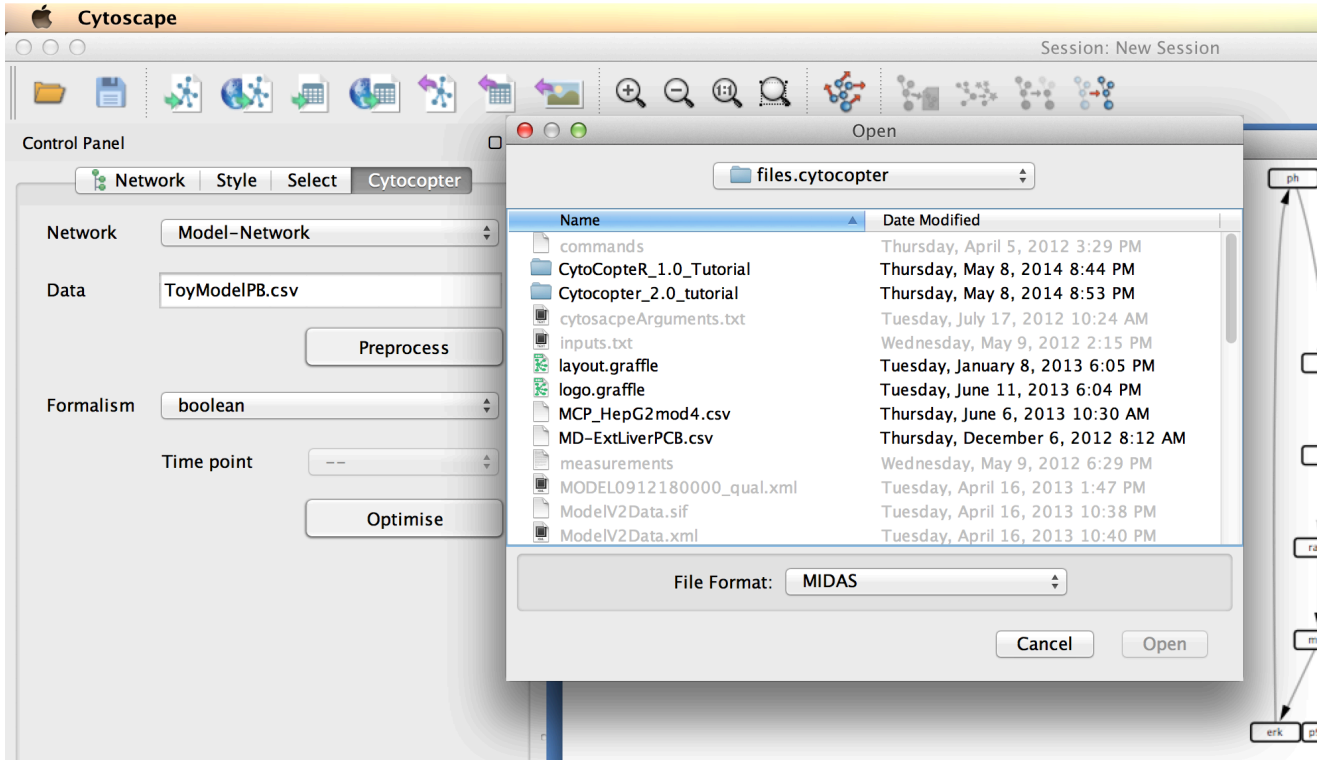
1. Select *Cytocopter* panel on the left-sided panel

