Representation, Learning and Inference in Models of Cellular Networks

BMI/CS 776
www.biostat.wisc.edu/bmi776/
Mark Craven
craven@biostat.wisc.edu
Spring 2009

Various Subnetworks within Cells

- *metabolic*: describe reactions through which enzymes convert substrates to products

- *regulatory* (genetic): describe interactions that control expression of particular genes

- *signaling*: describe interactions among proteins and (sometimes) small molecules that relay signals from outside the cell to the nucleus

- note: these networks are linked together and the boundaries among them are not crisp

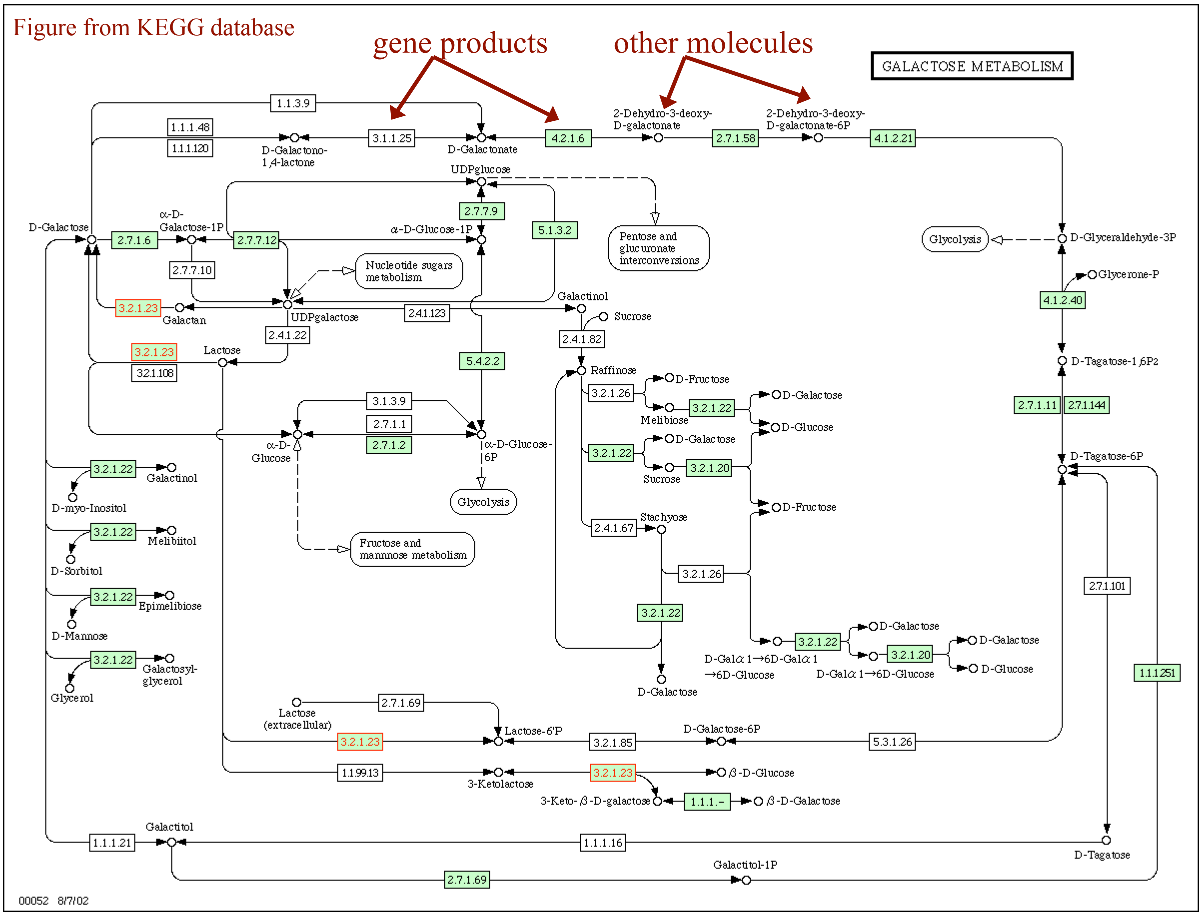


Figure from KEGG database

gene products

other molecules

Part of the E. coli Regulatory Network

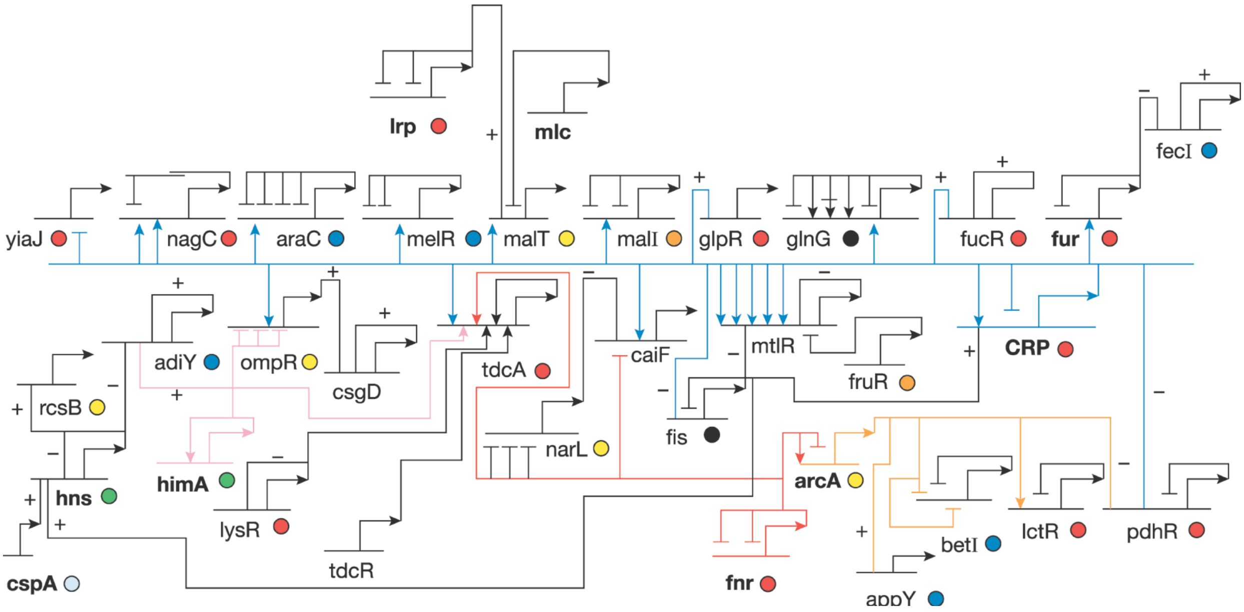
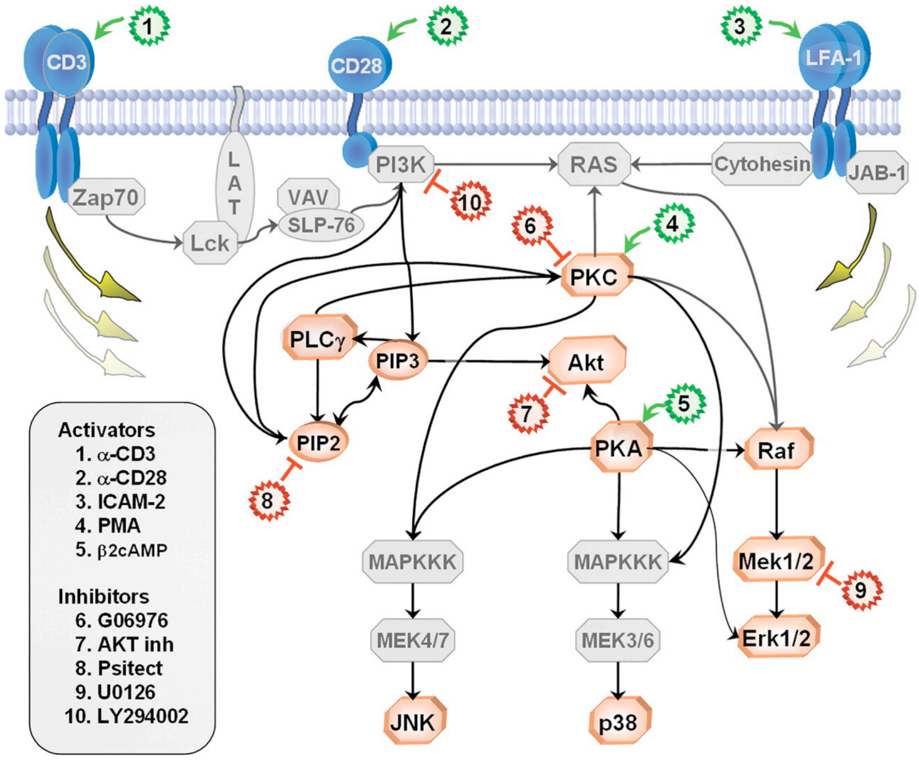


