

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

Химический факультет

УТВЕРЖДАЮ:

И.о декана химического факультета

А.С. Князев

Рабочая программа дисциплины

Практикум по биоинформатике

по направлению подготовки

04.04.01 Химия

Направленность (профиль) подготовки:

«Трансляционные химические и биомедицинские технологии»

Форма обучения

Очная

Квалификация

Магистр

Год приема

2023

СОГЛАСОВАНО:

Руководитель ОП

И.А. Курзина

Председатель УМК

Л.Н. Мишенина

1. Цель и планируемые результаты освоения дисциплины (модуля)

Целью освоения дисциплины является формирование следующих компетенций:

– ПК-1. Способен планировать работу и выбирать адекватные методы решения научно-исследовательских и/или производственных задач в выбранной области химии, химической технологии или смежных с химией науках.

– ПК-3. Способен к решению профессиональных производственных задач

Результатами освоения дисциплины являются следующие индикаторы достижения компетенций:

ИПК-1.1. Разрабатывает стратегию научных исследований, составляет общий план и детальные планы отдельных стадий.

ИПК-1.3. Использует современное физико-химическое оборудование для получения и интерпретации достоверных результатов исследования в выбранной области химии, химической технологии или смежных с химией науках, применяя взаимодополняющие методы исследования.

ИПК-3.1. Анализирует имеющиеся нормативные документы по системам стандартизации, разработки и производству химической продукции и предлагает технические средства для решения поставленных задач.

ИПК-3.2. Производит оценку применимости стандартных и/или предложенных в результате НИР технологических решений на применимость с учетом специфики изучаемых процессов.

2. Задачи освоения дисциплины

– формирование базовых знаний о различных ресурсах и их организации содержащих информацию и программы для анализа биологических данных;

– практическое освоение методов биоинформатического анализа на примере одного из участков генома человека;

– формирование основных навыков и приобретение практического опыта, необходимого для проведения самостоятельных научных исследований в области биоинформатического анализа.

3. Место дисциплины (модуля) в структуре образовательной программы

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

Дисциплина относится к обязательной части образовательной программы. Дисциплина входит в модуль Дисциплины по выбору 3 (ДВ.3).

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

Семестр 3, зачёт с оценкой

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: Химия, Биология, Математический анализ.

6. Язык реализации

Русский

7. Объем дисциплины (модуля)

Общая трудоемкость дисциплины составляет 3 з.е., 108 часов, из которых:

– лекции: 12 ч.;

– практические занятия: 20 ч.;
в том числе практическая подготовка: 20 ч.
Объем самостоятельной работы студента определен учебным планом.

8. Содержание дисциплины (модуля), структурированное по темам

Тема 1. Функциональные элементы генома Геном.

Доля транскрибируемой ДНК. Мусорная ДНК. Информационная емкость. Реализация ДНК как генетического материала. Анализ первичной структуры ДНК и её функции. Проект ENCODE. Гены человека. Псевдогены, их классификация. Процессированные псевдогены. Механизмы функционального действия процессированных псевдогенов. Регуляторные участки в геноме: промотор, ТАТА-бокс, энхансер, сайленсер, инсулятор. Повторяющиеся последовательности в ДНК. Тандемные повторы: микросателлиты, минисателлиты и сателлиты. Болезни экспансии тринуклеотидных повторов. Диспергированные повторы: транспозоны и ретротранспозоны. Открытие мобильных элементов. Полиморфизм ДНК. Функционирование вторичной структуры ДНК. G-квадруплексы в промоторах и теломерах.

Тема 2. Основные ресурсы для анализа функциональных элементов генома.

Сервер NCBI. Основные базы данных расположенные на нём: GenBank, Genome, Gene, dbSNP, dbEST, UniGene, PubMed, OMIM, Protein, и другие. Запросы онлайн. Перекрёстный анализ. Скачивание данных для локального анализа.

Тема 3. Геномные баузеры.

Геномные баузеры: MapView, USCS, ensemble. Устройство и пользование баузером USCS. Создание индивидуальных сессий. Обмен информацией в сессиях. Сборки геномов. Трэки, управление и настройки трэков. Трэк гены. Сравнение различных аннотаций генов человека. Трэки мРНК и EST. Совместный анализ трэков гены, мРНК и EST.

Тема 4. Анализ структуры локуса генома человека.

