

MINISTRY OF EDUCATION AND SCIENCE OF UKRAINE
Odesa I. I. Mechnikov National University
Department of Microbiology, Virology and Biotechnology

Syllabus of the course

METAGENOMIC ANALYSIS OF MICROORGANISMS

Course volume:	4.0 credits, 120 hours
Semester, Year	Full-time – 1 semester, 1 year of master's degree
Days, Time, Place	according to the schedule
Lecturer	Vasilyeva Nataliia Job title: Associate Professor Scientific degree:: Doctor of Philosophy Academic title:: Associate Professor of the Department of Microbiology, Virology and Biotechnology
Contact phone	+380634915899
E-mail:	takamic@onu.edu.ua
Workplace	Faculty of Biology, Department of Microbiology, Virology and Biotechnology 2, Champagne Lane, Room 31
Consultations	Face-to-face consultations: Monday 16:00-17:00 Online consultations: on Zoom, Viber platforms

COMMUNICATION

Communication with students will be carried out via email, social networks (Viber group), e-mail and Zoom platforms, as well as on the "Face to face" principle during face-to-face consultations in the classroom.

COURSE SUMMARY

The course will provide an overview of the main methods, online tools, and software used in metagenomic analysis.

Subject of study of the discipline “METAGENOMIC ANALYSIS OF MICROORGANISMS”. The subject of study of the discipline is a complex of theoretical knowledge, modern methodologies and practical skills necessary for the analysis of metagenomes of microbial communities without the need for their laboratory cultivation

Course prerequisites: required course knowledge "Methods of Biotechnological Research", "Bioinformatics", "Fundamentals of Information Technologies in Biotechnology", "Microbiology", "Molecular Biology", "Genetics"»

Course post-course requirements. The theoretical knowledge obtained, as well as practical skills and abilities, can be used in the preparation of a dissertation for the degree of Doctor of Philosophy and in future professional activities.

The **purpose** of the course is to form in the students a system of knowledge about the diversity of the microbial population of various ecological groups, as well as to familiarize them with the practical aspects of metagenomics. When conducting the course, students, using modern information and computing technologies, analyze the quality of sequences, prepare them for further analysis, consider the issues of sequencing and assembly of metagenomes, their annotation and application.

The gradual achievement of the set goal forms professional competencies in the students of higher education in specialty 162 "Biotechnology and Bioengineering"

The objectives of the discipline:

- 1) to familiarize students with the tasks that the need to study complex microbial communities poses to physicians, biologists, programmers and mathematicians;
- 2) to familiarize students with software products and analytical platforms created for working with metagenomic data;
- 3) to familiarize students with the skills of working with metagenomic data, to teach them to properly plan an experiment, to assess the complexity of the task and the resources necessary for its solution (laboratory and computer), to assess the quality of the generated data from the point of view of the task set;
- 4) to teach students to correctly choose software products to solve the tasks that need to be solved.

As a **result** of studying the academic discipline, a higher education applicant should **know**:

- theory and principles of metagenomics of microorganisms;
- principles of sampling, metadata collection and their analysis;
- methods of metagenomics;
- analyze the obtained experimental data;

be able to:

- independently conduct experiments according to a given scheme;

- analyze the obtained experimental data;
- independently acquire new knowledge in this area and apply the acquired knowledge in practice and when studying other disciplines

COURSE DESCRIPTION

The course will be taught in the form of lectures (20 hours), practical classes (10 hours), and organization of independent work of students (90 hours).

In the process of studying the academic discipline, the following teaching methods are used:

- verbal (narration, explanation, lecture, instruction);
- visual (demonstration, presentation);
- practical (individual performance of practical work, the method of traditional analysis of specific situations).

Content module 1. General concepts and postulates of metagenomics

Topic 1. Metagenomics

What is metagenomics. Description of metagenomic communities of various natures, where microorganisms live. Methods of their analysis. Basic approaches to metagenomic analysis and features of analytical approaches used for the analysis of natural microbial communities. The concept of metagenomic analysis of 16S amplicons. The concept of complete metagenomic analysis. Use of metagenomic analysis results in biotechnology.

Topic 2. Sequencing technologies. Data formats

Sequencing technologies. Use of different platforms for sequencing. Data formats after sequencing. Fasta, fstaq, stff and others. Programs for reading formats.

Content module 2. Using software products to solve metagenomic analysis problems.

