

Area 4: GENES and GENOMES
19:30-20:00, September 29 (Thursday), 2005
ECCB '05 - Madrid, Spain

**Prediction of Regulatory Modules
Comprising
MicroRNAs and Target Genes**

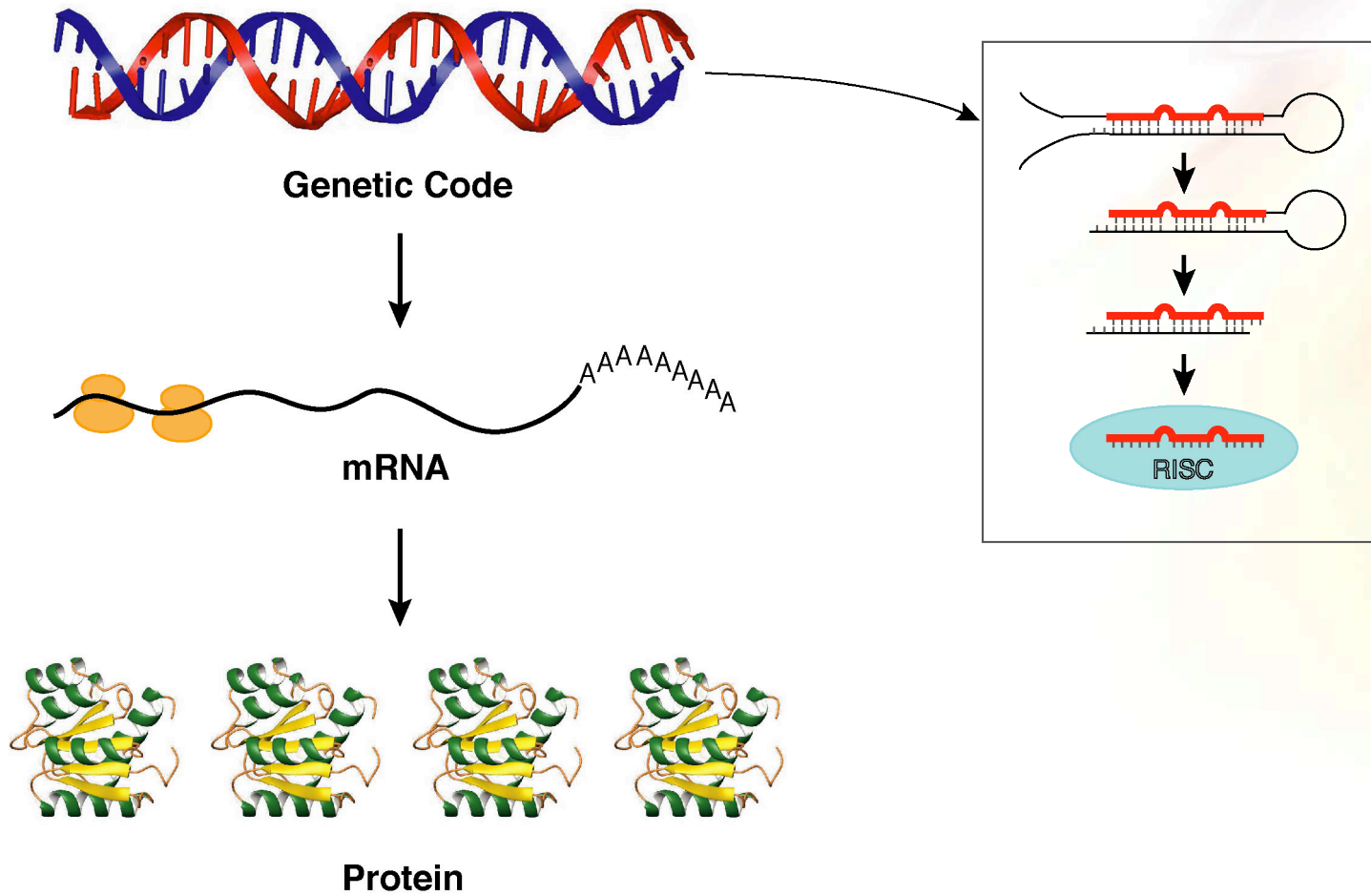


Sungroh Yoon, Stanford University, USA
Giovanni De Micheli, EPF Lausanne, Switzerland

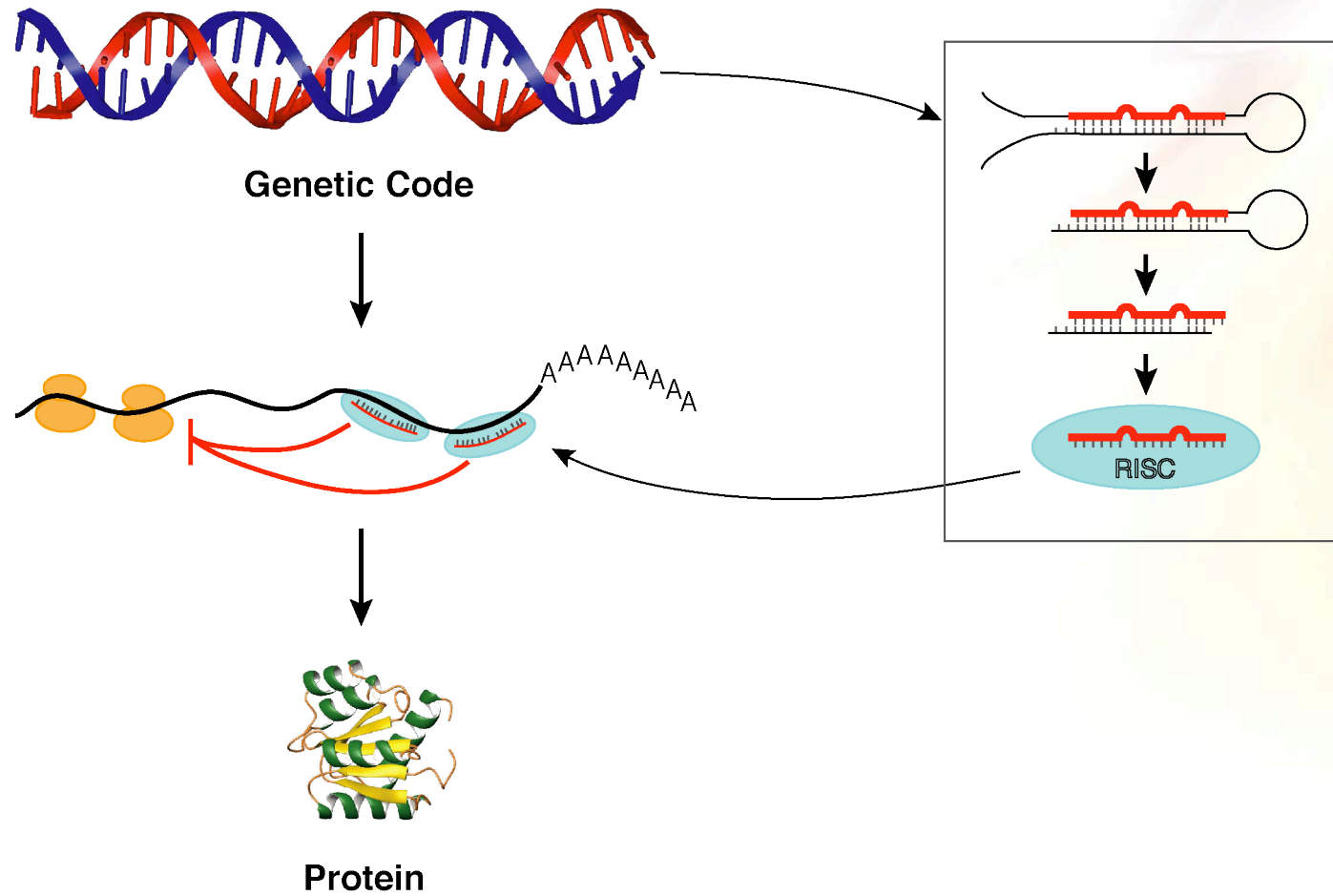
- **MicroRNA (miRNA)**
- **Representation of miRNA-target interactions**
- **Finding miRNA regulatory modules**
- **Experiments and validation**

