CB207V Computational Analyses of mRNA-seq, ChIP-seq and ATAC-seq 7.5 credits

Analys av mRNA-seq, ChIP-seq och ATAC-seq data

This is a translation of the Swedish, legally binding, course syllabus.

Establishment

Course syllabus for CB207V valid from Spring 2024

Grading scale

P, F

Education cycle

Second cycle

Main field of study

Biotechnology

Specific prerequisites

BSc or related competence in cell biology, molecular biology, genomics, biotechnology or bio(medical) sciences. Adequate skills in English, corresponding to English B, are required.

Language of instruction

The language of instruction is specified in the course offering information in the course catalogue.

Intended learning outcomes

After completion of the course, the students shall have knowledge to:

- Describe genomics, epigenomics and how distinct proteins and RNAs coordinate chromatin state and transcription.
- Describe readers, writers, transcription factors, co-factors, general transcription factors, RNA Polymerases, DNA Polymerases, negative elongation factors, positive elongation factors, RNA binding proteins and Topoisomerases.
- Give characteristics of functional genomic regions, including enhancers, promoters, insulators, gene bodies and termination windows.
- Explain how histone modifications and chromatin compartments can influence chromatin accessibility, DNA repair, DNA synthesis and transcription.
- Obtain ChIP-seq, mRNA-seq and ATAC-seq datasets from public repositories.
- Understand the process of sequencing techniques from biochemistry to dataanalyses.
- Map protein binding sites and open chromatin regions.
- Quantify mRNA levels and perform differential gene expression analyses
- Perform gene ontology analyses and investigate transcriptional responses.
- Understand the concept of large-scale data and how machine learning can be integrated into genomics.

Course contents

- **Introduction to genomics** - Genomics, epigenomics and analyses of chromatin.
- **Introduction to transcriptomics** - Capturing the distinct RNAs in the cell.
- **The genome regulators** - Readers, writers, insulators, TFs, Pols, helicases.
- **Biochemical foundations** - from molecular capture to libraries and reads.
- **Sequencing** - Adapters, UMIs, clusters, sequencing-by-synthesis.
- **Mapping the reads** - from fastq to bigwigs and peak files.
- **Computational foundations** - Basics in shell scripting and R.
- **ChIP-seq and ATAC-seq data** - Protein-DNA interactions and accessible regions.
- **MEME-ChIP** - DNA elements directing protein binding.
- **mRNA-seq** - Quantifying gene expression and tracking changes in mRNA levels.
- **DESeq2 and Gene Ontology** - Differential expression and transcriptional responses.
- **Integrating genomic datasets** - Integrating and visualizing large-scale data.
- **Open access for data and code** - GEO, SRA, ENCODE, GitHub.

- **Systems on systems** - Epigenetic pattern recognition, identification of functional genomic regions. Functional analyses of DNA-binding proteins.
- **Challenges and future directions** - Normalization, resolution, purity. Machine learning in genomics.

Examination

- DEL1 - 80% presence and completed exercises, 4.0 credits, grading scale: P, F
- SEM1 - Group work and seminar presentation, 1.0 credits, grading scale: P, F
- PRO1 - Group project completed, 2.5 credits, grading scale: P, F

Based on recommendation from KTH's coordinator for disabilities, the examiner will decide how to adapt an examination for students with documented disability.

The examiner may apply another examination format when re-examining individual students.

If the course is discontinued, students may request to be examined during the following two academic years.

Ethical approach

- All members of a group are responsible for the group's work.
- In any assessment, every student shall honestly disclose any help received and sources used.
- In an oral assessment, every student shall be able to present and answer questions about the entire assignment and solution.