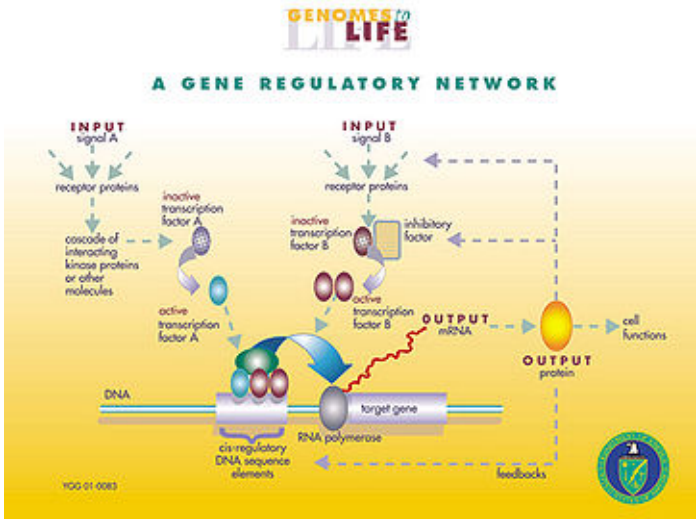
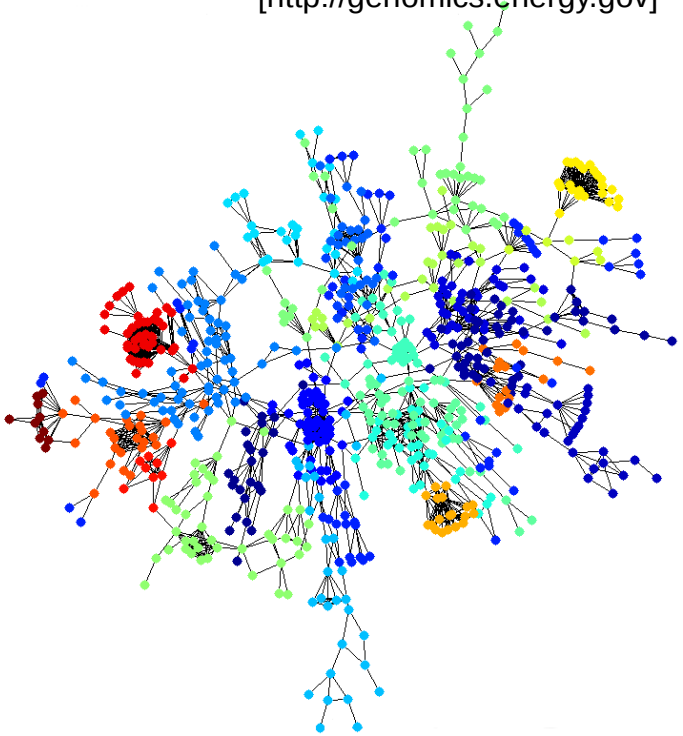
Combining topological priors with gene expression and interaction data for the inference of gene regulatory networks

Sumeet Agarwal
IIT Delhi

Reverse engineering the circuitry of biology



[http://genomics.energy.gov]

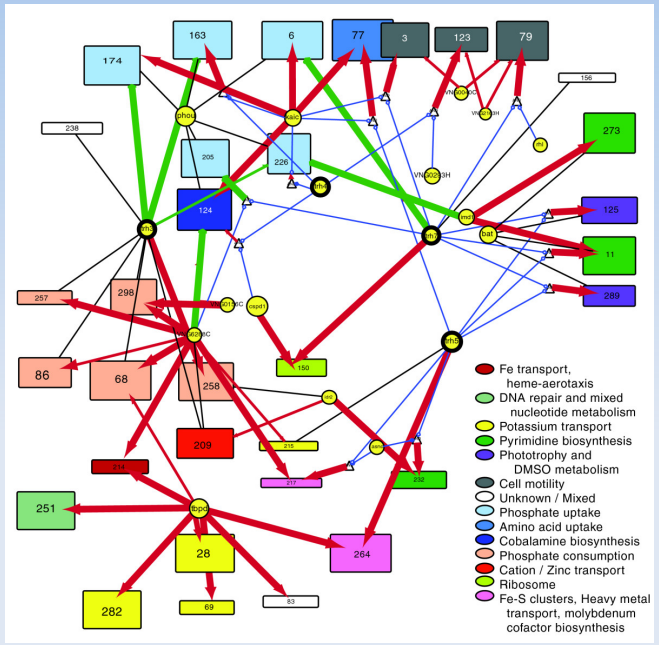


- Biological cells are essentially bags of interacting genes/proteins, which combine to carry out the various processes of life
- Given experimental data about how the concentration levels of proteins respond to various kinds of stimuli, can we try to recover the relationships of regulation and control between different genes/proteins?
- This can be thought of as learning the structure of a dynamical system, given some input/output characteristics
- We are looking at a range of approaches for mathematically modelling and learning these regulatory networks, such as Petri Nets, ODEs, and Markov Nets

