

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

Практикум по геномике

по направлению подготовки

06.04.01 Биология

Направленность (профиль) подготовки:
Генетика, геномика и синтетическая биология

Форма обучения
Очная

Квалификация
Магистр

Год приема
2024

СОГЛАСОВАНО:
Руководитель ОП
Г. Н. Артемов

Председатель УМК
А. Л. Борисенко

1. Цель и планируемые результаты освоения дисциплины

Целью освоения дисциплины является формирование следующих компетенций:

ОПК-2 Способен творчески использовать в профессиональной деятельности знания фундаментальных и прикладных разделов дисциплин (модулей), определяющих направленность программы магистратуры.

ОПК-5 Способен участвовать в создании и реализации новых технологий в сфере профессиональной деятельности и контроле их экологической безопасности с использованием живых объектов.

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

Результатами освоения дисциплины являются следующие индикаторы достижения компетенций:

ИОПК-2.2 Демонстрирует понимание методологических основ дисциплин, определяющих направленность программы магистратуры

ИОПК-2.3 Использует фундаментальные знания, практические наработки и методический базис специальных дисциплин, определяющих направленность программы магистратуры, при планировании и реализации профессиональной деятельности

ИОПК-5.1 Понимает теоретические принципы и современный практический опыт использования биологических объектов в сфере профессиональной деятельности

ИОПК-6.1 Описывает разнообразие, пути и перспективы применения компьютерных технологий в современной биологии

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

2. Задачи освоения дисциплины

- Освоить методы секвенирования ДНК
- Научиться применять инструменты анализа данных секвенирования для решения научных и практических задач
- Уметь искать и обрабатывать данные, доступные в сети интернет

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.

Для успешного освоения дисциплины требуются результаты обучения по следующим дисциплинам: «генетика», «филогенетика», «молекулярная биология», «компьютерные технологии»

6. Язык реализации

Русский

7. Объем дисциплины

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

-лекции: 4 ч.

-практические занятия: 20 ч.

Объем самостоятельной работы студента определен учебным планом.

8. Содержание дисциплины, структурированное по темам

Тема 1. Секвенирование ДНК. История, методы и современные технологии.

Методы секвенирования ДНК, включая классические подходы, такие как метод Сэнгера и методы Максама-Гилберта, а также современные технологии следующего поколения (NGS), такие как пиросеквенирование и секвенирование Illumina. Технологии третьего поколения, включая PacBio и Oxford Nanopore Technologies (ONT), их влияние на геномные исследования.

Тема 2. Форматы файлов в геномике: структурирование и хранение данных.

Обзор основных форматов файлов, используемых в геномике, таких как FASTA, FASTQ, SAM, BAM и VCF. Обсуждение структуры форматов, их предназначения и значимости для хранения, передачи и анализа биоинформатических данных, а также их роль в обеспечении совместимости между различными программными инструментами.

Тема 3. Программирование в геномике: языки и инструменты для анализа данных.

Анализ ключевых языков программирования, используемых в геномике, таких как Python и R, основы работы с Bash. Основные библиотеки и инструменты для анализа геномных данных, примеры применения программирования для решения задач в области биоинформатики.

Тема 4. Оценка качества данных и методы выравнивания последовательностей.

Методы оценки качества прочтений после секвенирования, подходы к выравниванию последовательностей. Парное и множественное выравнивание, глобальное и локальное выравнивание, различия между картированием и выравниванием прочтений. Обсуждение программных инструментов, используемых для этих задач.

Тема 5. Геномная сборка и анализ геномных вариантов.

Методы сборки генома из коротких и длинных прочтений, анализ геномных вариантов. Концепции контигов и скэффолдов, преимущества парноконцевых прочтений, использование графов де Брюйна и направление исследования геномных вариантов, включая их классификацию.

Тема 6. Метагеномика и анализ экспрессии генов: современные подходы и приложения.

