Session 11.3.2

Computational Genomics, Gene Expression, and Genetic Networks II

10:45 – 12:00, Sept 3, Exhibition Hall I

Prediction and Analysis of Human MicroRNA Regulatory Modules

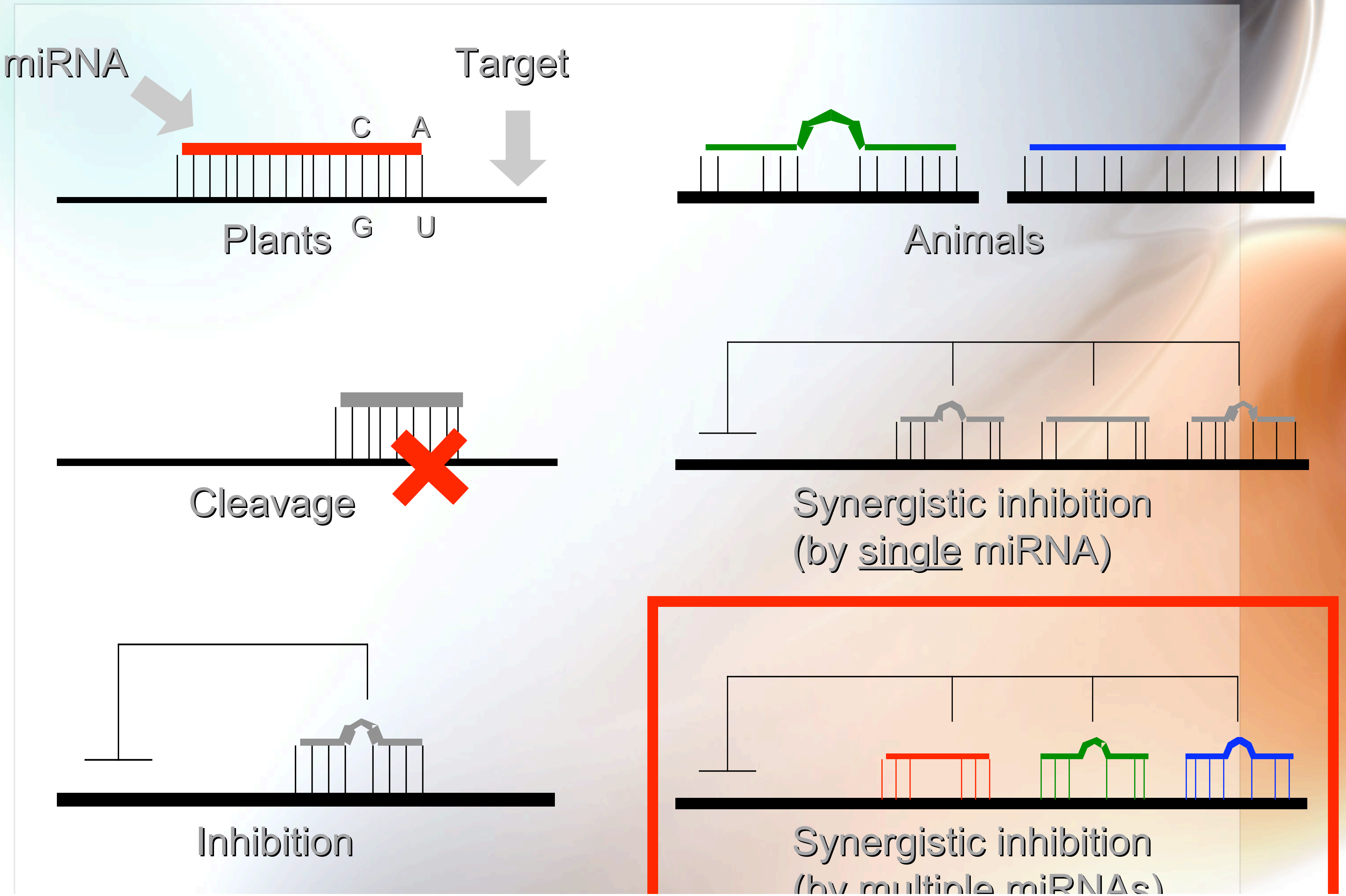
Sungroh Yoon, *Stanford University*

Giovanni De Micheli, *EPF Lausanne*

MicroRNA (miRNA)

