Nature Precedings : doi:10.1038/npre.2012.7118.1 : Posted 3 Apr

From Machine Learning to Learning Machines - A Perspective toward Personalized Medicine

Malay Bhattacharyya

malay_r@isical.ac.in
Machine Intelligence Unit, Indian Statistical Institute, Kolkata

Presented at C2B2, Columbia University

November 17, 2011

Outline of the talk

- Introduction
- Learning Machines: A Bottom-up Approach with a Network
- Analysis of Networks
- From Networks to Personalized Medicine

Precedings : doi:10.1038/npre.2012.7118.1 : Posted 3 Apr

Precedings : doi:10.1038/npre.2012.7118.1 : Posted 3 Apr

Introduction

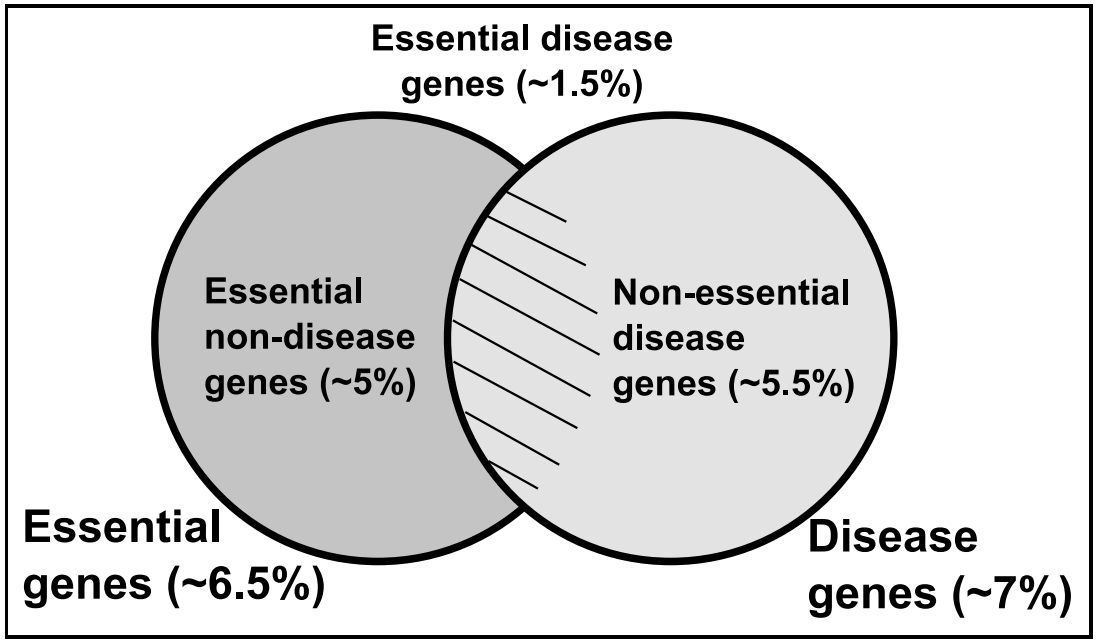
The Interactomes

- Molecular Networks
  - Protein-protein Interaction Networks
  - Metabolic Networks
  - Regulatory Networks - TF-gene Networks
  - Post-translational Networks - Kinase-substrate Networks
  - RNA Networks - TF-miRNA Networks, miRNA-gene Networks
- Phenotypic Networks
  - Co-expression Networks
  - Genetic Networks
- What is a Disease Network?

Nature Precedings : doi:10.1038/npre.2012.7118.1 Posted 3 Apr

Disease Genes (A Snapshot from 2007)

Precedings : doi:10.1038/npre.2012.7118.1 : Posted 3 Apr



- Less than 10% of human genes are known to have association with specific diseases [Barabási, 2011].
- Of these ∼19% are known to be oncogenes (November 16, 2011; Cancer Genome Project).

