COURSE DESCRIPTION CARD - SYLLABUS

Course name
Genomic Data Analysis in Bioconductor [S2Bioinf2>BIOCON]

Course

Field of study Bioinformatics	Year/Semester 2/3
Area of study (specialization) –	Profile of study general academic
Level of study second-cycle	Course offered in Polish
Form of study full-time	Requirements elective

Number of hours

Lecture 30	Laboratory classes 30	Other (e.g. online) 0
Tutorials 0	Projects/seminars 0	

Number of credit points

4,00

Coordinators	Lecturers
dr hab. inż. Aleksandra Świercz aleksandra.swiercz@put.poznan.pl	

Prerequisites

A person undertaking a degree in the second degree of 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 http://www.cat.put.poznan.pl/ The student starting this module should have basic knowledge of molecular biology, genomics and transcriptomics, as well as programming, statistics and bioinformatic analysis of biological sequences. It should have the ability to solve basic biological and bioinformatics problems, to test and correct errors in the programmes he/she implements, and to obtain information from the indicated sources and use databases. In addition, in terms of social competence, the student must present such attitudes as honesty, responsibility, perseverance, cognitive curiosity, creativity, personal culture, respect for other people.

Course objective

1. Providing students with basic knowledge of the Bioconductor environment, in particular: data structure and types, available packages, data analysis and visualisation capabilities 2. Familiarize students with the current set of Bioconductor packages, their purpose and examples of use, with a focus on discussing analytical approaches in genomics 3. Develop students' skills to solve genomic data analysis problems in the Bioconductor environment, select and apply Bioconductor tools and packages to analyze selected types of biological data, and apply graphical data visualization to support analysis and interpretation of results.

Course-related learning outcomes

Knowledge:
1. He/she has structured, theoretically built-up knowledge of the Bioconductor environment and the use of selected packages for genomic and population-level analysis
2. Is aware of the applications of Bioconductor packages for bioinformatic analysis and integration of different types of genomic data and visualization of biological data
3. Is familiar with the use of Bioconductor packages for the analysis of genetic data in medicine
4. Is familiar with the methods, techniques and tools used in the process of solving complex bioinformatics tasks, mainly of an engineering nature
5. Is familiar with the principles of planning bioinformatics research

Skills:
1. He/she can obtain the source data described in scientific publications (e.g. results of genomic analyses) and in an integrated manner to obtain and interpret information on different sources (e.g. functional annotations, relationship paths)
2. Can identify the appropriate Bioconductor package for typical genomic data analyses and apply them in practice
3. Can use advanced IT techniques and tools to solve biological problems and assess their usefulness
4. Uses the Bioconductor environment to process genomic data and statistically analyze the results
5. Prepares in polish and english a presentation of the results of research work, discusses the results obtained in the context of existing scientific knowledge
6. Formulates and tests hypotheses related to bioinformatics problems.

Social competences:
1. Develops an attitude of openness to cooperation in the academic and industrial environment, working to respect the principles of ethics and good practices in the analysis of biological and biomedical data.
2. Is aware of the importance of protecting sensitive data and complies with the principles related to their processing, in particular in the context of genomic data, which may have a direct impact on the health and privacy of patients.

Methods for verifying learning outcomes and assessment criteria

Learning outcomes presented above are verified as follows:

Formative assessment
(a) Lectures, verification of the intended learning outcomes shall be based on your activity in discussing the material in question;
(b)Laboratories, verification of the intended learning outcomes shall be based on an assessment of the current progress of the tasks;
Summary assessment
(a) in the scope of lectures, verification of assumed learning outcomes is carried out by: written colloquium consisting of 5 questions/problem tasks - each task scored 0-4 points (tasks may consist of several sub-points - then a partial score is assigned for each sub-point). To pass, you must score at least 11 points.
b) in the scope of laboratories, verification of the assumed learning outcomes is carried out by:
the final grade is the average of the grade for the development and presentation of the selected package/group of packages and the grades for the performance of individual practical exercises in the scope of data analysis in Bioconductor during the semester. A maximum of 10 points can be obtained for the presentation of the selected package. A maximum of 5 points can be obtained for each exercise/report (8-10 exercises). A pass can be obtained after obtaining more than 50% of the total

number of points, provided that all required reports are submitted.

Programme content

The subject is devoted to the theoretical and practical presentation of the Bioconductor packages, dedicated to the analysis and visualization of biological data, available free of charge under an open GPL. The lectures will discuss the following topics: introduction to Bioconductor / consolidating knowledge about variable types and operations on them; loading and saving data from/ to various external formats and programs; data processing and reporting; integration with other programmes, selected statistical procedures in R; review and install available editors; an overview of useful web resources; discuss the current resource of available packages and their applications, learn about and work with selected libraries for genomic data analysis and visualization. The issues will be illustrated with examples from the scientific literature. Laboratories rely on the analysis of scientific work and the analysis of biological and biomedical data, with a focus on understanding the possibilities of the packages discussed in the lecture section and their applications. The scope of these packages is updated to take account of the latest developments and trends in genomics. In addition, each student selects, develops and presents a package / group of packages and its capabilities, according to the scenario agreed with the subject

Course topics

Familiarization with the R environment and with access to and the diverse topics of Bioconductor packages.
Preparation of graphs summarizing bioinformatic analyses, ready for publication using the ggplot2 package.
Discussion of selected Bioconductor packages.
Preparation of presentations of selected packages, according to students' interests.

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:
Biecek P. ,,Przetwarzanie danych w programie R''; epub; http://biecek.pl/R/
Biecek P. ,,Wizualizacja i modelowanie''; epub; http://biecek.pl/R/

Additional:
1.Biecek P. Przewodnik po pakiecie R. Oficyna Wydawnicza GiS 2017 (wydanie 4), ISBN 978-83-89020-79-6
Górecki T. Podstawy statystyki z przykładami w R. Wydawnictwo BTC 2011, ISBN 978-83-60233-69-6.

Breakdown of average student's workload

	Hours	ECTS
Total workload	105	4,00
Classes requiring direct contact with the teacher	60	2,50
Student's own work (literature studies, preparation for laboratory classes/ tutorials, preparation for tests/exam, project preparation)	45	1,50