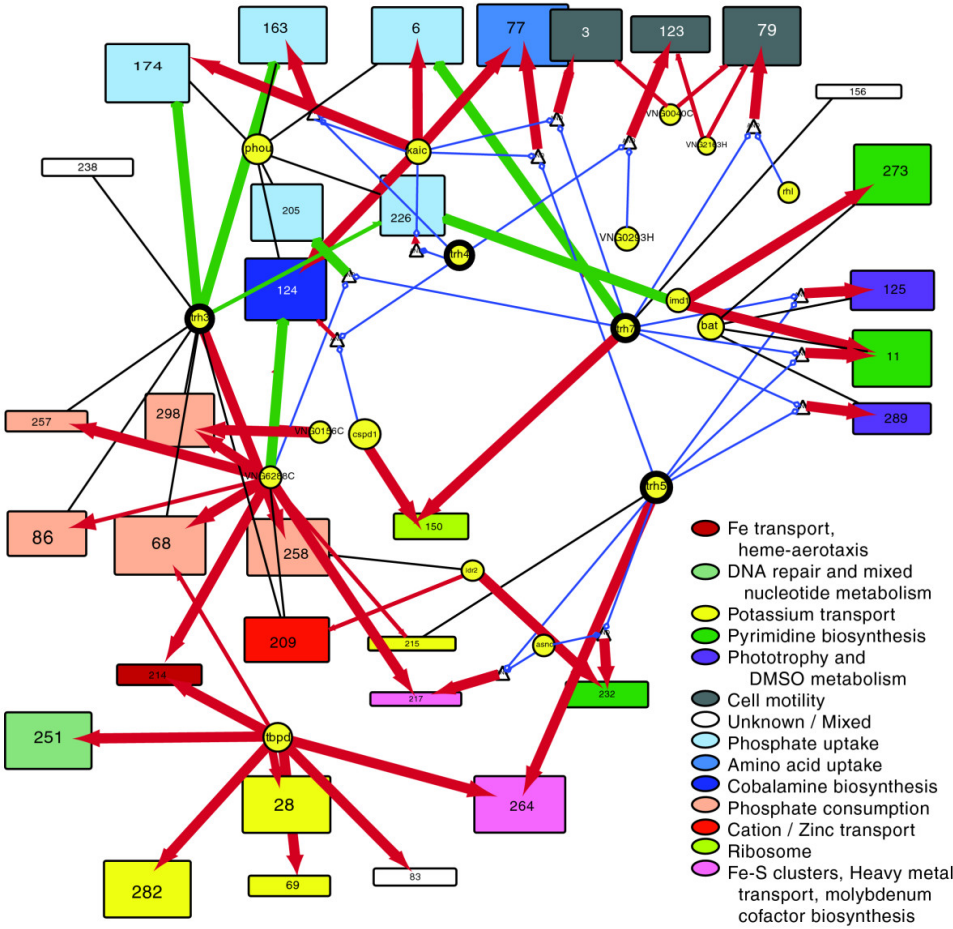
# Dynamics and Inference on Biological Networks




[Bonneau *et al.* 2006]

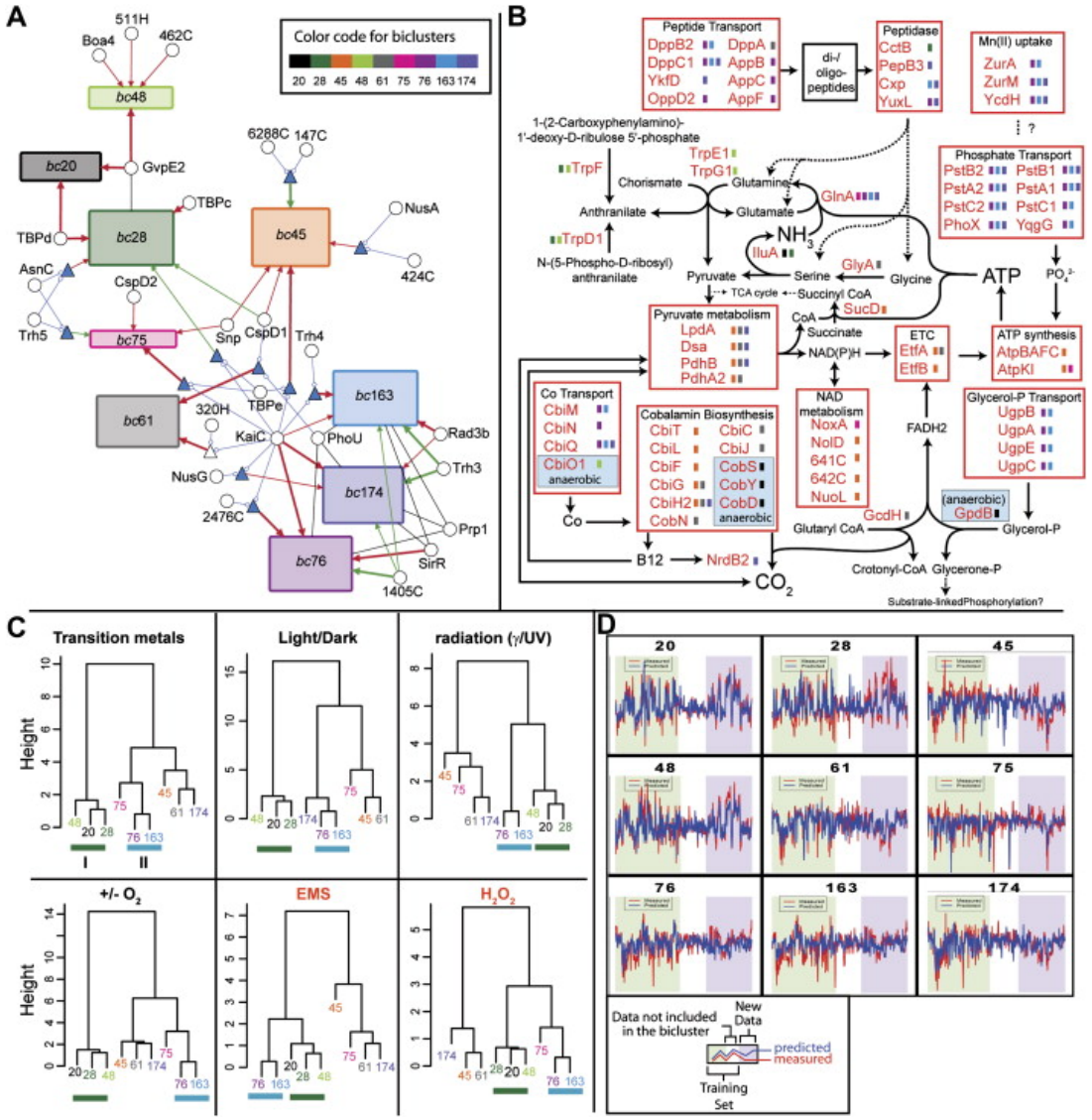
Sumeet Agarwal
(sumeet@iitd.ac.in)

