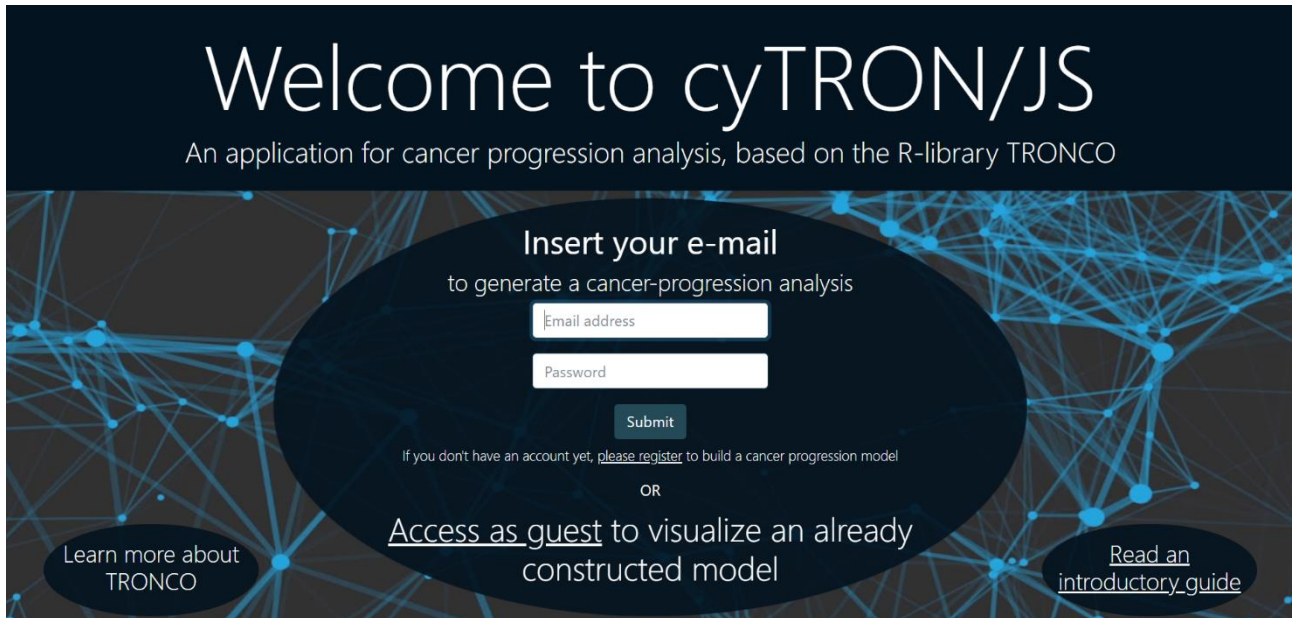


cyTRON/JS: Tutorial guide

Welcome to cyTRON/JS, a tool for the reconstruction of cancer progression models based on the R library TRONCO. This tutorial is a guide which shows examples about how to use this tool.



1) Exploring TRONCO graphs

By clicking on the *Access as guest* option in the home page displayed above, it is possible to start exploring some mutational graphs produced by TRONCO.

After selecting this option you will be redirected to this page:



Here, you have **two** options:

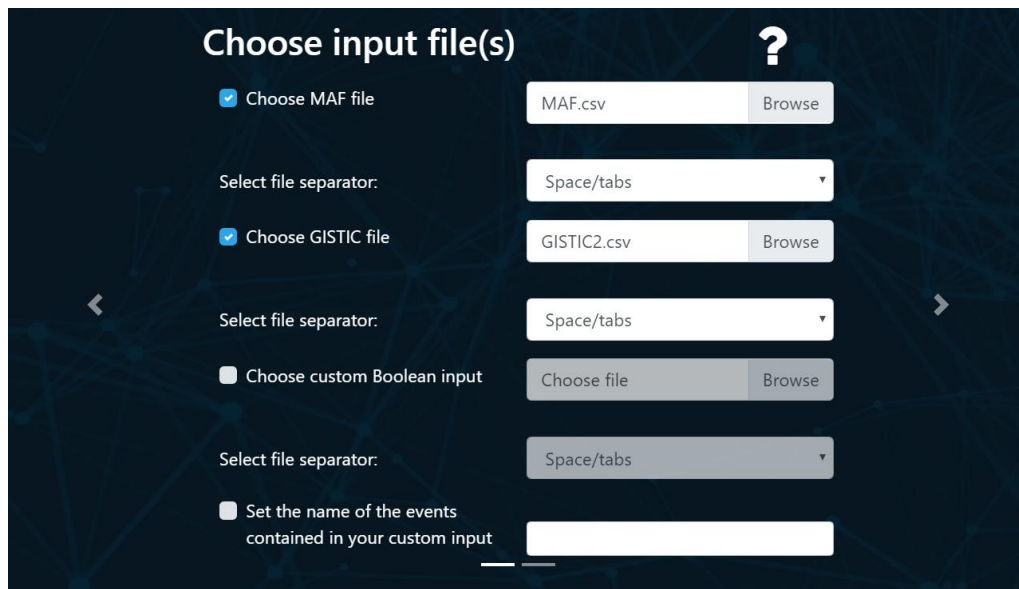
- On the **left** panel you can choose to visualize one of the publicly available results of a TRONCO analysis.
- On the **right** panel you can upload your local graphml file, which should contain the result of a TRONCO analysis.

2) TCGA-prostate example

The files needed for every step can be found here: <https://github.com/BIMIB-DISCO/cyTRON-js/tree/master/examples/TCGA-prostate>

First, in order to proceed further you need to create an account or to sign-in from the home page. After this step is complete, you will be redirected to a page where you need to insert a name for the analysis you are about to run.

1: Selecting input

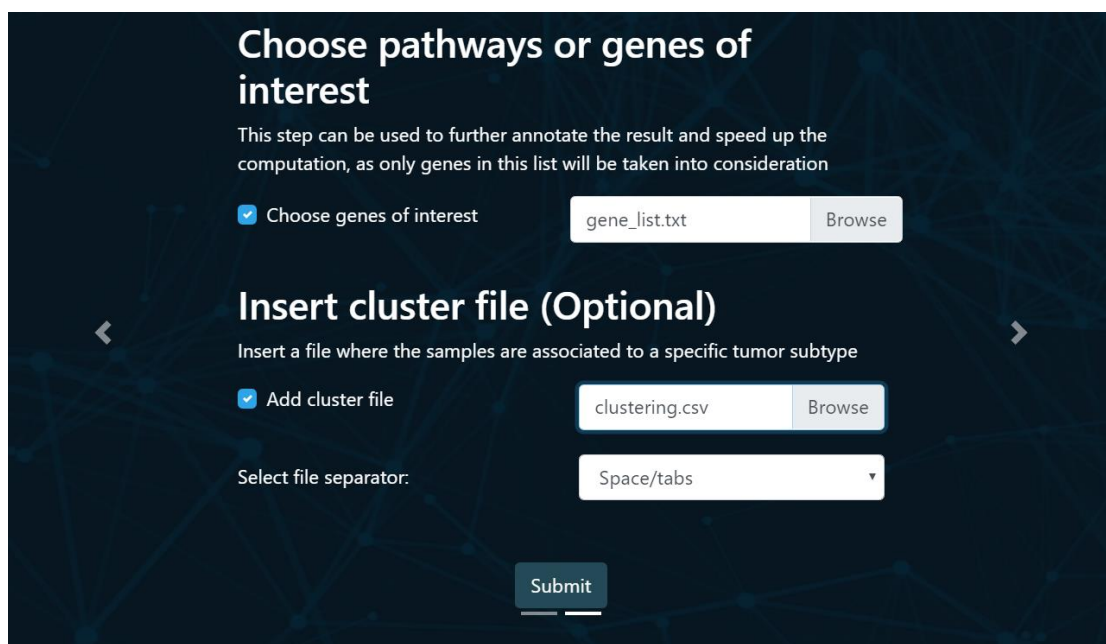


The screenshot shows a web form titled "Choose input file(s)" with a question mark icon. It contains three main sections for file selection:

- MAF file:** A checked checkbox "Choose MAF file" is followed by a text input field containing "MAF.csv" and a "Browse" button.
- GISTIC file:** A checked checkbox "Choose GISTIC file" is followed by a text input field containing "GISTIC2.csv" and a "Browse" button.
- Custom Boolean input:** An unchecked checkbox "Choose custom Boolean input" is followed by a text input field containing "Choose file" and a "Browse" button.

Below each file selection section is a "Select file separator:" dropdown menu, all of which are set to "Space/tabs". At the bottom, there is an unchecked checkbox "Set the name of the events contained in your custom input" followed by an empty text input field.

The image above shows the page which is displayed after inserting the study name. There are three different possible input files: MAF, GISTIC and a custom Boolean file, which contains a matrix indicating which mutation is present in each sample. For this example we are going to use the first two types of input.



The screenshot shows two sections of a web form:

Choose pathways or genes of interest

This step can be used to further annotate the result and speed up the computation, as only genes in this list will be taken into consideration

- A checked checkbox "Choose genes of interest" is followed by a text input field containing "gene_list.txt" and a "Browse" button.

Insert cluster file (Optional)

Insert a file where the samples are associated to a specific tumor subtype

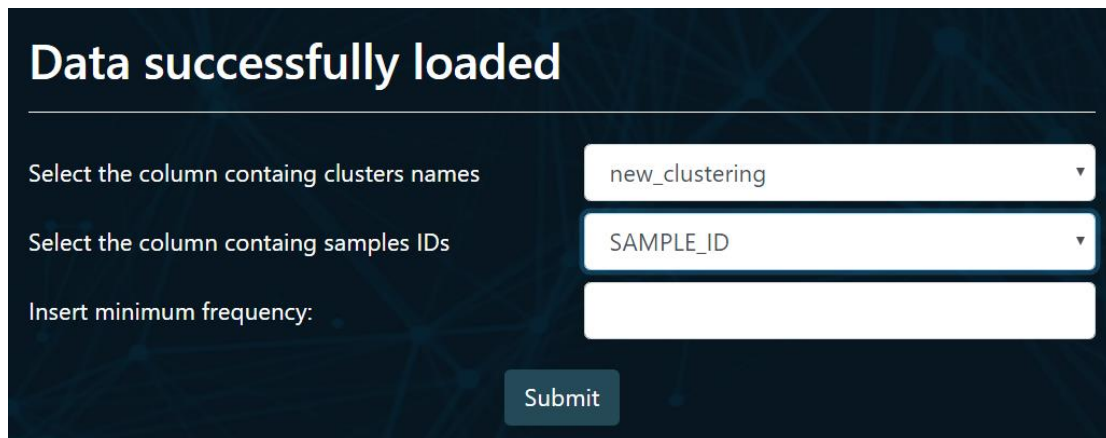
- A checked checkbox "Add cluster file" is followed by a text input field containing "clustering.csv" and a "Browse" button.
- A "Select file separator:" dropdown menu is set to "Space/tabs".

At the bottom of the form is a "Submit" button.

The image above displays the second slide of the input selection. Here a user can decide to upload a **list of genes** to take into consideration during the analysis. This step can speed up the computation, as only genes contained in this list will be taken into consideration.

The cancer type involved in the study might also be divided into different subtypes: to achieve this goal, a user can upload a **cluster file**, which should contain a list of samples with an associated subtype.