Topic 3. Getting to know the Galaxy server. General information about the server capabilities. Data loading. Creating a workflow for data processing.

Topic 4. Data preprocessing.

Data preprocessing. Quality control. FastQC program. Program features. Program arguments and attributes. Results.

NGS data correction. Adapter removal. Cutadapt program. Program features. Program arguments and attributes. Results.

Trimming, sequencing error correction, normalization. Trimmomatic program. Program features. Program arguments and attributes. Results.

Updated programs for sequence quality analysis and processing. Fastp program. Program features. Program arguments and attributes. Results.

Topic 5. Analysis of metagenomic data on the galaxy server (mothur algorithm)

Data loading. Formation of a common data set from individual sequences. Statistical analysis of amplicons. Sequence alignment. Data cleaning - formation of a set of unique sequences. Removal of chimeric sequences. Taxonomic classification. Clustering of sequences into OTUs. The concept of "OTU". Diversity analysis. Alpha-diversity indices, beta-diversity distance matrices. Visualization using Krona

Educational and methodological support

Methodological support: working program of the academic discipline; syllabus; multimedia support of lecture and seminar materials.

Recommended literature

Basic

1. Achudhan AB, Kannan P, Gupta A, Saleena LM. A Review of Web-Based Metagenomics Platforms for Analysing Next-Generation Sequence Data. *Biochem Genet.* 2024 Apr;62(2):621-632. doi: 10.1007/s10528-023-10467-w.
2. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014 Aug 1;30(15):2114-20. doi: 10.1093/bioinformatics/btu170.
3. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics.* 2018 Sep 1;34(17):i884-i890. doi: 10.1093/bioinformatics/bty560. PMID: 30423086
4. Garfias-Gallegos D, Ziri6n-Mart6nez C, Bustos-D6az ED, Arellano-Fern6ndez TV, Lovaco-Flores JA, Espinosa-Jaime A, Avelar-Rivas JA, S6lem-M6jica N. Metagenomics Bioinformatic Pipeline. *Methods Mol Biol.* 2022;2512:153-179. doi: 10.1007/978-1-0716-2429-6_10.
5. Liu YX, Qin Y, Chen T, Lu M, Qian X, Guo X, Bai Y. A practical guide to amplicon and metagenomic analysis of microbiome data. *Protein Cell.* 2021 May;12(5):315-330. doi: 10.1007/s13238-020-00724-8.
6. L6pez-Garc6a A, Pineda-Quiroga C, Atxaerandio R, P6rez A, Hern6ndez I, Garc6a-Rodr6guez A, Gonz6lez-Recio O. Comparison of Mothur and QIIME for the Analysis of Rumen Microbiota Composition Based on 16S rRNA Amplicon Sequences. *Front Microbiol.* 2018 Dec 13;9:3010. doi: 10.3389/fmicb.2018.03010.
7. Nooij S, Schmitz D, Vennema H, Kroneman A, Koopmans MPG. Overview of Virus Metagenomic Classification Methods and Their Biological Applications. *Front Microbiol.* 2018 Apr 23;9:749. doi: 10.3389/fmicb.2018.00749.
8. Odom AR, Faits T, Castro-Nallar E, Crandall KA, Johnson WE. Metagenomic profiling pipelines improve taxonomic classification for 16S amplicon sequencing data. *Sci Rep.* 2023 Aug 26;13(1):13957. doi: 10.1038/s41598-023-40799-x.

9. Schmieder, R. and Edwards, R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics*, 27(6):863–864, 2011
10. Sudarikov K, Tyakht A, Alexeev D. Methods for The Metagenomic Data Visualization and Analysis. *Curr Issues Mol Biol*. 2017;24:37-58. doi: 10.21775/cimb.024.037.
11. Thang MWC, Chua XY, Price G, Gorse D, Field MA. MetaDEGalaxy: Galaxy workflow for differential abundance analysis of 16s metagenomic data. *F1000Res*. 2019 May 23;8:726. doi: 10.12688/f1000research.18866.2.
12. Wang D. Amplicon Sequencing Pipelines in Metagenomics. *Methods Mol Biol*. 2023;2649:69-83. doi: 10.1007/978-1-0716-3072-3_4.
13. Wingett SW, Andrews S. FastQ Screen: A tool for multi-genome mapping and quality control. *F1000Res*. 2018 Aug 24;7:1338. doi: 10.12688/f1000research.15931.2.