2. On the network dropdown box select the imported network

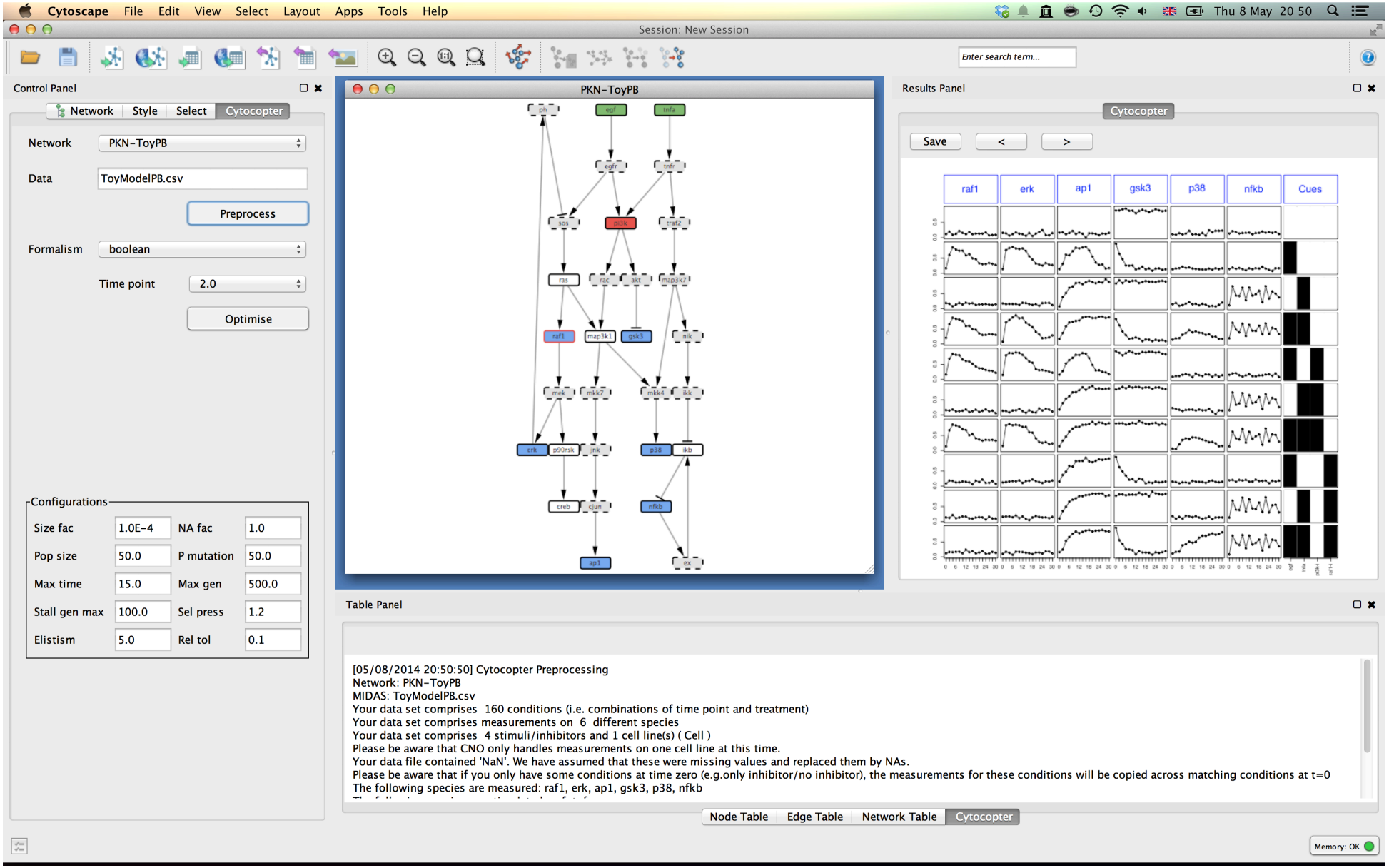
3. Click on the data text field a window will pop-up to browse the previously downloaded experimental (MIDAS file)

4. After the Data file is selected press the *Preprocess* button

   This procedure may take a few minutes; in particular if it is the first time all the necessary R packages will be installed



5. The Preprocess function will annotate automatically the network: Green nodes are stimulated species; Red nodes are inhibited species; Blue nodes are measured species; Grey nodes with dashed border can be removed to simplify the network; Blue nodes with Red border are measured and inhibited nodes; White nodes are not measured or perturbed nodes that can not be simplified