- **Gene regulation**
 - By protein (conventional view)
 - By protein and RNA (recent view)
- **MicroRNAs (miRNAs)**
 - ~22 nucleotides long
 - Mediate post-transcriptional gene regulation
 - Bind target mRNAs for cleavage/translational repression
 - Nearly 1% of the genes in human genome
 - Add a new dimension to our understanding of complex gene regulatory networks

miRNA-mediated Posttranscriptional Regulation

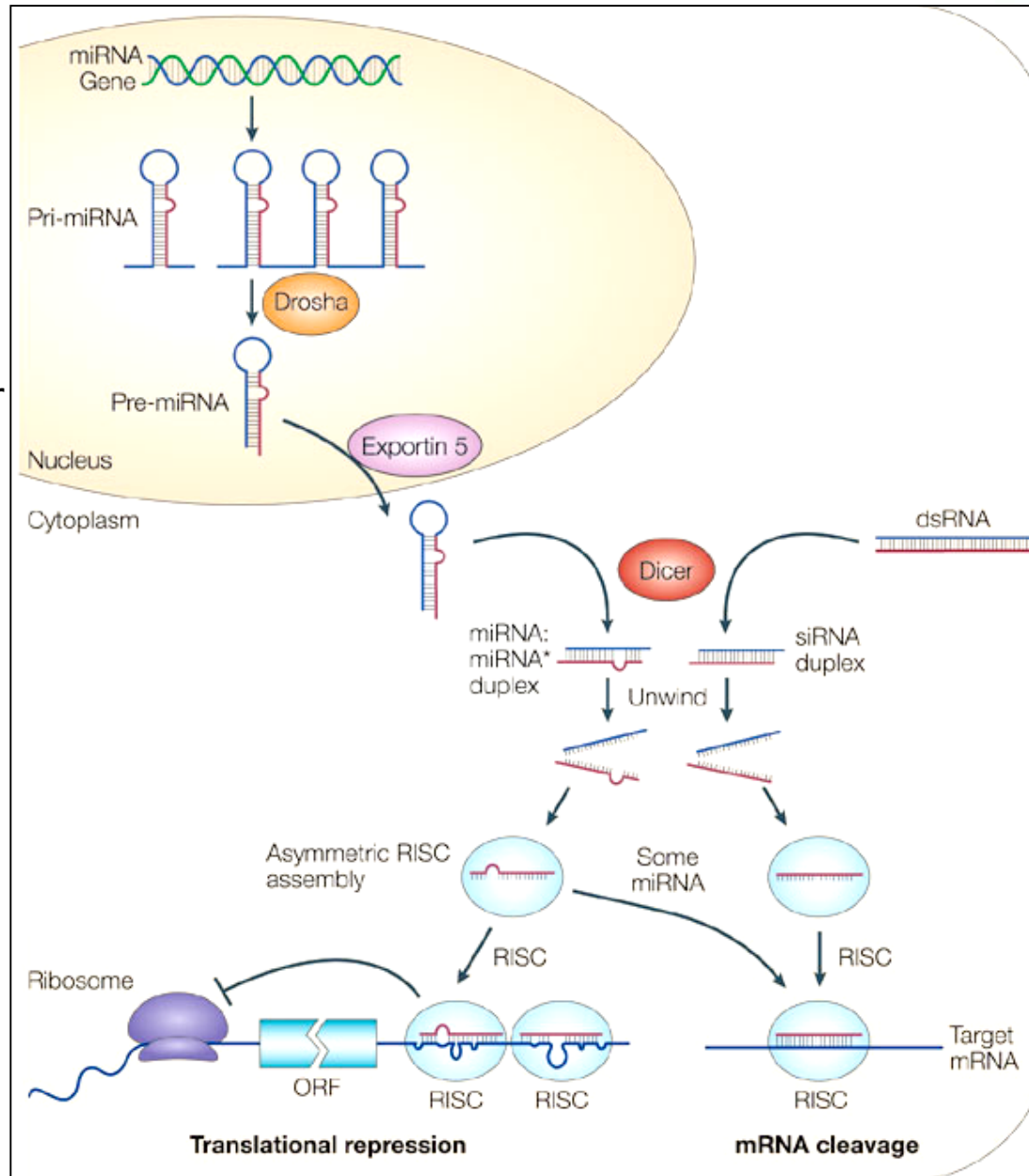


miRNA-mediated Posttranscriptional Regulation



* Primary
miRNA

* Precursor
miRNA



* small
interfering
RNA

* RNA
Induced
Silencing
Complex

miRNA Functions in Flies

- **Bantam:** Brennecke et al., *Cell*, 2003, 113:25
 - Prevents apoptosis by targeting Hid
 - Regulate fly growth during development
- **miR-14:** Xu et al., *Current Biology*, 2003, 1:790
 - Suppresses cell death
 - Regulates fat metabolism