Precedings : doi:10.1038/npre.2012.7118.1 : Posted 3 Apr

Learning Machines: A Bottom-up Approach with a Network

A Framework of Co-expression Analysis

- Co-expression
- Differential Expression
- Differential Co-expression
- Co-expression Dynamics

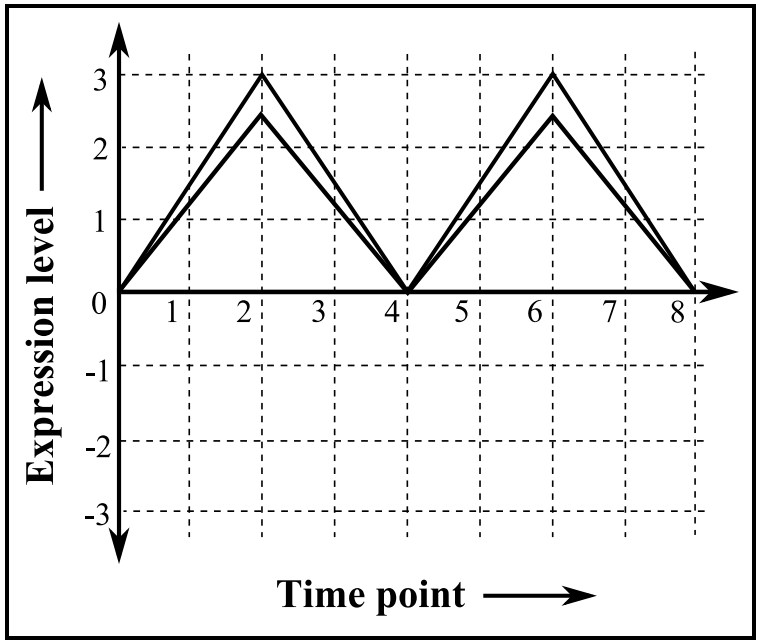
Nature Precedings : doi:10.1038/npre.2012.7118.1 : Posted 3 Apr

Co-expression

Precedings : doi:10.1038/npre.2012.7118.1 : Posted 3 Apr

Definition

Pairwise similarity pattern (spatial or temporal) of expression vectors.



- Different statistical measures [Bandyopadhyay, 2011].

Differential Expression

Nature Precedings : doi:10.1038/npre.2012.7118.1 : Posted 3 Apr

Definition

Varying patterns (spatial or temporal) of expression vectors in different phenotypes.

- SAM, t-test, etc.

Differential Co-expression

Nature Precedings : doi:10.1038/npre.2012.7118.1 : Posted 3 Apr

Definition

Pairwise varying dependence (spatial or temporal) between expression vectors in different phenotypes.

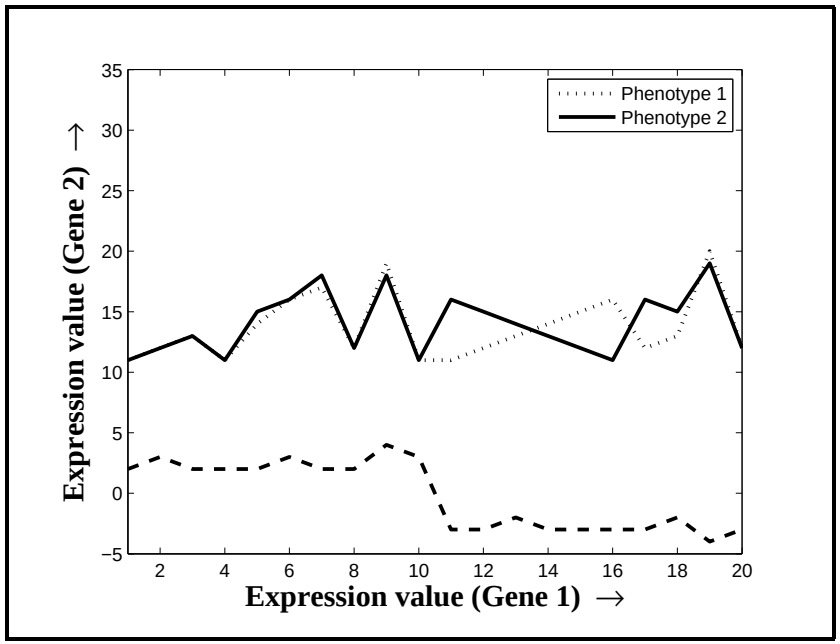
- On/off case and gap/subsitution case [Dettling, 2005].