Cytoscape File Edit View Select Layout Apps Tools Help
Thu 8 May 20 50
Session: New Session
Enter search term...
Control Panel
Network Style Select Cytocopter
Network PKN-ToyPB
Data ToyModelPB.csv
Preprocess
Formalism boolean
Time point 2.0
Optimise
Configurations
Size fac 1.0E-4
NA fac 1.0
Pop size 50.0
P mutation 50.0
Max time 15.0
Max gen 500.0
Stall gen max 100.0
Sel press 1.2
Elistism 5.0
Rel tol 0.1
PKN-ToyPB
Results Panel
Cytocopter
Save
<
>
raf1
erk
ap1
gsk3
p38
nfkb
Cues
Table Panel
[05/08/2014 20:50:50] Cytocopter Preprocessing
Network: PKN-ToyPB
MIDAS: ToyModelPB.csv
Your data set comprises 160 conditions (i.e. combinations of time point and treatment)
Your data set comprises measurements on 6 different species
Your data set comprises 4 stimuli/inhibitors and 1 cell line(s) (Cell)
Please be aware that CNO only handles measurements on one cell line at this time.
Your data file contained 'NaN'. We have assumed that these were missing values and replaced them by NAs.
Please be aware that if you only have some conditions at time zero (e.g.only inhibitor/no inhibitor), the measurements for these conditions will be copied across matching conditions at t=0
The following species are measured: raf1, erk, ap1, gsk3, p38, nfkb
Node Table Edge Table Network Table Cytocopter
Memory: OK

Optimising network with experimental data

In this step we will use the optimise features of CellNOptR R package to optimise the topology of the network against the previously loaded experimental data. For more details please consult *Terfve et al.* [1].

1. Select a formalism to model the network. Select the Boolean Formalism. Boolean formalism will consider that the states of each node will be either active or inactive

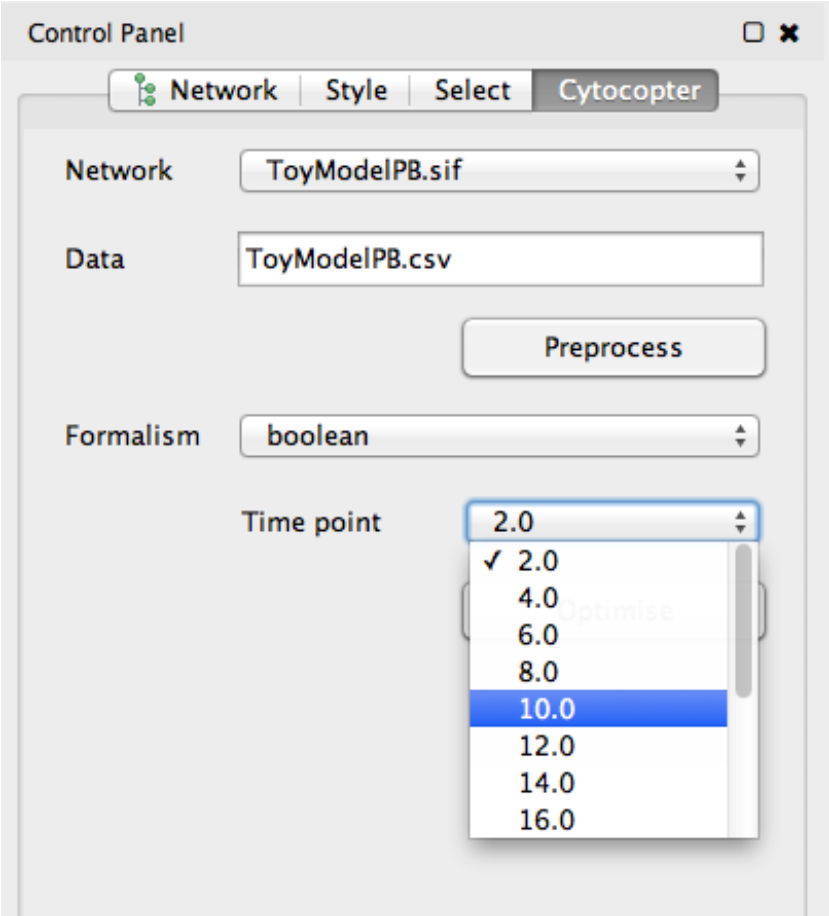
   Because the Boolean formalism requires only two time points and the given data-set has 16, the first time point is automatically set to 0 and the second one is selected by the user