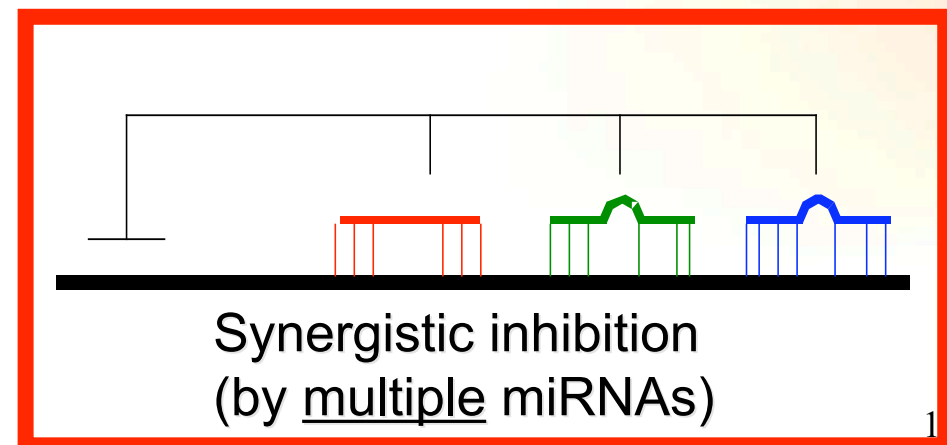
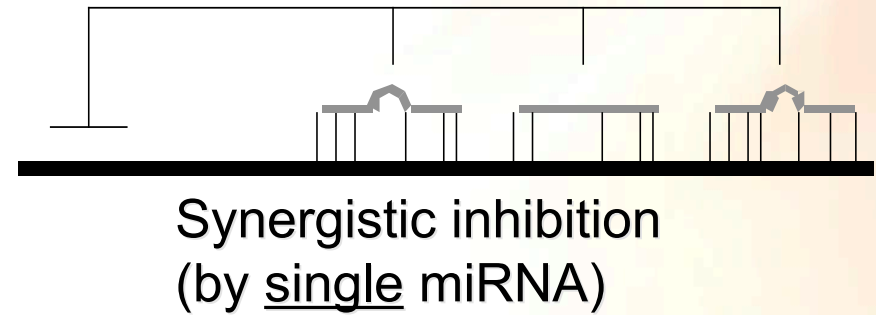
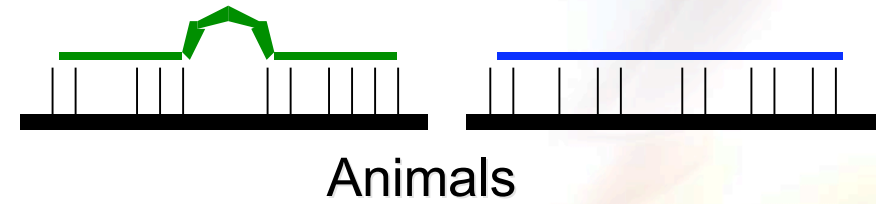
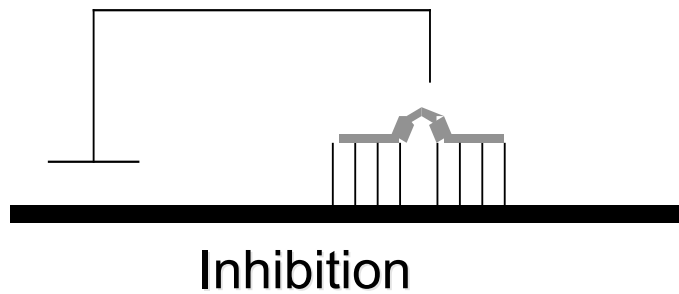
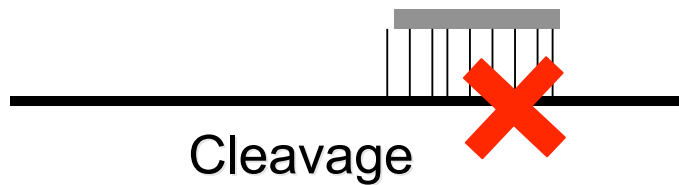
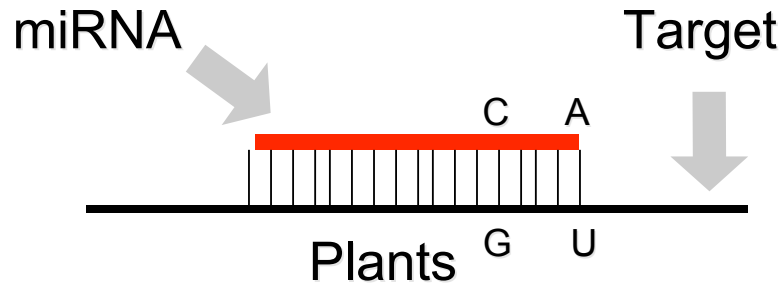
miRNA Functions in Vertebrate

- **miR-181 (Mouse):** Chen et al., *Science*, 2004, 303:83
 - Isolated from mouse bone marrow
 - Increased B cell development in vitro and in vivo
- **miR-375 (Mouse):** Poy et al., *Nature*, 2004, 432:226
 - Pancreatic islet-specific microRNA
 - Suppressed glucose-induced insulin secretion
- **miR-15, miR-16 (Human):** Calin et al., *PNAS*, 2002, 99:15524
 - Cloned from Hela cells
 - Deleted or down-regulated in B-CLL, and prostate cancer
 - Tumor suppressor gene?

- **MicroRNA (miRNA)**
- **Representation of miRNA-target interactions**
- **Finding miRNA regulatory modules**
- **Experiments and validation**

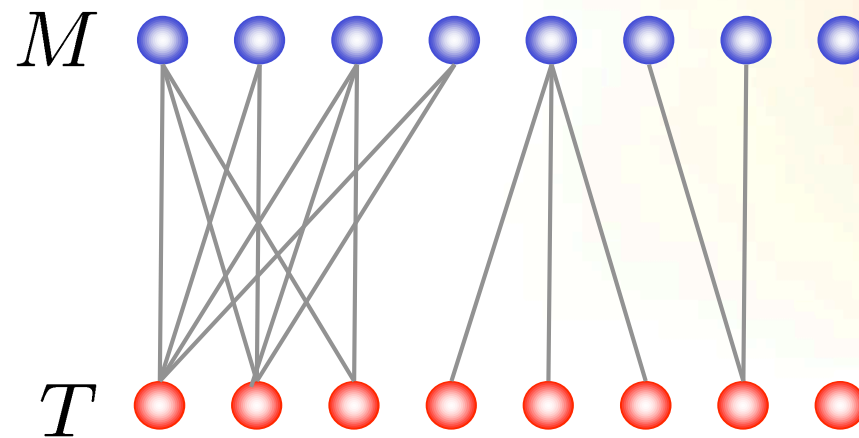
miRNA-target Duplex Configurations

(Lai, 2004)



