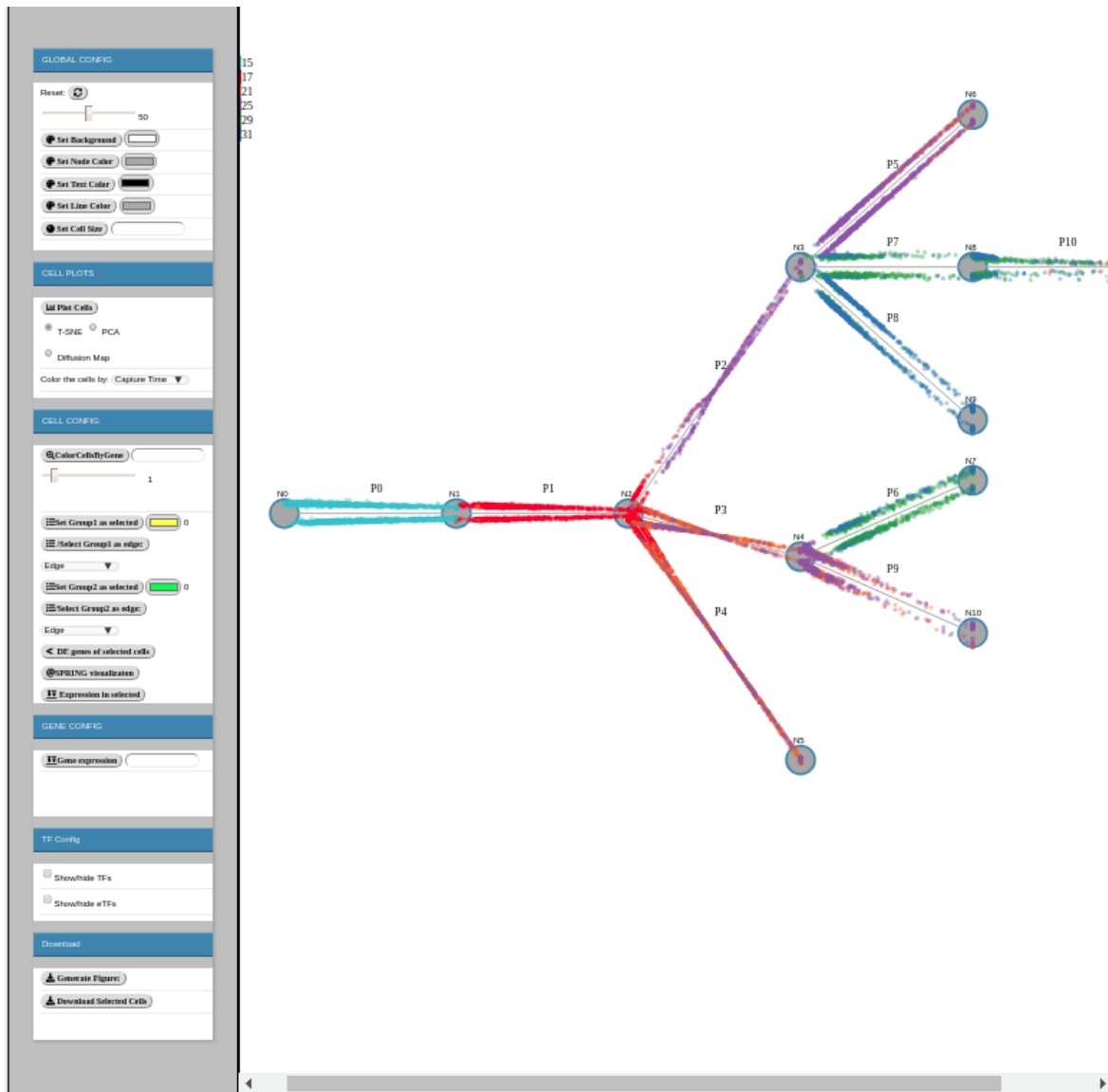