Основы метагеномики, включая методы анализа 16S рРНК и статистический анализ метагеномных данных. Подходы к анализу дифференциальной экспрессии генов и их применение для изучения взаимодействия микроорганизмов с окружающей средой. Современные технологии, такие как секвенирование одиночных клеток и применение машинного обучения в биологии.

9. Текущий контроль по дисциплине

Текущий контроль по дисциплине проводится путем контроля посещаемости, проведения контрольных работ, тестов по лекционному материалу, рефератов и практических заданий 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 устной форме по билетам.

Зачетный билет состоит из двух частей. Продолжительность зачета 3 часа.

Первая часть содержит один вопрос, проверяющий ИОПК-2.2, ИОПК-2.3, ИОПК-5.1. Ответ на вопрос первой части дается в развернутой форме. Вторая часть содержит один вопрос, проверяющий ИОПК-6.1, ИОПК-6.2. Ответ на вопрос второй части дается в развернутой форме.

Если студент сдал тесты и сделал доклады на общую сумму баллов, равную 80 % от максимально возможной суммы баллов, то он получает зачет.

Вопросы к зачету по дисциплине «Практикум по геномике»

ИОПК-2.2 Демонстрирует понимание методологических основ дисциплин, определяющих направленность программы магистратуры;

1. Направления исследований в прикладной геномике.
2. Основные технологии секвенирования
3. Место геномики в современной науке
4. Форматы fasta и fastq
5. Форматы SAM, BAM.
6. Скриптовые языки программирования. Как применяются в геномике, какие наиболее часто используются и при решении каких задач.
7. Основные характеристики языка программирования Python.
8. Основные характеристики языка программирования R.
9. Язык Bash и основы работы в командной строке.

ИОПК-2.3 Использует фундаментальные знания, практические наработки и методический базис специальных дисциплин, определяющих направленность программы магистратуры, при планировании и реализации профессиональной деятельности;

10. Оценка качества прочтений. Программы для оценки качества прочтений.
11. Выравнивание последовательностей. Парное и множественное выравнивание.
12. Выравнивание последовательностей. Глобальное и локальное выравнивание.
13. Отличие картирования прочтений от выравнивания прочтений.
14. Картирование РНК-секвенирования на референсный геном, особенности картирования.
15. Программы для работы с SAM, BAM.
16. Поиск геномных вариантов. Основные направления изучения.
17. Классификация геномных вариантов.

ИОПК-5.1 Понимает теоретические принципы и современный практический опыт использования биологических объектов в сфере профессиональной деятельности

18. Сборка генома из коротких прочтений.
19. Контиги, скэффолды. Преимущество парноконцевых прочтений при сборке генома.
20. Сборка de novo.
21. Сборка уже собранного генома. В каких задачах имеется такая необходимость.

22. Графы де Брюйна, особенности применения.
23. Преимущества и недостатки сборки генома длинными прочтениями.
24. Гибридная сборка генома. Метод Hi-C в приложении сборки генома.

ИОПК-6.1 Описывает разнообразие, пути и перспективы применения компьютерных технологий в современной биологии;

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

25. Формат VCF, программы для работы с VCF.
26. Метагеномика. Приложение и методы.
27. Метагеномный анализ на основе 16S рРНК.
28. Статистический анализ метагеномных данных.
29. Анализ дифференциальной экспрессии генов. Основы.
30. Секвенирование одиночных клеток. Приложение, преимущества и недостатки.
31. Машинное обучение в биологии. Общий обзор.

Критерии оценивания:

Оценка	Критерии оценки
Не зачтено	Нет ответа даже на общие вопросы
Зачтено	Неполный ответ на все вопросы, полный развернутый или частично неполный ответ на все вопросы

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 среде электронного обучения iDO - <https://lms.tsu.ru/course/view.php?id=32735>

б) Оценочные материалы текущего контроля и промежуточной аттестации по дисциплине.

