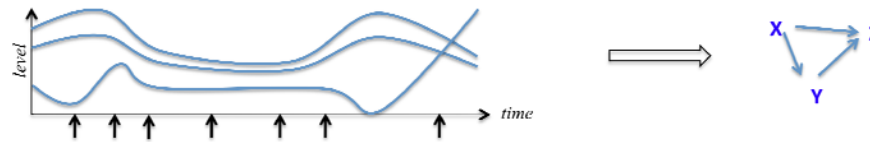


# Inference of Gene Regulatory Networks from Expression Data

14.6.10

Objective: To give a presentation of about 60 minutes at the end of the week covering the key aspects of the inference of gene regulatory networks from expression data.



The expression levels of mRNA is a function of the inner wirings of a cell and in favourable circumstances allow inference of the regulatory networks of the cell. However, there are many hurdles to be overcome. Expression data is very noisy and hard to get under the ideal controlled circumstances. Measurements are often the average over a large number of cells that might be in different conditions and not corresponding to the same timepoints in a time series. A large set of data has been collected since the late 90s and a series of models have been developed to interpret these data. There are many additional facets besides the pure “expression data  $\rightarrow$  network” inference problem: Additional knowledge might be available that could constrain the possible networks proposed and networks might be known in related species. In the unconstrained version of this problem, the number of possible networks can be counted by standard combinatorial techniques and is enormous. This should lend scepticism to claims that the networks have been found for a larger set of genes in a given experiment.

## The Big Questions Are:

- Which data allows network inference?
- What are the hidden assumptions when you try to infer a network?
- What are the major classes of network models used?
- How many networks are there?
- How reliable is such inference?
- Can comparative approaches help?

## Possible Contents of Presentation

1. History of Expression Data
2. Networks in Cellular Biology
3. Which types of data allow network inference
4. Mathematical/statistical models of expression data
5. Principles of inference
6. Counting Networks
7. Examples of successful network inference

## Recommended literature

- Aijo et al. (2009). Learning gene regulatory networks from gene expression measurements using non-parametric molecular kinetics. *Bioinformatics*
- Bansal et al. (2007). How to infer gene networks from expression profiles. *Molecular Systems Biology*, 3, 1-10.
- Bumgarner, R. & Yeung, K. (2009). Methods for the inference of biological pathways and networks. *Computational Systems Biology*, 541, 225-245.
- Camacho et al. (2007). Comparison of reverseengineering methods using an in silico network. *Annals of New York Academy of Science*, 1115, 7389.
- Cao, J. & Zhao, H. (2008). Estimating dynamic models for gene regulation networks. *Bioinformatics*, 24, 1619-1624.
- de Hoon et al. (2003). Inferring gene regulatory networks from timeordered gene expression data using differential equations *Pac Symp Bioc*, 267274.
- De Jong, H. (2002). Modeling and simulation of genetic regulatory systems: A literature review. *Journal of Computational Biology*, 9, 67103.
- Friedman et al. (2000). Using bayesian networks to analyze expression data. *Journal of Computational Biology*, 7, 601620.
- Margolin et al. (2006). ARACNE: An algorithm for the reconstruction of gene Networks in a Mammalian Cellular Context *BMC Bioinf*, 7(Suppl 1);S7doi:10.1186/1471-2105-7-S1-S7
- Perrin et al. (2003). Gene networks inference using dynamic Bayesian networks. *Bioinformatics*, 19, ii138ii148.
- Ratray et al. (2008). Gaussian process modelling of latent chemical species: applications to inferring transcription factor activities. *Bioinf* 24,i70i75.
- Zou et al. (2004) A new dynamic bayesian network (DBN) approach for identifying gene regulatory networks from time course microarray data. *Bioinf* 21, 7179.

“Big Questions”, “Contents” and “Recommended Literature” are only suggestions from which the student is welcome to depart from or completely ignore.