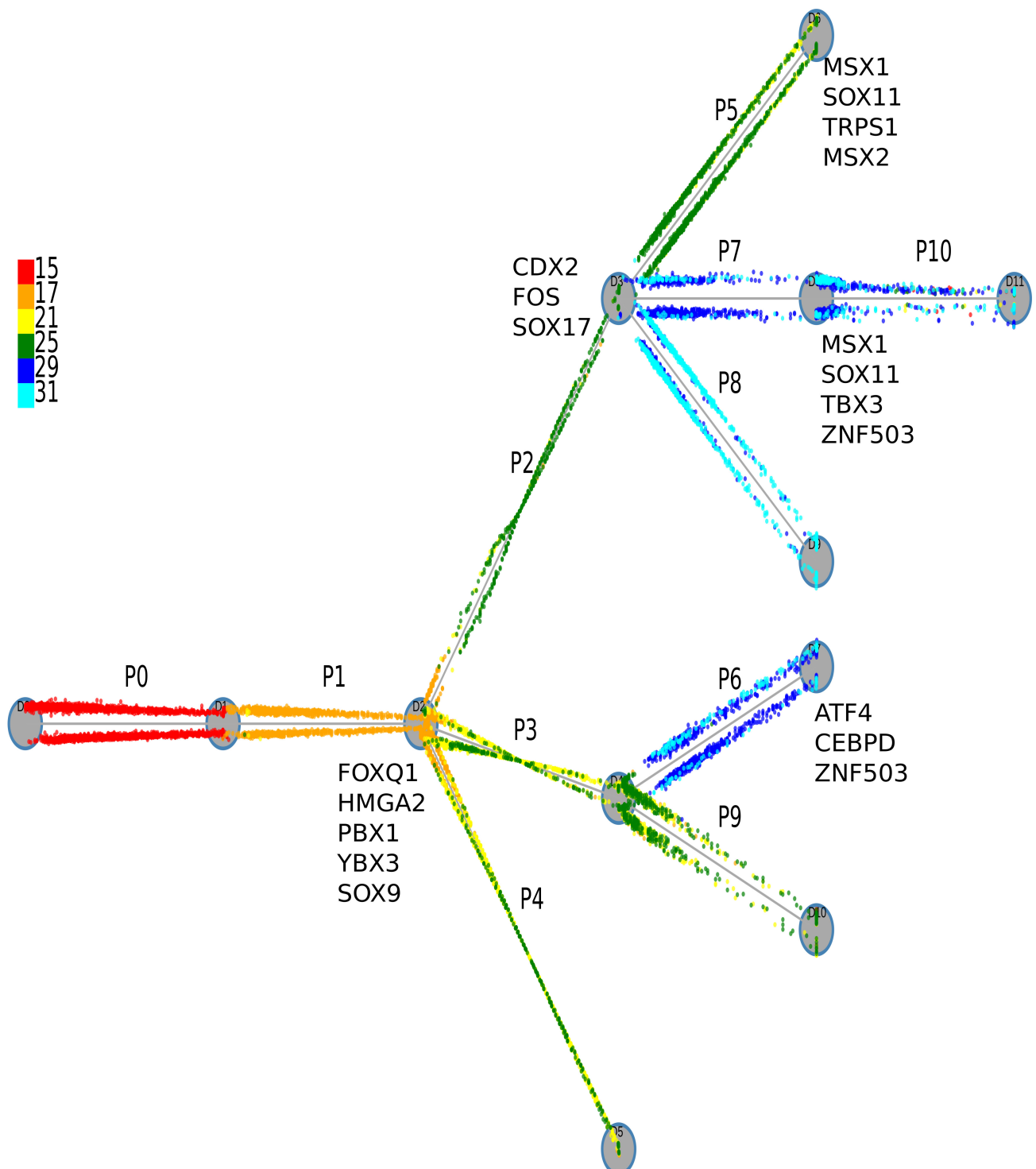


