COURSE DESCRIPTION CARD - SYLLABUS

Course name
High-throughput Data Analysis

Course

Field of study
Bioinformatics

Year/Semester
2/3

Area of study (specialization)

Profile of study
general academic

Level of study
Second-cycle studies

Course offered in
Polish

Form of study
full-time

Requirements
compulsory

Number of hours

Lecture	Laboratory classes	Other (e.g. online)
15	15	
Tutorials	Projects/seminars	

Number of credit points

3

Lecturers

Responsible for the course/lecturer:
dr hab. inż. Aleksandra Świercz
email: aleksandra.swiercz@cs.put.poznan.pl
tel: 616653030
Wydział Informatyki i Telekomunikacji
ul. Piotrowo 3 60-965 Poznań

Responsible for the course/lecturer:

Prerequisites

A person undertaking a second degree in Bioinformatics should have achieved the results of education from the 1st degree of this field of study, as defined in the Resolution of the Senate PUT – these effects are presented on the website of the Faculty of www.cat.put.poznan.pl. The student starting this module should have a basic knowledge of statistics, algorithms and genomics. The student must present attitudes such as honesty, responsibility, perseverance, cognitive curiosity, creativity, personal culture, respect for other people.

Course objective

1. Providing students with basic knowledge of new high-pass sequencing technologies.
2. Familiarize students with various problems and issues that can be solved with next-generation sequencing. Familiarise students with alternative splicing, determining differences between genomes of individuals of the same species, de novo assemby and resequencing.

3. Develop students' ability to apply the knowledge they have learned to solve the above-mentioned problems.

Course-related learning outcomes

Knowledge

1. Is familiar with the methods and tools used in the process of solving complex bioinformatics tasks, mainly of an engineering nature
2. Knows and can use specialized IT and bioinformatics tools
3. Has knowledge of bioinformatic analysis on the genome scale based on a statistical background
4. Is familiar with basic sequencing technologies, has knowledge and is able to use basic genomic sequence analysis methods and tools in the context of these issues
5. Knows the development trends of bioinformatics, understands and can draw conclusions from scientific publications

Skills

1. Is able to use the methods and IT tools learned to solve biological problems, is able to assess their usefulness and, if necessary, propose an alternative solution
2. Formulates and tests hypotheses related to the bioinformatics issues discussed, e.g. studying differences in gene expression levels, studying alternative splicing, etc.

Social competences

1. Understands the need to systematically seek new solutions, to familiarity with scientific journals, including in english, in order to improve bioinformatic knowledge
2. Systematically update its biology and computer science knowledge and see opportunities for its practical application

Methods for verifying learning outcomes and assessment criteria

Learning outcomes presented above are verified as follows:

Formative assesment:

(a) Lectures, verification of the intended learning outcomes shall be carried out:
• based on your activity in discussing the material in question;

(b) laboratories: verification of the intended learning outcomes shall be carried out:
• on the basis of the current progress of the tasks;
• assessment for activity in class or lack of preparation for classes
• evaluation of reports prepared partly during the classes and partly after their completion
• assessment of the student's laboratory exercises

Summary assessment:

(a) Lectures, verification of the intended learning outcomes shall be carried out by:

• Written colloquium consisting of 6 questions / problem tasks - maksimum nuber of points: 5. He/she must earn at least 2.7 points to get a positive mark.

b) laboratories: verification of the intended learning outcomes shall be carried out:

• assessment of knowledge and skills related to the content transmitted in the exercises through the final colloquium
• a summary of the assessments issued during the semester in weighted average form. The assessments obtained from the exercises carried out in the classes and from the reports are taken to the average with a weight of 1, while the assessment from the colloquium with a weight of 2. In addition, activity in classes can raise the final rating by 0.5 (provided that the rating is not exceeded 5.0).
Activity during classes is rewarded with additional points, in particular for:
• an overview of additional aspects of the issue,
• the effectiveness of applying the acquired knowledge when solving a given problem,
• comments leading to the improvement of teaching materials or the teaching process

Programme content

The lecture programme covers the following topics: Familiarize students with new next-generation sequencing technologies. Different ways to generate libraries for the Illumina sequencer. Sequencing single reads and paired end reads. DNA sequencing; de novo sequencing approaches and methods; resequencing, i.e. mapping readings to the reference genome; Smith-Waterman algorithm; Burrows-Wheeler transformation. Graphs in the context of assemble algorithms; graphs of assumptions and DNA graphs. RNA sequencing; search for new splicing sites for transcriptomes, study of gene expression levels for several samples, analysis of short RNA (e.g. miRNA, piRNA). Different approaches and algorithms to solve these problems. Ways to analyse data for genomic differences (CNV, SNP).

Laboratory exercises are conducted in the form of seven/eight two-hour classes held in a computer laboratory. The first classes are designed to familiarize students with the rules of laboratory use and the reckoning of exercises. The exercises are carried out independently by each student. The laboratory program covers the following topics: familiarize students with the available tools for analyzing data from next-generation sequencers, browsing databases for experimental data. Use of available tools for problems discussed in lectures and laboratory activities: pre-filtering data by quality, finding and cutting off adapters in sequencer readings, DNA and RNA mapping to reference genome, gene expression level testing, alternative splicing gene search, short RNA analysis, de novo transcriptome assembly, visualization of mapping results in IGV

Teaching methods

A lecture illustrated with a multimedia presentation containing the programme content in question, enriched with examples;

Laboratories: practical exercises in data analysis, presentations, discussion, group work

Bibliography

Basic

1. N. Rodriguez-Ezpelta, M. Hackenberg, A.M. Aransay eds. „Bioinformatics for high throughput sequencing", Springer, 2012

2. TA Brown, „Genomy", PWN

Additional

1. M.Zhao, Q. Wang, Q. Wang, P. Jia, Z. Zhao "Computational tools for copy number variation (CNV) detection using next-generation sequencing data: features and perspectives", BMC Bioinformatics, 2013, 14:S1

2. Briefings in Bioinformatics vol. 11 issue 5, 2010, Special Issue: Second generation sequencing.

Breakdown of average student's workload

	Hours	ECTS
Total workload	75	3,0
Classes requiring direct contact with the teacher	30	1,5
Student's own work (literature studies, preparation for laboratory classes/tutorials, preparation for tests/exam, project preparation) [1]	45	1,5

[1] delete or add other activities as appropriate