в) План семинарских / практических занятий по дисциплине.

Практические занятия проводятся по единому плану:

1. Доклады обучающихся по темам, соответствующим содержанию дисциплины
2. Обсуждение представленной информации.
3. Знакомство с информационными источниками по теме семинара

Темы практических занятий (ПЗ):

- 1) Введение в операционную систему Linux (4 часа)
- 2) Языки программирования, используемые в биоинформатике и геномике (4 часов)
- 3) Первичная обработка результатов секвенирования (2 часа)
- 4) Сборка генома (4 часа)
- 5) Выравнивание прочтений на геном (2 часа)
- б) Метагеномика (4 часа)

д) Методические указания по организации самостоятельной работы студентов.

Целью самостоятельной работы обучающихся является:

- закрепление фундаментальных знаний в области прикладной геномики, изучение работы современных программ и их применение в генетики и геномике;
- развитие умения самостоятельно работать с учебным материалом;
- приобретение навыков поиска и реферирования доступной научной информации в области прикладной геномики.

Самостоятельная работа студентов предусматривает:

- повторение лекционного материала, подготовку к семинарским занятиям;
- подготовку к зачету.

Во время самостоятельной работы для подготовки к семинарским занятиям обучающийся может использовать рекомендованные литературные источники и интернет-ресурсы, а также иные источники информации (статьи в периодических изданиях и др.), позволяющие получать современную информацию об исследованиях в прикладной геномике.

12. Перечень учебной литературы и ресурсов сети Интернет

а) основная литература:

- Compeau, Phillip, and Pavel Pevzner. Bioinformatics algorithms: an active learning approach. Vol. 1. La Jolla, California: Active Learning Publishers, 2015..
- Anisimova, Maria, ed. Evolutionary Genomics: Statistical and Computational Methods, Volume 2. Vol. 855. Humana Press, 2012.
- Felsenstein, Joseph. "Theoretical evolutionary genetics joseph felsenstein." University of Washington, Seattle (2005).
- Bleidorn, Christoph. Phylogenomics: an introduction. Springer, 2017.

б) дополнительная литература:

- Стефанов, Василий Евгеньевич, Александр Александрович Тулуб, and Григорий Ростиславович Мавропуло-Столяренко. "Биоинформатика." (2019).
- Клетки по Льюину / Окс Реймонд, Джоуклин Кребс Е., Дэвид Бир Дж. [и др.] ; под редакцией Л. Кассимерис [и др.] ; перевод И. В. Филиппович. — 3-е изд. — Москва : Лаборатория знаний, 2018. — 1057 с. — ISBN 978-5-00101-587-1. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: <https://www.iprbookshop.ru/88935.html> (21.02.2023).

в) ресурсы сети Интернет:

- Курс на stepik.org «Bioinformatics Algorithms: An Active Learning Approach» <https://stepik.org/course/55789/info>
- Курс на stepik.org «Введение в NGS» <https://stepik.org/course/1142/info>

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 (Яндекс диск и т.п.).

б) информационные справочные системы:

- Электронный каталог Научной библиотеки ТГУ – <http://chamo.lib.tsu.ru/search/query?locale=ru&theme=system>
- Электронная библиотека (репозиторий) ТГУ –

<http://vital.lib.tsu.ru/vital/access/manager/Index>

– ЭБС Лань – <http://e.lanbook.com/>

– ЭБС Консультант студента – <http://www.studentlibrary.ru/>

– Образовательная платформа Юрайт – <https://urait.ru/>

– ЭБС ZNANIUM.com – <https://znanium.com/>

– ЭБС IPRbooks – <http://www.iprbookshop.ru/>

в) профессиональные базы данных:

– Ensembl – <https://www.ensembl.org/index.html>

– UniProt – <https://www.uniprot.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

Кириленко Кирилл Михайлович, ассистент каф. генетики и клеточной биологии