Cosimo-Continuous State Hidden Markov Model



Cosimo is an interactively visualized Continuous hidden Markov model to reconstruct the cell differentiation trajectories. With the model, users can generate many visualizations based on their specific needs. They can also make queries and predictions with their own interests. The model (right) can be configured by multiple panels (left), which were explained in details below.

Cosimo model



cell: a dot in the model, which is color-coded by the capture time of the cell.

P0-P10: paths in the model. Cells are differentiating into different cell fates via different paths.

N0-N11: Nodes in the model, which represent specific cell states (e.g. divergence states or end states).

TFs/eTFs: The regulators which dictate the cell differentiation in the path ending at that specific node.

The TFs are inferred based on their target genes. The target genes of the TFs are differentially expressed in the regulated paths. eTFs are predicted based on the expression of those TFs. To be more specific, the expression of the eTFs in the regulated paths are significantly different compared with the expression in the parent and sibling paths.

Global Config Panel

Set background, Node color, Text color, Line color, and Cell Size of the model

Cell Plots Panel

CELL PLOTS

Plot Cells

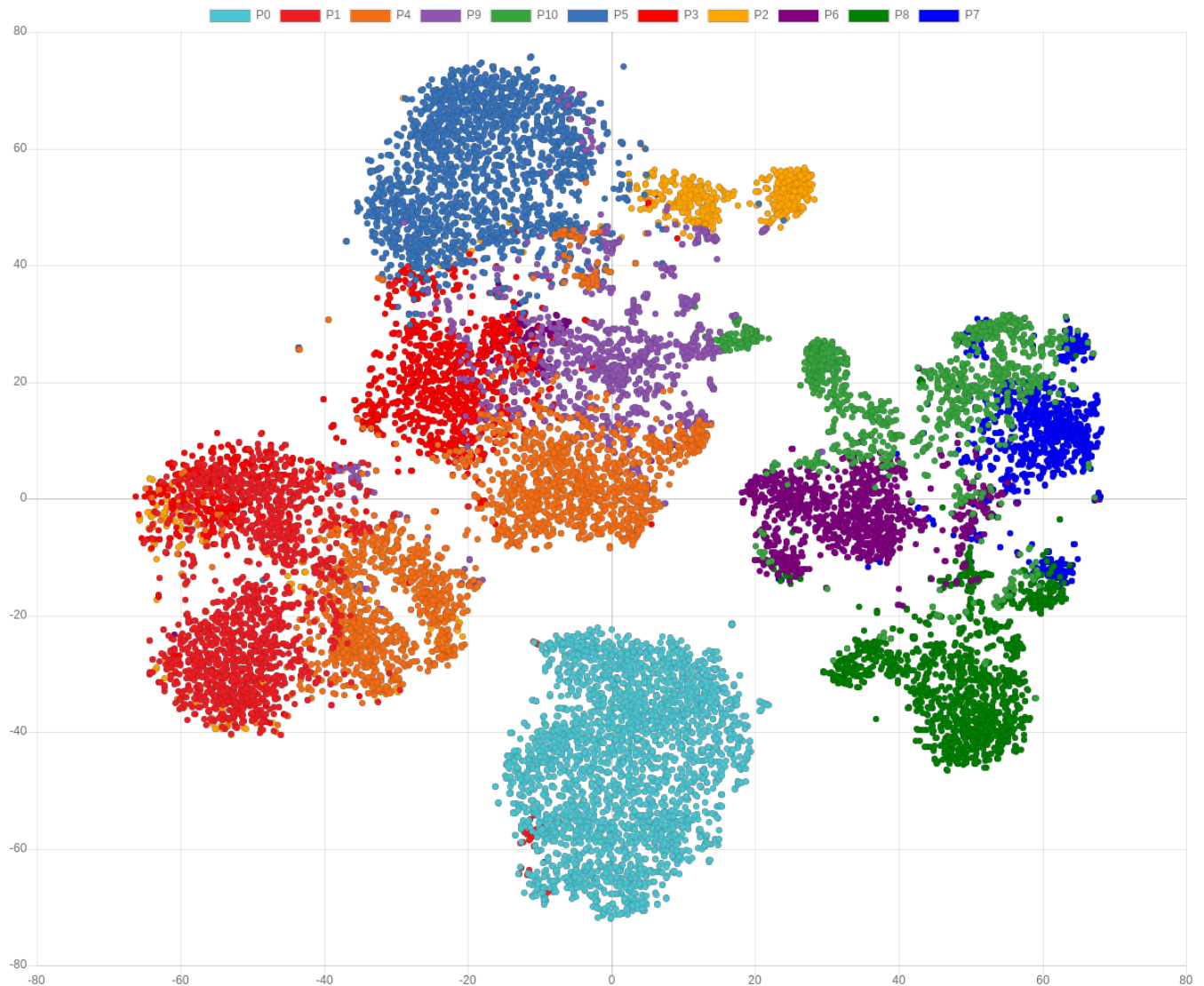
T-SNE PCA

Diffusion Map

Color the cells by: Capture Time

Get the t-SNE (<https://lvdmaaten.github.io/tsne/>), PCA (<https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.PCA.html>) and Diffusion map (<https://academic.oup.com/bioinformatics/article/31/18/2989/241305>) for all the cells in the dataset. You can also choose how would you want to color the cells using the "Color the cells by:" drop-down menu (capture time/Model path of the cell)

For example, t-sne plot and color-coded by the Model path (cells at different paths will be colored differently).



