

Lecture IV

Demonstration of KDA Package

Bin Zhang, PhD

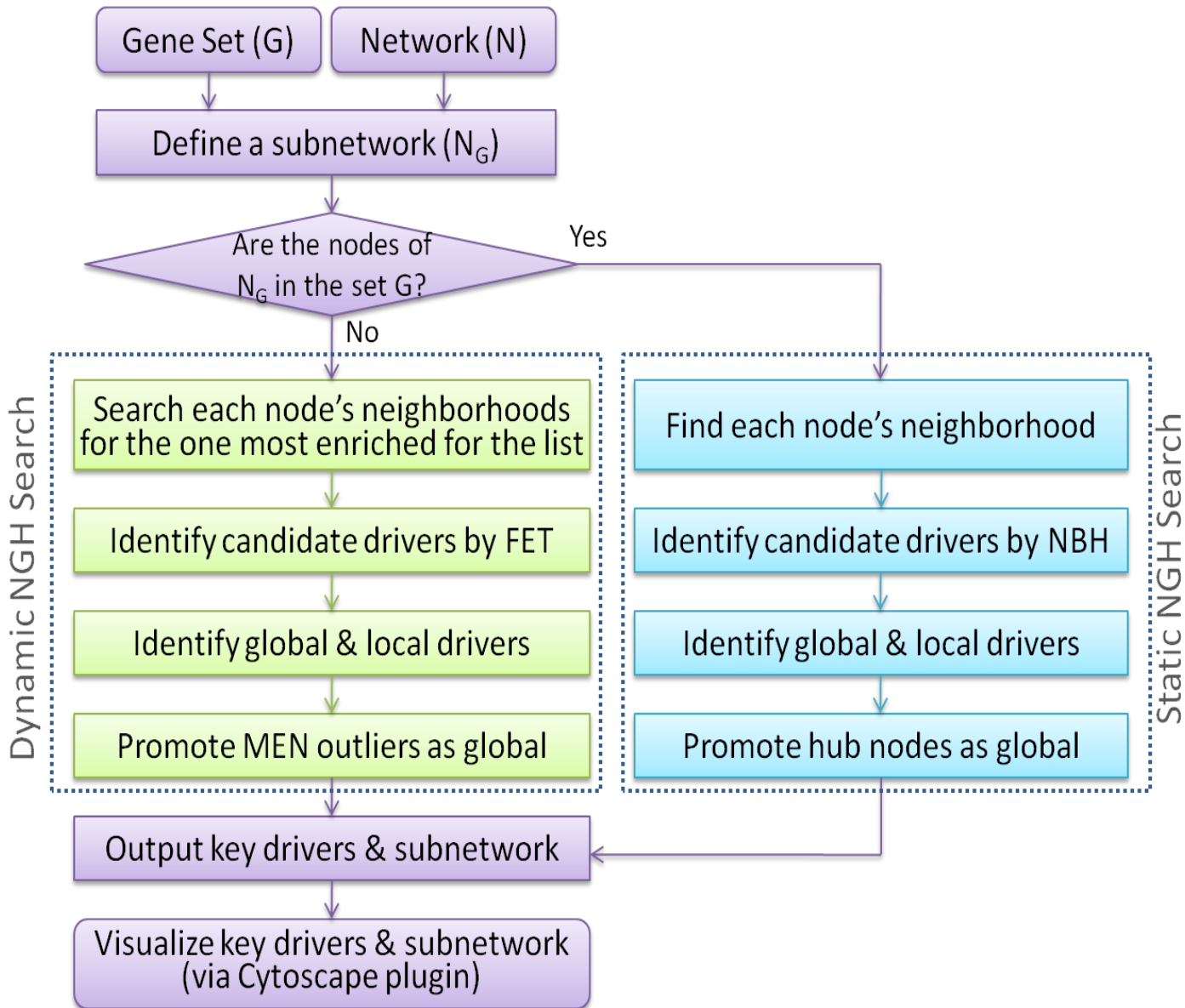
Department of Genetics & Genomic Sciences
Icahn Institute of Genomics and Multiscale Biology
Icahn School of Medicine at Mount Sinai, New York, USA

Email: bin.zhang@mssm.edu

Web: <http://research.mssm.edu/multiscalenetwork>



Key Driver Analysis



Dynamic Neighborhood Search (DNS)

- **Input**

- a set of nodes (**G**)
- a directed/undirected network (**N**)

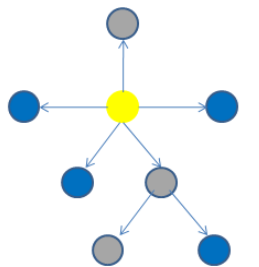
- **Procedure**

- Generate a sub-network N_G , defined as the set of nodes in N that are no more than h -layers away from the nodes in G .
- Search the h -layer neighborhood ($h=1,\dots,H$) for each gene in N_G ($HLN_{g,h}$) for the optimal h^* , such that

$$ES_{h^*} = \max(ES_{h,g}) \forall g \in N_g, h \in \{1..H\}$$

where $ES_{h,g}$ is the computed enrichment statistic for $HLN_{g,h}$

- A node becomes a candidate driver if its HLN is significantly enriched for the nodes in G
- Candidate drivers without any parent node (i.e., root nodes in directed networks) are designated as global drivers and the rest are local drivers.



