Differential Dependency Network Analysis
MATLAB Toolbox v 1.1

1. Introduction

The differential dependency network (DDN) analysis is proposed to detect statistically significant topological changes in the transcriptional networks between two biological conditions. The detailed description of the method and experiments can be found in the paper "Differential Dependency Network Analysis to Identify Condition-Specific Topological Changes in Biological Networks" by Bai Zhang, Huai Li, Rebecca Riggins, Ming Zhan, Jianhua Xuan, Zhen Zhang, Eric P. Hoffman, Robert Clarke, and Yue Wang, 2008.

2. Usage

2.1 Syntax

To call differential dependency network analysis, use

ddn(data1, data2, max_K, geneNames, threshold, p_value_cutoff, expTitle)

2.2 Description

The <u>inputs</u> to the ddn function are:

data1	data under condition 1, $M \times N_1$ matrix, M : No. of features, N_1 : No. of samples
data2	data under condition 2, $M \times N_2$ matrix, M : No. of features, N_2 : No. of samples
max_K	the maximum size of set $\mathbf{Z}_i$, an integer between 1 and $M-1$
geneNames	gene names in a $M \times 1$ cell array
threshold	threshold of the fit of the local dependency model
p_value_cutoff	p value cutoff for selecting the statistically significant local structures
expTitle	the experiment title, a string.

The <u>outputs</u> of this program are three files saved in the folder results.

One output file is a ".sif" file, which specifies the network topology of the differential dependency network and can be imported in Cytoscape.

The second file is a ".NA" file, which is the node attribute file for visualization in Cytoscape.

The third file is a ".txt" file, which includes the "hot spots" identified by the DDN analysis, together with their fold-changes and *p*-values.

2.3 View the differential dependency network in Cytoscape

The ".sif" file can be imported and displayed in Cytoscape (http://www.cytoscape.org/), which is an open source bioinformatics software platform for visualizing molecular interaction networks.

3. An example

demo_simulation_study.m gives an example of the DDN analysis on a simulation dataset. The dataset can be found under data\SimulationData\, which contains the dataset generated by SynTReN (Van den Bulcke, et al., 2006), based on a sub-network of 20 genes drawn from an existing signaling network in *S. cerevisiae*. Under each condition, there are 50 samples.

In MATLAB command window, input

>> demo_simulation_study

Two files are saved under the folder results. Their file names (which include experiment title, date, and parameters in their file names) are:

simData20081028T174032_DDN_pvalue0.01_threshold0.25_K2.sif
simData20081028T174032_node_attribute0.01_threshold0.25_K2.NA
simData20081028T174032_HotSpot_pvalue0.01_threshold0.25_K2.txt

 "simData20081028T174032_HotSpot_pvalue0.01_threshold0.25_K2.txt" gives the "hot spots" identified by the DDN algorithm:

Hot spots	Fold change	p-value for t-test
CLB5	0.892426	6.920837E-043
CLB6	0.047876	1.995793E-001
FLO1	-0.050397	4.442320E-001
FLO10	0.773264	3.516430E-024
MBP1_SWI6	0.037370	4.575129E-001
PHO2	0.154190	1.338633E-005
SWI4	-0.123162	3.641910E-007
TRP4	0.082590	1.002472E-002

"simData20081028T174032_DDN_pvalue0.01_threshold0.25_K2.sif" can be viewed in Cytoscape. To download Cytoscape, please visit http://www.cytoscape.org/download.php?file=cyto2_6_0. To install Cytoscape, please refer to http://www.cytoscape.org/manual/Cytoscape2_6Manual.html.

After launching Cytoscape, select **File** -> **Import** -> **Network (multiple file types) …**, and then select the file "simData20081028T174032_DDN_pvalue0.01_threshold0.25_K2.sif" and import. Now the differential dependency network shows up. The menu "Layout" implements different layout methods, or simply drag the nodes in the network to obtain a proper layout.

"simData20081028T174032_node_attribute0.01_threshold0.25_K2.NA" is a node attribute file, which contains the fold changes of the gene expressions between two conditions (after base 2 logarithm).

In Cytoscape, select **File** -> **Import** -> **Node Attributes …**, and then select the file "simData20081028T174032_node_attribute0.01_threshold0.25_K2.NA" and import.

To better visualize the topological changes under two conditions, the Vizmap Property File "ddn.props" is designed to differentiate the network changes and represent expression changes under two conditions. Select **File** -> **Import** -> **Vizmap Property File …**, and then select the file "ddn.props" and import.

