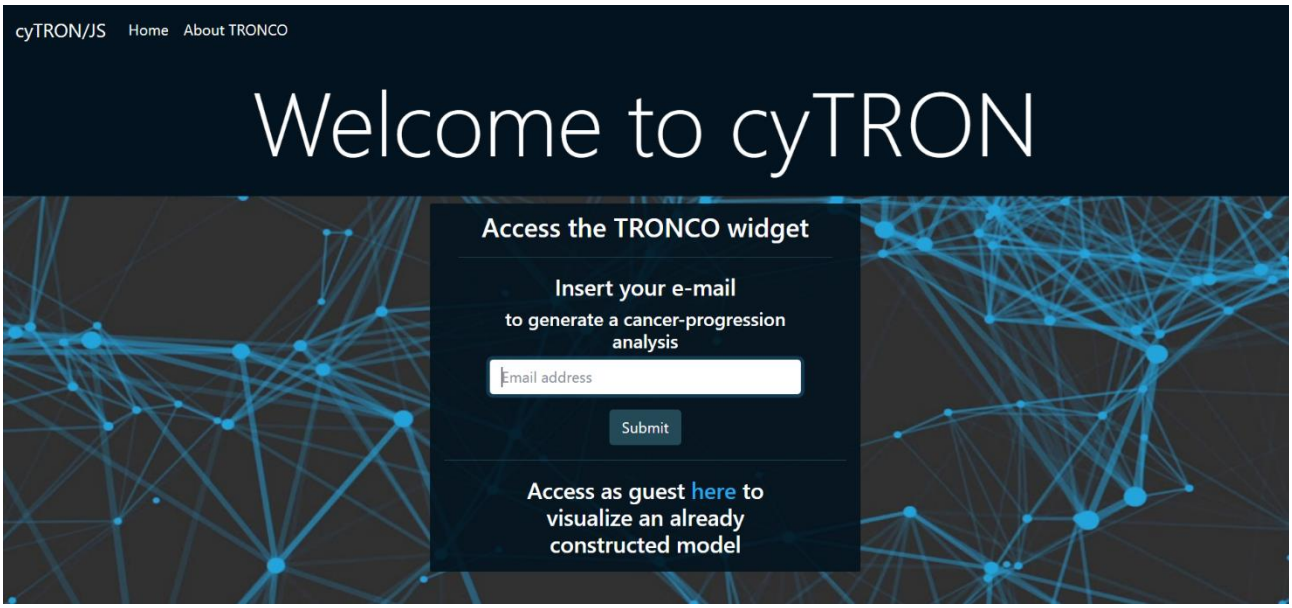
Example

The files needed for every step can be found in the examples/TCGA-prostate folder.

