TieDIE Tutorial

Version 1.0

Evan Paull

June 3, 2013

Introduction

TieDIE is an algorithm that finds subnetworks connecting genomic perturbations to transcriptional changes in large gene interaction networks. To do this, the algorithm generates a 'diffusion" kernel describing the flow of information in the master network, using code written in SciPy (`http://www.scipy.org/`) or MATLAB (`http://www.mathworks.com/products/matlab/`). The computation of the diffusion kernel is computationally intensive–particularly for a large pathway–but, once computed, the kernel can be saved to a file and re-used when running the TieDIE algorithm on the same master pathway. It is recommended that kernels for large networks be generated using MATLAB, which is considerably faster and more memory-efficient, but a free SciPy implemention is provided for those without a MATLAB license. The entire program is written in python 2.7.X and requires the numpy-1.7+ package to be installed before running. TieDIE should run on any UNIX system, and has been tested on Linux and MacOS, at this point. Windows compatability is not supported at this point.

GBM Test Signaling Pathway Example

An example network and input is shown under "examples/GBM.test". Three input files are in this directory:

- pathway.sif: A signaling pathway extracted from the 2008 TCGA GBM Nature paper supplement [1]).
- upstream.input: Input heats for the 'upstream" set of genes, weighted by frequency of mutation or amplification.
- downstream.input: Input heats for the 'downstream" set of transcriptional responses.

To run the test, change directory to this folder and type 'make". The program should run in a few seconds and report to stderr, a subdirectory with the output will be created. Important output files are:

- report.txt: Network statistics and a summary report for the TieDIE subnetwork. Also includes the results of the permutation test.
- tiedie.sif: Connecting TieDIE subnetwork.
- tiedie.cn.sif: Connecting subnetwork after filtering for logically consistent paths.
- heats.NA: Node-attribute file formatted for Cytoscape input (visualization).

There are many software packages available for network visualization, but we've found the Cytoscape package [2] to be well suited for network visualization. The TieDIE output network (tiedie.sif) can be directly imported to cytoscape 2.8 or later, and the node attribute file (heats.NA) along with the supplied properties file (vizmap.props) can be used to color nodes by linker heat, as shown in the figure below.