Then in the **Control Panel**, select Tab **VizMapper™**, and choose "ddn" for **Current Visual Style**, as shown in Figure 1.
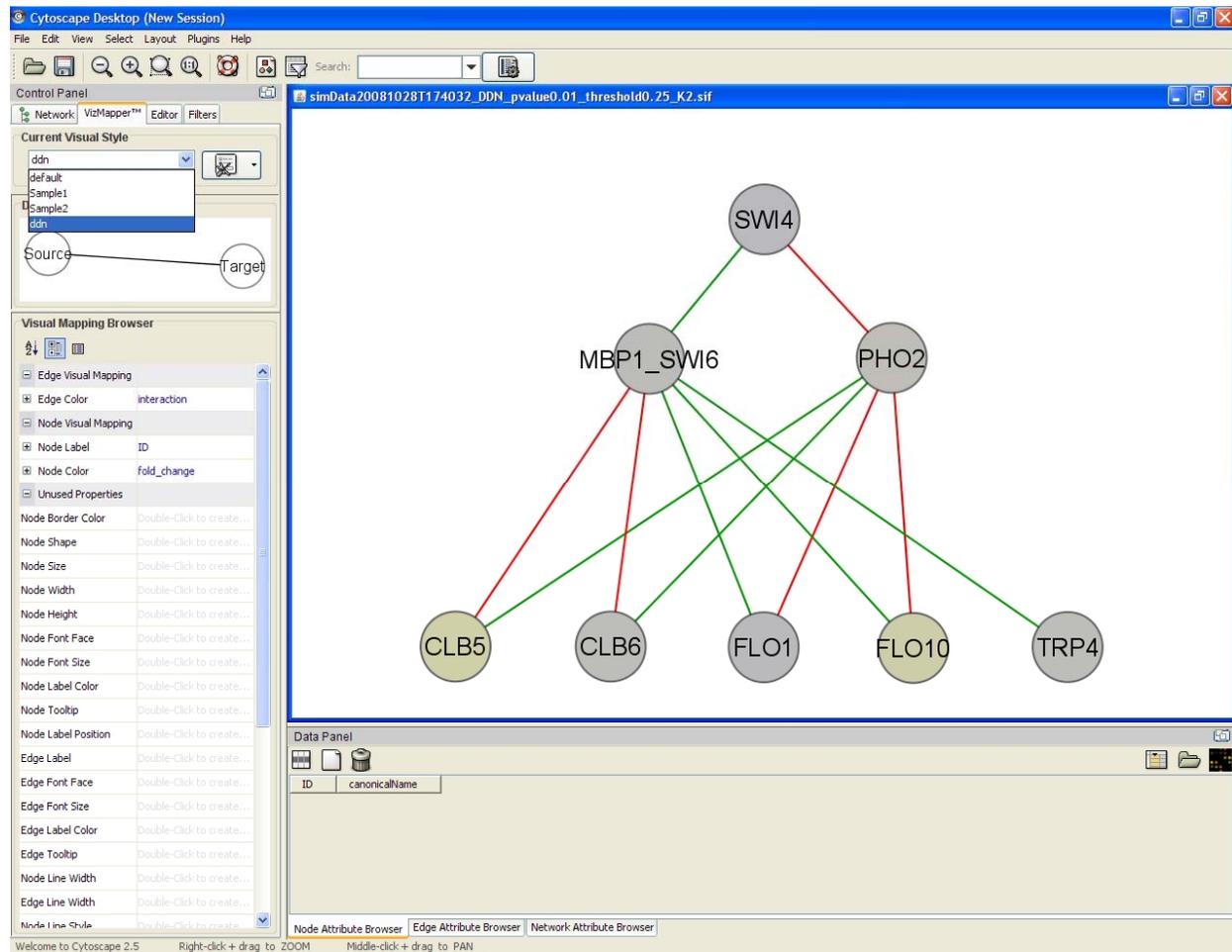

Figure 1. Use Cytoscape to visualize the differential dependency network.

In the network, the red lines represent the connections (dependencies) that only exist under condition 1, and the green lines represent the connections (dependencies) that only exist under condition 2. The colors on the nodes represent the fold changes of gene expressions under two conditions, calculated by

$$\log_2\left(\frac{mean(X_i^{(2)})}{mean(X_i^{(1)})}\right), \ i = 1, 2, ..., N.$$

Yellow implies up-regulation and blue implies down-regulation.