Challenges To Address

- Scientific
  - To understand cellular circuitry, we need to look at dynamics and control in the interactions between genes/proteins: not merely static networks
  - Regulation and information flows; how systems-level behaivour emerges from individual components
- Technical
  - Much work on gene regulatory networks; but hard to choose from a plethora of models, as data leaves the problem underspecified
  - Integrating gene regulatory networks with protein interaction networks
- Medical/Sociological
  - Feeding back into wet lab and medical applications. I plan to collaborate with experimentalists at JNU/NCBS, focusing on under-studied organisms such as the amoebiasis parasite *E. histolytica* (responsible for ~100,000 deaths/year)

ODE-based model



(Bonneau et al., *Cell* 2007)

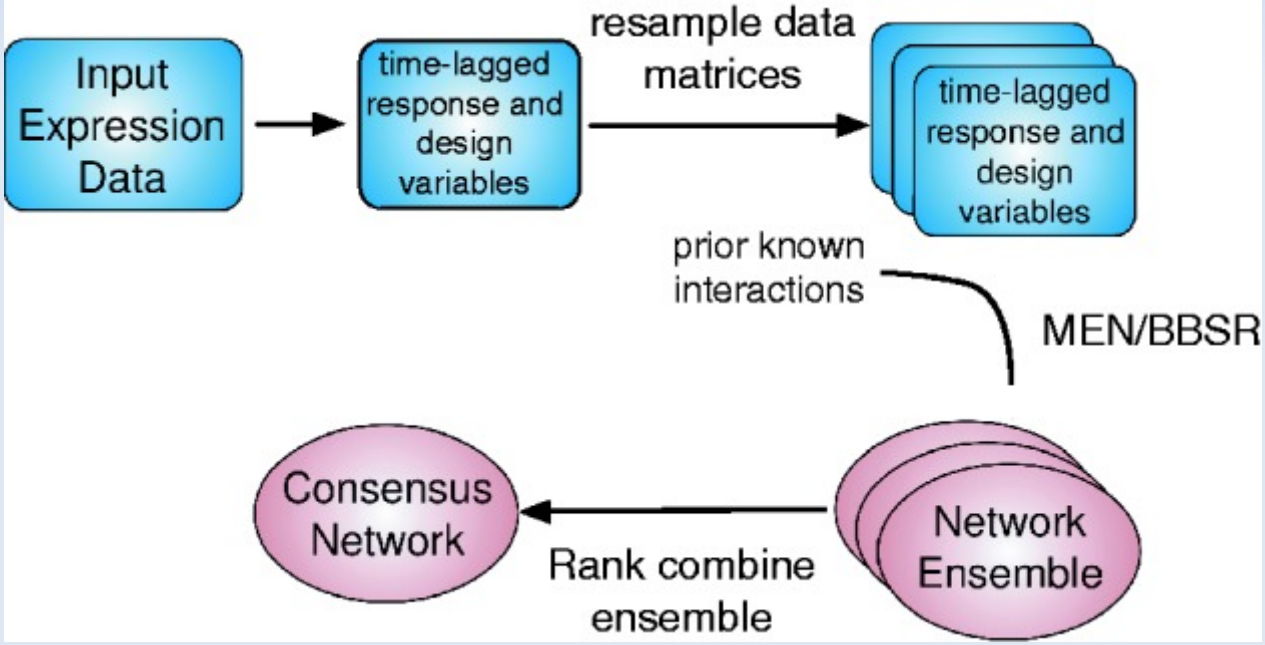
$$\tau \frac{dy}{dt} = -y + g(\beta \bullet Z)$$

$$g(\beta Z) = \begin{cases} \beta Z: & \text{if } \min(y) < \beta Z < \max(y) \\ \max(y): & \text{if } \beta Z > \max(y) \\ \min(y): & \text{if } \beta Z < \min(y) \end{cases}$$