- Gene regulation
  - By protein (conventional view)
  - By protein and RNA (recent view)
- MicroRNAs (miRNAs)
  - 21-22 nucleotides long
  - Post-transcriptional gene regulation
  - Bind target genes for cleavage/translational repression in sequence-specific manner
  - Nearly 1% of the genes in human genome

MiRNA-target Duplex (Lai, 2004)



Plants

Animals

Cleavage

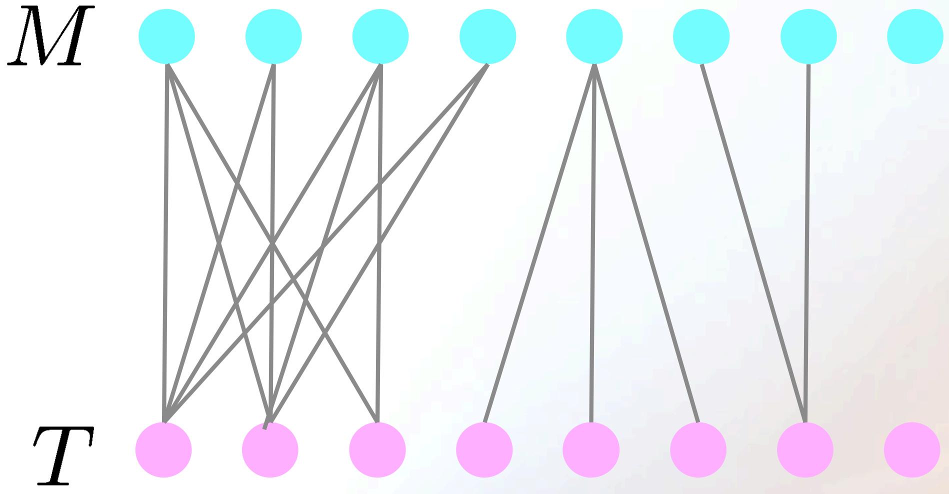
Synergistic inhibition
(by <u>single</u> miRNA)

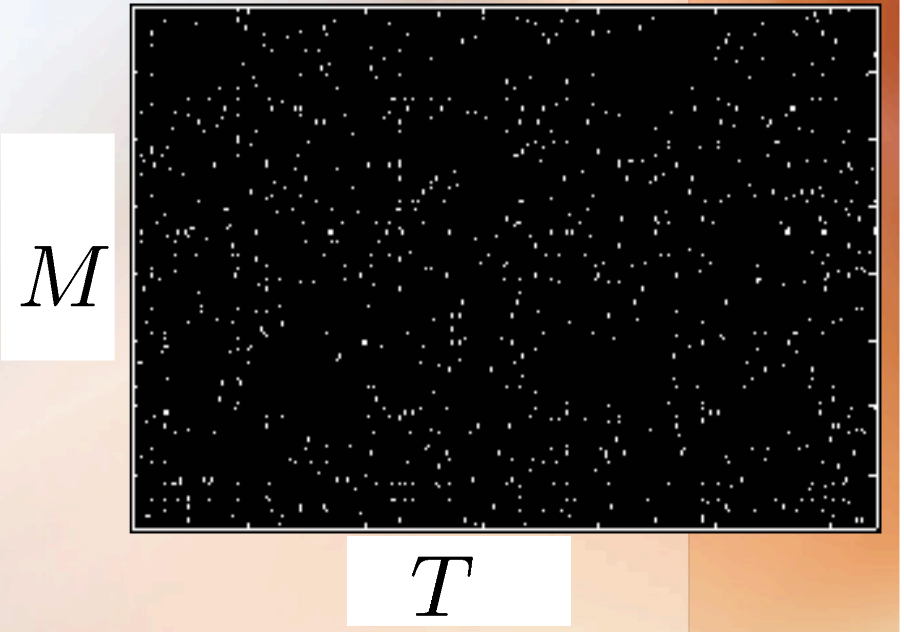
Inhibition

Synergistic inhibition
(by multiple miRNAs)