HOME PAGE

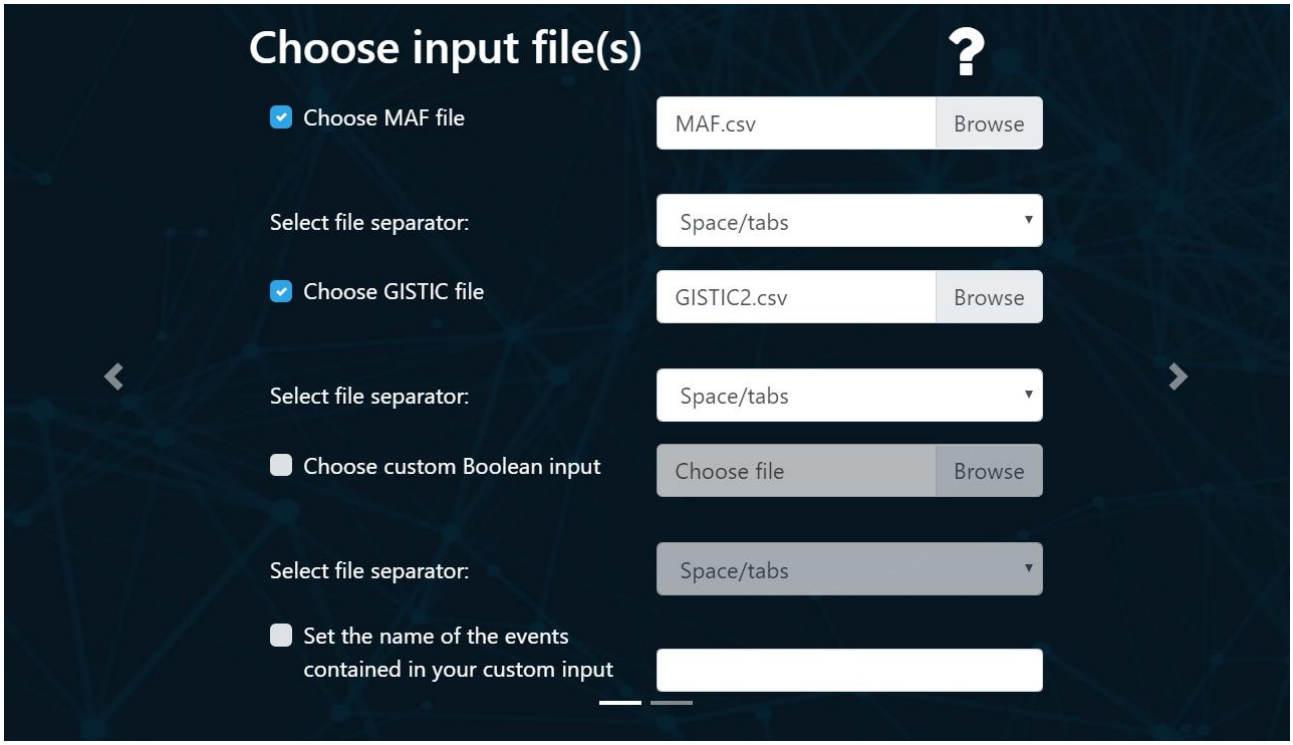


The image above displays the application home page: a user can decide whether to create a progression analysis or visualizing a graph which has already been constructed.

