

Министерство науки и высшего образования Российской Федерации
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ (НИ ТГУ)

Институт биологии, экологии, почвоведения, сельского и лесного хозяйства (Биологический институт)

УТВЕРЖДЕНО:

Директор
Д. С. Воробьев

Рабочая программа дисциплины

Биоинформатика

по направлению подготовки

06.04.01 Биология

Направленность (профиль) подготовки:

Biodiversity (Биоразнообразие)

Форма обучения

Очная

Квалификация

Магистр

Год приема

2025

СОГЛАСОВАНО:

Руководитель ОП
И.И. Волкова

Председатель УМК
А.Л. Борисенко

Томск – 2025

1. Цель и планируемые результаты освоения дисциплины

Целью освоения дисциплины является формирование следующих компетенций:

ОПК-6 Способен творчески применять и модифицировать современные компьютерные технологии, работать с профессиональными базами данных, профессионально оформлять и представлять результаты новых разработок.

ОПК-8 Способен использовать современную исследовательскую аппаратуру и вычислительную технику для решения инновационных задач в профессиональной деятельности.

Результатами освоения дисциплины являются следующие индикаторы достижения компетенций:

ИОПК-6.2 Использует компьютерные технологии и профессиональные базы данных при планировании профессиональной деятельности, обосновывает их выбор

ИОПК-8.2 Применяет современную исследовательскую аппаратуру и вычислительную технику при решении стандартных и инновационных задач в профессиональной деятельности

2. Задачи освоения дисциплины

– получить и усвоить базовые знания о базах данных биоинформатики и их назначении, о методах проведения филогенетического анализа, его основных этапах и ограничениях;

– научиться вносить информацию в базы данных и находить алгоритмы поиска информации в области биологии и биотехнологии.

3. Место дисциплины в структуре образовательной программы

Дисциплина относится к Блоку 1 «Дисциплины (модули)».

Дисциплина относится к обязательной части образовательной программы.

4. Семестр(ы) освоения и форма(ы) промежуточной аттестации по дисциплине

Второй семестр, зачет

5. Входные требования для освоения дисциплины

Для успешного освоения дисциплины требуются компетенции, сформированные в ходе освоения образовательных программ предшествующего уровня образования. Обучающиеся должны иметь общие знания по статистике и биологии.

6. Язык реализации

Английский

7. Объем дисциплины

Общая трудоемкость дисциплины составляет 2 з.е., 72 часов, из которых:

-лекции: 10 ч.

-семинар: 14 ч.

Объем самостоятельной работы студента определен учебным планом.

8. Содержание дисциплины, структурированное по темам

Тема 1. Introduction to Bioinformatics

Historical and Introduction overview of DNA sequence databases. DNA barcode for evolution and conservation. Sequence from public databases. Sequence analysis programs.

Тема 2. Bioinformatics Database and Software Resources

Collecting and Storing of Sequences. Databases (NCBI, Protopedia, Worldwide Protein Data Bank, EMBL). DNA sequence entry. FASTA sequence format.

Тема 3. Genome organization and evolution

DNA sequencing, Genomic sequencing. Sequence formats. Sequencing cDNA libraries of expressed genes. Submission of sequences to the databases. Sequence accuracy. Computer storage of sequences.

Тема 4. Alignments and phylogenetic trees

Alignment of sequences. Finding local alignments between sequences. Multiple sequence alignment. Pair-Wise sequence analysis. Phylogenetic tree. Clustering methods. Cladistic methods.

Тема 5. Methods in bioinformatics and computer sciences programming

Bayesian statistics to sequence analysis and Markov Chains. Installation, features, and programming in R, Jupyterlab and MEGA.

Тема 6. Next Generation Sequencing (NGS)

Reads From Sequencing and Calling of Genetic Variants. Transcriptome Analysis.

Тема 7. Conservation biology: quantifying species richness and evolutionary diversity within and among communities.

DNA barcode library. Comparative measures of phylogenetic diversity as standard metrics for conservation assessment.

Тема 8. Metagenomics: the collection of genomes in a coherent environmental sample DNA sequence-based insights into “inaccessible” organisms and different microbiomas as soil, seawater and others.

9. Текущий контроль по дисциплине

Текущий контроль по дисциплине проводится путем контроля посещаемости, тестов по лекционному материалу и фиксируется в форме контрольной точки не менее одного раза в семестр.

Оценочные материалы текущего контроля размещены на сайте ТГУ в разделе «Информация об образовательной программе» - <https://www.tsu.ru/sveden/education/eduop/>.

10. Порядок проведения и критерии оценивания промежуточной аттестации

Зачет во втором семестре проводится в письменной форме по билетам. Экзаменационный билет состоит из двух частей. Продолжительность зачета 1 час.