Genes clustered first, using domain knowledge like protein-protein interactions

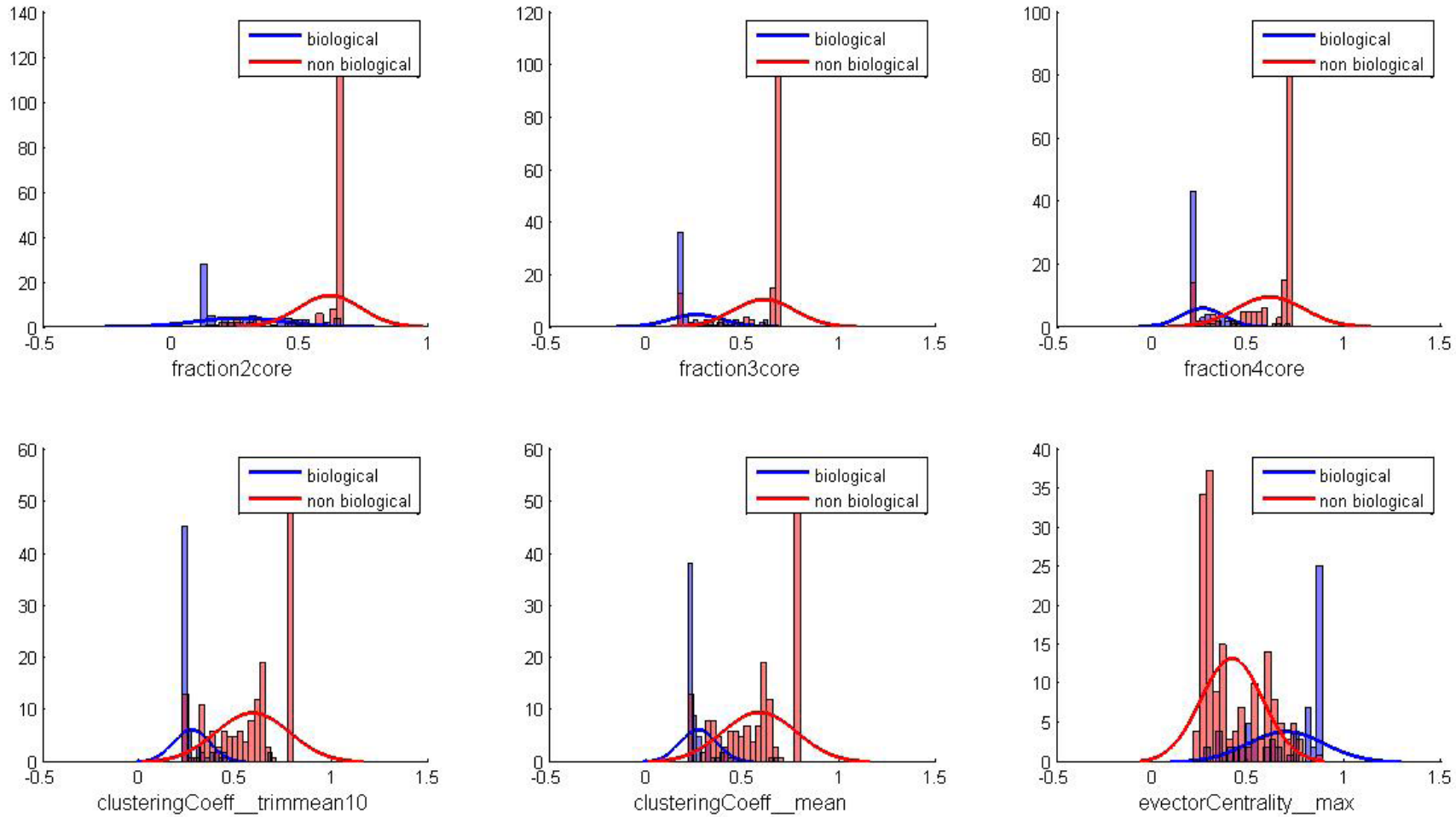
Edgewise priors

[Greenfield et al. 2013]