# Outline

- Part I: Network Dynamics
  - What do we mean by ‘dynamics’?
  - Why study dynamics?
  - Mathematical representation
  - Link with time series
  - Example: Expression dynamics on protein interaction networks, date/party hubs
- Part II: Network Inference from Time Series Data
  - Causality in networks
  - Inference techniques
  - Example: Learning gene regulatory networks from microarray data
- Applications: Predictive Systems Biology

# What do we mean by ‘dynamics’?

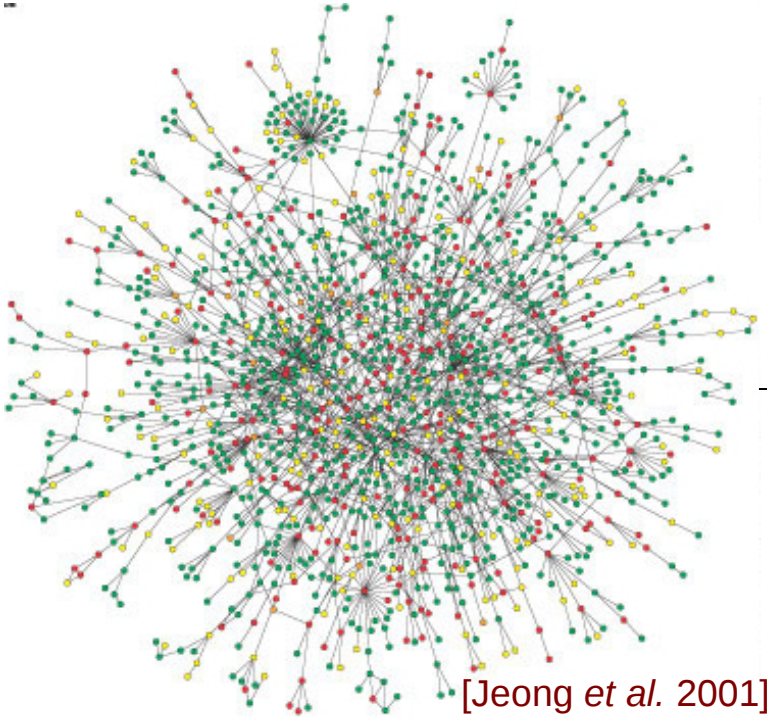
- Basically: things are changing with time
- For networks, each node may represent a time-varying quantity, such as gene expression levels
- Links may also change with time, or have a weight/strength dependent on the endpoints



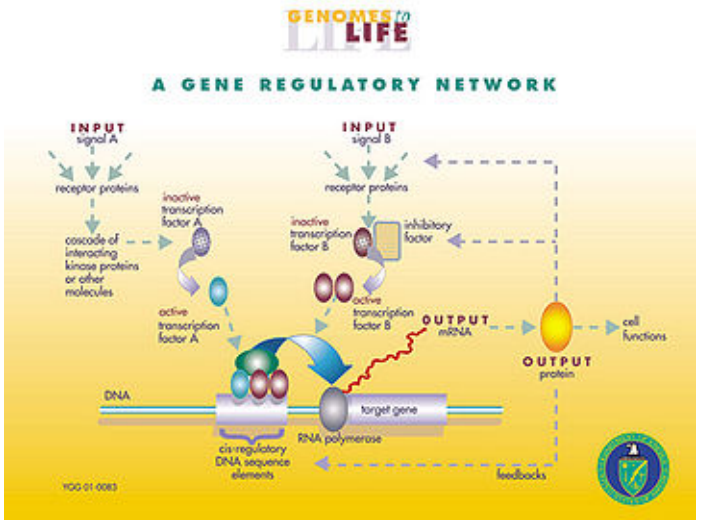

[Bonneau *et al.* 2007]