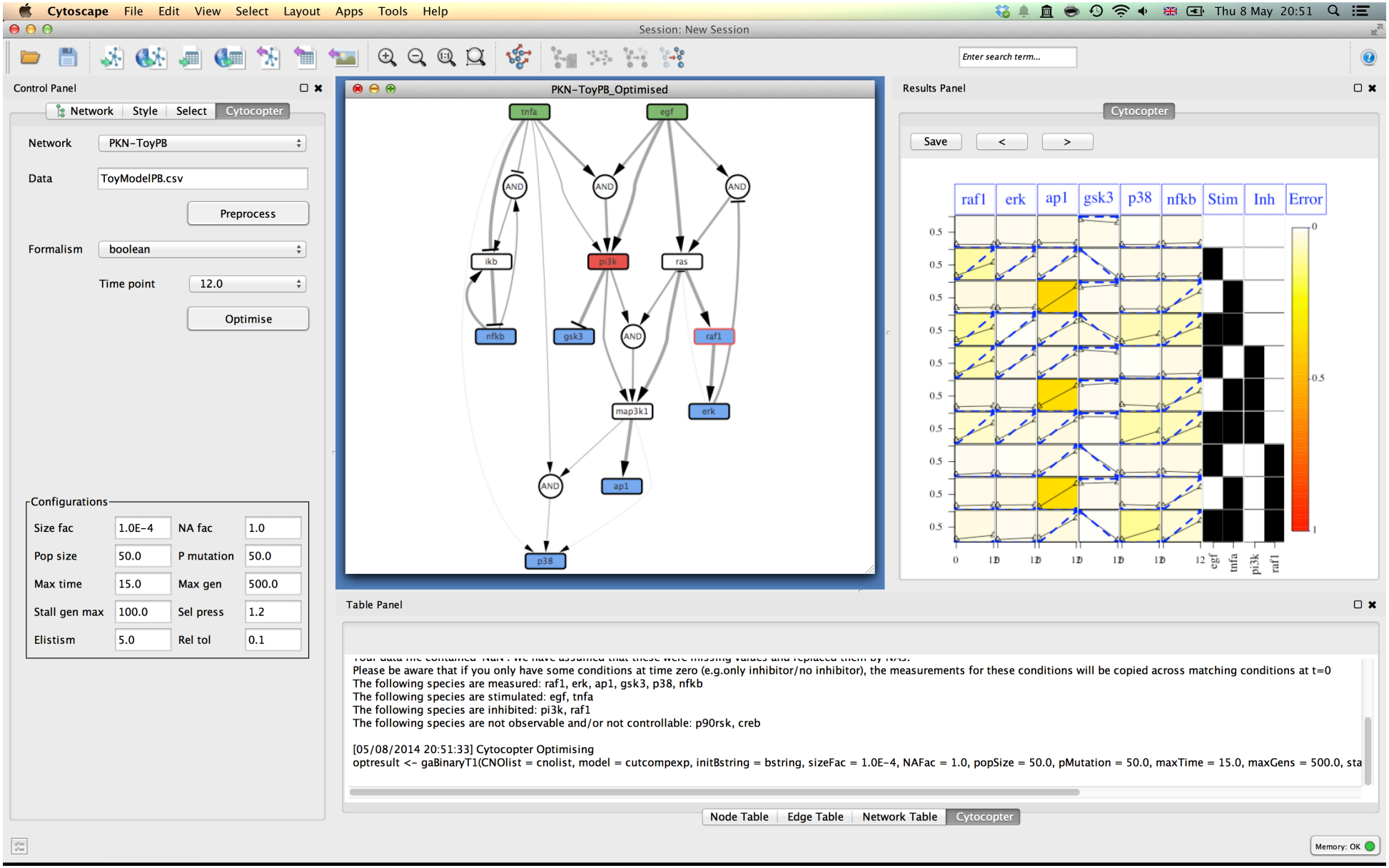
2. Select time point 10 from the *Time point* dropdown box