CELL Config Panel

CELL CONFIG

ColorCellsByGene

1

Set Group 1 as selected 0

Select Group 1 as edge: Edge

Set Group 2 as selected 0

Select Group 2 as edge: Edge

< DE genes of selected cells

@SPRING visualizaton

Expression in selected

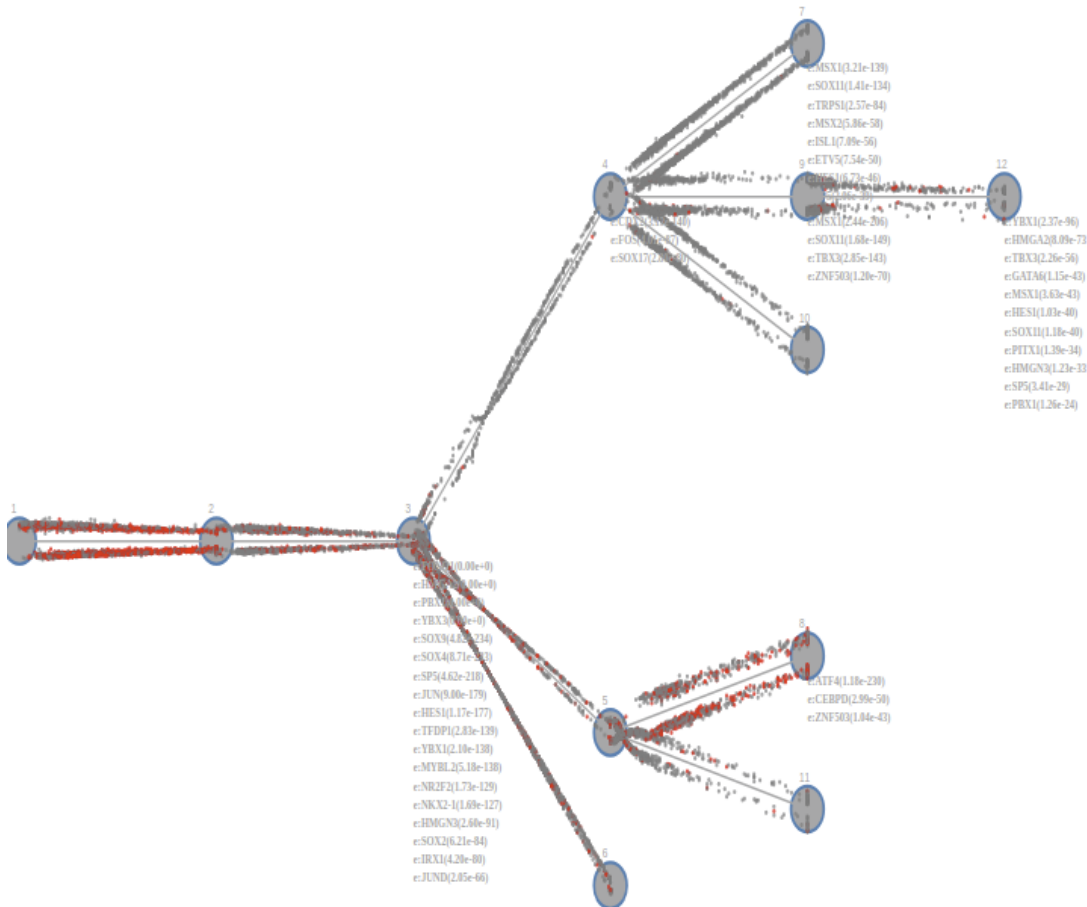
Clear all selected

(1) ColorCellsByGene:

ColorCellsByGene

1

Color all the cells based on the expression level of input genes.
 If > the level specified by the slider below => red
 Otherwise => gray
 The cell size can also be set by the Global config panel.



You can use the slider to customize the threshold. The coloring process usually takes a few secs but it varies based on the connection.

(2) Cell Group analysis:

☰ Set Group1 as selected 0

☰ /Select Group1 as edge: Edge ▼

☰ Set Group2 as selected 0

☰ /Select Group2 as edge: Edge ▼

< DE genes of selected cells

@SPRING visualization

⏏ Expression in selected

🗑 Clear all selected

(2.1) Set Group1 as selected:

While SHIFT pressed, use mouse to drag select the cells on the right. The selected cells will turn into gray.



Then, press the "Set Group 1 as selected" to appoint the selected cells as group 1.

The appointed cells will be colored as the specified group1 color.