Modeling miRNA-target interactions

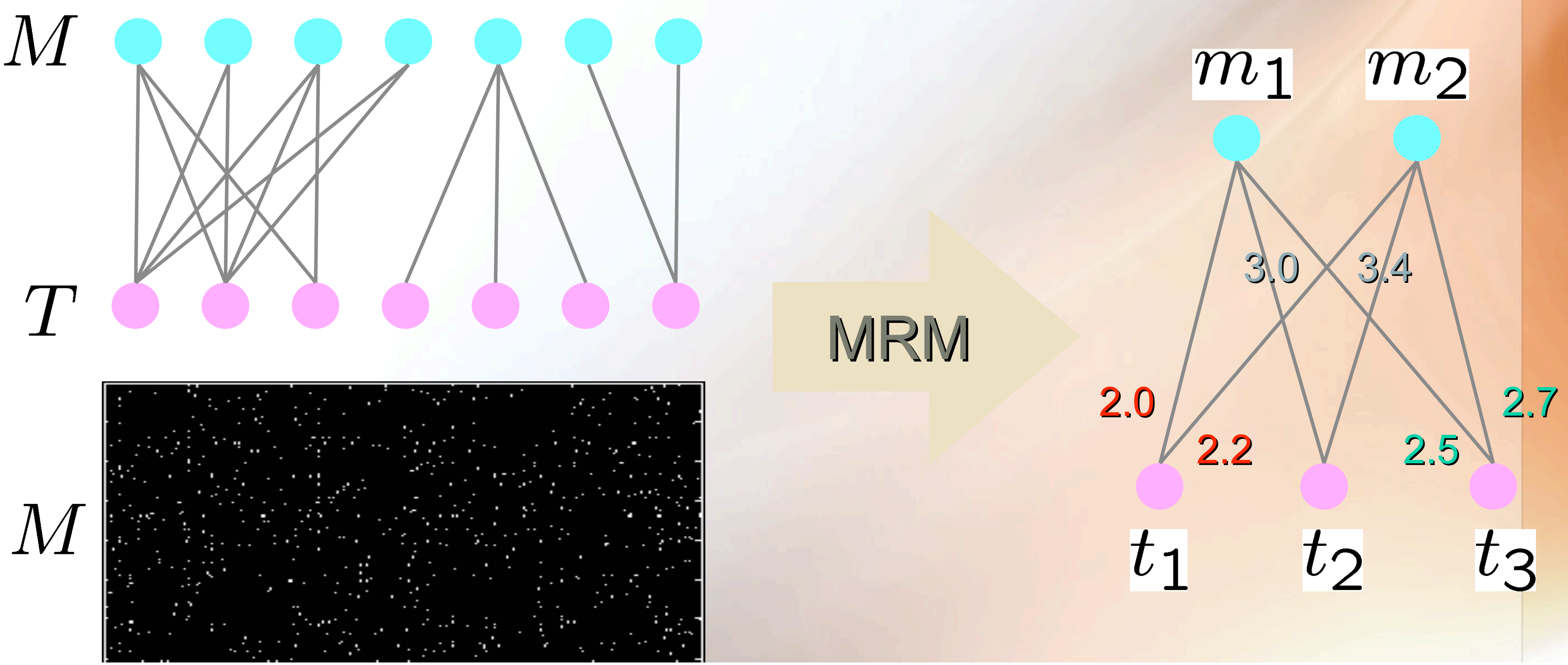
- **Weighted bipartite graph** $G = (M \cup T, E, W)$