3. After selecting the time point press *Optimise* button

   Note this will take several minutes. The optimisation is done using a genetic algorithm and its parameters can be configured in the Configuration panels, such as maximum time allowed for the algorithm to run in the Max time parameter.

4. The optimisation outputs an optimised network and a plot displaying how well the simulated models fit with the given experimental data. The colour gradient quantifies the error levels.



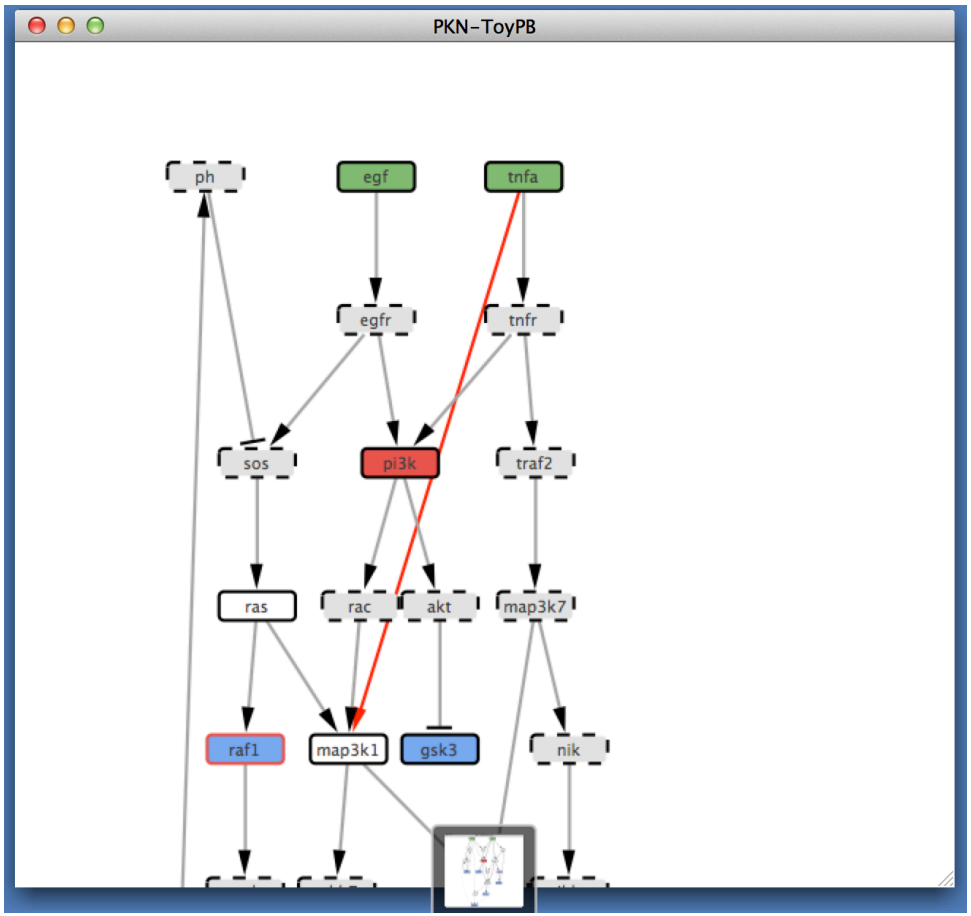

Cytoscape File Edit View Select Layout Apps Tools Help
Thu 8 May 20:51
Session: New Session
Enter search term...
Control Panel
Network Style Select Cytocopter
Network PKN-ToyPB
Data ToyModelPB.csv
Preprocess
Formalism boolean
Time point 12.0
Optimise
Configurations
Size fac 1.0E-4
NA fac 1.0
Pop size 50.0
P mutation 50.0
Max time 15.0
Max gen 500.0
Stall gen max 100.0
Sel press 1.2
Elistism 5.0
Rel tol 0.1
PKN-ToyPB_Optimised
tnfa
egf
AND
ikb
pi3k
ras
nfkb
gsk3
raf1
map3k1
erk
ap1
p38
Results Panel
Cytocopter
Save
<
>
raf1 erk ap1 gsk3 p38 nfkb Stim Inh Error
Table Panel