# Why study dynamics?

- Most networks/systems in the real world do change over time
- Studying dynamics tells us about certain properties of the system: both global (steady states, attractors) and local (causality)
- A model of dynamics can be used to make predictions about the future, and also about how the system responds to perturbation



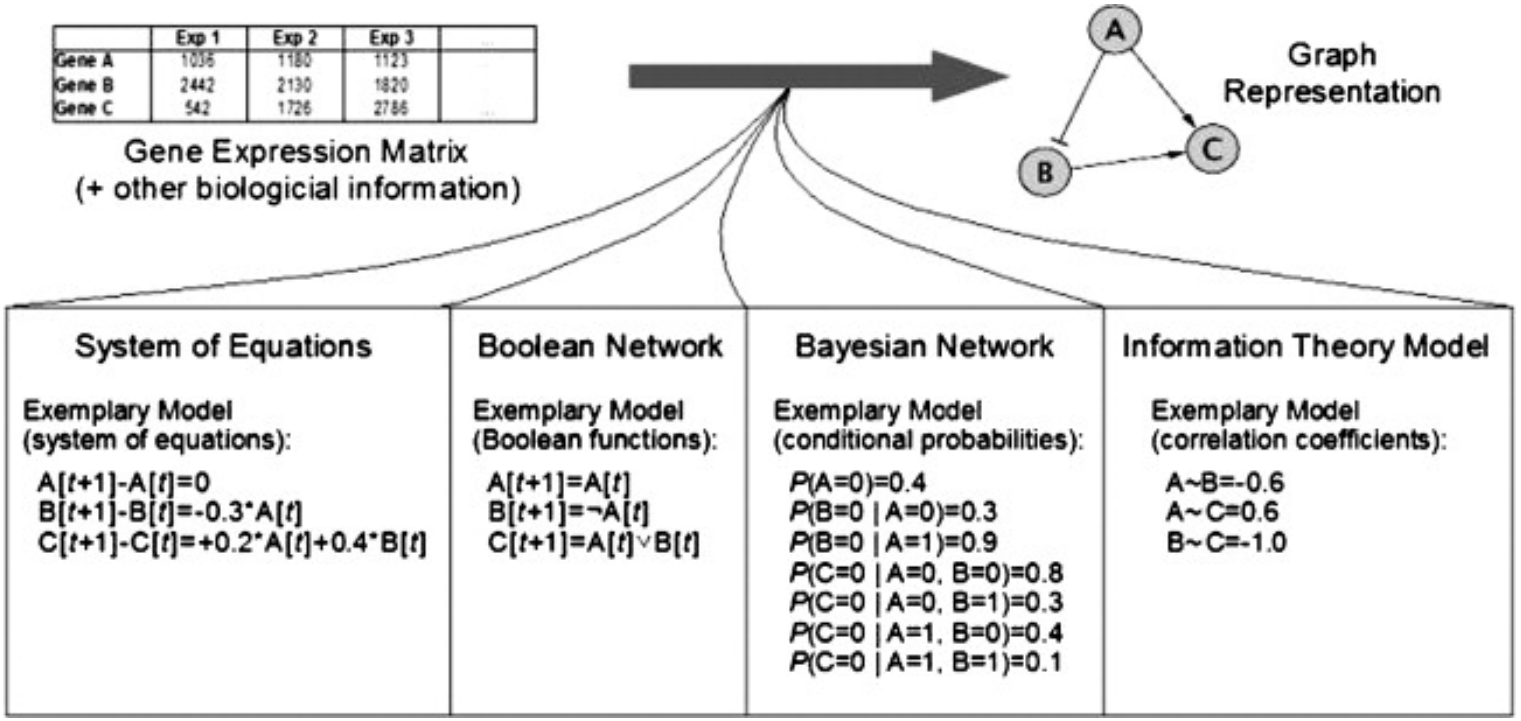
[Jeong *et al.* 2001]




[https://genomics.energy.gov]

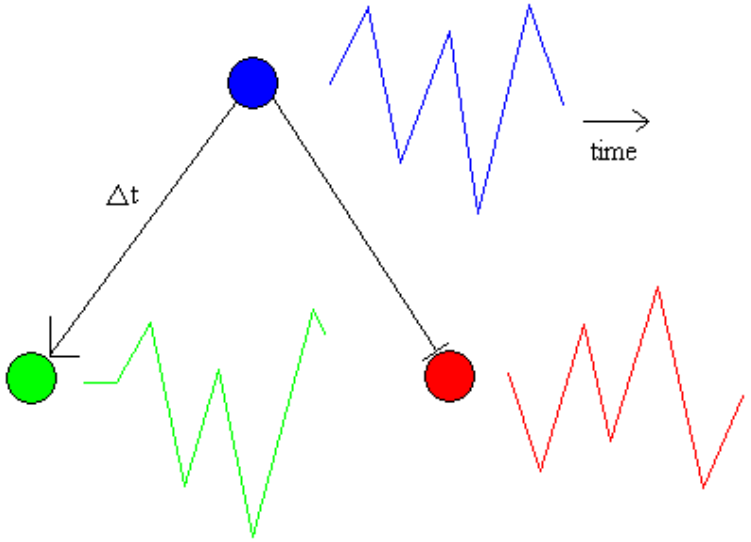
# Mathematical Representation

- The most preferred method, if feasible, is ordinary differential equations (ODEs), but for most large-scale systems we have to use appropriate simplifications for tractability