MiRNA Regulatory Modules

- MiRNA regulatory modules (MRMs)
  - Maximal bicliques in G
  - Similar weights for edges incident on $t \in T$



Overview of our approach

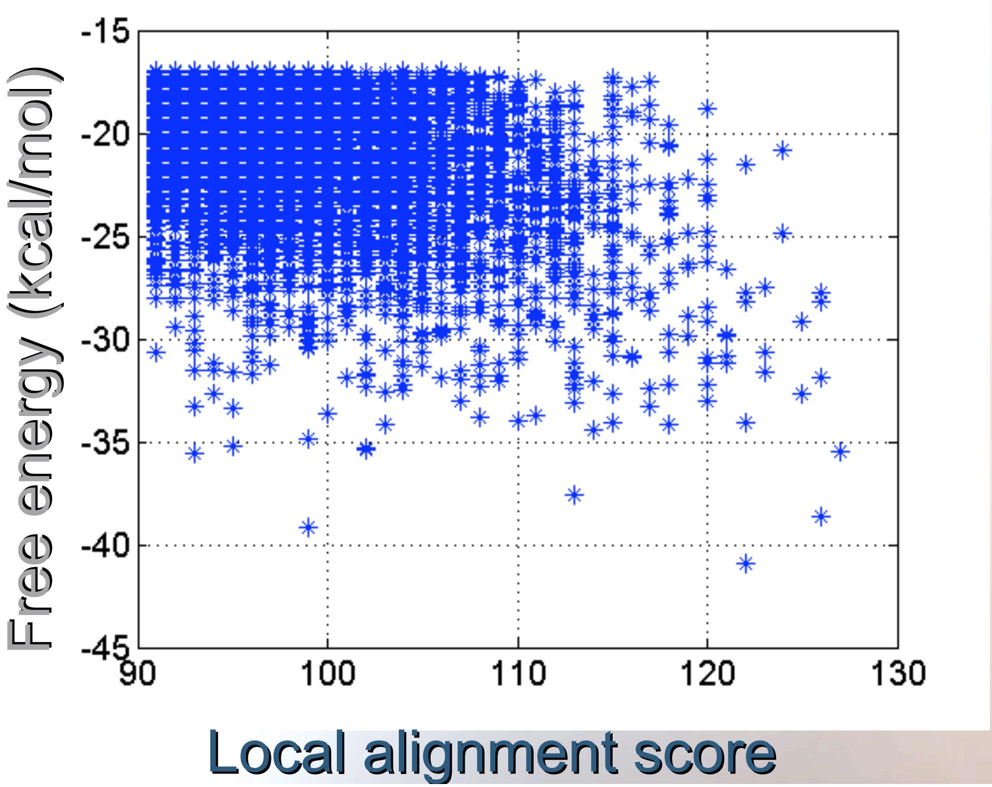
1. Target identification
2. Weighted bipartite graph representation
3. Finding “seed” bicliques
    - For each target, find a group of miRNAs that bind with similar binding strength
4. Merging “seeds”
    - Use a technique similar to frequent itemset mining or co-clustering
5. Post-processing
    - Identify statistically significant modules