Co-expression Dynamics

Precedings : doi:10.1038/npre.2012.7118.1: Posted 3 Apr

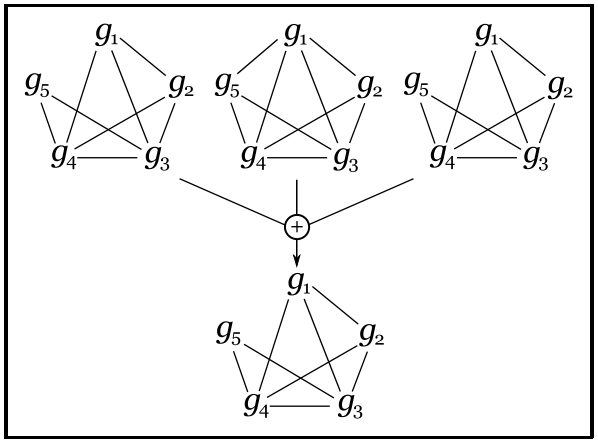
Definition

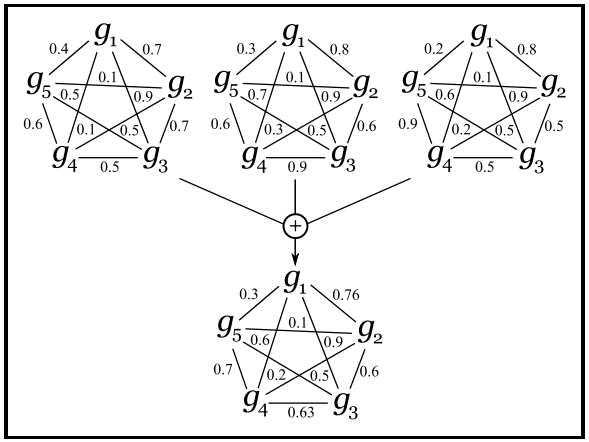
Pairwise varying dependence (spatial or temporal) between expression vectors based on another expression vector.



[Li, 2002].

Making Networks Robust by Integration





Supervised Approach

$$LS = \ln \frac{P(L|E)/\sim P(L|E)}{P(L)/\sim P(L)}$$

$P(L|E)$ and $\sim P(L|E)$ are the observed linkage frequencies of annotated genes operating in the same and in different pathways, respectively, whereas $P(L)$ and $\sim P(L)$ are the prior annotations [Lee, 2004].

