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)

MiRNA

Target

Plants

Animals

Cleavage

Inhibition

Synergistic inhibition (by single miRNA)

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 MRMNs

• 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)

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

**Kidney** (Loeb and Sukumar, 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)

<table>
<thead>
<tr>
<th>Target</th>
<th>Description</th>
<th>mir15a</th>
<th>mir16</th>
<th>mir195</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAK7</td>
<td>P21-activated kinase 7</td>
<td>1.609</td>
<td>-0.789</td>
<td>0.676</td>
</tr>
<tr>
<td>RAB9B</td>
<td>Ras-associated oncogenic protein 9b</td>
<td>1.303</td>
<td>-0.746</td>
<td>-0.956</td>
</tr>
<tr>
<td>BTG2</td>
<td>B-cell translocation gene 2</td>
<td>-0.162</td>
<td>-0.816</td>
<td>-1.259</td>
</tr>
<tr>
<td>PPM1D</td>
<td>Protein phosphatase 1D</td>
<td>0.487</td>
<td>0.817</td>
<td>1.143</td>
</tr>
</tbody>
</table>
Validation with Gene Ontology

<table>
<thead>
<tr>
<th>Item</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO term</td>
<td>Negative regulation of cell proliferation</td>
</tr>
<tr>
<td>Corrected p-value</td>
<td>0.0184</td>
</tr>
</tbody>
</table>
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