[Hecker *et al.* 2009]

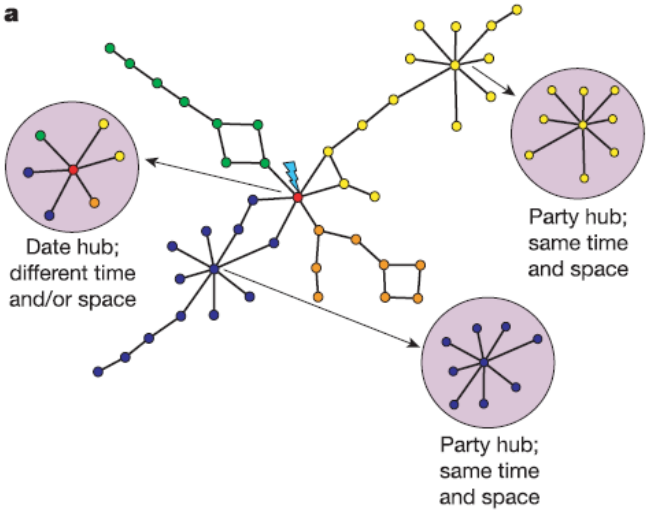
# Link with time series

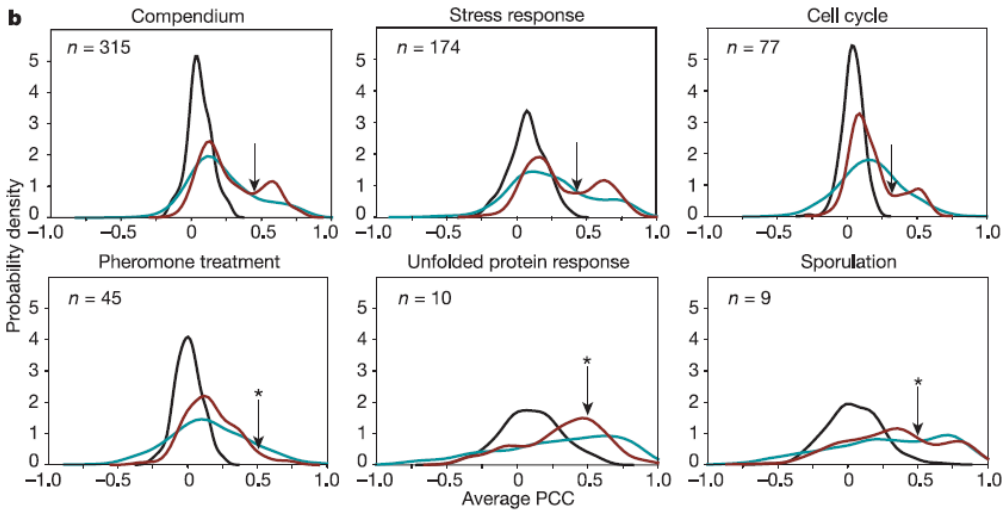



- The behaviour at each node can be described by a time series tracking an appropriate quantity (body weight, gene expression,...)
- In practice, raw observations are often in the form of time series data, which can be used (possibly combined with other information) to construct a plausible network
- Statistical techniques for time series (such as correlations) are widely used for this purpose---more on this later

# Example: date and party hubs

For a yeast protein-protein interaction network, hubs, defined as nodes with degree 5 or greater, were found to fall into two classes based on average co-expression with their interaction partners (Han *et al*. 2004):



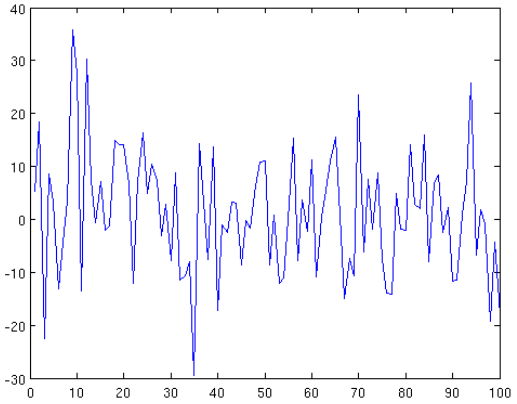




Date hubs were proposed to be high-level coordinators or linkers, whereas party hubs were thought to be local organisers for particular functional groups

