

**SGN-6156, Lecture 12**  
**Review lecture**

**Harri Lähdesmäki, harri.lahdesmaki@tut.fi**

**Department of Signal Processing,  
Tampere University of Technology**

**07.05.2008**

## The main topics

- Biological sequence analysis
- Modeling biological regulatory networks

## Important points by lecture

- Lecture 1
  - Basic building blocks and basic regulatory mechanisms (from the point of view what was studied later during the course)
  - Motivation for sequence alignment: why it is useful?
  - A probabilistic alignment scoring model: substitution matrices, gap penalties, likelihood ratios, etc.
- Lecture 2
  - Dynamic programming algorithms for finding alignments (developed separately from the scoring model)
  - Global alignments, local alignments, and the ideas of more complex alignment procedures
  - Algorithmic complexity
  - Motivation and the idea of heuristic alignment methods (BLAST)

- Bayesian model comparison (applications throughout the course)
- Standard hypothesis testing also non-parametric (applications throughout the course)
- Lecture 3
  - Alignment with the affine gap penalty
  - Substitution matrices: estimation
  - Markov chains: definition, parameter estimation, simple applications, significance of a Markov model
  - Hidden Markov models (HMM): definition/basics, Viterbi for path estimation
- Lecture 4
  - HMMs: the probability of a sequence/forward algorithm, the probability of a state/forward-backward algorithm, posterior decoding
  - Pairwise alignment using HMMs (truly probabilistic interpreta-

tion): Viterbi, forward algorithm, posterior of alignments

- Lecture 5
  - Profile HMMs for sequence families: definitions, searching (Viterbi and forward algorithms)
  - Multiple alignments: scoring methods, dynamic programming for finding alignments, iterative (heuristic) methods
- Lecture 6
  - DNA sequence motifs
  - Transcription factor binding sites: representation, prediction, motif discovery
- Lecture 7
  - Stochastic simulation methods: benefits, stochastic vs. deterministic simulation methods

- Gillespie's simulation algorithm
- Lecture 8
  - Simulation of biological processes: general, the level of approximation
  - Quantitative models: biochemical reaction systems
  - Modeling of biological processes: parameter estimation, model selection, cross-validation
- Lecture 9
  - Bayesian networks and dynamic Bayesian networks
  - Motivation for probabilistic graphical models, definition, probability factorization, discrete model, parameter learning, structure learning, search strategies for structure
- Lecture 10

- Stoichiometric analysis of metabolic networks and pathways
- Metabolism, metabolic networks and pathways
- Reconstruction of metabolic networks, stoichiometry and fluxes
- Metabolic pathway analysis: steady-state, flux balance analysis, minimization of metabolic adjustment, elementary modes, extreme pathways, applications
- Lecture 11
  - Gene ontology: ontologies and annotations
  - Gene set enrichment analysis

## Exam

- The first exam on 14-05-2008
- You can use a (graphical) calculator in the exam (you may or may not need it...)