Unsupervised Approach

$$W_c(i,j) = \sqrt[\alpha]{\sum_{k=1}^{n} \xi_k(i,j) W_k(i,j)^\alpha}$$

Preprints: doi:10.1038/npre.2012.7118.1 : Posted 3 Apr

Nature Precedings : doi:10.1038/npre.2012.7118.1 : Posted 3 Apr

Analysis of Networks

Understanding the Diseasome

Nature Precedings : doi:10.1038/npre.2012.7118.1 : Posted 3 Apr 2012

Hypotheses

Degree: Disease biomolecules avoid hubs

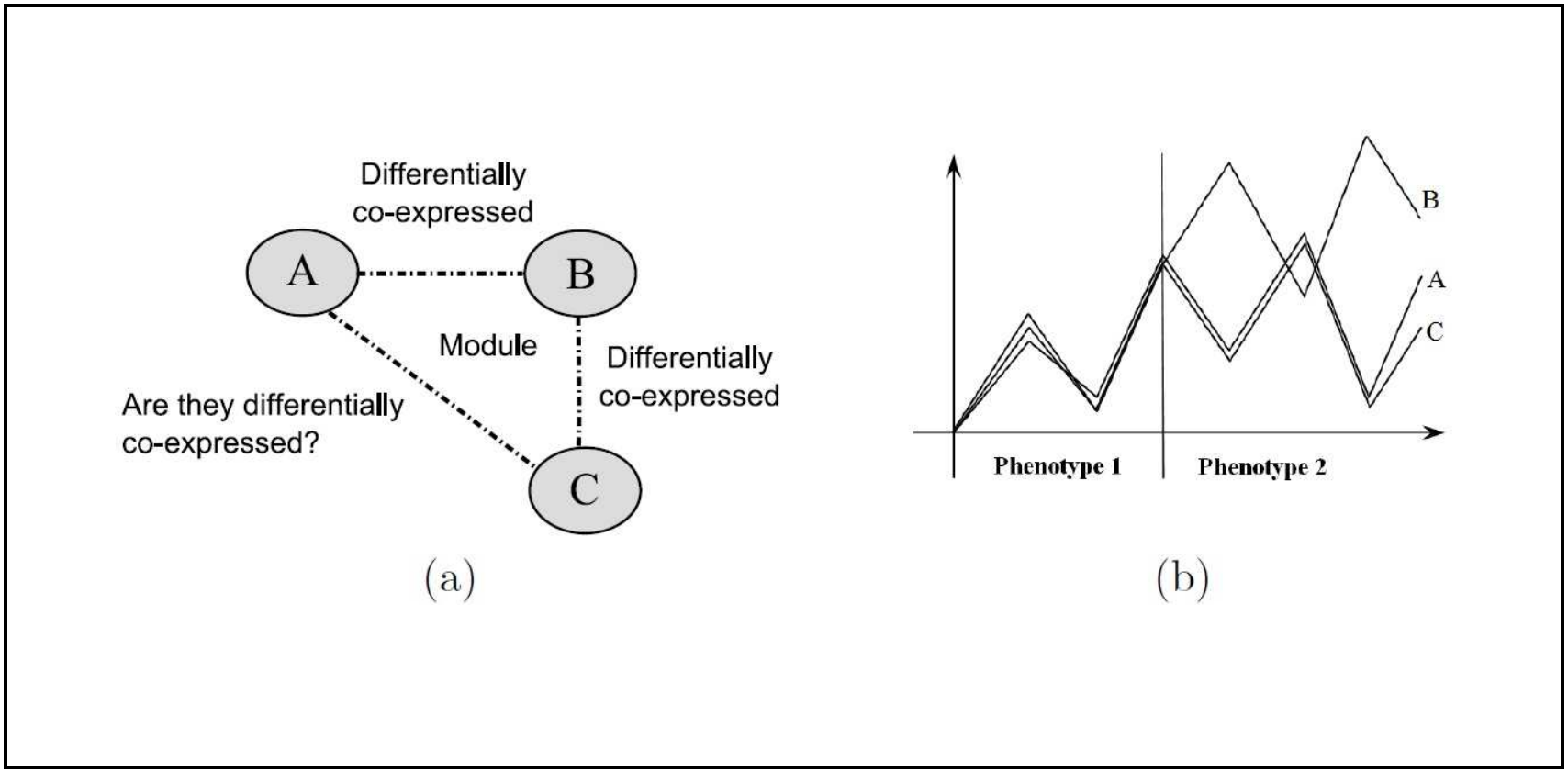
Modularity: Biomolecules specific to a disease form modules