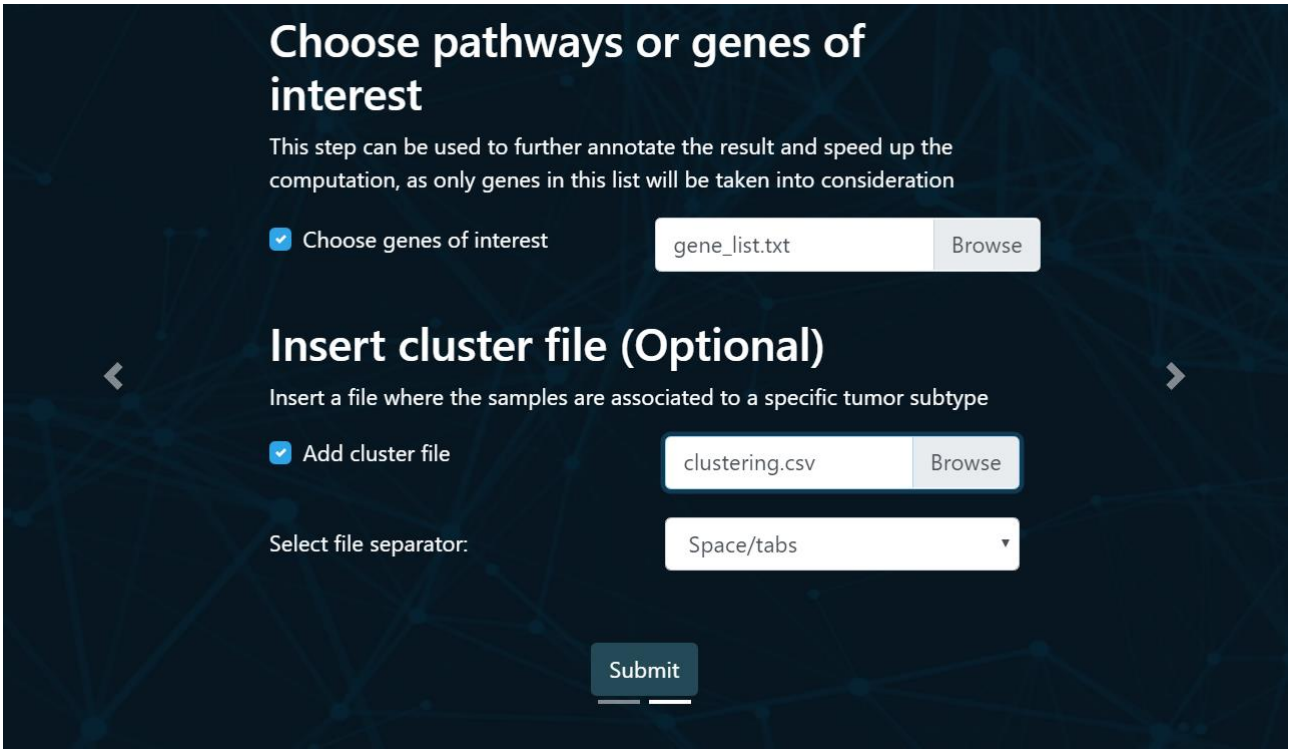
PIPELINE FOR CANCER PROGRESSION

1: Selecting input

After inserting the e-mail, the user is redirected to a page where he/she will insert the name for the study.



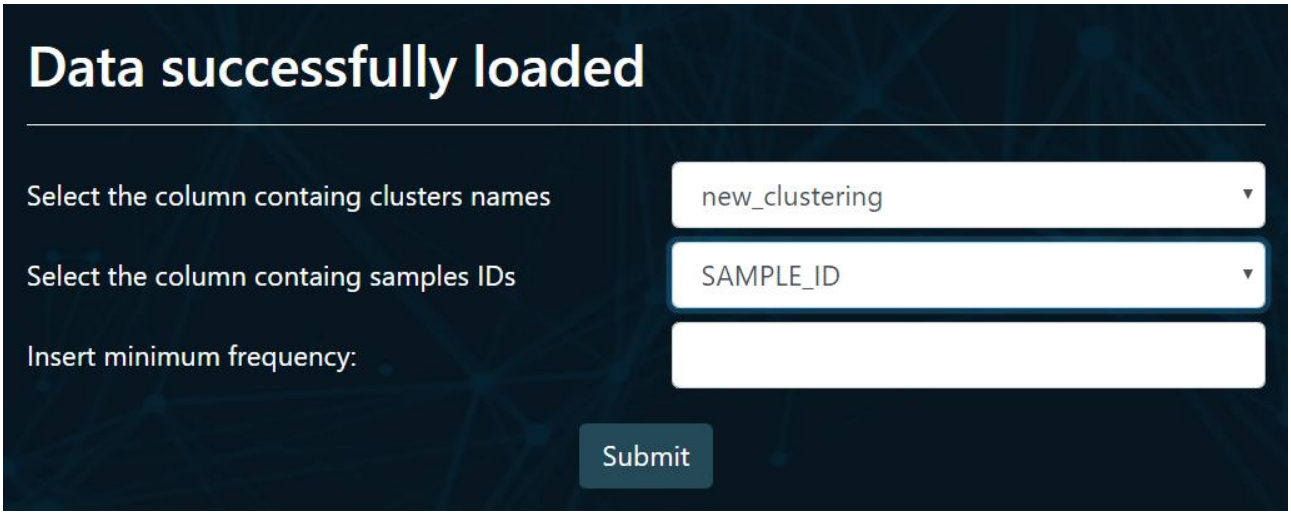
The image above shows the page which is displayed after inserting the study name and is the beginning of the analysis pipeline. There are three different possible input files: MAF, GISTIC and a custom Boolean file, which contains a matrix indicating which mutation is present in each sample. For this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



The 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

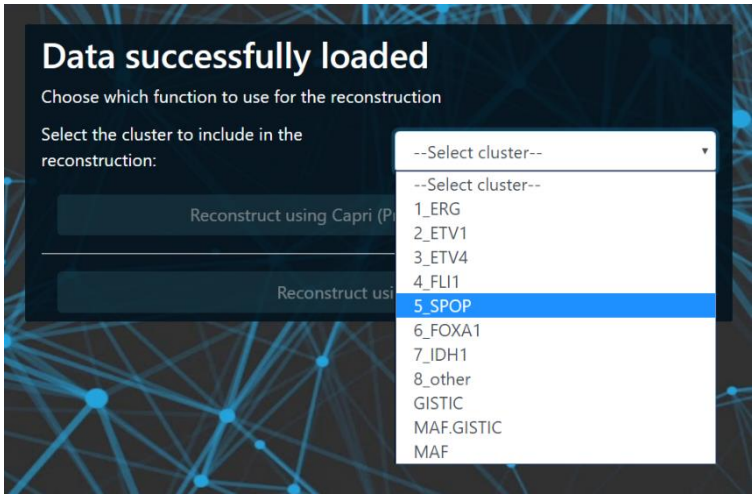


This is the page displayed after the input selection. Here the user can access to the names of the columns contained in the cluster file and can indicate which columns contain the samples IDs and the subtype names relative to each sample.