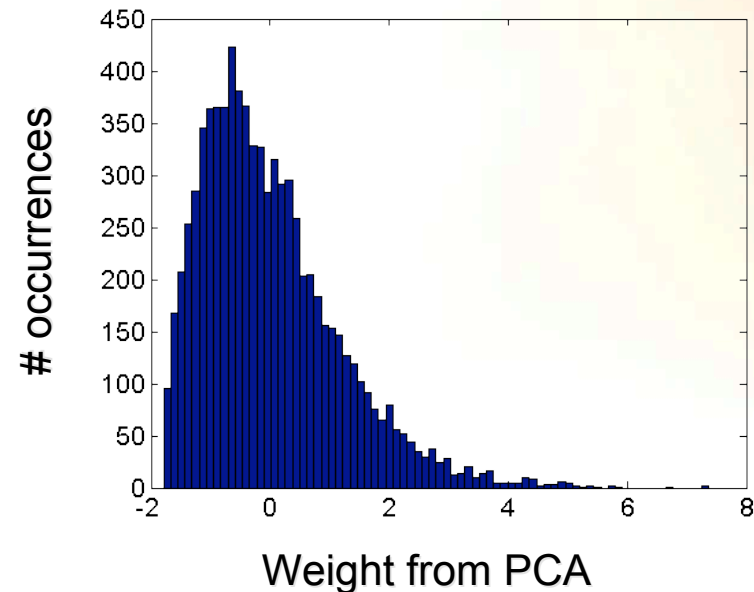
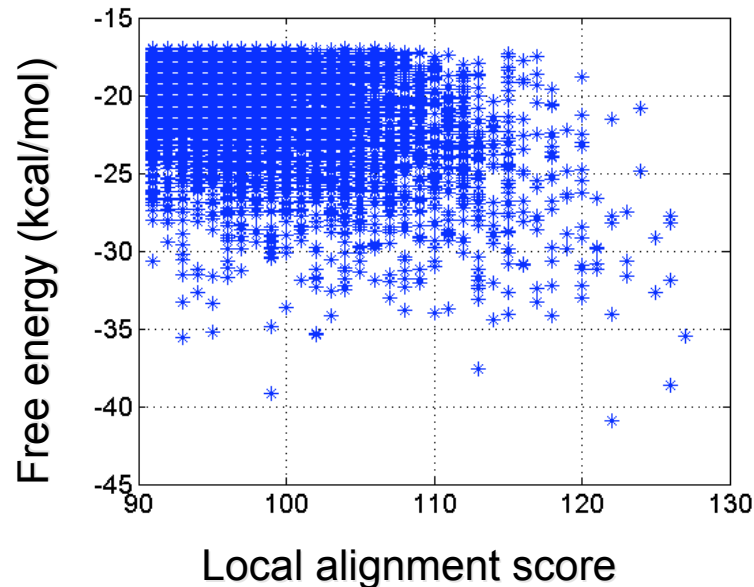
Modeling miRNA-target Interactions

- **Weighted bipartite graph** $G = (M \cup T, E, W)$
 - **M-vertex: miRNA**
 - **T-vertex: target mRNA**
 - **Edge weight: binding strength**
 - **“Interaction graph”**



miRNA Target Identification

- Guided by sequence complementarity
- Lewis *et al.*, 2003; John *et al.* 2004
 - Local alignment score
 - Free energy
 - Conserved sequences



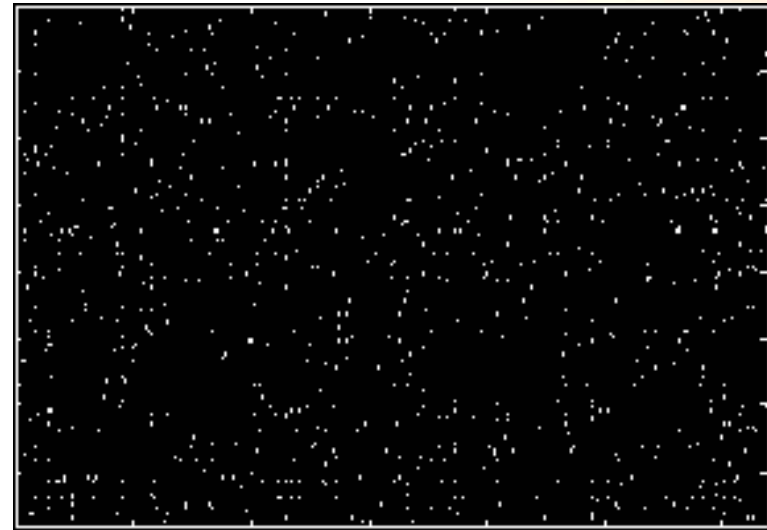
miRNA Target Identification

- **Lewis *et al.*, 2005**
 - Focus on the nucleotides 2-7 of miRNA
 - Targets: conserved complementarity to this region
 - ✓ Overrepresentation of conserved adenosines (flanking target sites that are complementary to the nt 2-7 of miRNA) indicates that primary sequence determinants can supplement base pairing to specify miRNA target condition

Alternative Representation

- Matrix of real numbers
 - Row: miRNA
 - Column: target mRNA
 - Values: binding strength
 - “Interaction matrix”