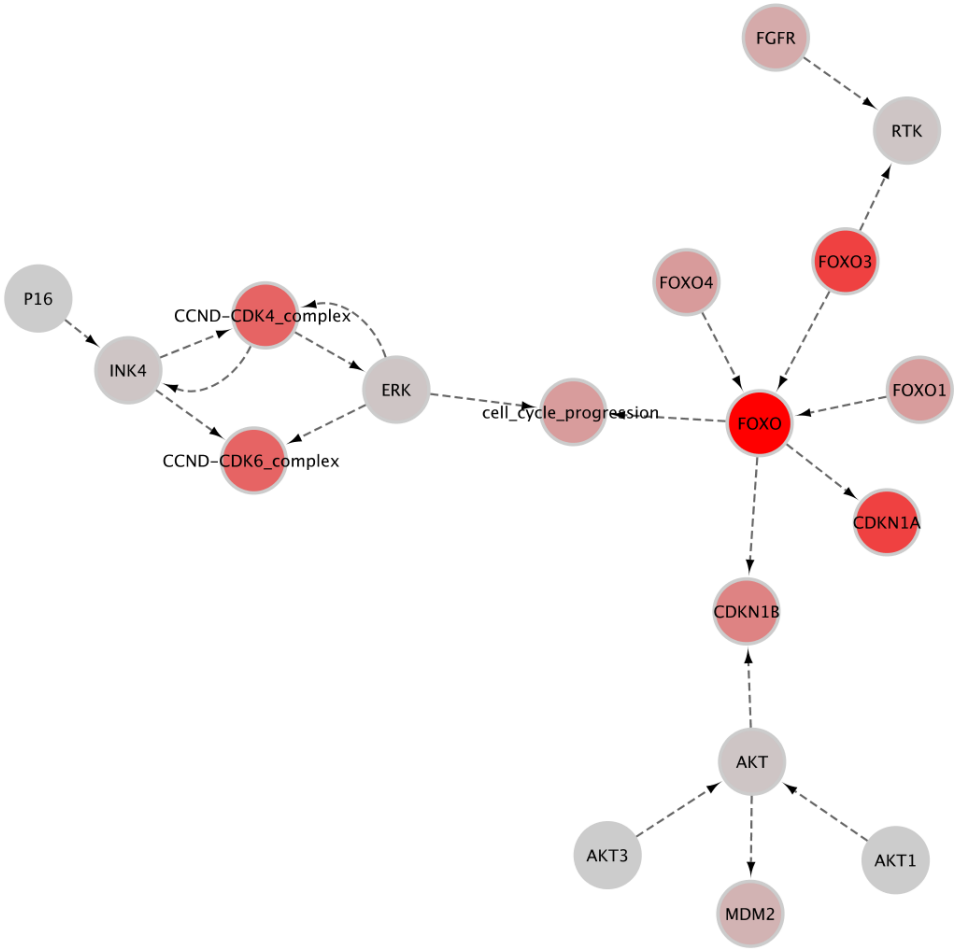

Figure 1: An small network example: sources are colored in blue, targets are red. Nodes and edges captured in the TieDIE Solution are outlined in green.

More detailed visualizations can also be performed in Cytoscape: for example, nodes can be colored or shaped based on source or target node status, and nodes/edges can be highlighted, as shown in the visualization below.

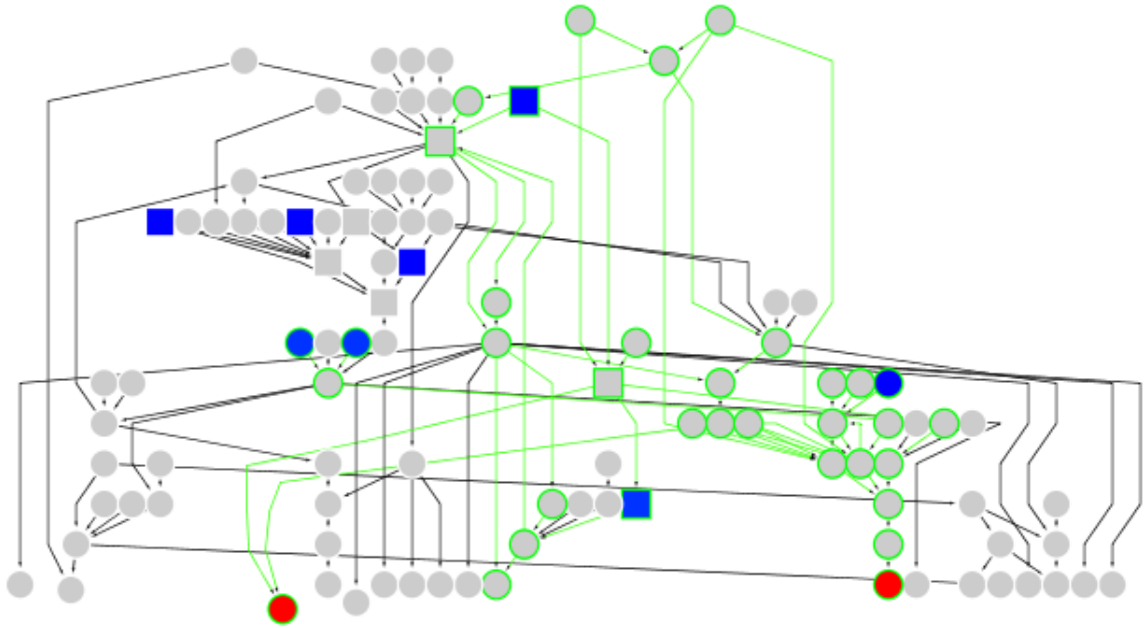

Figure 2: An small network example: sources are colored in blue, targets are red. Nodes and edges captured in the TieDIE Solution are outlined in green.

Bibliography

[1] The Cancer Genome Atlas Research Network*. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature*, 455, October 2008.

[2] Paul Shannon and et. al. Ideker, Trey. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research*, 2003.