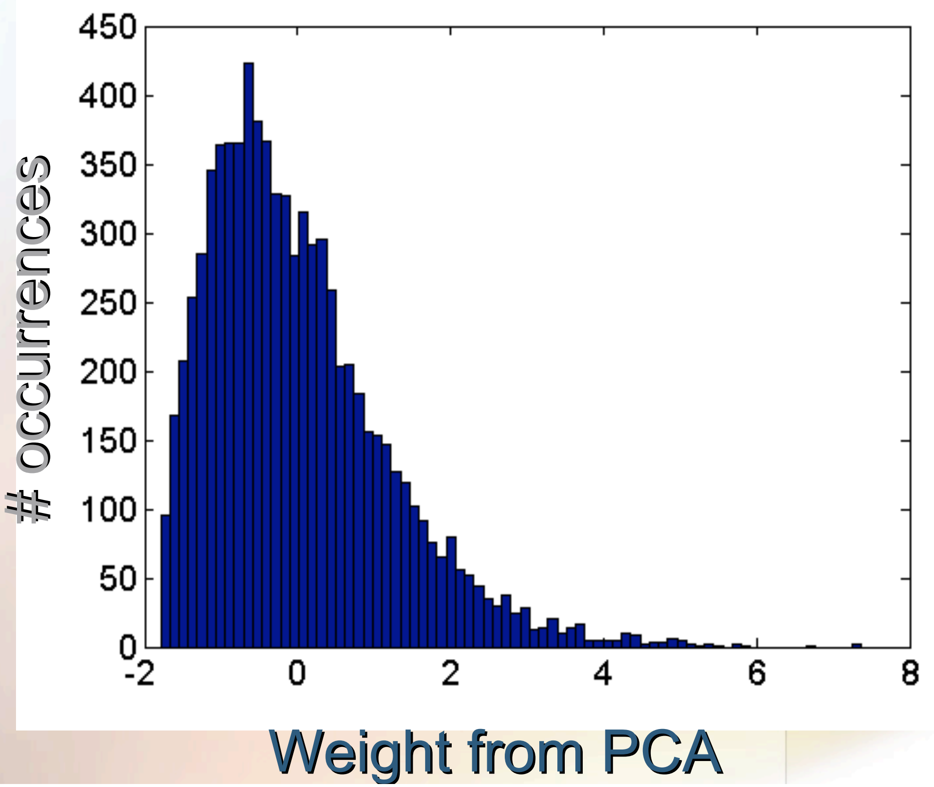
MiRNA Target Identification Methods

- Lewis *et al.*, 2003; John *et al.* 2004
  - Local alignment score
  - Free energy
  - Conserved sequences
- Lewis *et al.*, 2005
  - Identify targets with conserved complementarity to 2-7 nt of miRNA
  - Often flanked by adenosines

Binding Strength Estimation

- We adopted the first option
  - Measured free energy and local alignment score of each miRNA-target duplex
- Distribution of binding strength





Experimental Results

- Input
  - Human genes and miRNA sequences
- Procedure
  - Estimate binding strength
  - Build weighted bipartite graph

$$(|T|, |M|, |E|) = (2888, 156, 7886)$$

  - Find MRMs
- Output
  - 431 miRNA regulatory modules ($P < 0.01$)
  - Average size: 6.74 targets and 3.58 miRNAs

A Predicted Oncogenic Module

Breast; Renal; Prostate cancer (Struckman *et al.*, 2004; Kawakubo *et al.*, 2004; Ficazzola *et al.*, 2001)

Clustered within 0.5 kb on chromosome 13q14. This region is deleted in B cell chronic lymphocytic leukemia (B-CLL), mantle cell lymphoma, multiple myeloma, and prostate cancer cases. (Stilgenbauer *et al.*, 1998; Migliazza *et al.*, 2000; Calin *et al.*, 2002)

Breast cancer; Neuroblastoma (Li *et al.*, 2002; Saito-Ohara *et al.*, 2003)