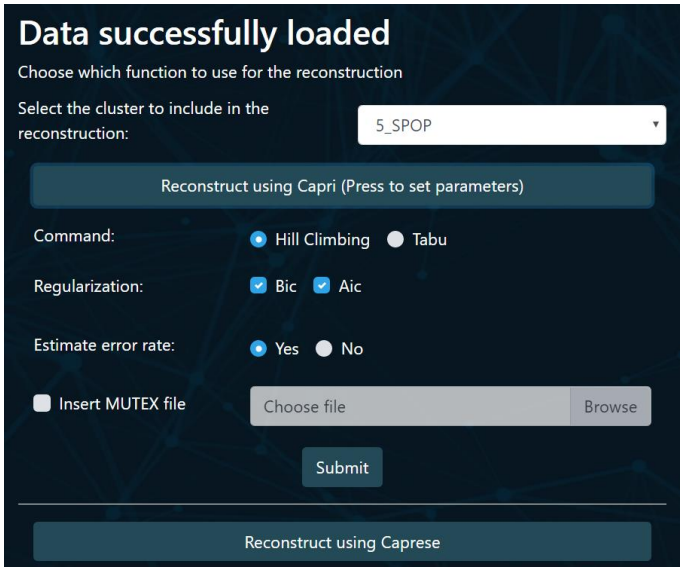
3: Model reconstruction

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

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



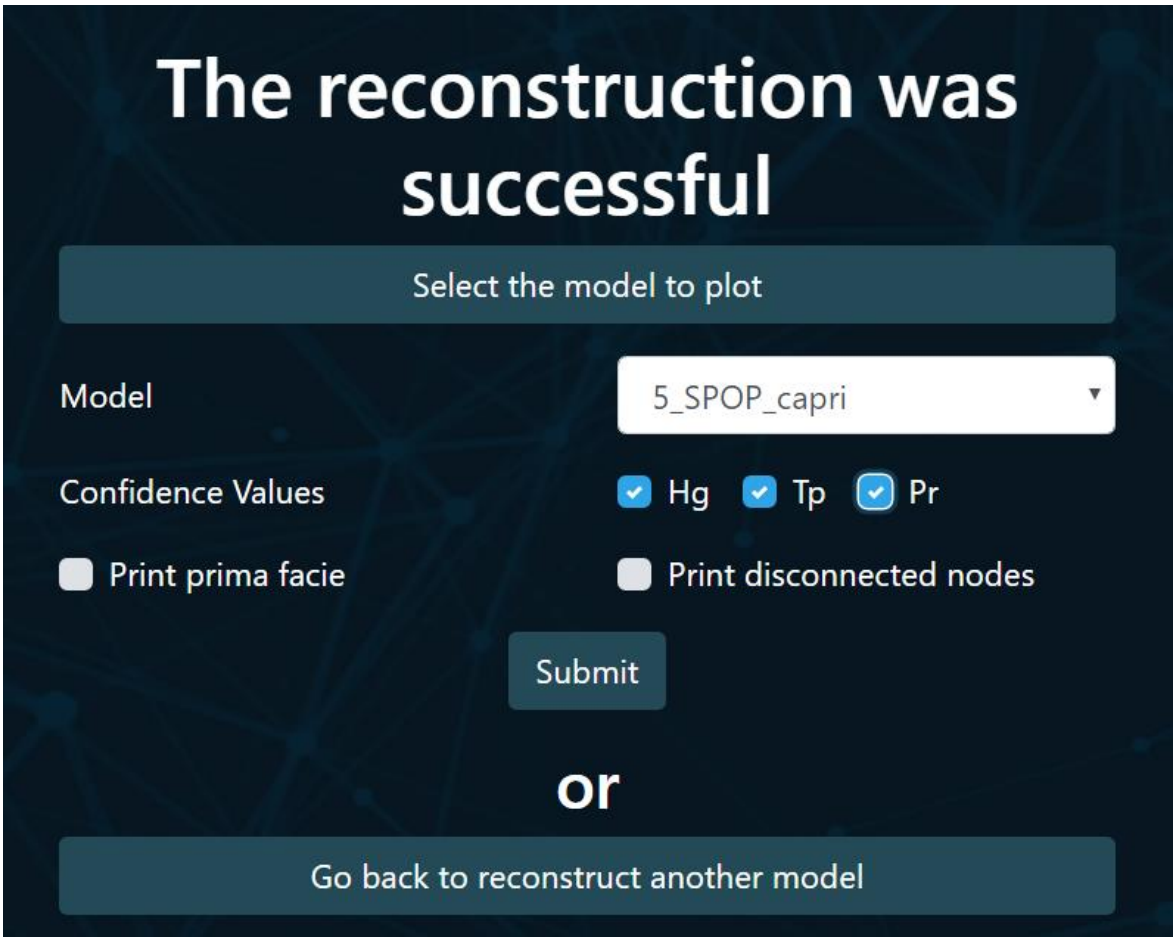
First, the user needs to select which one among the cluster will be used.