Sharing: Diseases having common biomolecules show phenotypic similarity

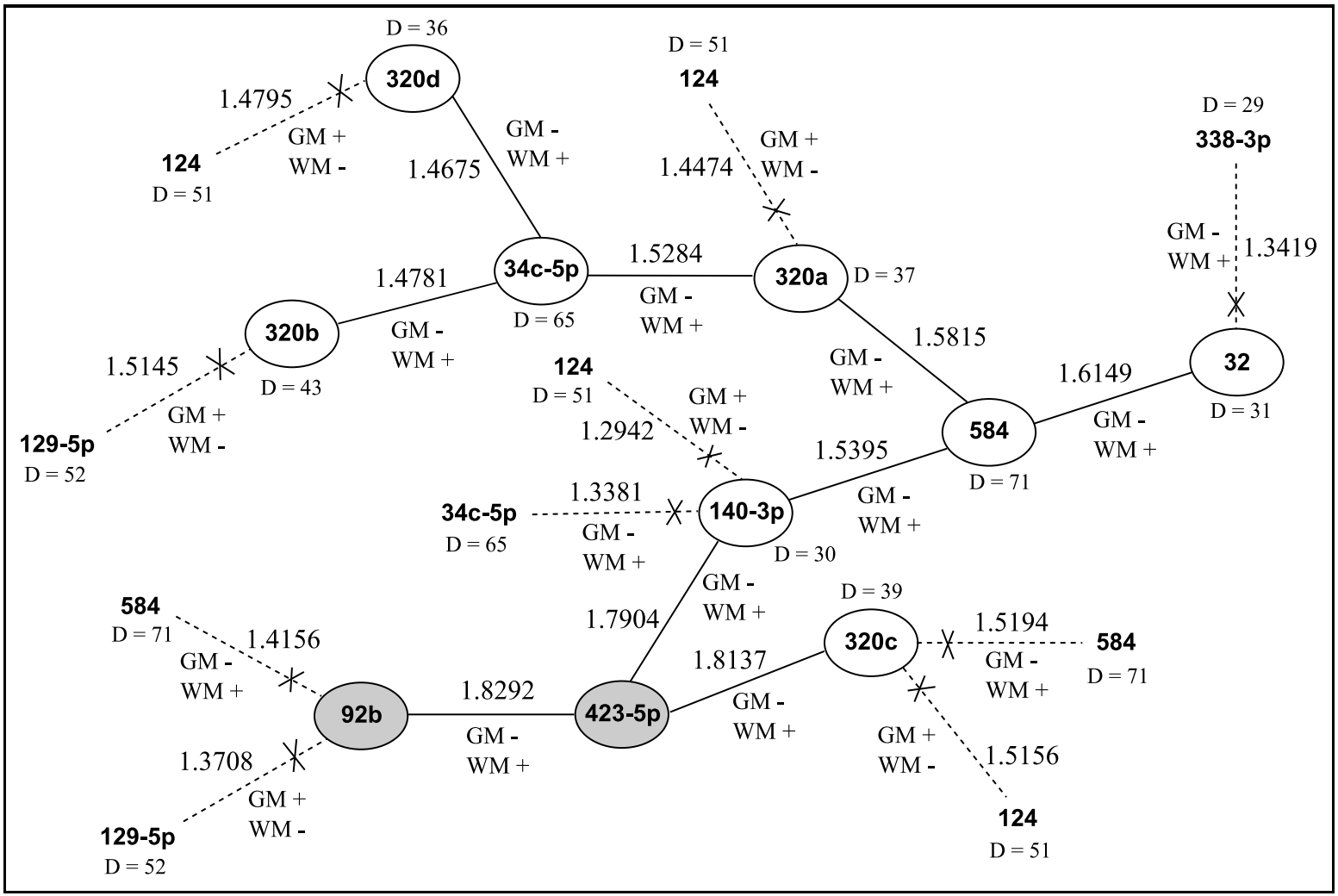
Closeness: Causal pathways coincide with the connectors of known disease-subnetworks