Please be aware that if you only have some conditions at time zero (e.g.only inhibitor/no inhibitor), the measurements for these conditions will be copied across matching conditions at t=0
The following species are measured: raf1, erk, ap1, gsk3, p38, nfkb
The following species are stimulated: egf, tnfa
The following species are inhibited: pi3k, raf1
The following species are not observable and/or not controllable: p90rsk, creb
[05/08/2014 20:51:33] Cytocopter Optimising
optresult <- gaBinaryT1(CNOlist = cnolist, model = cutcompexp, initBstring = bstring, sizeFac = 1.0E-4, NAFac = 1.0, popSize = 50.0, pMutation = 50.0, maxTime = 15.0, maxGens = 500.0, sta
Node Table Edge Table Network Table Cytocopter
Memory: OK

Adding the missing link

This type of analysis is very helpful to identify possible missing links. For instance, if we focus on the fit plot generated by the optimisation function we can see that ap1 has an increased error when compared to the other measured species. Lets see if we can improve this.

1. Minimise the optimised network and select the network used to run the optimisation

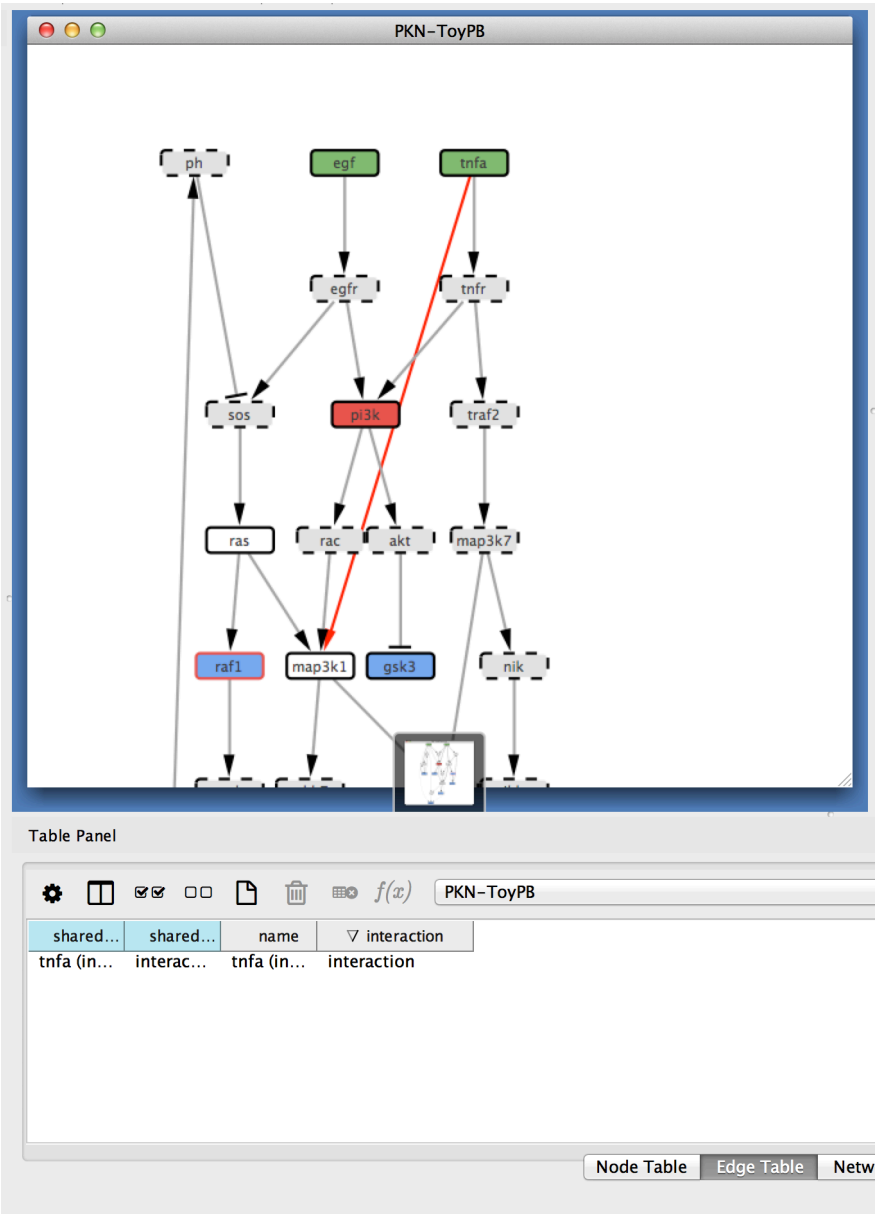
2. Right-click on the **tnfa** then click *Add* and then *Edge*. Drag the edge and connect **tnfa** to **map3k1** node



