

Application of Biclustering To Data Integration

T. M. Murali

August 19, 2008

Integrating Heterogeneous Data

- ▶ Gene expression data sets
- ▶ Protein-DNA binding data
- ▶ Protein interactions
- ▶ Knock-out phenotypes

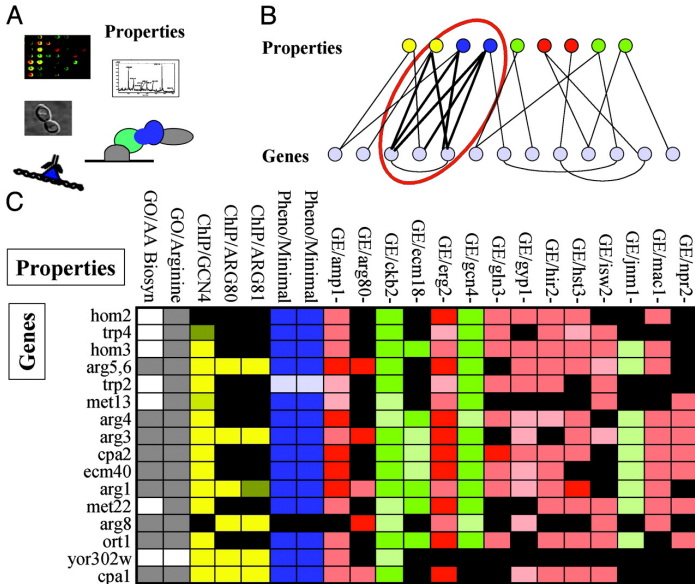
Integrating Heterogeneous Data

- ▶ 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?

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?
- ▶ How do we use the data types to find gene modules?

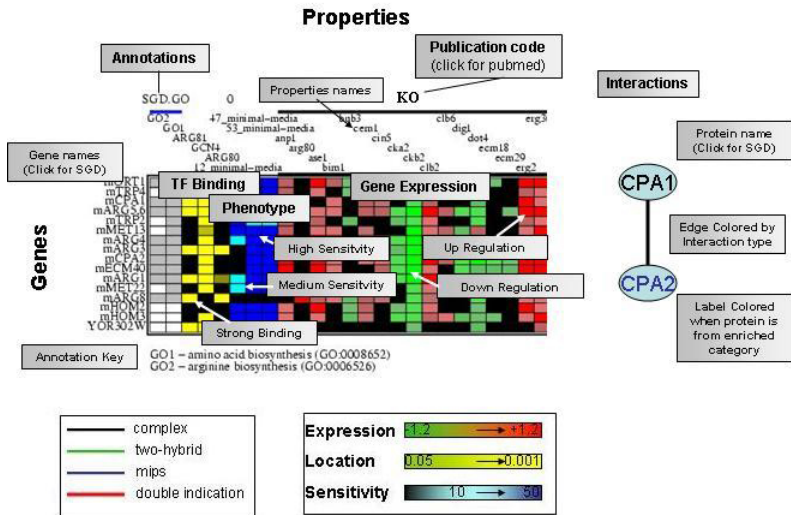
Analysis Pipeline



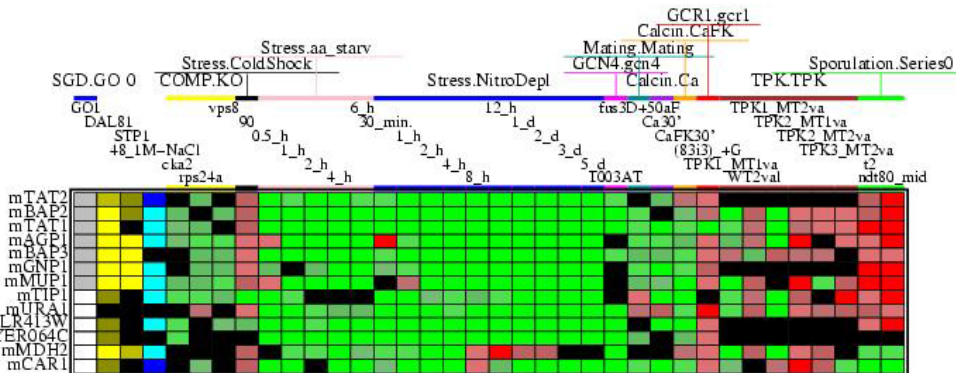
Analysis After Computing Biclusters/Modules

