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 → 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 at 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

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