☰ Set Group1 as selected 637

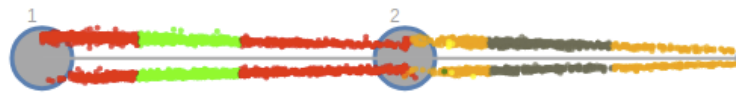
(2.2) Set Group2 as selected:

Use similar method to select Group2 cells.

☰ Set Group1 as selected 637

☰ /Select Group1 as edge: Edge ▼

☰ Set Group2 as selected 908



Or, you can use select all the cells in an edge as group1/group2 using the "/Select Group1(2) as edge" button.

(2.3) Get different genes between selected group1 and group2

You can get the DE genes between two selected groups using "DE genes of selected cells" button.

< DE genes of selected cells

Using these button, you can get all the up-regulated DE genes between group 1 and group 2.

If you want to get down-regulated DE genes between two groups, please switch the order of group1 and group2.

At the bottom of the DE gene list, you can see a button "functional analysis", which can be used to perform GO enrichment analysis for those genes.

Untitled - Google Chrome

about:blank

choose your species:

			0.6807894193329387
			0.6723550481455419
			0.6716514432640858
			0.6692529676983476
			0.6682654468547511
			0.665709563973447
			0.6629567434063695
			0.6610382486198866
			0.6597906195273513
			0.6539394149796162
			0.648260987730245
			0.6430195458658199
			0.6360546292347086
			0.6355245687180224
			0.6279527967615681
NACA3P	0.8686065070595466	1.4945385996409333	0.6259320925813867
CFDP1	0.29723756906077353	0.9188581687612194	0.6216205997004458
HMG3	0.35696746470227153	0.9768833034111308	0.6199158387088592
ID1	0.9388888888888897	1.558438061041292	0.6195491721524023
COTL1	0.44507673419275606	1.0642154398563726	0.6191387056636165
ZFAND5	0.3558133824432167	0.9711454219030514	0.6153320394598347
COL6A1	0.723787599754451	1.334337522441651	0.6105499226872
GDI2	0.8726580724370782	1.482653500897665	0.6099954284605867
ZFP36L1	0.24947820748925728	0.8545996409335729	0.6051214334443156
TMSB15A	0.3853161448741562	0.9854434470377024	0.6001273021635462

Table: Differentially expressed genes (UP-regulated) between Group1 and Group2

(2.4) Map to SRPING result

You can map the model to SPRING visualization using the "SPRING Visualization" button.

(2.5) Expression in selected groups

You can check the expression of input genes in selected cells.

(2.6) Clear all selected

You can unset all the groups by using this button.

Gene Config Panel

You can explore the expression for specific genes using this Panel.

(1) TFs are predicted based on their targets. Basically, target genes of a specific TF are differentially expressed in a specific path.

This inference is relying on the TF-gene interaction annotation we have. If we don't have good annotation for some TFs, we might miss them in the prediction.

That's why we also need eTF as a complement. eTF stands for expressed-based TF. In other words, if the expression of TF in a specific path/edge is significantly different compared with both parents and siblings.

In other words, it's unique. We also list those TFs as eTFs.

For the eTFs listed on N7, basically, it means the TFs regulating the path ending at N7 (P6).

To be more accurate, the expression of ATF4, CEBPD and ZNF503 are "different" in P6 compared with the parent (P3) and the sibling (P9). z

TF Config Panel



You can show/hide TFs and eTFs using this Panel.

(1) TFs: are predicted based on their targets.

Basically, target genes of a specific TF are differentially expressed in a specific path. This inference is relying on the TF-gene interaction annotation we have.

If we don't have good annotation for some TFs, we might miss them in the prediction. That's why we also need eTF as a complement.

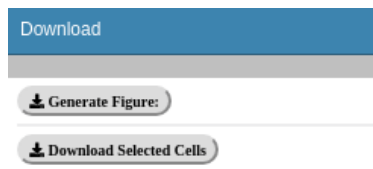
(2) eTF stands for expressed-based TF.

If the expression of a TF in a specific path/edge is significantly different compared with both parents and siblings, we regard it as an eTF.

For example the eTFs listed on N7 means TFs regulating the path ending at N7 (P6), which was inferred based on their expression.

To be more specific, the expression of ATF4, CEBPD and ZNF503 (at N7)s are "different" in P6 compared with the parent (P3) and the sibling (P9).

Download Panel



You can generate/download a figure for current visualization using the "Generate Figure" Button.

With the "Download Selected Cells", you will be able to download the cells you selected using the "Cell Config Panel" (Group1 and/or Group2 cells).