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

Module Visualisation

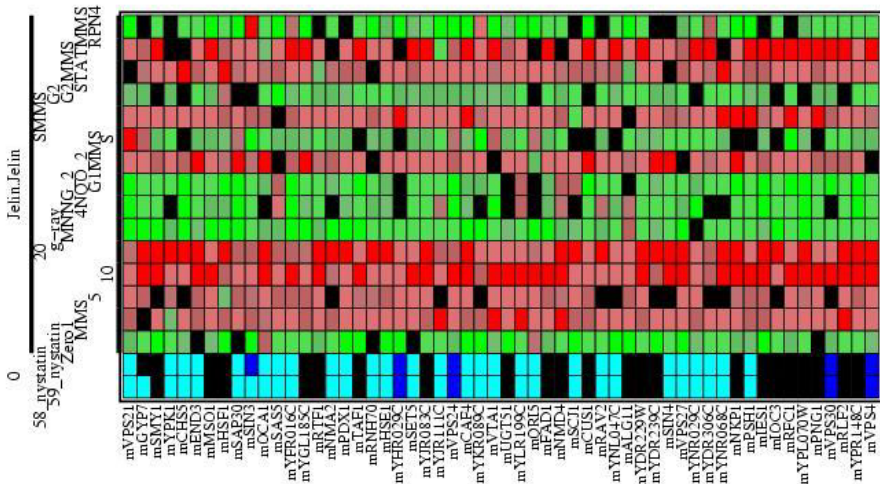


Amino Acid Transport Module

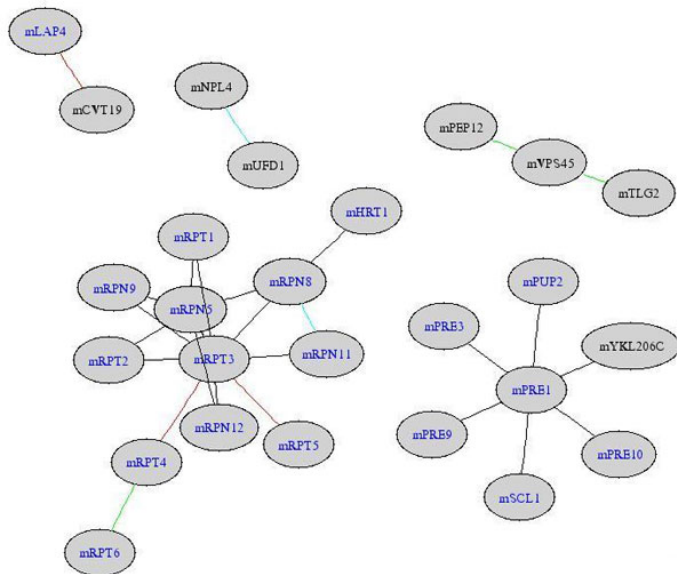


GO1 – amino acid transport (GO:0006865)

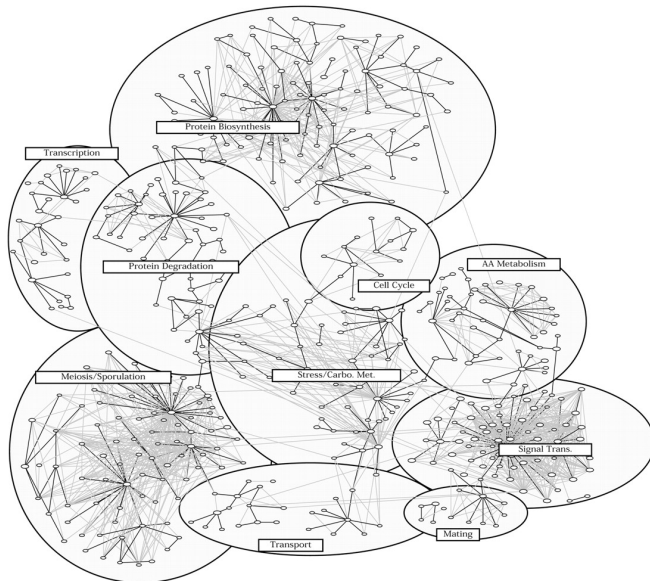
Vesicle Transport Module



Ubiquitin-dependent Protein Degradation Module



Hierarchical Organisation of Modules



Functional Annotation