● Gene in the list
● Key driver

Enrichment Test

Static Neighborhood Search (SNS)

- **Input**

- a set of nodes (**G**)
- a directed/undirected network (**N**)

- **Procedure**

- Compute the size of the h-layer neighborhood (HLN) for each node.
- Let μ be the sizes of HLN and d be the out-degrees for all the nodes
- Nodes with HLN sizes greater than $\bar{\mu} + \sigma(\mu)$ are considered as candidate drivers
- Candidate drivers without any parent node (i.e., root nodes in directed networks) are designated as global drivers and the rest are local drivers.
- Promote hub nodes nodes with out-degrees above $\bar{d} + 2\sigma(d)$ as global drivers

KDA Package

- Permanent Link to KDA Package

- <http://research.mssm.edu/multiscalenetwork/Resources.html>

- Installation

```
install.packages("C:/ZSinai/KDA/Release/KDA_0.1.zip", lib="C:/ZSinai/RLocalLibs", repos=NULL)
```

- Test

```
library("KDA", lib.loc="C:/ZSinai/RLocalLibs/")  
demo("breastCancer") # SNS  
demo("yeast") # DNS
```

Inputs

Input:

1. gene list: finputlist

2. causal network: fcausalnet

3. directed or undirected network: directed = TRUE

4. expand subnetwork based on L-layer neighbors: layer=0

5. gene annotation file(NULL if not available): fgeneinfo=NULL

#

- Gene list should be a matrix in which the first column includes node names and the last column includes each node's group (module) membership
- Drivers will be identified for each group (module) of nodes
- Causal network should be a matrix in which the nodes in the 1st column are causal for or associated with the nodes in the 2nd column.
- "layer" can be set as 2 or 3

Outputs (1)

- # 1. keydrivers for each subnetwork: "*_keydriver.xls"
- # 2. Cytoscape network: *_cys.txt
- # 3. Cytoscape node properties: *_cys-nodes.txt
- # 4. combined key drivers for all runs: "_KDx_combined.xls"
- # 5. parameters used for key driver analysis: "_KDx_parameters.xls"
- #

1-global driver
0-local driver

1. keydrivers for each subnetwork: "*_keydriver.xls"

keydriver	is_signature	hits	downstream	signature	subnetwork	signature	network	signature	optimal_length	fold_change	pvalue_whole	fold_change	pvalue_subnet	pvalue_corrected_subnet	keydriver
YNL083W	1	98	117	267	461	267	461	277	2	1.446205	9.16E-12	1.44620506	9.16E-12	4.22E-09	1
YIL093C	1	142	195	267	461	267	461	277	4	1.257313	1.76E-08	1.25731297	1.76E-08	8.11E-06	1
YDR175C	1	49	56	267	461	267	461	277	2	1.510768	4.08E-07	1.51076779	4.08E-07	0.000188193	1
YMR024W	1	77	102	267	461	267	461	277	2	1.303408	2.59E-05	1.30340751	2.59E-05	0.011951523	1

Enrichment test P value

Outputs (2)

- # 1. keydrivers for each subnetwork: "*_keydriver.xls"
- # 2. Cytoscape network: *_cys.txt
- # 3. Cytoscape node properties: *_cys-nodes.txt
- # 4. combined key drivers for all runs: "_KDx_combined.xls"
- # 5. parameters used for key driver analysis: "_KDx_parameters.xls"
- #

2. Cytoscape network: *_cys.txt

V1	V2
YAL016W	YGL161C
YAL016W	YML127W
.....	

3. Cytoscape node properties: *_cys-nodes.txt

nodename	color	shape	size	font_size
YAL009W	grey	circle	40	12
YAL016W	lightgreen	circle	40	12
YAL039C	grey	circle	40	12
YBL007C	grey	circle	40	12



Legend of node shape, size and color

label	shape	color	border	node_size	font_size
SIG_GKD	square	red	red	120	36
SIG_LKD	square	blue	blue	80	24
SIG_NKD	circle	lightgreen	lightgreen	40	12
NSIG_GKD	circle	red	red	120	36
NSIG_LKD	circle	blue	blue	80	24
NSIG_NKD	circle	grey	grey	40	12

SIG – in the gene list
NSIG – not in the gene list

GKD – Global Driver
LKD – Local Driver
NKD – Not A Driver

Visualization

KDA outputs as input of Cytoscape

- # 2. Cytoscape network: *_cys.txt
- # 3. Cytoscape node properties: *_cys-nodes.txt

Acknowledgements

Sage Bionetworks

- Justin Guinney
- Dave A Henderson