CS 5984: Application of Biclustering To Data Integration

T. M. Murali

February 13, 2007
Integrating Heterogeneous Data

- Gene expression data sets
- Protein-DNA binding data
- Protein interactions
- Knock-out phenotypes
Integrating Heterogeneous Data

- Gene expression data sets
- Protein-DNA binding data
- Protein interactions
- Knock-out phenotypes
- How do we incorporate all these data types into a bipartite graph?
Analysis After Computing Biclusters/Modules

- Visualisation of modules.
- Functional enrichment of modules.
- Assessing utility of data integration.
- Hierarchical organisation of modules.
- Functional annotation.
Module Visualisation

[Diagram showing gene expression and properties with annotations and interactions]

- Annotations
- Publication code (click for PubMed)
- Gene names (Click for SGD)
- KO
- Interactions
  - Protein name (Click for SGD)
  - Edge Colored by Interaction type
  - Label Colored when protein is from enriched category

- Expression: Up regulation
- Location: 0.05 → 0.001
- Sensitivity: 10 → 50

- Complex
- Two-hybrid
- MIPS
- Double indication

T. M. Murali: CS 5984: Application of Biclustering To Data Integration
Amino Acid Transport Module

GO1 - amino acid transport (GO:0006865)
Vesicle Transport Module
Ubiquitin-dependent Protein Degradation Module
Functional Enrichment of Modules
Hierarchical Organisation of Modules
Functional Annotation