Differential Co-expression to Module Finding

Nature Precedings : doi:10.1038/npre.2012.7118.1 : Posted 3 Apr



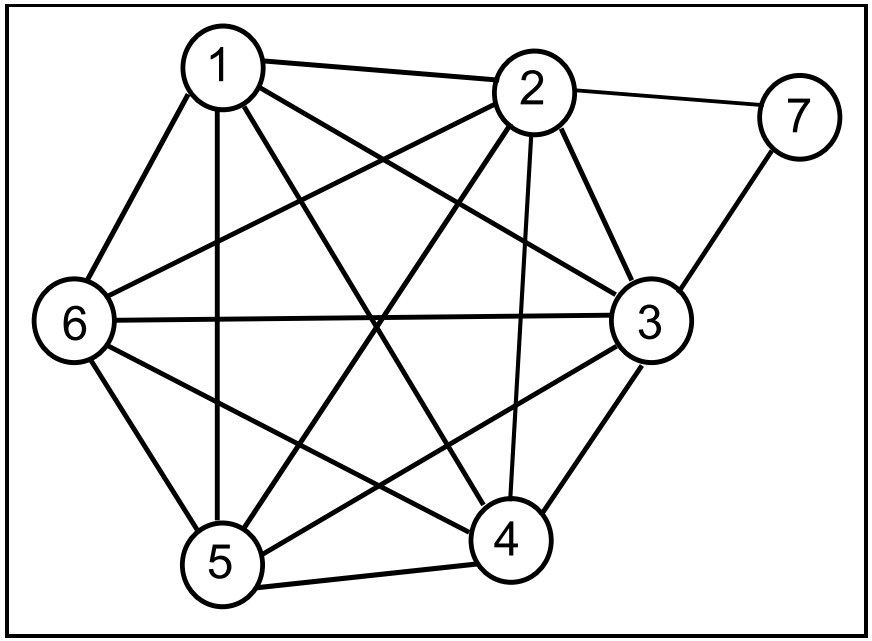
Differentially Co-expressed Switching Tree



Preceedings : doi:10.1038/npre.2012.7118.1 : Posted 3 Apr

How to Define an Association as Dense?

Precedings : Doi:10.1038/npre.2012.7118.1 : Posted 3 Apr



Density $= 17/21 \sim 0.81$

Minimum participation density $= 2/6 \sim 0.33$

A Dense Module

Preceedings : doi:10.1038/npre.2012.7118.1 Posted 3 Apr

Association density of a vertex

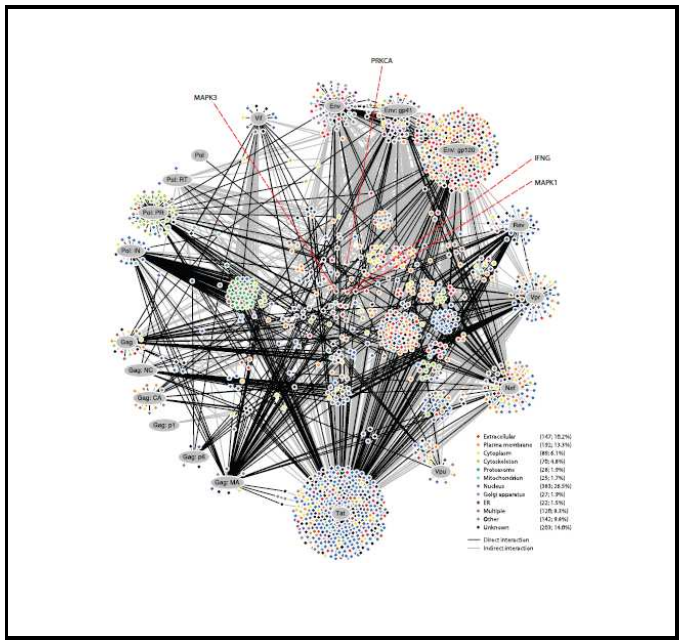
Given a weighted *graph* $G = (V, E, \Omega)$, the association density, μ_{v_i/V^N}, of a vertex v_i of G with respect to a vertexlet V^N (such that $v_i \notin V^N$), is defined as the ratio of the sum of the edge weights between v_i and each of the vertices belonging to V^N, and the cardinality of the set V^N. Thus, the association density of a vertex v_i with respect to the vertexlet V^N is computed as $\mu_{v_i/V^N} = \frac{\sum_{v_j \in V^N} \Omega_{v_i v_j}}{N}$.