M

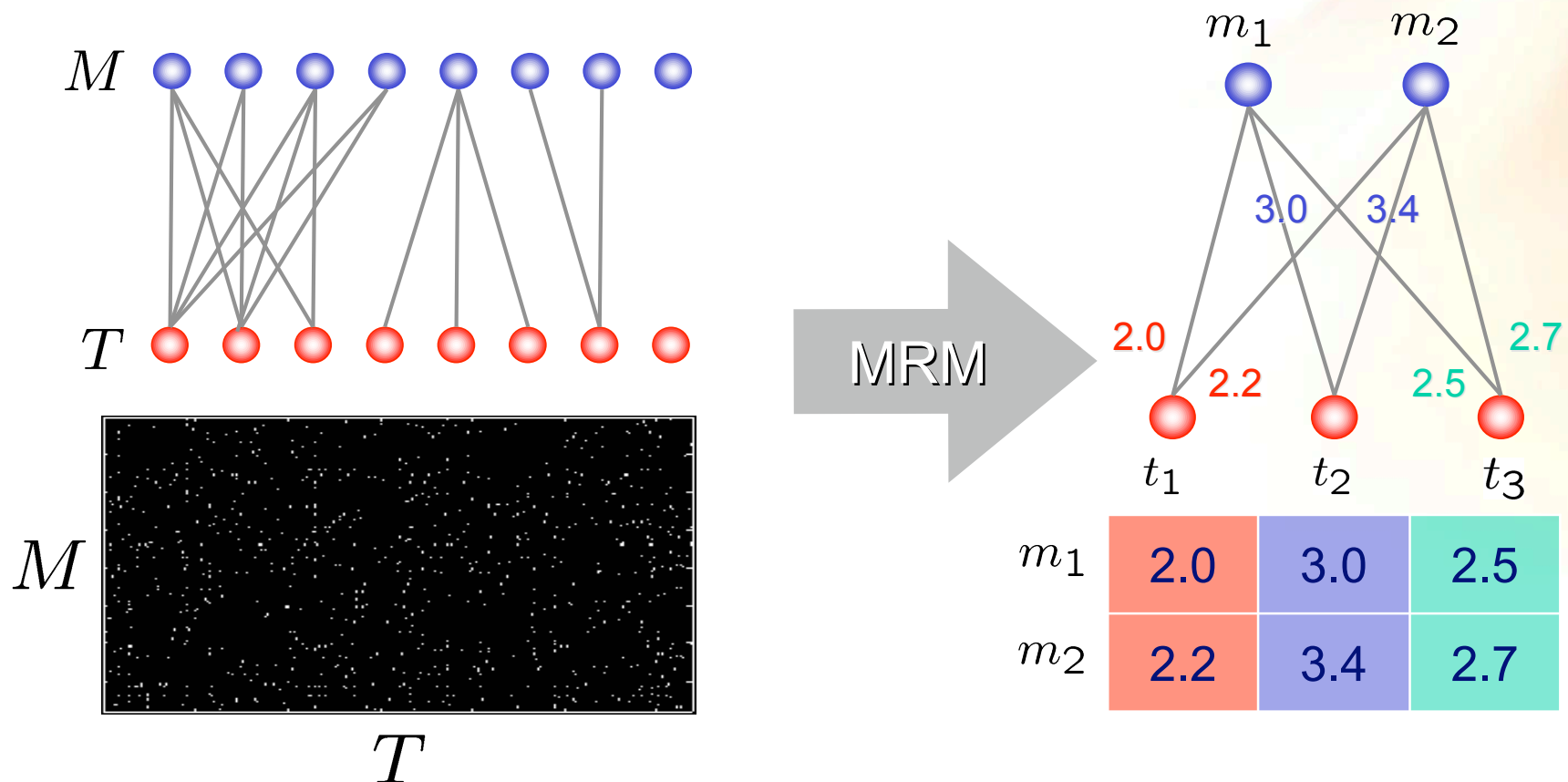


T

miRNA Regulatory Modules

miRNA regulatory module (MRM)

- (Maximal) biclique in G
- Similar weights for edges incident on $t \in T$



- **MicroRNA (miRNA)**
- **Representation of miRNA-target interactions**
- **Finding miRNA regulatory modules**
- **Experiments and validation**

Algorithm Overview

- **Input**
 - Weighted bipartite graph representing miRNA-target interactions
- **Output**
 - MRM: maximal biclique with similar edge weight for each target vertex
- **Algorithm**
 - Step 1: finding “seeds”
 - Step 2: merging seeds

- **Seed: $S(t)$**
 - Maximal set of miRNAs that bind target t with similar strength
- **Algorithm ($\delta = 0.5$)**

	t1	Sort
m1	0.4	m3 0.0
m2	2.6	m6 0.1
m3	0.0	m5 0.3
m4	0.9	m1 0.4
m5	0.3	m4 0.9
m6	0.1	m7 1.5
m7	1.5	m2 2.6