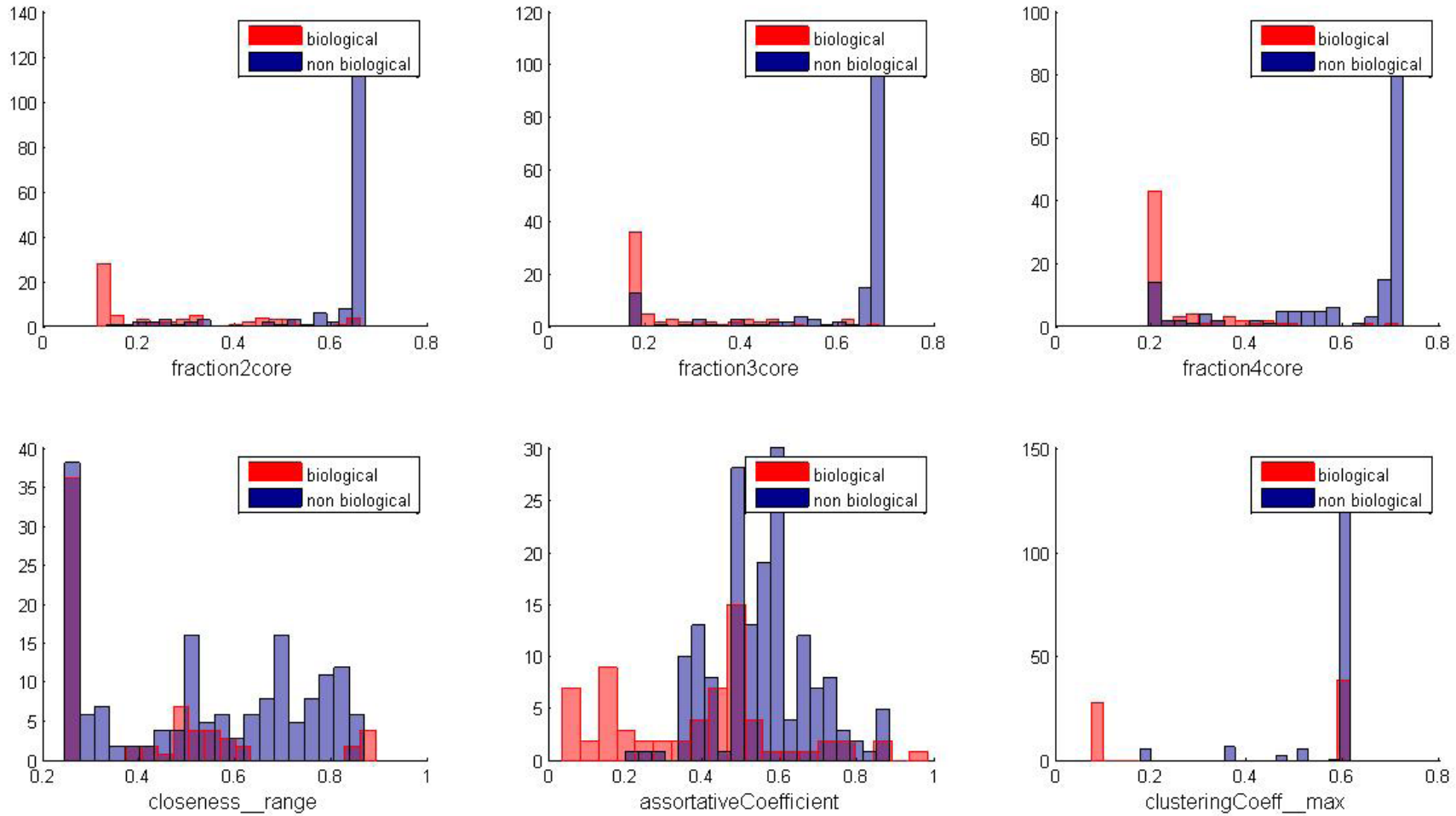
Biological network structure

Differing network features



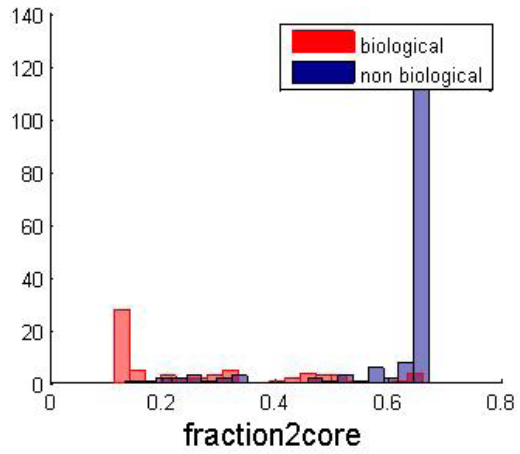
Biological network structure

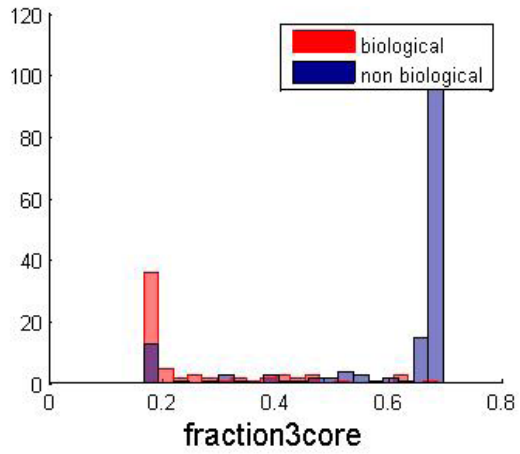
Density-controlled

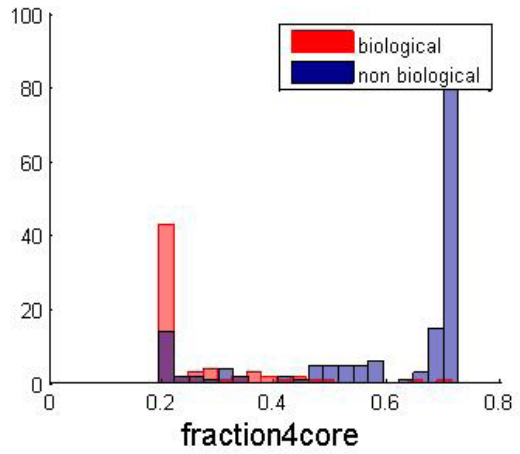


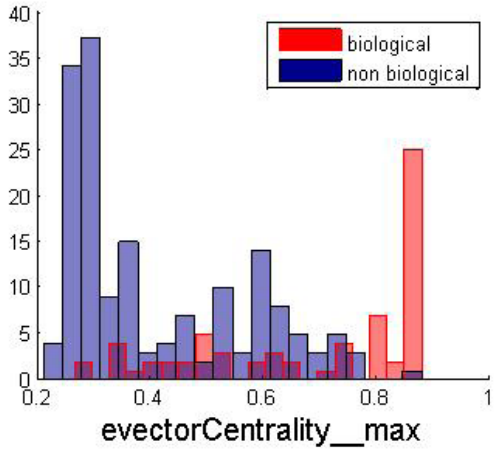