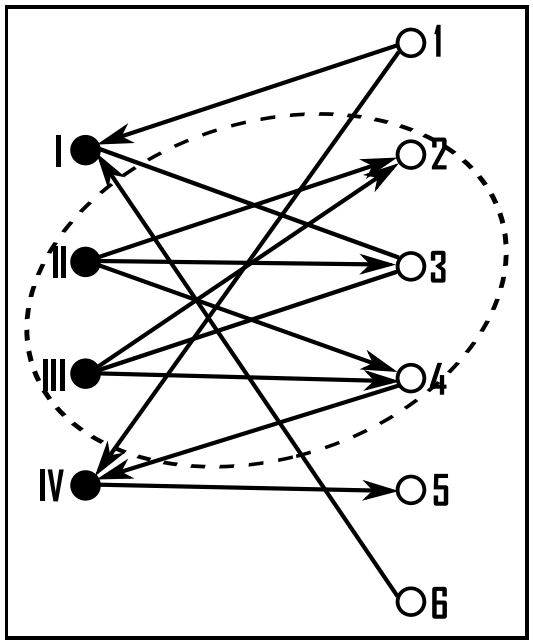
Association density of a vertexlet

The association density of a vertexlet V^N is defined to be the minimum of the association density of every vertex belonging to the vertexlet with respect to the vertexlets of order $(N-1)$. So, the association density of a vertexlet V^N is given by $\mu_{V^N} = \min_{v_i \in V^N} \left(\mu_{v_i/V^N - \{v_i\}} \right)$.

Regulatory Network Analysis

Nature Precedings : doi:10.1038/npre.2012.7118.1 : Posted 3 Apr





[Ptak, 2008].

DBClique

A DBClique is a fully connected subgraph $G' = (V_1', V_2', E') \subseteq G$ of a directed bipartite graph G such that either $i \in V_1', j \in V_2', \forall (i,j) \in E'$ or $i \in V_2', j \in V_1', \forall (i,j) \in E'$.

Nature Precedings : doi:10.1038/npre.2012.7118.1 : Posted 3 Apr

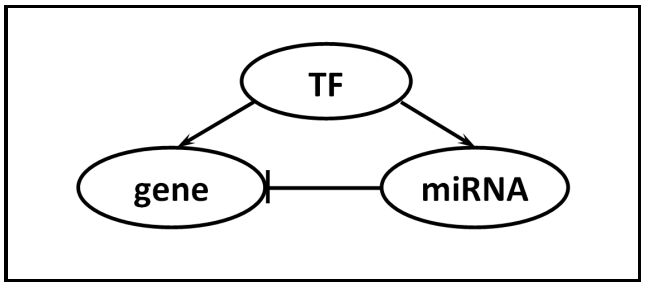
	Endogenous	**Exogenous**
TF combinations	complex	simple
Path length	long	short
TF outdegree	Low	High
Significant motifs	FFLs	single input
TF interconnectedness	High	Low

Table: Topological properties of Networks.

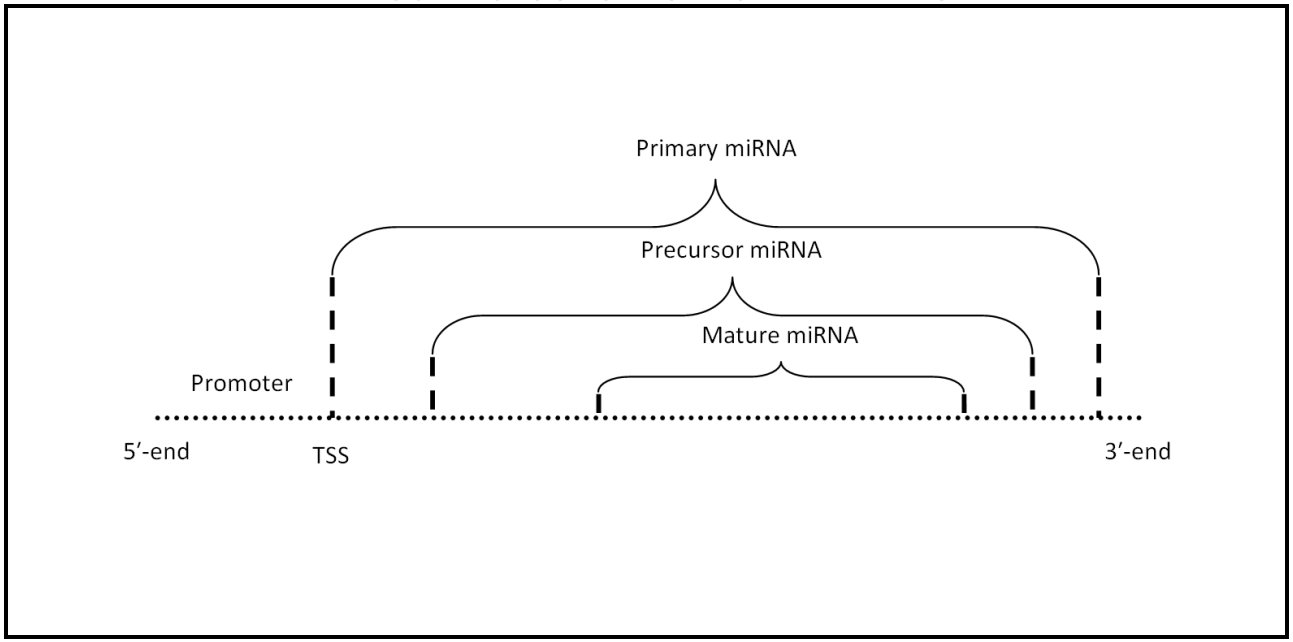
[Luscombe, 2004].