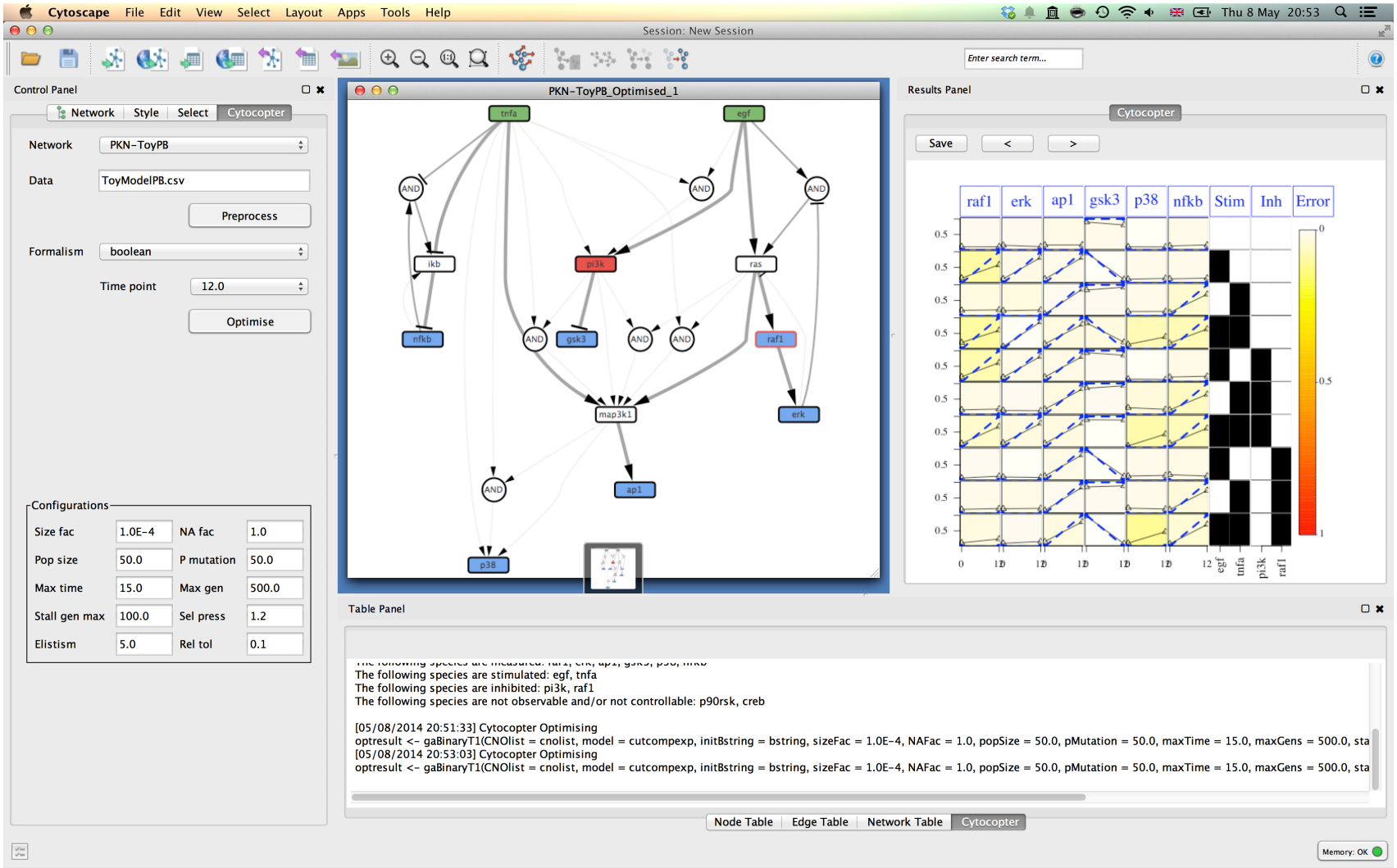
3. After adding the edge we need to define the type of interaction, i.e. if **tnfa** activates (1) **map3k1** or inhibits (-1) its activity. Click on the Edge Table panel in Cytoscape bottom panel then select the previously added edge.