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

- **Command**: it defines the heuristic search to be performed.
- **Regularization**: select the regularization for the likelihood estimation.
- **Error rate**: enable or disable the estimation of the error rates give the reconstructed model.
- **Mutex**: 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

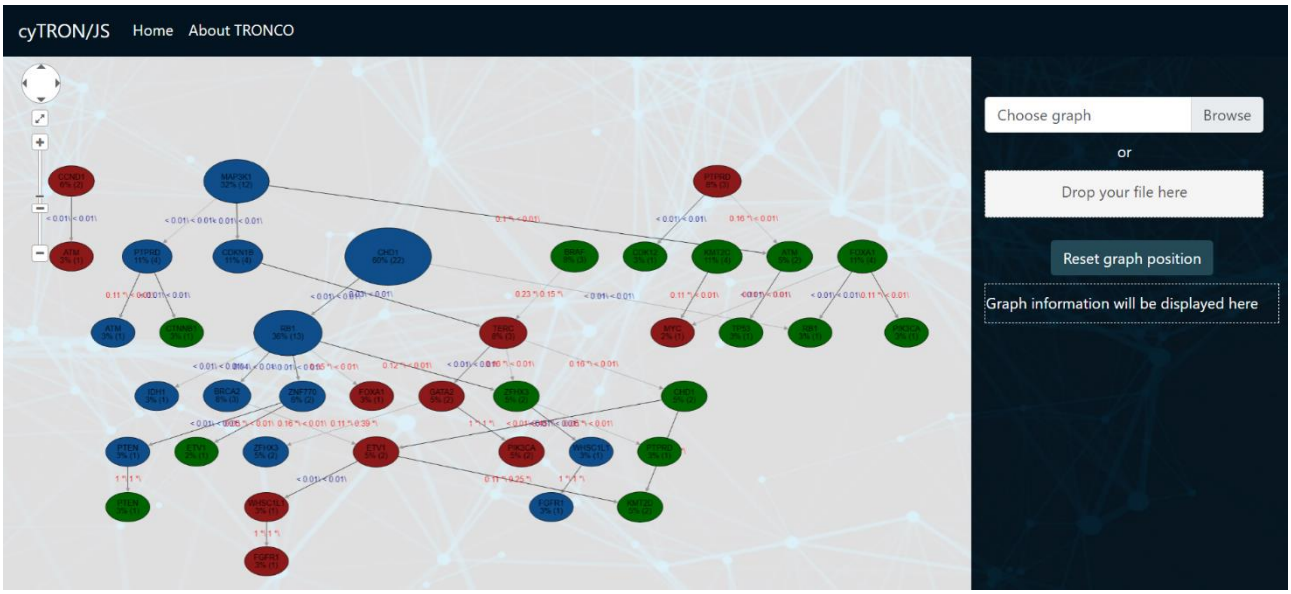


This is the page displayed after the reconstruction has finished. Here the user can decide to either reconstruct another model or plot one which has already been constructed.

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.