Strengthening the Network Analysis

Precedings : doi:10.1038/npre.2012.7118.1 : Posted 3 Apr



Activities of small RNAs



Transcriptional regulation of miRNAs

Nature Precedings : doi:10.1038/npre.2011.7118.1 : Posted 3 Apr

From Networks to Personalized Medicine

Drugs and Biomarkers

Nature Precedings : doi:10.1038/npre.2012.7118.1 : Posted 3 Apr

How?

- Prepare integrative network models
- Correlate network dynamics and states to the phenotype and patient disease data
- Identify potential multi-node signatures

A test case of cancer metastasis is proposed pursuing the 'seed and soil' principle [Erler, 2009].

More focus on aberrations and pathways

- Identify the genetic aberrations and the master regulators that drive proliferation, survival, metastasis, and drug resistance
- Model the adaptive/feedback mechanisms that thwart the efficacy of potent drugs
- Predict additional target pathways for combinatorial drug treatment

[Ray, 2010, Pe'er, 2011].

References

S. Bandyopadhyay and M. Bhattacharyya (2011)
A Biologically Inspired Measure for Coexpression Analysis.
IEEE/ACM Trans Comput Biol Bioinform 8(4), 929 – 942.

A. Barabási, N. Gulbahce and J. Loscalzo (2011)
Network medicine: a network-based approach to human disease.
Nat Rev Genet 12, 56 – 68.

M. Dettling, E. Gabrielson and G. Parmigiani (2005)
Searching for differentially expressed gene combinations.
Genome Biol 6, R88.

J. T. Erler and R. Linding (2009)
Network-based drugs and biomarkers.
J Pathol 220(2), 290 – 296.

I. Lee *et al.* (2004)
A probabilistic functional network of yeast genes.
Science 306, 1555 – 1558.

Preceedings : doi:10.1038/npre.2012.7118.1 : Posted 3 Apr

References (continued)

Preceedings : doi:10.1038/npre.2012.7118.1 : Posted 3 Apr

K. Li (2002)

Genome-wide coexpression dynamics: Theory and application.

PNAS USA 99(26), 16875 – 16880.

N. M. Luscombe *et al.* (2004)

Genomic analysis of regulatory network dynamics reveals large topological changes.

Nature 431, 308 – 312.

D. Pe'er and N. Hacohen (2011)

Principles and Strategies for Developing Network Models in Cancer.

Cell 144, 864 – 873.

R. G. Ptak *et al.* (2008)

Cataloguing the HIV type 1 human protein interaction network.

AIDS Res Hum Retroviruses 24(12), 1497 – 1502.

M. Ray and W. Zhang (2010)

Analysis of Alzheimers disease severity across brain regions by topological analysis of gene co-expression networks.

BMC Syst Biol 4, 136.

Precedings : doi:10.1038/npre.2012.7118.1 : Posted 3 Apr

THANK YOU