$$S(t_1) = \{m_3, m_6, m_5, m_1\}$$

$$S(t_1) = \{m_1, m_4\}$$

Step 2-1

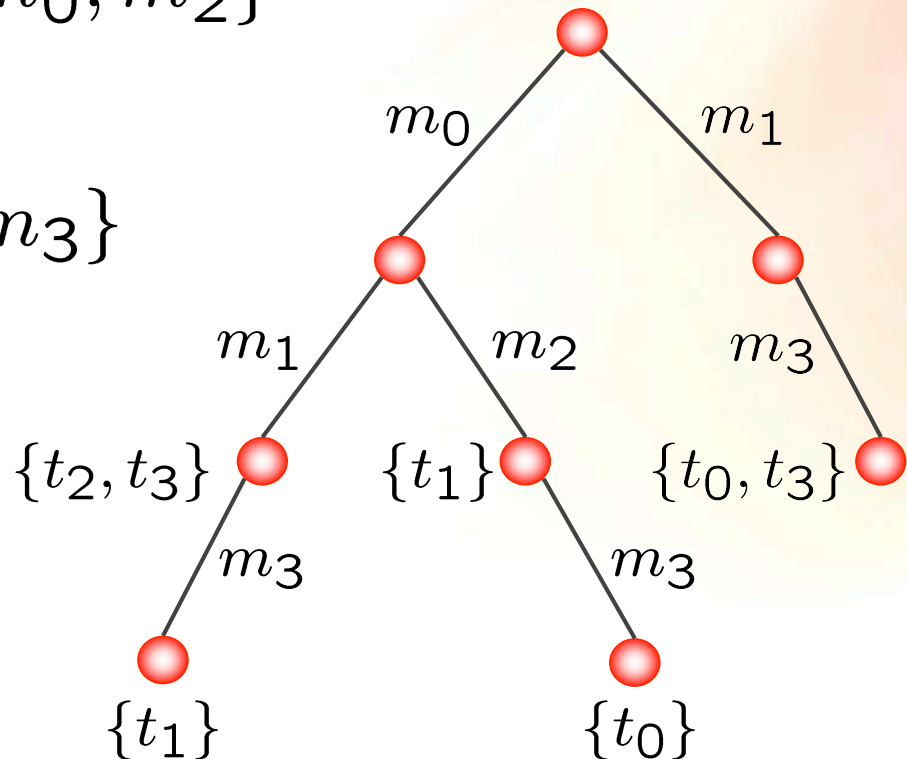
- Collect seeds in a trie

$$S(t_0) : \{m_0, m_2, m_3\}, \{m_1, m_3\}$$

$$S(t_1) : \{m_0, m_1, m_3\}, \{m_0, m_2\}$$

$$S(t_2) : \{m_0, m_1\}$$

$$S(t_3) : \{m_0, m_1\}, \{m_1, m_3\}$$



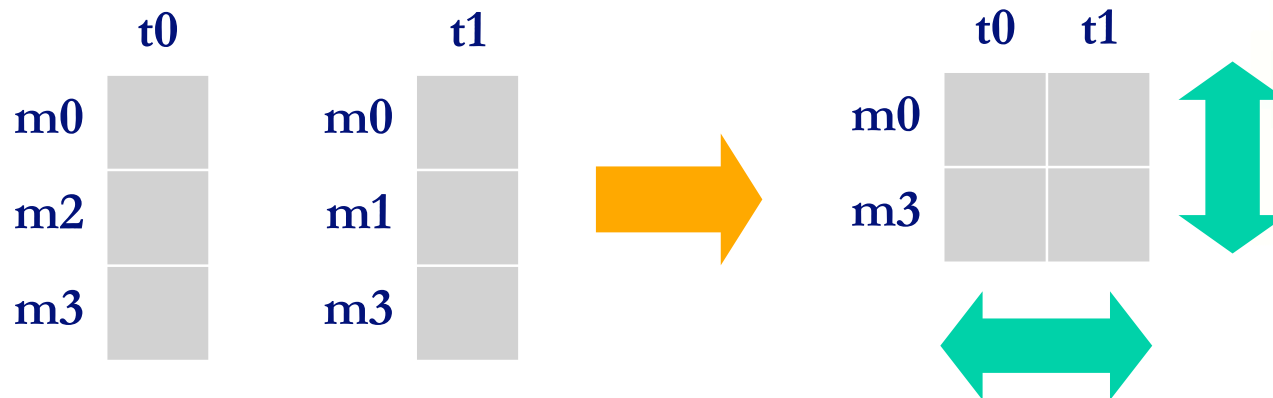
Merging Seeds

$$S(t_0) : \{m_0, m_2, m_3\}, \{m_1, m_3\}$$

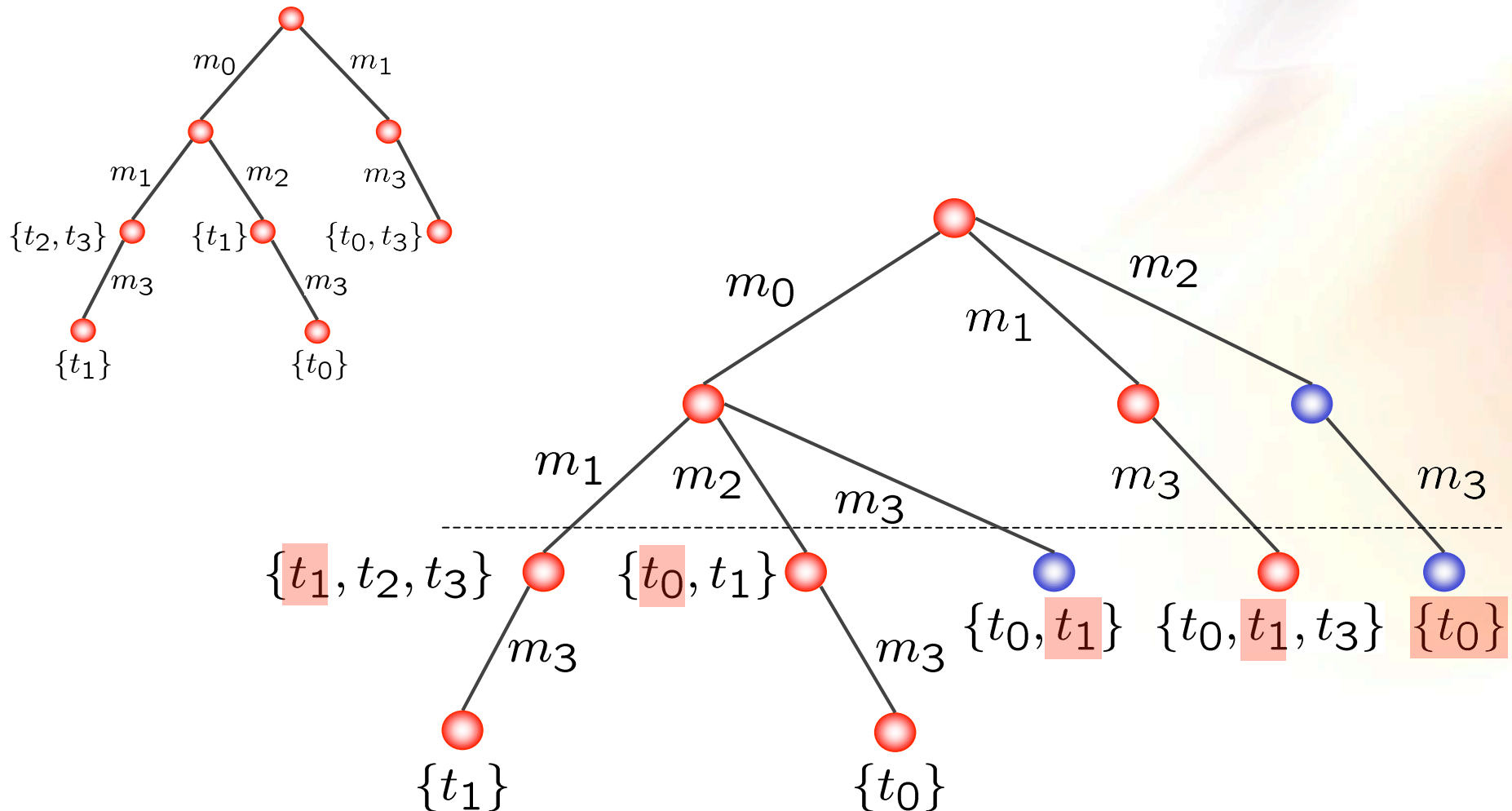
$$S(t_1) : \{m_0, m_1, m_3\}, \{m_0, m_2\}$$