Примерный список теоретических вопросов:

1. Databases and information systems in bioinformatics, their classification and purpose (NCBI GenBank, EMBL, SWISS-Prot, UniProt, PDB)
2. 2. Methods of data and text information analysis in biology.
3. 3. Algorithms for the analysis of genetic sequences and their adaptation to high-performance computing systems.
4. 4. Algorithms for structural and functional annotations of genomic sequences. Primary work with sequences (alignment, correction of erroneous data), search for homologous sequences (BLAST)
6. 5. Development of primers for amplification (Primer Blast),

7. 6. Search for restriction sites and open reading frames within the sequence.
8. 7. Depositing nucleotide sequences in the international GenBank database (using the Sequin program)
9. 8. Methods for aligning sequences and constructing dendrograms, their statistical evaluation.
10. 9. Algorithms of molecular evolution. Building a phylogenetic tree. Advantages and limitations of some methods. Interpretation of the received data.
11. 10. Protein databases (Uniprot, PDB, TrEMBL).
12. 11. Modeling of the three-dimensional structure of proteins based on homology, visualization of the obtained structures.
13. 12. Working with different bioinformatic databases - phylogenetic analysis

Критерии и шкалы оценивания устного ответа:

Критерий	Описание	Шкала оценивания
Знание теоретической части курса.	В процессе ответа студент демонстрирует теоретические знания по теме билета.	Да – 3 балла. Частично – 1–2 балла. Нет – 0 баллов.
Связь теории с практикой.	При ответе на практическую часть вопроса студент обосновывает выбор метода теоретическими знаниями и на их основе приводит алгоритм решения практической задачи.	Да – 3 балла. Частично – 1–2 балла. Нет – 0 баллов.
Владение основными понятиями.	Студент грамотно использует в своей речи основные определения и термины, изученные в курсе.	Да – 2 балла. Частично – 1 балл. Нет – 0 баллов.
Решение практической задачи	Студент демонстрирует решение практической задачи, обосновывает этапы ее выполнения, аргументирует ответ.	Да – 3 балла. Частично – 2–1 балл. Нет – 0 баллов.

Оценочные материалы для проведения промежуточной аттестации размещены на сайте ТГУ в разделе «Информация об образовательной программе» - <https://www.tsu.ru/sveden/education/eduop/>.

11. Учебно-методическое обеспечение

- а) Электронный учебный курс по дисциплине в электронном университете «iDO» - <https://lms.tsu.ru/course/view.php?id=19121>
- б) Оценочные материалы текущего контроля и промежуточной аттестации по дисциплине.
- в) План семинарских занятий по дисциплине.
- г) Методические указания по организации самостоятельной работы студентов.

12. Перечень учебной литературы и ресурсов сети Интернет

- а) основная литература:
 - Arthur Lesk. Introduction to Bioinformatics (fifth edition) — Oxford University Press, 2019. — 432 pages.
 - Pevsner, J. (2015). Bioinformatics and functional genomics. John Wiley & Sons.

– National Research Council (US) Board on Biology; Pool R, Esnayra J, editors. Bioinformatics: Converting Data to Knowledge: Workshop Summary. Washington (DC): National Academies Press (US); 2000. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK44939/> doi: 10.17226/999

13. Перечень информационных технологий

а) лицензионное и свободно распространяемое программное обеспечение:
– Microsoft Office Standart 2013 Russian: пакет программ. Включает приложения: MS Office Word, MS Office Excel, MS Office PowerPoint, MS Office On-eNote, MS Office Publisher, MS Outlook, MS Office Web Apps (Word Excel MS PowerPoint Outlook);
– публично доступные облачные технологии (Google Docs, Яндекс диск и т.п.).
– Ubuntu 20.04, Microsoft Visual Code, Anaconda, R, RStudio, GATK 4, samtools, veftools, bwa, FastQC, Fastp, MultiQC, Annovar, ClinCNV, CNVKit, IGV, Loupe Browser;
– публично доступные веб-технологии (BLAST: <https://blast.ncbi.nlm.nih.gov/Blast.cgi>; IGV Web App: <https://igv.org/>; Genome Browser Gateway: <https://genome-euro.ucsc.edu/cgi-bin/hgGateway>).

б) информационные справочные системы:
– CATK Best Practices Workflows — <https://gatk.broadinstitute.org/hc/enus/sections/360007226651-Best-Practices-Workflows>
– IGV User Guide — <https://software.broadinstitute.org/software/igv/UserGuide>

в) профессиональные базы данных (*при наличии*):
– GenBank — <https://www.ncbi.nlm.nih.gov/genbank/>
– Набор различных баз данных для аннотации найденных мутаций с помощью инструмента Annovar — <https://annovar.openbioinformatics.org/en/latest/user-guide/download/>
– Genome Aggregation Database — <https://gnomad.broadinstitute.org/>

14. Материально-техническое обеспечение

Аудитории для проведения занятий лекционного типа.

Аудитории для проведения занятий семинарского типа, индивидуальных и групповых консультаций, текущего контроля и промежуточной аттестации.

Помещения для самостоятельной работы, оснащенные компьютерной техникой и доступом к сети Интернет, в электронную информационно-образовательную среду и к информационным справочным системам.

15. Информация о разработчиках

Волкова И.И., канд. биол. наук, каф. ботаники Биологического института ТГУ, доцент.

Калдерон Руэда Симена Соледад, ассистент, кафедра ихтиологии и гидробиологии БИ ТГУ.