

## Genome Informatics

---

**Subject area:** Biology/Biomedical Engineering

**University:** TalTech  
**Level:** BA3, BA4, MA all years, PhD  
**Teaching mode:** completely online, not time-specific  
**Instructor(s):** Professor Olli-Pekka Smolander

### Short description

Genome informatics is an M.Sc. level basic course with an emphasis on developing hands-on knowledge in executing and critically assessing genome related bioinformatic studies. Different from standard courses in bioinformatics and genetics, this course is more focused on practical implementation of research assignments in the field of genome biology studies.

### Full description

#### 1. OVERVIEW

Genome informatics is a basic course with an emphasis on developing hands-on knowledge in executing and critically assessing genome related bioinformatic studies. Different from standard courses in bioinformatics and genetics, this course is more focused on practical implementation of research assignments in the field of genome biology studies.

The objective of the course is to advance understanding of various principles and practices of genome informatics research, and to improve the students' readiness to execute a research project in their own discipline. The course covers the fundamentals of the sequencing approaches, sequencing data quality control and processing, sequence data alignment, genome assembly and annotation, gene expression analysis and analytical and statistical evaluation of the results, all with a focus on issues specific to genome informatics and emphasis on practical experience with computing cluster and R environment and other relevant applications used in the field.

Topics covered in the course include the fundamentals of genome-based research design, such as genome assembly and differential gene expression; preparation of data, such as quality control and filtering, downstream analyses such as gene ontology and pathway enrichment, and evaluation of the results using statistical methods, publicly available tools and databases. Practical examples and exercises are embedded in the course structure.

#### 2. TO WHOM

The course offers basic skills for applying bioinformatics methods in data driven biosciences. Therefore it's especially designed for bioscience and life science students, but it also offers it-students a way to enter the bioinformatics application area.

The course is designed for students who want to gain practical knowledge about genome informatics. This involves carrying out multiple computer assignments, taking part in on-site/online class

discussions, and working independently and in a group on the provided case-study assignment. The course offers basic skills for applying bioinformatics methods in data driven biosciences. Therefore it's especially designed for bioscience and life science students, but it also offers it-students a way to enter the bioinformatics application area.

The main learning outcomes are to gain an understanding of the uses and limitations of common tools for analysing bioinformatic data and to develop competence in collecting, evaluating, and analysing sequencing-based data. Students will improve their capacity to interpret and critically assess methodological execution of empirical genome related bioinformatic studies. On completion of this course, students are prepared to utilise obtained methodological skills in their master's theses, doctoral dissertations, and/or other research projects.

### 3. PASSING THE CLASS

There are 5 assignments and 2 project works on which your final grade will be based on. In addition there are courseworks that are graded passed / not passed.

A successful study strategy involves keeping up with the readings as we go. Your schedule should allow for this. The workload is calculated to be a little less than five hours per class, so that means a student should reserve about 10 hours of course working time every week.

The course reading package consists of scientific articles and other relevant material. Unavoidably some parts are demanding, technically as well as language wise. When reading the articles, please pay particular attention to their Research questions, Methods, and Result sections.

The course assignments require analysis skills and judgement which reflect the knowledge gained in the class and through independent work on the assigned reading. The assignments may not have a single correct answer and are graded for quality, as reflected in the rigorous execution, strong argumentation of the choices you make, and clarity of reporting.

### 4. A NOTE ON COURSE IMPLEMENTATION

The course is fully available online. There is also a parallel implementation available for TalTech students which also includes contact teaching. Both implementations are identical in content and learning objectives. There may be interaction between students from the two implementations through exercises. If needed (e.g. due the Covid-19 situation) the whole class will be moved online. Links to all pre-recorded lectures will be provided to you through Moodle.

## Learning outcomes

### 5. COURSE GOALS AND LEARNING OUTCOMES

This course is designed to introduce students in various areas of genome informatics to the principles and practice of bioinformatics research. The aim is to deepen the insights and knowledge of each learner to support their individual future career path. The students are encouraged to reflect on the topics of bioinformatics through their specification area but also to understand how other learners' see the same phenomena and what's important for them. Widening the understanding of different methods, theories and applications used in the field shall strengthen the needed cooperation skills in diverse teams.

The exercises are designed to support individual learning but at the same time to help growing students' professional identities and various academic skills such as reflective and critical thinking, collaboration & communication abilities, reporting and presentation techniques, scientific language as well as reading and analysing academic papers.

Course will provide students with the understanding of modern sequencing methods, their specific properties and general applications from the bioinformatics perspective. The course will also introduce

students to genome wide bioinformatics approaches and methods so that the student will be able select correct methods to address biological research questions.

After completing the course, the student is able to:

explain how the data is produced by modern sequencing methods, what are the basic applications regarding genome informatics and how the data is presented, stored and processed;

use sequencing data to address simple research questions using concepts such as de novo genomics, gene prediction and annotation, and gene expression analysis;

review and select suitable tools and approaches for above mentioned research questions.

## General information

**Contact hours per week:** 0-4

**Total workload:** 156 (in student hours for the whole course)

**ECTS credits:** 6

**Language:** English

**Course start date:** 29 August 2022

**Course end date:** 31 December 2022

**Add. info about start date:** flexible starting possibility until 1.10., please ask more info

**Weekly teaching day/time:** Course is implemented in hybrid learning and can be taken fully online.

**Time zone:** CET +1 (Estonia, Israel)

**Further information:** Provider: Department of chemistry and biotechnology, TalTech, Estonia

**Prerequisites:** IT and computer skills: access to a up to date computer, comfortable with using computer as a tool, ability to install programmes

Basic knowledge of molecular biology and nucleic acids

**Activities and methods:** Pre-recorded lectures, exercises and self-study

**Presence on campus:**

## Final examination

**Form:** no examination

**Date:**

**Location/format:**

**Re-sit possibility:**

**Transcript available:**

**Add. info/requirements:** There is no final examination. There are 5 assignments and 2 project works on which your final grade will be based on. In addition there are compulsory courseworks that are graded passed / not passed.

## Registration

To register for this course, follow the registration requirements of your **home university** as specified here: [www.euroteq.eu/courses-registration](http://www.euroteq.eu/courses-registration).

## Administration

**Number of places:** 100  
**Minimum participants:** This is a mandatory TalTech course and will be organised in any matter  
**Internal course code:** LKG0050  
**Contact:** [olli-pekka.smolander@taltech.ee](mailto:olli-pekka.smolander@taltech.ee)

---

*This course is part of the EuroTeQ Engineering University joint course catalogue 2022/2023. This is a collaborative activity of the partner universities DTU, L'X, TU/e, TalTech, CTU, TUM as well as Technion. Students from these universities can participate in the offered courses. It is the responsibility of the student to check if you fulfil the requirements to participate in a specific course. Students are also advised to check with their home institution how to get recognition of the ECTS credits gained in courses of the EuroTeQ course catalogue. For further information about EuroTeQ Engineering University, visit [www.euroteq.eu](http://www.euroteq.eu) or get in touch with the above-mentioned point of contact.*