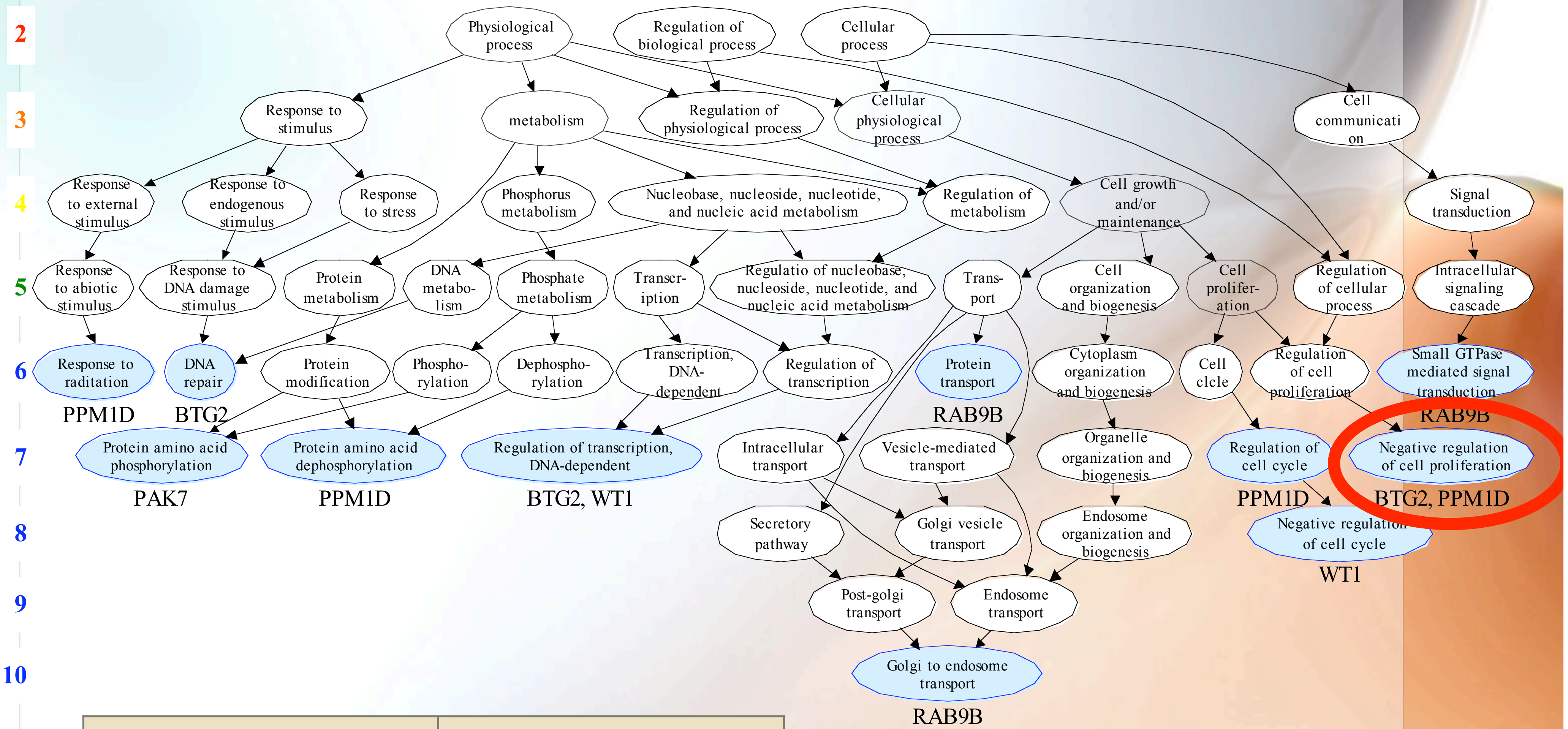
BTG2

Kidney (Loeb and Sukumar, 2002)

PPM1D

Target	Description	mir15a	mir16	mir195
PAK7	P21-activated kinase 7	*1.609*	*-0.789*	*0.676*
RAB9B	Ras-associated oncogenic protein 9b	*1.303*	*-0.746*	*-0.956*
BTG2	B-cell translocation gene 2	*-0.162*	*-0.816*	*-1.259*
PPM1D	Protein phosphatase 1D	*0.487*	*0.817*	*1.143*

Validation with Gene Ontology



Item	Value
GO term	Negative regulation of cell proliferation
Corrected p-value	0.0184

Summary

- MicroRNA
  - Regulates target genes in a sequence-specific manner
- Finding miRNA modules
  - Can contribute to gene regulatory network reconstruction
  - Modeled by finding weighted bicliques
- Experimental results
  - Successfully identify 431 human miRNA regulatory modules
  - Validated with GO and the literature