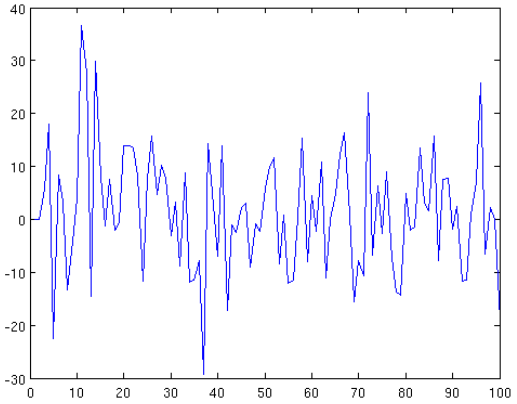
Intermission

# How to do Inference in Two Easy Steps

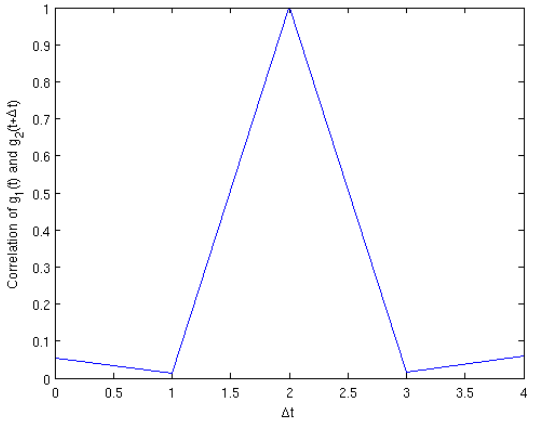
Gene 1

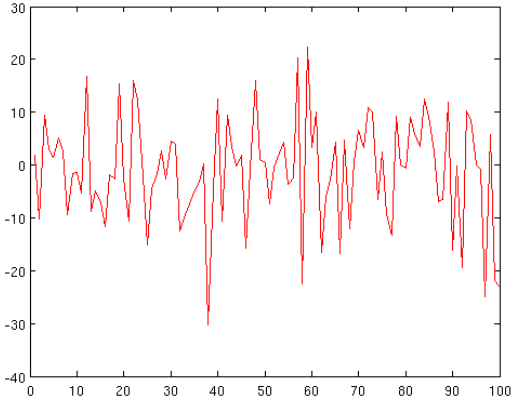


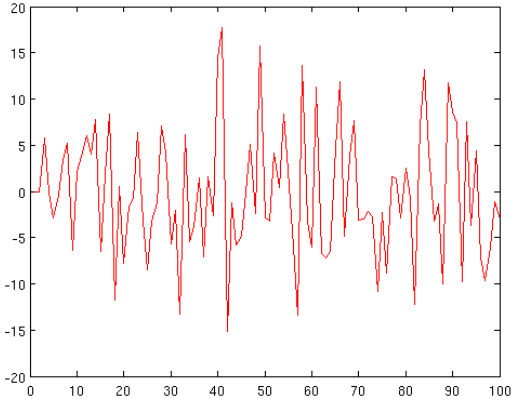
Gene 2

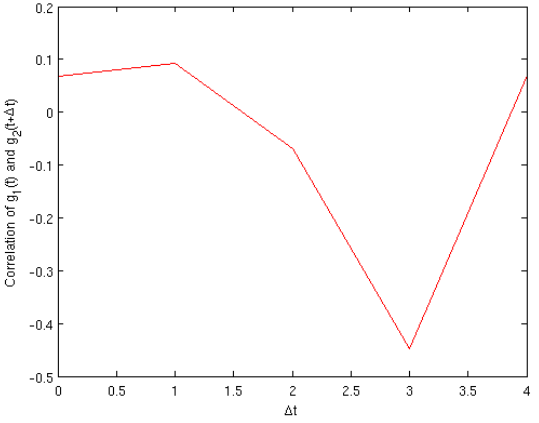


Cross-Correlograms



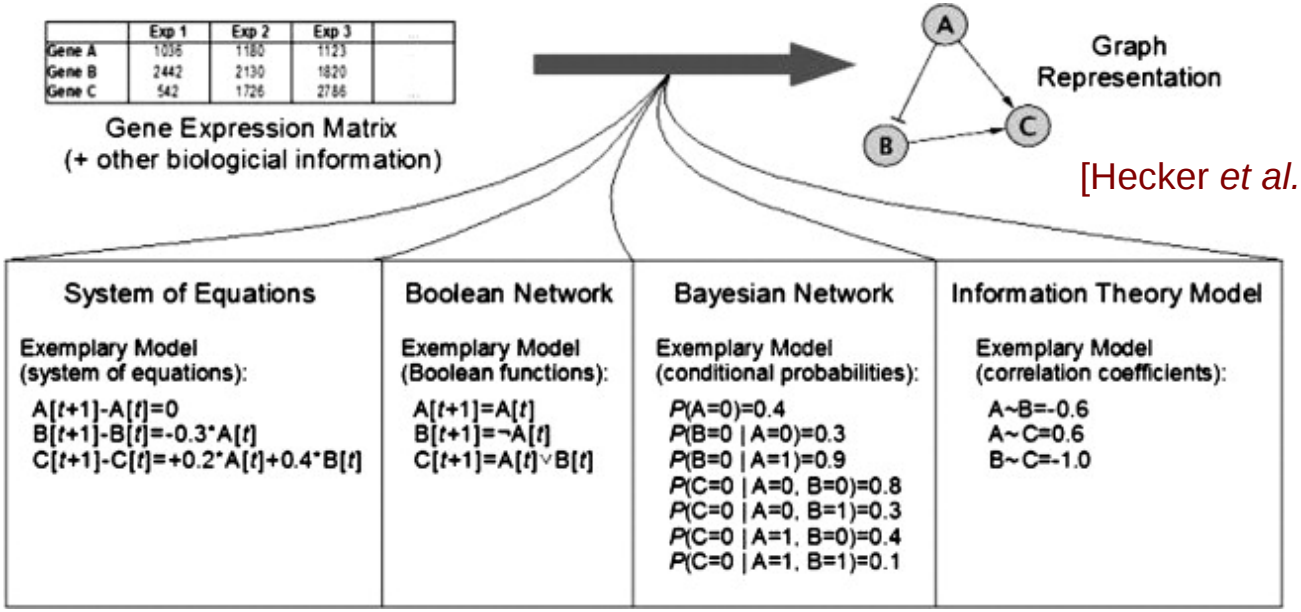





What can you infer about the relation between the two genes in each case?

