

1. **Genome sequencing.**

Bioinformatic pipelines and algorithms for analysing genome sequencing data. Different methods for next-generation DNA sequencing, short- and long-read sequencing. Background: The molecular mechanism of replications. Algorithms: Euler and Hamilton path, de Bruijn graphs

2. **Transcriptional regulation.**

How do transcription-factor motif search algorithms work? Introduce the Chip-Seq analysis and the analysis of Chip-seq data. Background: The regulation of pro- and eukaryotic gene expression, transcription factors. Algorithms: Random motif search, Gibbs sampling

3. **Transcriptomics.**

Different high-throughput methods and the bioinformatic pipeline to investigate the transcriptome. Background: The molecular mechanisms of transcription and translation. The characteristics of the genetic code. Algorithms: Differential expression analysis with pairwise models and with generalised linear models.

4. **Proteomics.**

Methods and the bioinformatic pipeline to analyse mass spectrometry data. Background: Amino acids and peptides. The proteome and the proteomics. Various types of posttranslational modifications and their roles in regulating protein function. Antibiotic sequencing Algorithms: Branch and Bound algorithms

5. **Proteins.**

Prediction of protein sequence features, evaluation of prediction methods Background: Structure-function relationships of proteins Algorithms: PSSM, Shannon entropy, pseudocount

6. **Protein structure.**

Domain classifications, Structural comparisons, structure prediction methods, macromolecular modelling Background: Levels of protein structure. Structure determination methods. The protein folding problem. Algorithms: Hidden Markov Models, Homology Modelling

7. **Classical bioinformatic methods.**

Sequence comparison, alignment, and sequence similarity search algorithms. Dynamic programming, substitution matrix, gap penalty, local/global alignment, BLAST, e-value Background: Homology, orthology, paralogy, gene evolution, sequence evolution.

8. **Phylogenetics.**

The conceptual basis of phylogenetics, different phylogenetic methods. Reconstructing evolutionary history from molecular sequences. Background: How genes and genomes evolve? Algorithms: Maximum Parsimony, distance based methods, Neighbor-joining,

9. **Network biology.**

Algorithms to analyse networks and graphs. Basic characteristics of graphs, protein interaction networks, Background: Protein protein interaction networks: physical and functional interactions, pathways, Gene ontology

## **10. Metagenomics.**

Bioinformatic methods of metagenomics to study bacterial species diversity. The concepts of amplicon and shot-gun sequencing based metagenomic approaches. The concepts of Operational Taxonomic Units (OTUs) and amplicon sequence variants (ASVs). Background: Bacterial species diversity.

## **11. Unsupervised learning algorithms**

Algorithms: Hierarchical clustering, K-means clustering, Lloyd algorithm, PCA, applications in biology

## **12. Supervised learning algorithms**

Algorithms: Linear and logistic regression, shallow and deep neural networks, loss and activation functions, forward and back propagation, applications in biology