4. Now double-click in the table cell with the value "*interaction*". Replace the *"interaction"* text with 1

5. Now click on the *Optimise* button to re-run exactly the previous optimisation with the exception of the extra edge that we just added.



6. After the optimisation we can see that the **ap1** simulated values fit much better the existing data which can possibly indicate that the link that we added before is missing from the network that we imported.



Links

- ***Cytocopter*** - http://www.ebi.ac.uk/saezrodriguez/cno/cytocopter
- ***Cytocopter (Github)*** - https://github.com/EmanuelGoncalves/cytocopter
- ***Cyrface*** - http://www.ebi.ac.uk/saezrodriguez/cyrface/
- ***CellNOptR*** - http://www.cellnopt.org/
- ***Cytoscape*** - http://www.cytoscape.org/
- ***R*** - http://www.r-project.org/

References

1. Terfve, C. *et al.* CellNOptR: a flexible toolkit to train protein signaling networks to data using multiple logic formalisms. *BMC Syst Biol* **6,** 133 (2012).

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

3. Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P. L. & Ideker, T. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* **27,** 431–432 (2011).

4. Gonçalves, E. & Saez-Rodriguez, J. Cyrface: An interface from Cytoscape to R that provides a user interface to R packages. *F1000Res* **2,** 192 (2013).

5. MacNamara, A., Terfve, C., Henriques, D., Bernabé, B. P. & Saez-Rodriguez, J. State-time spectrum of signal transduction logic models. *Phys Biol* **9,** 045003 (2012).

6. Saez-Rodriguez, J. *et al.* Flexible informatics for linking experimental data to mathematical models via DataRail. *Bioinformatics* **24,** 840–847 (2008).

7. Chaouiya, C. *et al.* SBML qualitative models: a model representation format and infrastructure to foster interactions between qualitative modelling formalisms and tools. *BMC Syst Biol* **7,** 135 (2013).