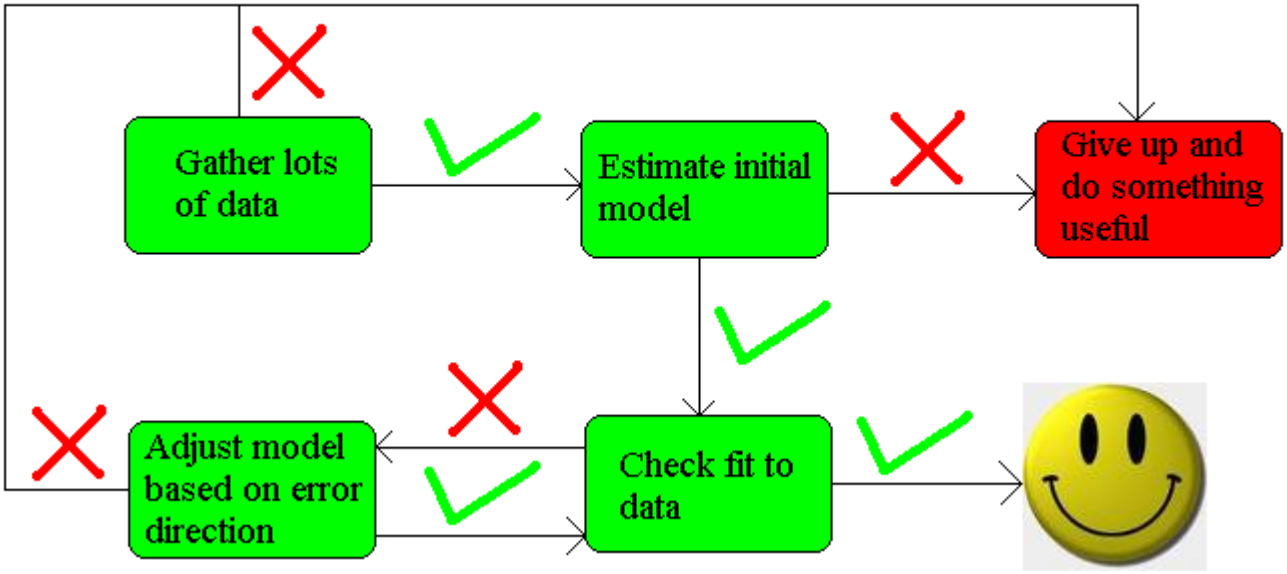
# Causality in networks

- Is a somewhat vexed philosophical notion; the classic 'correlation vs. causation' conundrum
- For our purposes, it just means we're trying to understand, at some given level (say proteins in a cell) whether changes in one entity lead to changes in another
- We can represent causal relationships in a network, and model them more precisely using a beautiful language called mathematics