Figure from Wei et al., *Biochemical Journal* 2004

A Signaling Network



Activators
1. α-CD3
2. α-CD28
3. ICAM-2
4. PMA
5. β2cAMP

Inhibitors
6. G06976
7. AKT inh
8. Psitect
9. U0126
10. LY294002

Figure from Sachs et al., *Science* 2005

Two Key Tasks

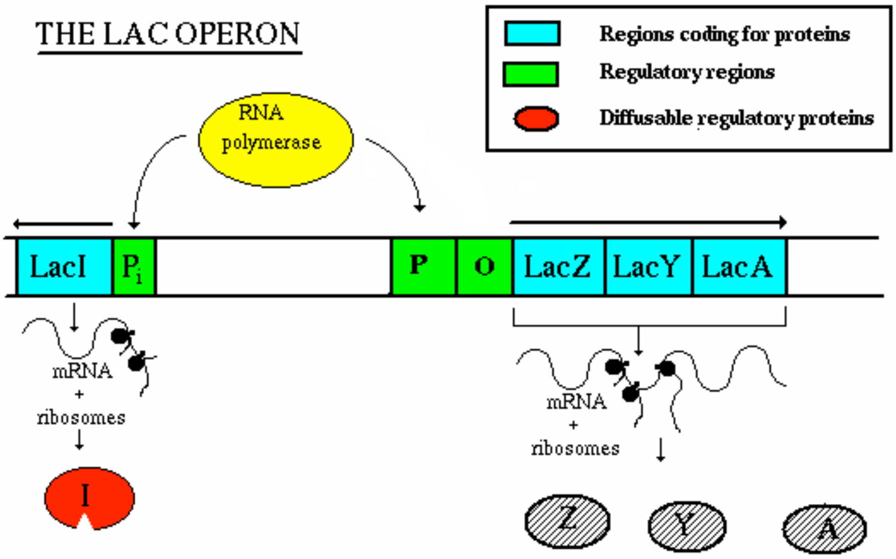
- *learning*: given background knowledge and high-throughput data, try to infer the (partial) structure/parameters of a network
- *inference*: given a (partial) network model, use it to predict an outcome of biological interest (e.g. will the cells grow faster in medium *x* or medium *y*?)
- both of these are challenging tasks because typically
  - data are noisy
  - data are incomplete – characterize a limited range of conditions
  - important aspects of the system not measured – some unknown structure and/or parameters

Representations for Network Models

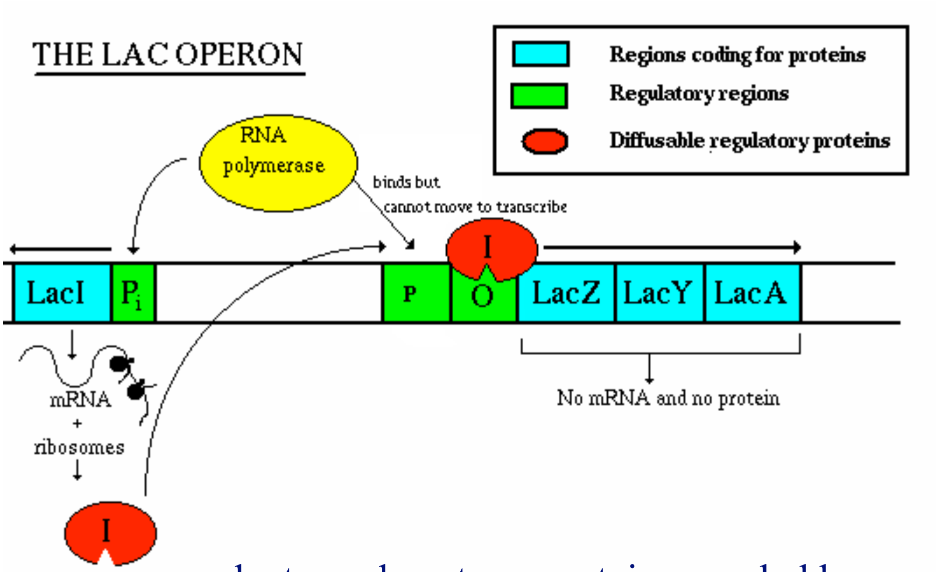
- directed graphs
- Boolean networks
- Bayesian networks
- differential equations
- systems of constraints + linear programs
- logic programs
- etc.