Кодирующие и некодирующие РНК. lncRNA, miRNA. Псевдогены. Повторяющиеся последовательности. Семейство Alu повторов.

Тема 5. Экспрессионный анализ.

Анализ данных микрочипов Affymetrix. Устройство пробсета. База данных BioGPS. Анализ данных RNAseq разных проектов. Количественный и качественный анализ. Выбор вида представления данных. Экспрессионный анализ по клеточным культурам, по локализации транскриптов, различным условиям обработки клеточных культур, фракции полиА+ или полиА-. Сравнение структуры генов, РНК и EST. Данные Fantom5.

Тема 6. Изучение регуляции локуса.

Регуляция с помощью miRNA. CLASH data. Антисмысловые взаимодействия. Транскрипционные факторы. Гистоновые модификации. Метилирование. Чувствительность к ДНКазам.

Тема 7. Исследование кодирующего потенциала гена.

Поиск открытой рамки считывания. Определение белковых изоформ продукта. Анализ доменной структуры белка. Онлайн инструменты: CDS, CDART. База данных Pfam.

Тема 8. Анализ функции белка.

Выявление белковых партнёров. Пост-трансляционные модификации белка. Предсказание функции белка на основе его взаимодействий с белками-партнерами. Анализ

данных базы «IntAct Molecular Interaction Database». Анализ клеточных путей. База данных neXtProt. Данные проекта «The human protein atlas». PeptideAtlas. Uniprot.

Тема 9. Сравнительно-геномный анализ локуса.

Консервативность гена, эволюция локуса. Ортологи, паралоги, гомологи. Анализ генов паралогов в геноме человека.

Тема 10. Особенности анализа экспрессии генов.

Особенности анализа экспрессии генов в опухолях и при исследовании хронических заболеваний человека База данных GEO. Датасеты, их устройство. Депонирование данных. Инструменты для анализа данных. GEO2R.

Тема 11. Анализ мутации в генах.

Мутации, их классификация. Синонимичные и несинонимичные замены. Полиморфизм в гене. вариации дозы гена. Функциональный анализ мутации. Предсказательные подходы: консервативность, программы Mutation taster, PolyPhen-2, SIFT+Provean. Экспериментальные методы.

9. Текущий контроль по дисциплине

Текущий контроль по дисциплине проводится путем контроля посещаемости, контроля введения конспектов, проверки самостоятельной работы студента, решения задач и фиксируется в форме контрольной точки не менее одного раза в семестр.

10. Порядок проведения и критерии оценивания промежуточной аттестации

Зачет с оценкой в третьем семестре проводится устной форме по билетам. Продолжительность зачета 1,5 часа.

Примерный перечень теоретических вопросов

1. Функциональные элементы генома
2. Анализ первичной структуры ДНК и её функции. Проект ENCODE.
3. Гены человека. Различные базы данных по генам.
4. Псевдогены, их классификация. Процессированные псевдогены. Механизмы функционального действия процессированных псевдогенов.
5. Регуляторные участки в геноме: промотор, ТАТА-бокс, энхансер, сайленсер, инсулятор.
6. Повторяющиеся последовательности в ДНК. Тандемные повторы: микросателлиты, минисателлиты и сателлиты. Болезни экспансии тринуклеотидных повторов.
7. Диспергированные повторы: транспозоны и ретротранспозоны. Открытие мобильных элементов.
8. Полиморфизм ДНК.
9. Функционирование вторичной структуры ДНК. G-квадруплексы в промоторах и теломерах.
10. Основные ресурсы для анализа функциональных элементов генома
11. Сервер NCBI. Основные базы данных расположенные на нём: GenBank, Genome, Gene, dbSNP,
12. Базы данных dbEST, UniGene, PubMed, OMIM, Protein, и другие. Запросы онлайн.
13. Геномные браузеры: MapView, USCS, ensemble.
14. Устройство и пользование браузером USCS. Создание индивидуальных сессий. Обмен информацией в сессиях. Сборки геномов.