Biological network structure

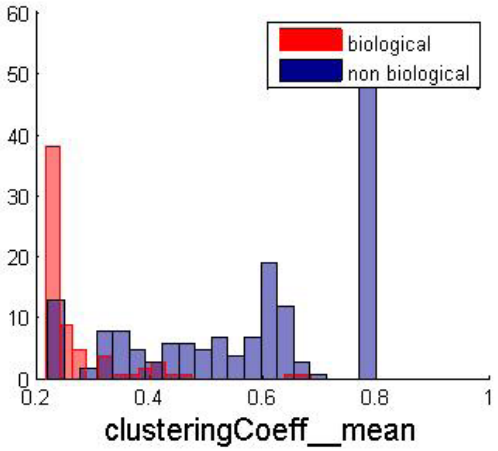
Size-controlled

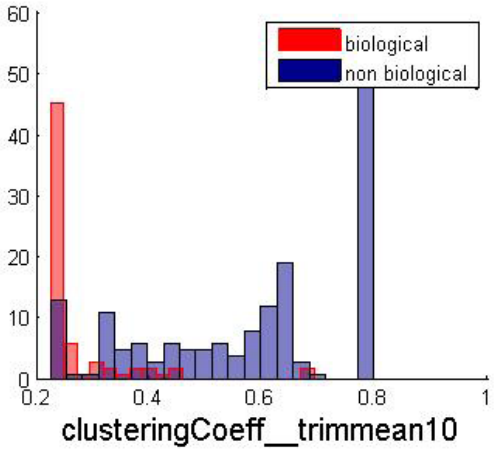












Other modelling approaches

- Extended Petri Nets [Durzinsky et al.]
- Markov Logic Networks
- Bayesian Networks

How can prior knowledge of higher-level network structure (beyond edges) be incorporated?