$$S(t_2) : \{m_0, m_1\}$$

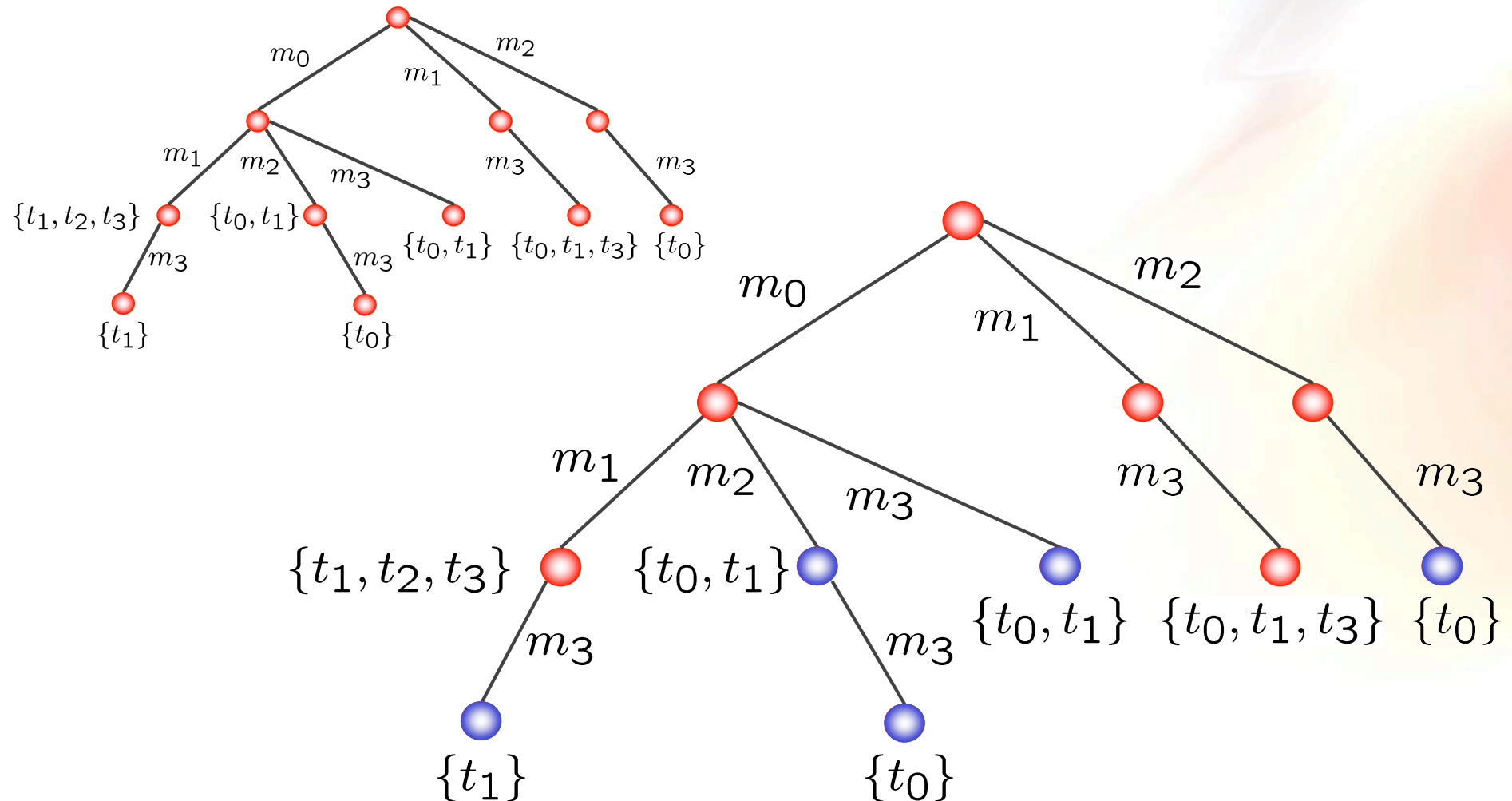
$$S(t_3) : \{m_0, m_1\}, \{m_1, m_3\}$$



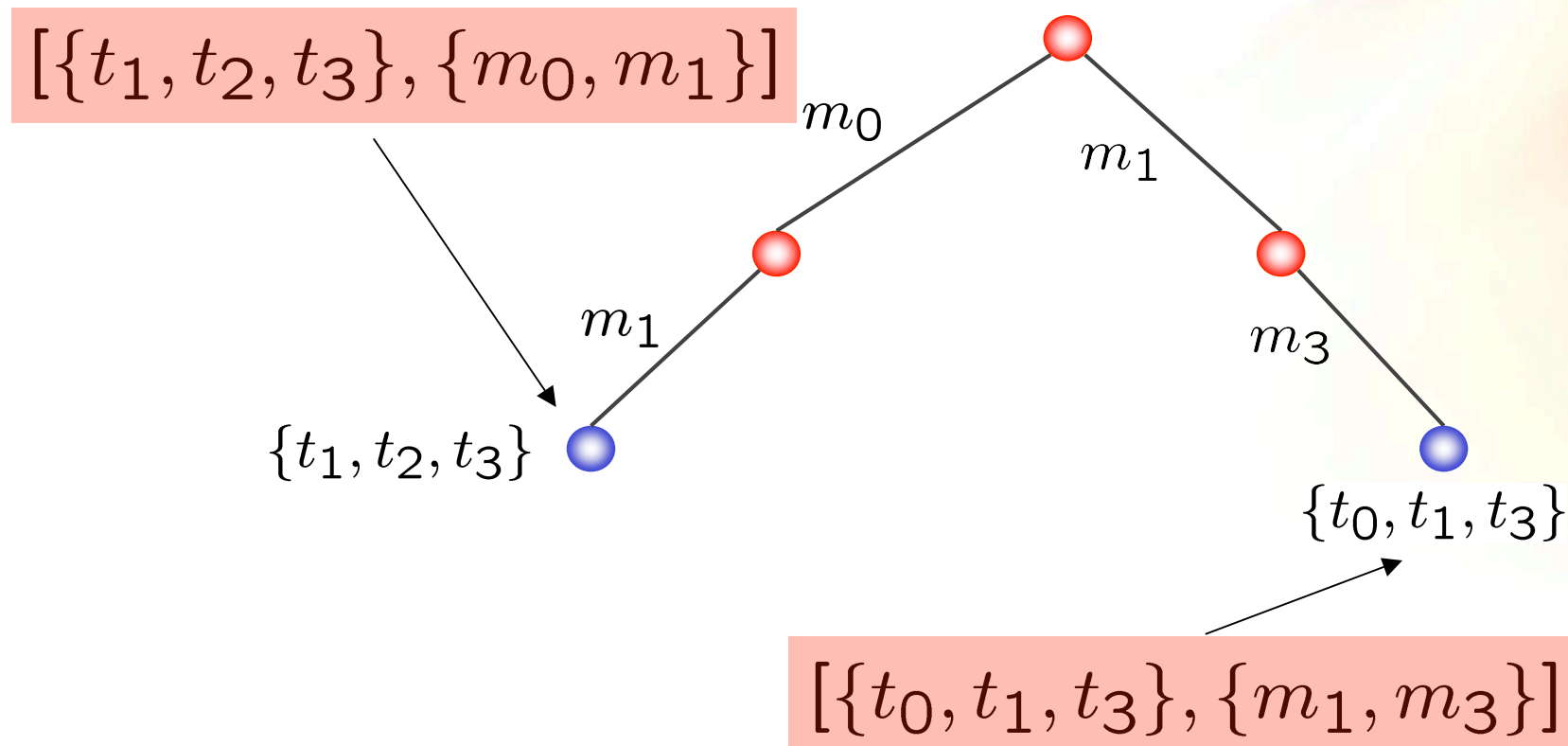
Expand the trie ($\min_M = 2$)