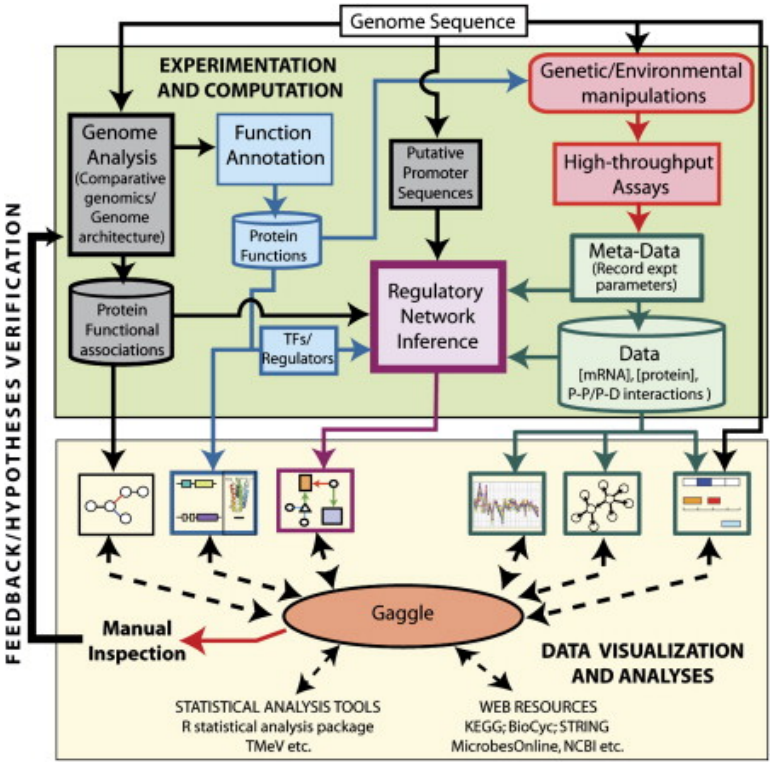
# Inference techniques

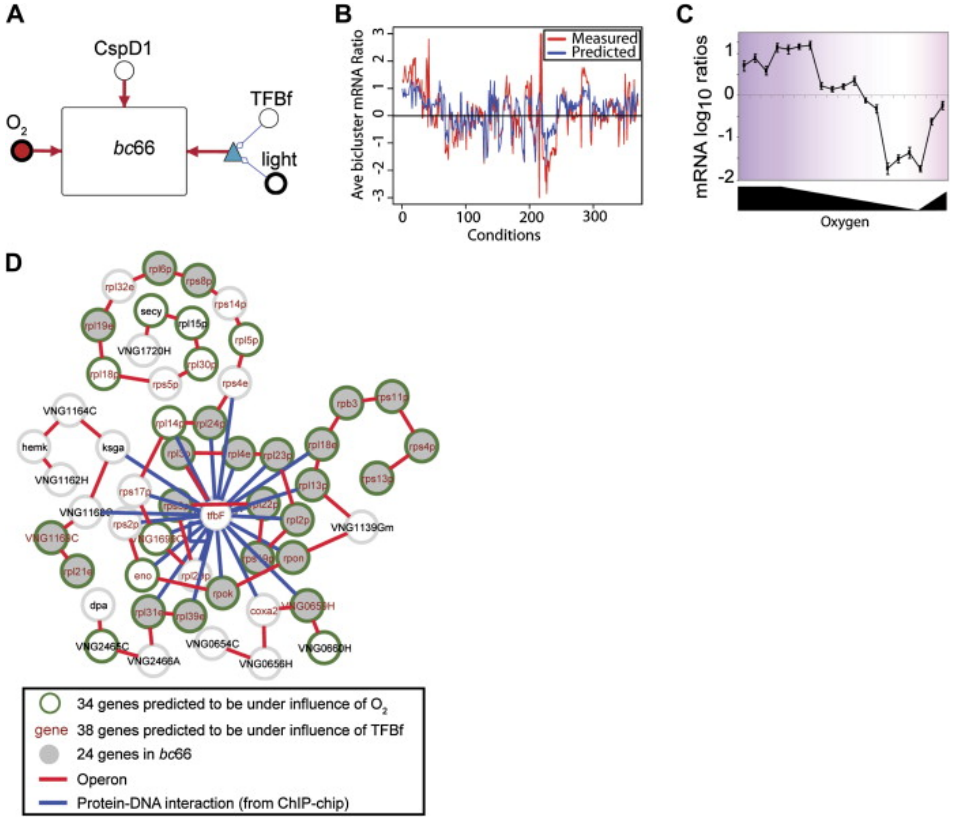
- We've already seen a simple example of how one might do this
- Two parts: inferring the network structure, and inferring the parameters or weights (which depend on the type of model used)
- Large number of techniques, but essential idea of all of them is to do statistical analysis of large quantities of experimental data
- Often involve iteration: infer a particular network, see how well it can reproduce your data, then attempt to adjust structure/parameters to improve this; stir and repeat until desired consistency is reached



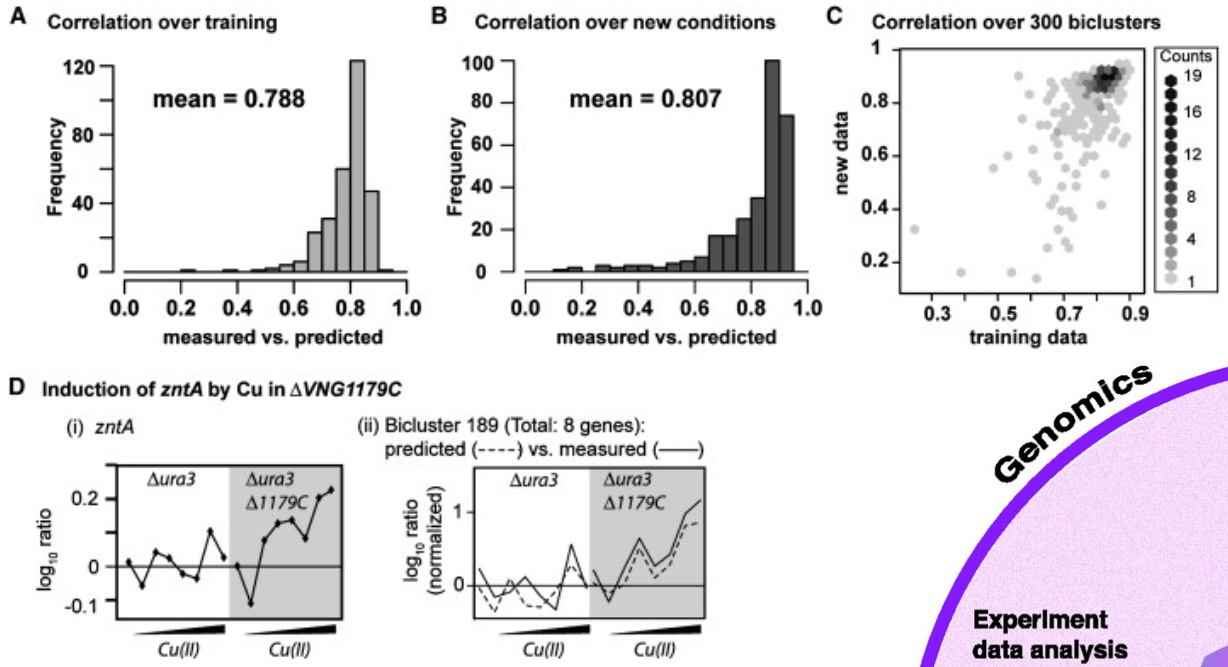
# Example: gene regulatory network for *H.salinarum*
(Bonneau *et al*. 2007)