2: Cluster selection



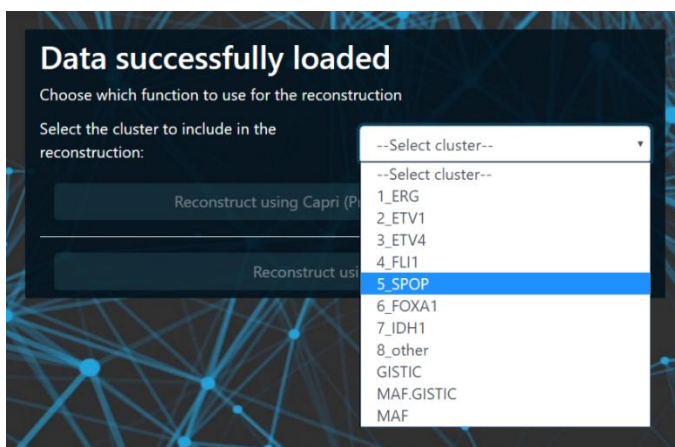
The screenshot shows a dark-themed web interface titled "Data successfully loaded". Below the title, there are three input fields on the left and their corresponding dropdown menus on the right. The first field is "Select the column containing clusters names" with a dropdown menu showing "new_clustering". The second field is "Select the column containing samples IDs" with a dropdown menu showing "SAMPLE_ID". The third field is "Insert minimum frequency:" with an empty text input box. A "Submit" button is located at the bottom right of the form area.

This is the page displayed after the input selection. Here the user can access the names of the columns contained in the cluster file, and can indicate which columns contain the samples IDs and the subtype names related to each sample.

3: Model reconstruction

After the clustering step, the user will be redirected to a page for the model reconstruction.

These images on the left show the interface for the reconstruction:



The screenshot shows a dark-themed web interface titled "Data successfully loaded". Below the title, there is a section titled "Choose which function to use for the reconstruction". Under this section, there is a label "Select the cluster to include in the reconstruction:" followed by a dropdown menu. The dropdown menu is open, showing a list of clusters: "--Select cluster--", "1_ERG", "2_ETV1", "3_ETV4", "4_FLI1", "5_SPOP" (which is highlighted in blue), "6_FOXA1", "7_IDH1", "8_other", "GISTIC", "MAF.GISTIC", and "MAF". Below the dropdown menu, there are two buttons: "Reconstruct using Capri (P)" and "Reconstruct using MAF.GISTIC".

First, the user needs to select which one among the clusters will be used.

Second, the user can decide which algorithm to use. In case 'Capri' is chosen, some parameters need to be set:

Data successfully loaded

Choose which function to use for the reconstruction

Select the cluster to include in the reconstruction:

Reconstruct using Capri (Press to set parameters)

Command: ☒ Hill Climbing ☐ Tabu

Regularization: ☒ Bic ☒ Aic

Estimate error rate: ☒ Yes ☐ No

☐ Insert MUTEX file

Reconstruct using Caprese

- **Command:** it defines the heuristic search to be performed.
- **Regularization:** it selects the regularization for the likelihood estimation.
- **Error rate:** it enables or disables the estimation of the error rates give the reconstructed model.
- **Mutex:** this can be used to insert a file containing a list of mutually exclusive mutations, created through the MUTEX tool.

For more information visit

<https://github.com/PathwayAndDataAnalysis/mutex>.

4: Plot

The reconstruction was successful

Select the model to plot

Model:

Confidence Values: ☒ Hg ☒ Tp ☒ Pr

☐ Print prima facie ☐ Print disconnected nodes

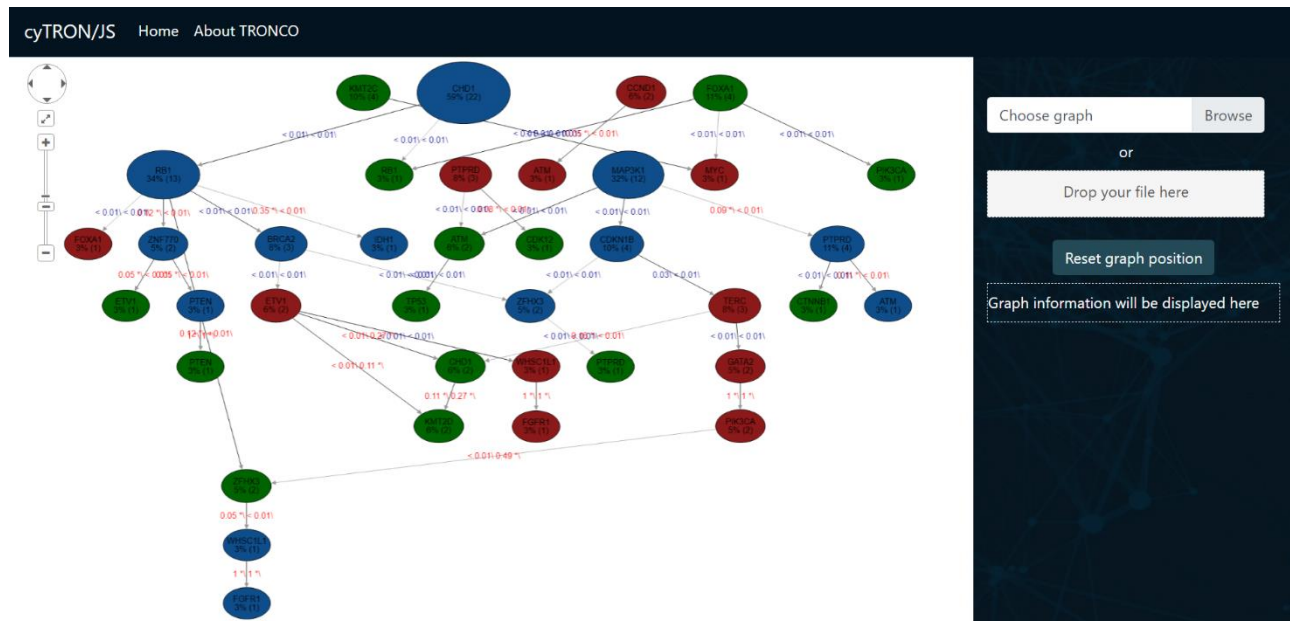
or

This is the page displayed after the reconstruction has finished. Here the user can decide to either reconstruct another model or plot one which has already been constructed.

To plot a model, the user needs to indicate which confidence values have to be displayed for every edge in the model, if *prima facies* needs to be displayed, and if disconnected nodes should be included in the visualization.

4: Visualization

After the previous step, a .graphml file containing the progression model is created, and it is displayed in the following page:



Here a user can click on the model's nodes to get more information about a gene and can read confidence information for each edge by clicking on them.

3) COADREAD example

Pre-requisites:

The input data for this example can be found here: <https://github.com/BIMIB-DISCO/cyTRON-js/tree/master/examples/COADREAD>

Before starting this example you should follow the link and download the data.

In order to complete the coadread analysis, you should also authenticate and select a new name for this study.

Input selection:

In the input page, you need to select the following files from the data folder you downloaded:

- MAF: for this file you need to select the “;” separator.
- GISTIC
- Genes_list: this is the file that needs to be selected in the *genes of interest* section.
- TCGA-clusters: select the “;” separator for this file

Cluster selection

After the previous step is complete, you will be redirected to a page where you need to select which columns of the clusters file contain the cluster names and the samples ID. Here you need to select the options as it is shown in the next figure:

Data successfully loaded

Select the column containing clusters names

MSI_status

Select the column containing samples IDs

patient

Submit

Model reconstruction

The next step is the reconstruction of the cancer progression model. Select the parameters as it is shown below:

Data successfully loaded

Choose which function to use for the reconstruction

Select the cluster to include in the reconstruction:

MSI_H

☐ Insert minimum frequency:

Reconstruct using Capri (Press to set parameters)

Command:

☒ Hill Climbing ☐ Tabu

Regularization:

☒ Bic ☒ Aic

Estimate error rate:

☒ Yes ☐ No

☒ Insert MUTEX file

Choose file

Browse

Submit

For this example we also consider the MUTEX file. For this reason you need to open the data folder you previously downloaded, open the *mutex* folder, and select the *msi_results* file.

Finally, click on the *submit* button, and then visualize the model.