Transcriptional Regulation Example: the lac Operon in *E. coli*

E. coli can use lactose as an energy source, but it prefers glucose. How does it switch on its lactose-metabolizing genes?

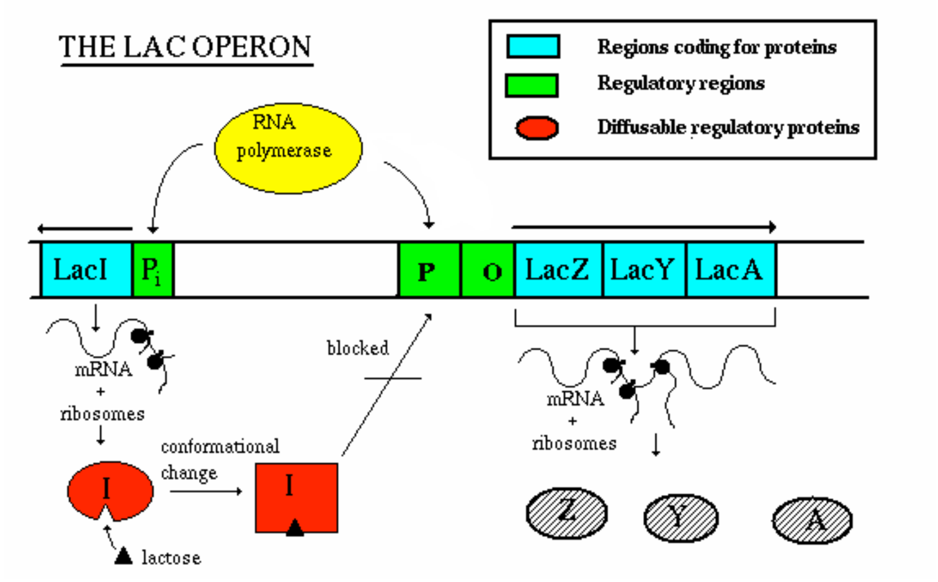


The lac Operon: Repression by LacI



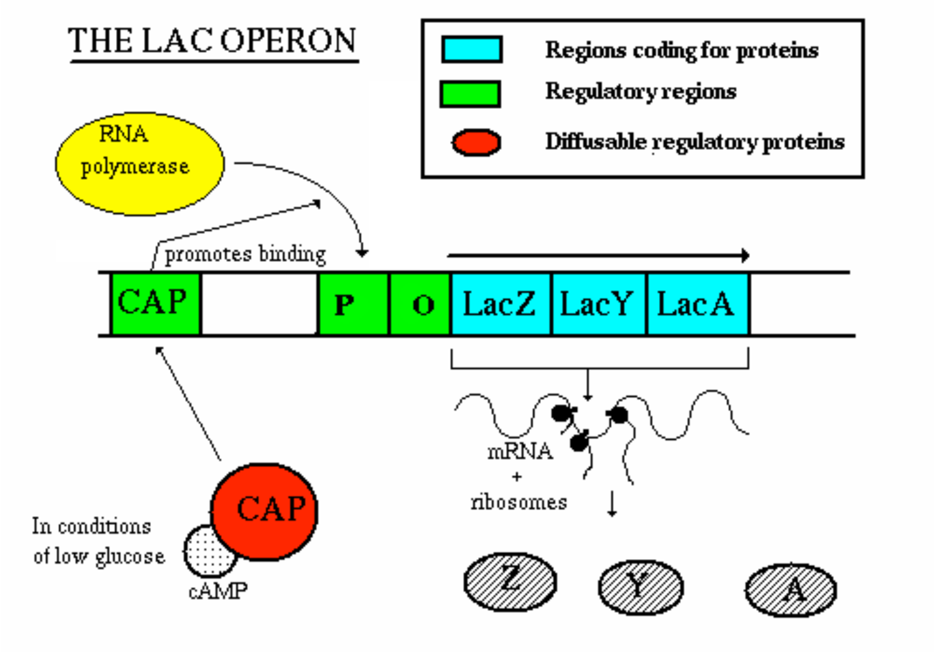
lactose absent ⇒ protein encoded by
lacI represses transcription of the lac operon

The lac Operon: Induction by LacI



lactose present ⇒ protein encoded by
lacI won’t bind to the operator (O) region

The lac Operon: Activation by Glucose



glucose absent ⇒ CAP protein promotes binding by RNA polymerase; increases transcription