For a novel, largely uncharacterised organism, using a large number of microarray experiments combined with homology and other information, a remarkably successful attempt at creating a network that explains the data



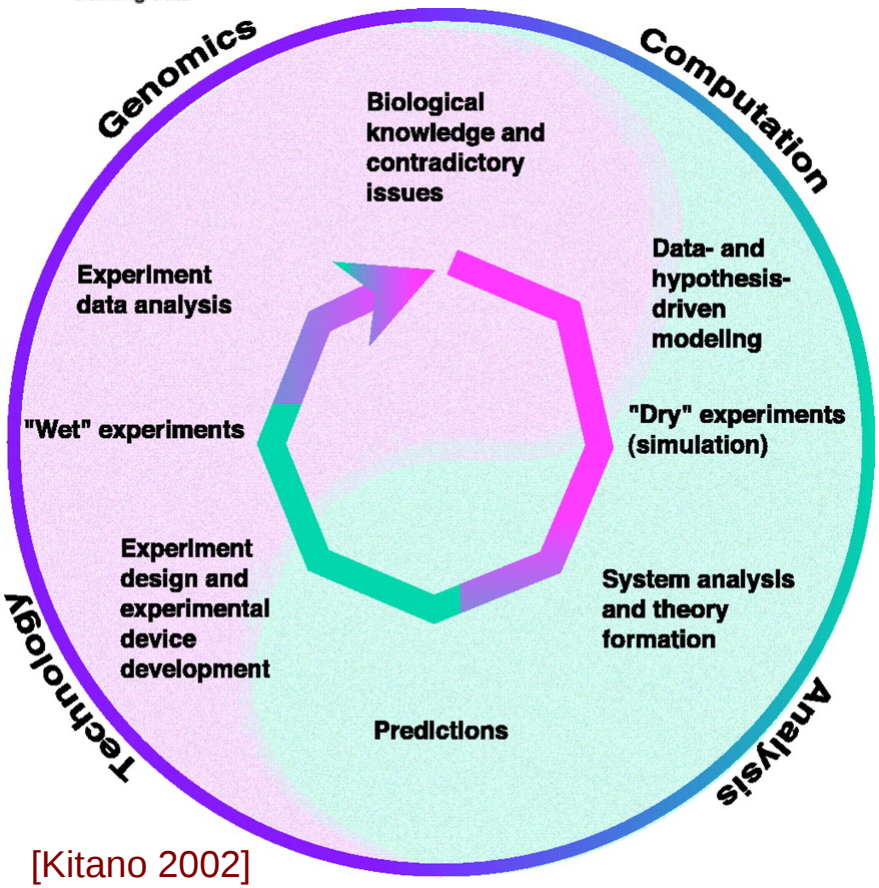



# Predictive Systems Biology




(Bonneau *et al*. 2007)

An important reality check for any model is: can it make accurate predictions of how the system will behave in novel circumstances? This can be regarded as a major goal of “systems biology”, and network-based models have a key role




[Kitano 2002]

# References


- Bonneau R, Reiss DJ, Shannon P, Facciotti M, Hood L, Baliga NS, Thorsson V. The Inferelator: an algorithm for learning parsimonious regulatory networks from systems-biology data sets *de novo*. *Genome Biology* 2006, **7:**R36.
- Bonneau R, Facciotti MT, Reiss DJ, Schmid AK, Pan M, Kaur A, Thorsson V, Shannon P, Johnson MH, Bare JC, Longabaugh W, Vuthoori M, Whitehead K, Madar A, Suzuki L, Mori T, Chang DE, Diruggiero J, Johnson CH, Hood L, Baliga NS. A Predictive Model for Transcriptional Control of Physiology in a Free Living Cell. *Cell* 2007, **131:**1354-1365.
- H. Jeong, S. P. Mason, A.-L. Barabási & Z. N. Oltvai. Lethality and centrality in protein networks. *Nature* 2001, **411**:41-42.
- Michael Hecker, Sandro Lambeck, Susanne Toepfer, Eugene van Someren and Reinhard Guthke. Gene regulatory network inference: Data integration in dynamic models—A review. *Biosystems* 2009, **96:**86-103.
- Jing-Dong J. Han, Nicolas Bertin, Tong Hao, Debra S. Goldberg, Gabriel F. Berriz, Lan V. Zhang, Denis Dupuy, Albertha J. M. Walhout, Michael E. Cusick, Frederick P. Roth & Marc Vidal. Evidence for dynamically organized modularity in the yeast protein–protein interaction network. *Nature* 2004, **430:**88-93.
- Hiroaki Kitano. Systems Biology: A Brief Overview. *Science* 2002, **295:**1662-1664.
- Sydney Brenner. Sequences and consequences. *Phil. Trans. R. Soc. B* 2010, **365:**207-212.
- Denis Noble. Biophysics and systems biology. *Phil. Trans. R. Soc. A* 2010, **368:**1125-1139.