



## SWEDISH UNIVERSITY OF AGRICULTURAL SCIENCES

Department of Plant Biology

### Syllabus

---

#### **Practical skills in sequence analysis**

2,0 Credits

Code: PNS0191

Finalized by: Research school Organism Biology, Mattias Thelander chair, 2023-10-30

Valid from: First half-year 2024 (2024-01-01)

Level within study regulation: Third cycle

---

#### **Subject**

Biology

#### **Entry requirements**

The course is primarily intended for PhD students within the SLU Organism Biology graduate school but will be open for other PhD students and researchers if there is space.

#### **Examination formats**

Students must attend at least 80 % (i.e. 4 out of 5) of group sessions and complete the final assignments, which consist of an oral presentation during the final day of the course as well as a written report to be handed in after the final group teaching session.

## Objectives

Today basic knowledge of sequence analysis methods such as sequence homology searches and phylogenetic analysis is an essential skill set to complement wet lab experimental work within the life sciences. Fairly simple analyses of biological sequences can reveal a lot about gene function and lead to the formulation of new and unexpected hypotheses for experimental verification. This course will demonstrate to the students how basic methods in sequence analysis can improve their understanding of one or more target genes or proteins that are relevant to their PhD projects.

The course is aimed at PhD students whose main research focus is not bioinformatics but whose PhD projects involve biological sequences (e.g. biochemistry, genetics, molecular biology, cell biology, structural biology, microbiology). The course strives to improve the bioinformatic literacy of students, which will enable them to get to know their favorite gene(s) or protein(s) better, provide them with the necessary skills to evaluate these types of analyses in the published scientific literature as well as allow the students to generate novel hypotheses for future wet lab experiments.

Upon completion of the course, students will be able to:

- Navigate the main NCBI sequence databases as well as selected conserved domain databases.
- Accurately define the concept of sequence homology.
- Execute and evaluate customized BLAST sequence similarity searches.
- Analyze non-annotated genomic sequences with regards to gene content.
- Perform, evaluate and visualize multiple sequence alignments.
- Perform, evaluate and visualize phylogenetic analyses.
- Identify conserved motives in non-coding DNA sequences and visualize these as sequence logos.
- Describe and discuss sequence analysis methods and results in sufficient detail for inclusion in scientific publications.
- Apply basic sequence analysis for generation of novel hypotheses.

## Content

The course will cover the fundamentals of sequence analysis including sequence databases, sequence homology searches, multiple sequence alignments, sequence motif searching and phylogenetic analysis. The course will also cover how to describe and visualize results of sequence analysis in an informative manner. Note that this course will not cover any programming, genome assembly or analysis of genome-scale datasets (e.g. transcriptomics, proteomics).

The course will consist of five scheduled afternoon group sessions with time for independent study in between. Course participants will view prerecorded lectures and perform their own analyses between sessions and then discuss their progress with the course leader and other course participants at each session. During the final session the course participants will summarize their work.

Course participants will select at least one gene or protein (but no more than two) that is relevant to their own PhD project, which will then be the subject of the individual student's course work. Any students unsure about which gene or protein to select for the course can consult the course leader prior to the start of the course.

The course will mainly rely on web-based tools although the section of visualization of bioinformatic data will involve some stand-alone software. The course will cover short-term license fees for Adobe Illustrator for those course participants that do not already possess this program.

### **Other information**

The course will be conducted remotely via Zoom. The course is organized by Tomas Linder (Department of Molecular Sciences, NJ faculty) on behalf of the SLU Organism Biology graduate school. The course can accommodate a maximum of 20 students.

### **Literature**

There is no set textbook for the course although recommendations will be provided to the students, should they wish to further explore the underlying theory of bioinformatic methods demonstrated during the course. Relevant articles will be handed out during the group teaching sessions.

### **Contact: Name**

Tomas Linder

### **Contact: Email**

Tomas.Linder@slu.se