15. Трэки, управление и настройки трэков. Трэк гены. Сравнение различных аннотаций генов человека.
16. Трэки мРНК и EST. Совместный анализ трэков гены, мРНК и EST.
17. Анализ структуры локуса генома человека.
18. Кодирующие и некодирующие РНК. lncRNA, miRNA.
19. Псевдогены. Повторяющиеся последовательности. Семейство Alu
20. Экспрессионный анализ. Анализ данных микрочипов Affymetrix. Устройство пробсета. База данных BioGPS.
21. Анализ данных RNAseq разных проектов. Количественный и качественный анализ. Выбор вида представления данных.
22. Данные проекта Fantom5.
23. Регуляция с помощью miRNA. CLASH data.
24. Антисмысловые взаимодействия.
25. Транскрипционные факторы.
26. Гистоновые модификации. Метилирование. Чувствительность к ДНКазам.
27. Поиск открытой рамки считывания. Определение белковых изоформ продукта. Анализ доменной структуры белка.
28. Онлайн инструменты: CDS, CDART. База данных Pfam.
29. Выявление белковых партнёров. Пост-трансляционные модификации белка.
30. Предсказание функции белка на основе его взаимодействий с белками-партнерами. Анализ данных базы «IntAct Molecular Interaction Database».
31. Анализ клеточных путей. База данных neXtProt. Данные проекта «The human protein atlas». PeptideAtlas. Uniprot
32. Сравнительно-геномный анализ локуса.
33. Консервативность гена, эволюция локуса. Ортологи, паралоги, гомологи. Анализ генов паралога в геноме человека.
34. Особенности анализа экспрессии генов в опухолях и при исследовании хронических заболеваний человека. База данных GEO. Датасеты, их устройство. Депонирование данных. Инструменты для анализа данных. GEO2R.

Оценка отлично (5 баллов) - выставляется студенту, показавшему всесторонние, систематизированные, глубокие знания учебной программы дисциплины, проявляющему интерес к данной предметной области, продемонстрировавшему умение уверенно и творчески применять их на практике при решении конкретных задач, свободное и правильное обоснование принятых решений.

Оценка хорошо (4 балла) - выставляется студенту, если он грамотно и по существу излагает материал, умеет применять полученные знания на практике, но недостаточно грамотно обосновывает полученные результаты.

Оценка удовлетворительно (3 балла) - выставляется студенту, показавшему фрагментарный, разрозненный характер знаний, недостаточно правильные формулировки базовых понятий, нарушения логической последовательности в изложении программного материала, но при этом он освоил основные разделы учебной программы, необходимые для дальнейшего обучения, и может применять полученные знания по образцу в стандартной ситуации.

Оценка неудовлетворительно (2 балла) - выставляется студенту, который не знает большей части основного содержания учебной программы дисциплины, допускает грубые ошибки в формулировках основных принципов и не умеет использовать полученные знания при решении типовых задач.

11. Учебно-методическое обеспечение

- а) Оценочные материалы текущего контроля и промежуточной аттестации по дисциплине.

12. Перечень учебной литературы и ресурсов сети Интернет

а) основная литература

– Computational Biology and Applied Bioinformatics. Edited by Heitor Silverio Lopes and Leonardo Magalhães Cruz, ISBN 978-953-307-629-4, 456 pages, Publisher: InTech, Chapters published September 02, 2011 under CC BY-NC-SA 3.0 license. DOI: 10.5772/772

– Applied Bioinformatics: An Introduction. Selzer, Paul Maria, Marhöfer, Richard, Rohwer, Andreas. 2008, XIV, 287 p.

б) дополнительная литература

– Bioinformatics for High Throughput Sequencing. Rodríguez-Ezpeleta, Naiara, Hackenberg, Michael, Aransay, Ana M. (Eds.). 2012, XI, 255 p.

– Bioinformatics. Tools and Applications. Edwards, David, Stajich, Jason, Hansen, David (Eds.). 2009, XII, 451 p.

– Systems and Synthetic Biology. Editors-in-Chief: R. Weiss; P.K. Dhar.

в) ресурсы сети Интернет:

– Научно-библиографические и патентные базы данных в области биологии, доступные по сети Интернет в бесплатном режиме - Science Citation Index (Web of Science), Medline (PubMed), Научная электронная библиотека (НЭБ), Российская патентная БД ФГУ ФИПС;

– Электронные адреса крупных научных издательств, предоставляющих доступ к полным текстам текущих и архивным выпускам этих журналов.

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 OneNote, 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.).

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/>

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

Воробьев Ростислав Сергеевич, лаборатория биоразнообразия и экологии, Обособленное структурное подразделение "Научно-исследовательский институт биологии и биофизики Томского государственного университета", младший научный сотрудник.