Electronic information resources

1. Resource with a description and overview of programs for advanced nutrition bioinformatics and metagenomics <https://omictools.com/>
2. Galaxy community hub server <https://galaxyproject.org/>
3. Galaxy community tool <https://usegalaxy.org/>

Forms of control and assessment methods

Current control of the knowledge of applicants is carried out by evaluating tasks during classroom lessons, answers to questions during practical lessons.

Periodic control - conducting module testing after studying the material of each content module.

Final control - exam (the exam is conducted orally).

The applicant's level of knowledge is assessed on a 100-point scale. The overall final score is determined as the sum of points for the results of current, periodic and final control (exam)

Distribution of points that have received by students

	Current and periodic control							Exa m	Total score
	Content module 1			Content module 2					
	T1	T2	Modular control	T3	T4	T5	Modular control	60	100
Share, ratings, %	0	0	10	4	4	12	10		

	Current and periodic control	Exam	Final control
--	------------------------------	------	---------------

	Content module 1		Content module 2			
Share, ratings, %	P	MC	P	MC	60%	100%
	0	10	20	10		

Note: P- practical; MC- modular control

Formative assessment

Types of study work	Score per lesson (maximum)	Number of lessons	Total score (maximum)
<i>Content module 1. General concepts and postulates of metagenomics</i>			
Surveys in practical classes	0	0	0
Testing for content module 1	10	1	10
Total for content module 1			10
Content module 2. Using software products to solve metagenomic analysis problems.			
Surveys in practical classes	4	5	20
Testing for content module 2	10	1	10
Total for content module 2			30
Exam			60
Total score			100

Independent work

Students' independent work during the study of this discipline includes theoretical preparation on individual topics of the special course, which are studied in the order of independent work, as well as preparation for the laboratory classes indicated above.

The student's independent work on the assimilation of the educational material from the course is carried out in classrooms, computer classes, and also at home.

Name of the educational platform, name of the online course, short description, link	Related to the topic
--	----------------------

Udemy Master Metagenomics and Microbiome Data Analysis Using Linux In this course, you will understand the fundamentals of metagenomics and microbiome research, including 16S rRNA and sequencing approaches; learn to perform quality control and pre-processing of raw sequences using command-line tools; and interpret biological significance from taxonomic analysis and diversity in real-world datasets. https://www.udemy.com/course/master-metagenomics-and-microbiome-data-analysis-using-linux/	Topic 1. Topic 2.
Coursera Gut Checkup: Exploring Your Microbiome This course will help you understand the concept of the human microbiome and its role in human health and functioning. https://www.coursera.org/learn/microbiome	Topic 1.

The applicant can independently choose resources that will correspond to the academic discipline. To be enrolled as a periodic, current or fragment of the final control, the applicant must submit a document confirming non-formal education, a program, and a description of the competencies and learning outcomes formed during non-formal education.

The teacher reserves the right to check the formation of the listed competencies and learning outcomes.

COURSE POLICY

Policy on deadlines and retakes: retakes or violation of deadlines are possible only for good reasons (medical reasons). In case of absence from a test, it can be retaken during a scheduled consultation. The list of questions for control is contained in the work program of the discipline, which is posted on Google Class and on the website of the Faculty of Biology - biologywiki.onu.edu.ua

The final control - the assessment is carried out in accordance with the requirements specified in paragraphs 3.12-3.14 of the Regulations on the organization and control of the learning outcomes of higher education applicants of the Odesa I. I. Mechnokov National University
http://onu.edu.ua/pub/bank/userfiles/files/documents/polozennya/poloz-org-kontrol_2022.pdf.

The possibility of taking the control measure early not exist.

Academic Integrity Policy: The higher education applicant and the lecturer must adhere to academic integrity in accordance with the Code of Academic Integrity

of Participants in the Educational Process of Odesa I. I. Mechnokov National University

<http://onu.edu.ua/pub/bank/userfiles/files/documents/acad-dobrochesnost.pdf>

Attendance policy: attendance at practical classes is mandatory, points for attending lecture classes are not awarded. A valid reason for absence from classes is considered to be illness, confirmed by a doctor's certificate (hospital).

Mobile devices: during work, students work on their own computers (laptops) with the necessary programs installed. Computers must have access to the Internet during a lecture or practical class to work with remote resources and information sources (determined by the lecturer).

Behavior in the classroom: businesslike and at the same time creative-emotional atmosphere during lectures and practical classes, during control activities - focused, without conversations and distractions