



CORE FACILITIES

SC00025 Unix with applications to Next Generation Sequencing (NGS) data, 1.5 credits

Unix with applications to Next Generation Sequencing (NGS) data, 1,5 högskolepoäng

Third-cycle level / Forskarnivå

Confirmation

This syllabus was confirmed by the Council for PhD Education at Sahlgrenska Academy on 2019-01-17, and was last revised on 2020-03-23. The revised course syllabus is valid from Autumn semester 2020.

Responsible Department

Core Facilities, Sahlgrenska Academy

Entry requirements

The course is open for PhD students who are in need to process and manage biological data.

To be able to follow the course one should have a background in genetics, cell biology, biomedicine, biochemistry, bioinformatics or similar.

The course is an elective course within the third cycle at Sahlgrenska Academy.

Learning outcomes

After completing the course the student is expected to be able to:

Knowledge and understanding

- Describe the UNIX system
- Describe the most common commands used for data manipulation
- Explain suitable algorithms for data analysis
- Identify repetitive tasks that can be automated
- Define a plausible workflow for data

Skills and abilities

- Navigate through the file structure

- Extract and process information with filters and pipes
- Develop custom workflows for data managing
- Develop shell script to simplify and automate frequent tasks
- Use biological tools

Judgement and approach

- Design and establish custom approaches to analyze and manage biological data
- Interpret others' scripts
- Describe the techniques introduced and be able to pick out the methods suitable for their own data.

Course content

This is a hands on course to learn the basic unix commands. The course consists of lectures introducing the different topics and practical sessions to practice the language. The exercises have applications in next generation sequencing data analysis. The topics covered are:

- The shell
- File system and permissions
- Text editors
- Handling files
- Regular expressions
- Piping
- For loops
- File compression
- Program installation
- Bash scripting

Types of instruction

Lectures, on-site exercises and practical homework.

Language of instruction

The course is given in English.

Grades

The grade Pass (G) or Fail (U) is given in this course.

To receive a passing grade, the PhD student is required to complete all practical sessions and demonstrate that the learning objectives have been reached.

Types of assessment

Assessment will be done through the practical sessions; these are designed to test the understanding of the different applications, so completion of all of them is mandatory. Active participation during group sessions and attendance for at least 80% is also mandatory.

A doctoral student who has failed a test twice has the right to change examiners, if it is possible. A written application should be sent to the Institute.

Course evaluation

The course evaluation will be done through a written questionnaire, available at the university's learning platform, where students will be asked to describe their opinions on the various stages of the course for future development. This information will be compiled and shared with students who participated in the course. Improvements are shared with students participating in the course to come.

Other information

This course can be the basis for the courses:

- Python programming for life science researchers
- Analysis of next generation sequencing data

Computer access with administration rights as well as internet access is required since all course communication and relevant documents such as lectures, exercises and literature will be posted at the virtual learning platform.

The syllabus was confirmed by the Council for PhD Studies on 13-09-2016 and was last revised on 26-02-2019 to be valid from autumn semester 2019 (reg.nr.: GU 2019/697). It was entered into FUBAS on 26-03-2019.