

Analysis of DNA Chips and Gene Networks

Xiaohui Xie

University of California, Irvine

Gene Expression

- Different tissues may express different genes. The same cell may express different genes under different circumstances (eg. stress, nutrition, stages of cell cycle etc.)
- to understand a genome, we need to understand
 - how active various genes are under different cell types and conditions
 - how genes interact with each other
- One way to do this is to take a 'snapshots' of the cells and measure
 - how much of each mRNA is there in the cell?
 - how much of each protein is expressed?

Microarrays

- Knowing how much of proteins is expressed would be most informative. But the technology is still not very well developed.
- Technology for measuring mRNA is much better developed. Microarrays enable us to measure mRNA levels for almost all genes simultaneously.
- One caveat: the level of mRNA in the cell is an indirect measurement of the protein level, because of post-transcriptional regulation.

Some questions microarrays intend to answer

The mRNA levels for all genes in a cell provide a good characterization of the *state* of the cell. One can ask questions such as

- How active are various genes in different cell/tissue types?
- How does the activity level change under different conditions?
 - nutrition
 - environmental conditions
 - stages of cell cycles
 - normal vs. disease tissues
 - knockout or over-expression of a gene
- to infer how gene expressions are regulated. What genes are likely regulated together?

Some computational tasks

- Clustering genes: which genes seem to be regulated together?
- Clustering samples: which conditions/treatments/individuals have similar profiles?
- Classifying genes: to which functional class does a given gene belong?
- Classifying samples: to which class does a given sample belong, eg. normal vs tumor tissues
- Reverse engineering: inferring regulatory networks from gene expression data