2. Visual Analysis of Multivariate Biological Networks

Information Visualization (186.141)
TU Vienna, Austria
June 13 & 14, 2013
2.1 Introduction

- **Correlations:** Genome – Chromosome – Gene – DNA

- **DNA:** Double helix with nucleotides from sugar, phosphate and four bases (Adenin, Cytosin, Guanin, Thymin)
2.1 Introduction

- **Genome:** only 5% genes consisting of Exons and Introns

![Genome Diagram](image)

<table>
<thead>
<tr>
<th>Organism</th>
<th>Base Pairs</th>
<th>Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>yeast</td>
<td>12 mill.</td>
<td>6000</td>
</tr>
<tr>
<td>fruit fly</td>
<td>137 mill.</td>
<td>13.601</td>
</tr>
<tr>
<td>human</td>
<td>3 billions</td>
<td>~30.000</td>
</tr>
</tbody>
</table>

- **RNA:** Uracil instead of Thymin, single strands
2.1 Introduction

- Proteins are built by amino acids
- **20 amino acids**: coded by nucleotide triplets (Codons)
  - e.g. GCA, GCC, GCG → Alanin

Prime-Secondary-Tertiary Structure

ADMVIKAPAG
AKVTKAPVAF
SHKGHASMDC

- Protein synthesis
  - **Transcription**: DNA → mRNA
  - **Translation**: mRNA → amino acids → protein

[Video]

http://de.wikipedia.org/wiki/Codon
2.1.1 Tasks of Bioinformatics

- Support of biologists within processing, analysis and interpretation of large data sets

- Approaches:
  - Biological data bases
  - Methods for comparison and function prediction of sequences
  - Function prediction of proteins
  - Locating of new correlations
  - Simulation of biological processes
2.1.2 Visualizations

- Visualizations of networks and related structures
  - Networks
  - RNA-secondary structure
  - Phylogenetic trees
- Visualizations of sequences
  - Sequence alignments
- Proteins
  - Structure prediction
  - Structure comparison
  - Protein dynamics

Here, we will focus on these issues!
2.2.1 Motivation

- **Network Biology**
  - Classical biology ⇒ Learning from the individual
  - Systems biology ⇒ Biological processes are integrated systems of many diverse interacting components
  - Starting with the human genome project, high-throughput methods generate a huge amount of data
  - Networks are the key concept to structure and combine that data

[Inspired by O. Kohlbacher's talk at Dagstuhl 08]
2.2.1 Motivation

There is a hierarchy of such biological networks

- (Molecular Graphs)
- Metabolic Networks
- Interaction Networks
- Regulatory Networks
- …
There is a hierarchy of such biological networks

- (Molecular Graphs)
- Metabolic Networks
- Interaction Networks
- Regulatory Networks
- …
2.2.1 Motivation

There is a hierarchy of such biological networks

- (Molecular Graphs)
- Metabolic Networks
- Interaction Networks
- Regulatory Networks
- …
Why is that important?

- Diseases can be explained in context of networks
  - Infection: a foreign network starts operating in our own network
  - Genetic defects: incorrect connectivity of an element

- Applications
  - Drug design, metabolic engineering, …
2.2.2 Metabolic Networks

- **Paths**: Sequence of reactions that provide all together for the translation of one substance into another one (pathways)

- One single reaction is defined by a catalyzing enzyme

- **Example**:

  - As a whole, the paths build a so-called metabolic network (MN) \[\rightarrow\text{Graph}\]
2.2.2 Metabolic Networks
2.2.2 Metabolic Networks

PathViewer [Uni Tübingen, 2003]
2.2.2 Metabolic Networks

2. BioNetVis

2.2 Networks

By Igor Rojdestvenski

VXML - interface to pathway databases

Linnaeus University

Erasmus Teaching Exchange 13
DV – Prof. Dr. Andreas Kerren
2.2.2 Metabolic Networks

- BioPath

by F. Schreiber
2.2.2 Metabolic Networks

Grid-based Visualization

[Uni Leipzig & Linnaeus Univ., 2009]
Visualization of Simulation Data

[Uni Leipzig & Linnaeus Univ., 2009]

Interrelationships between network components

[Linnaeus Univ. & IPG Gatersleben, 2012]
2.2.2 Metabolic Networks

- BioNetVis

2.2 Networks

[Video]

http://sourceforge.net/projects/gliep/

2.2.3 Protein-Protein Interactions

- Nodes are proteins and edges the interactions between them.
- Interactions between proteins are important for many biological functions.
- Example:
  - Signal Transduction: signals from the exterior of a cell are mediated to the inside of that cell by protein-protein interactions of the signaling molecules.
  - It plays a fundamental role in many biological processes and in many diseases.

http://en.wikipedia.org/wiki/Protein-protein_interactions
2.2.3 Protein-Protein Interactions

Cytoscape

[Schwikowski et al; 00]
2.2.4 Special Visualizations

Comparison of biochemical networks
[U. Brandes et al. Integrative Bioinformatics, 2004]
2.2.4 Special Visualizations

- Visualization of Network Centralities

[Video]

2.3 Phylogenetic Trees

- **Problem**
  - Analysis of the development history of today's species and their relations

- **Example:**

```
Siamang  Gibbon  Orangutan  Gorilla  Human  Chimpanzee
```

- hypothetical ancestors
- today's species
2.3 Phylogenetic Trees

- Objects
  - Species, populations, DNA, proteins

- Tree

There are two methods for the determination of the root
  - Outgroup method: with not-related objects
  - Hypothesis of a molecular clock: root has the same distance to all objects
2.3.1 Visualizations

- Example tool: PHYLODRAW
  - Presented by Choi et al., `00

http://pearl.cs.pusan.ac.kr/phylodraw
2.3.1 Visualizations

- **Example tool:** TreeJuxtaposer
  - Presented by T, Munzner et al, SIGGRAPH `03

- Focus & Context
- “Guaranteed visibility”
- Computing structural differences

[Video]