Prune the trie ($\min_T = 3$)



- 2 MRMs found



Assessment of Statistical Significance

$X_{m \times t}$: # of $(m \times t)$ MRMs

$P_{m \times t}$: Prob{a random $(m \times t)$ biclique is MRM}

$$P(X_{m \times t} = k) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad k = 0, 1, 2, \dots$$

$$\lambda = \binom{|M|}{m} \binom{|T|}{t} P_{m \times t}$$

$$P_{m \times t} \simeq \zeta^t [1 - \zeta]^{|T|-t} [1 - (1 + m^{-1})^t \delta^t]^{|M|-m}$$

(Califano *et al.*, 2000)

$$\zeta = m\delta^{m-1} - (m-1)\delta^m$$

$$p\text{-Value} = P(X_{m \times t} \geq 1) = 1 - P(X_{m \times t} = 0) = 1 - e^{-\lambda}$$

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 ($P < 0.01$)

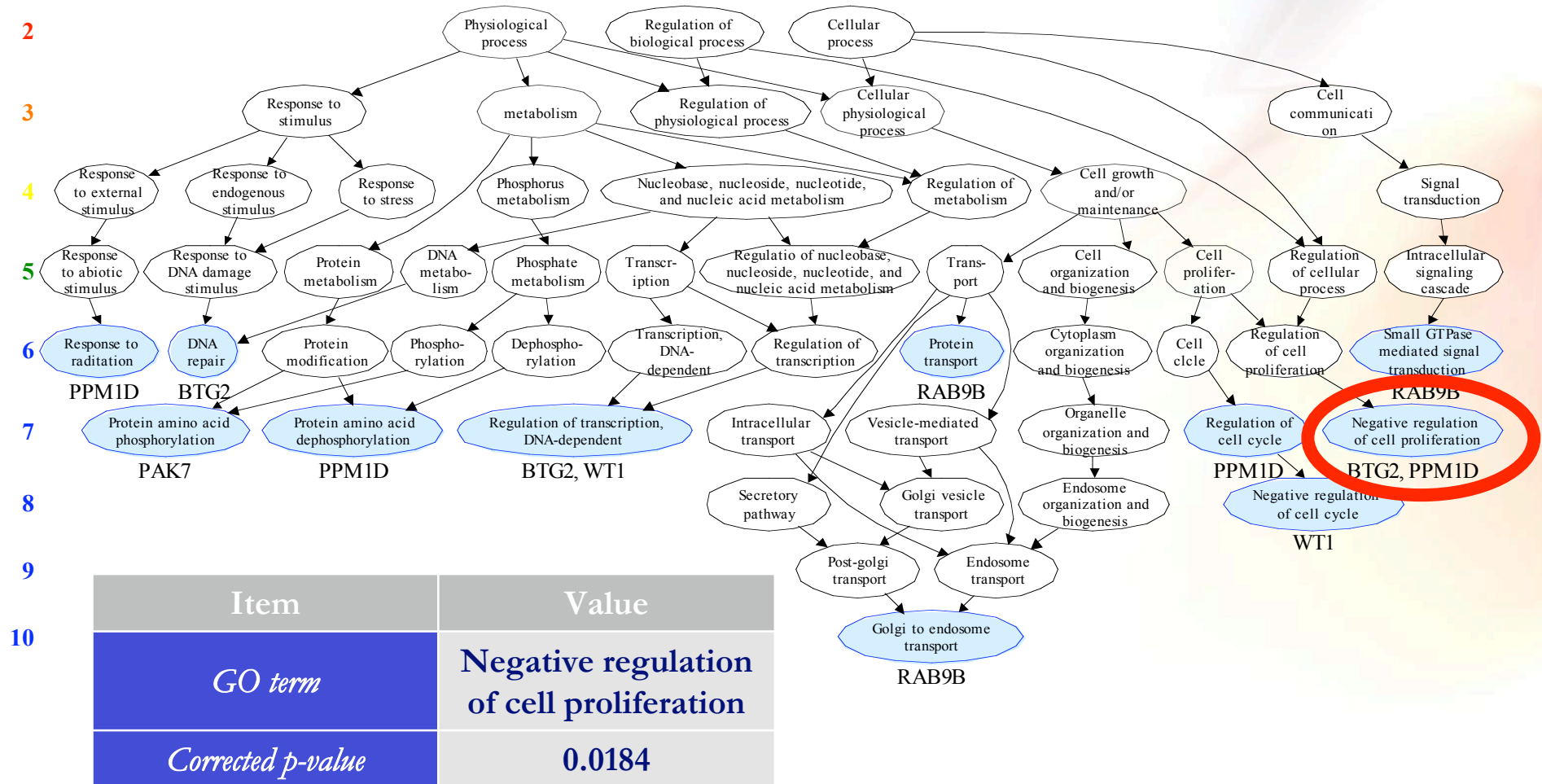
An Example Module

Breast; Renal;
Prostate cancer
(Struckman *et al.*,
2004; Kawakubo
et al., 2004;
Ficazzola *et al.*,
2001)
Sukuma (2002)

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)

		mir15a	mir16	mir195
	tion			
	ivated kinase 7	1.609	-0.789	0.676
	ociated oncogenic protein 9b	1.303	-0.746	-0.956
<i>BTG2</i>	B-cell translocation gene 2	-0.162	-0.816	-1.259
<i>PPM1D</i>	Protein phosphatase 1D	-0.487	-0.817	-1.143
<i>WT1</i>	Wilms' tumor	0.275	1.019	-0.514

An Example Module



- **MicroRNAs**
 - Mediate post-transcriptional gene regulation
- **MicroRNA regulatory modules**
 - Group of miRNAs and target genes
 - Can contribute to reconstruction of gene regulatory networks
- **Computational method to find MRMs**
 - 2-step data mining algorithm
 - Identify 431 human miRNA regulatory modules

■ Thank you!

