

**Федеральное государственное автономное образовательное
учреждение высшего образования
«Московский физико-технический институт
(национальный исследовательский университет)»**

УТВЕРЖДЕНО

**Директор физтех-школы
биологической и медицинской
физики**

Д.В. Кузьмин

	Рабочая программа дисциплины (модуля)
по дисциплине:	Биоинформатика
по направлению:	Биотехнология
профиль подготовки:	Биотехнология
	Физтех-школа Биологической и Медицинской Физики
	кафедра биоинформатики и системной биологии
курс:	3
квалификация:	бакалавр

Семестры, формы промежуточной аттестации:

6 (весенний) - Дифференцированный зачет

7 (осенний) - Экзамен

Аудиторных часов: 180 всего, в том числе:

лекции: 60 час.

семинары: 0 час.

лабораторные занятия: 120 час.

Самостоятельная работа: 105 час.

Подготовка к экзамену: 30 час.

Всего часов: 315, всего зач. ед.: 7

Программу составил: М.Ю. Скоблов, канд. биол. наук

Программа обсуждена на заседании кафедры биоинформатики и системной биологии 06.08.2022

Аннотация

Курс направлен на приобретение студентами наиболее важных представлений об основных подходах биоинформатического анализа, применяемых для изучения структуры и функционирования геномного локуса. Студент после освоения курса будет понимать основные геномные элементы, базы данных содержащих информацию о геномных элементах, инструменты для анализа геномных элементов, программы для представления данных по геномным элементам, различные форматы данных, методы с помощью которых получают биологические данные, методы исследования генома.

1. Цели и задачи

Цель дисциплины

дать студентам наиболее важные представления об основных подходах биоинформатического анализа применяемых для изучения структуры и функционирования геномного локуса.

Задачи дисциплины

- формирование базовых знаний о различных ресурсах и их организации содержащих информацию и программы для анализа биологических данных;
- практическое освоение студентами методов биоинформатического анализа на примере одного из участков генома человека;
- формирование у студентов основных навыков и приобретение ими практического опыта, необходимого для проведения самостоятельных научных исследований в области биоинформатического анализа.

2. Перечень формируемых компетенций

Освоение дисциплины направлено на формирование следующих компетенций:

Код и наименование компетенции	Индикаторы достижения компетенции
УК-6 Способен управлять своим временем, выстраивать и реализовывать траекторию саморазвития на основе принципов образования в течение всей жизни	УК-6.1 Определяет приоритеты профессиональной деятельности и способы ее совершенствования на основе самооценки
	УК-6.2 Способен планировать самостоятельную деятельность в решении профессиональных задач; подвергать критическому анализу проделанную работу; находить и творчески использовать имеющийся опыт в соответствии с задачами саморазвития
ОПК-1 Способен изучать, анализировать, использовать биологические объекты и процессы, основываясь на математических, физических, химических, биологических законах, закономерностях и взаимосвязях	ОПК-1.1 Способен анализировать поставленную задачу, намечать пути ее решения
	ОПК-1.3 Способен определять границы применимости полученных результатов
ПК-1 Способен планировать и проводить научные эксперименты (в избранной предметной области) и (или) теоретические (аналитические и имитационные) исследования	ПК-1.1 Владеет фундаментальными понятиями, законами и теориями современной физики и биологии
	ПК-1.7 Знает основные правила поведения и работы в современной научной лаборатории

3. Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны знать:

- основные геномные элементы;
- базы данных содержащих информацию о геномных элементах;
- инструменты для анализа геномных элементов;
- программы для представления данных по геномным элементам;
- различные форматы данных;
- методы с помощью которых получают биологические данные;
- методы исследования генома.

уметь:

- проводить биоинформатический анализ участка генома;
- вычленять информацию о функционировании участка генома;
- уметь собирать всю имеющуюся информацию из различных источников для анализа;
- проводить комплексный анализ различных видов данных;
- выстраивать гипотезы и планировать эксперименты для подтверждения наблюдаемых феноменов.

владеть:

- навыками освоения большого объёма информации;
- культурой постановки и моделирования биологических задач.

4. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий

4.1. Разделы дисциплины (модуля) и трудоемкости по видам учебных занятий

№	Тема (раздел) дисциплины	Трудоемкость по видам учебных занятий, включая самостоятельную работу, час.			
		Лекции	Семинары	Лаборат. работы	Самост. работа
1	Функциональные элементы генома	5		10	7
2	Основные ресурсы для анализа функциональных элементов генома	5		10	7
3	Геномные баузеры	5		10	7
4	Анализ структуры локуса генома человека	5		10	8
5	Экспрессионный анализ	5		10	8
6	Изучение регуляции локуса	5		10	8
7	Исследование кодирующего потенциала гена	5		10	10
8	Анализ функции белка	5		10	10
9	Сравнительно-геномный анализ локуса	5		10	10
10	Особенности анализа экспрессии генов	5		10	10
11	Анализ мутации в генах	5		10	10
12	Теория дизайна лабораторного эксперимента и представления данных	5		10	10
Итого часов		60		120	105
Подготовка к экзамену		30 час.			
Общая трудоёмкость		315 час., 7 зач.ед.			

4.2. Содержание дисциплины (модуля), структурированное по темам (разделам)

Семестр: 6 (Весенний)

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.

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.

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.

4. Анализ структуры локуса генома человека

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

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.

6. Изучение регуляции локуса

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

Семестр: 7 (Осенний)

7. Исследование кодирующего потенциала гена

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

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

9. Сравнительно-геномный анализ локуса

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

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.

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.

12. Теория дизайна лабораторного эксперимента и представления данных

Выбор объекта исследования. Постановка задач и целей. Выбор методического решения поставленной задачи. Определение сроков эксперимента и объёма работы. Расходные материалы. Выбор журнала для публикации работы. Импакт фактор. Анализ основных разделов статьи. Логика формирования материалов для разных разделов публикации. Черновой вариант. Ответы рецензенту. Представление результатов в виде устных и постерных сообщений. Оформление работы.

5. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю)

Аудитория, мультимедийное оборудование (проектор, звуковая система), доска для рисования, маркеры, компьютер.

6.Перечень рекомендуемой литературы

Основная литература

Литература предоставлена базовой кафедрой:

- 1) 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
- 2) Applied Bioinformatics: An Introduction. Selzer, Paul Maria, Marhöfer, Richard, Rohwer, Andreas. 2008, XIV, 287 p.

Дополнительная литература

Литература предоставлена базовой кафедрой:

- 1) Bioinformatics for High Throughput Sequencing. Rodríguez-Ezpeleta, Naiara, Hackenberg, Michael, Aransay, Ana M. (Eds.). 2012, XI, 255 p.
- 2) Bioinformatics. Tools and Applications. Edwards, David, Stajich, Jason, Hansen, David (Eds.). 2009, XII, 451 p.
- 3) Systems and Synthetic Biology. Editors-in-Chief: R. Weiss; P.K. Dhar.

7. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины (модуля)

Не используются

8. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень необходимого программного обеспечения и информационных справочных систем (при необходимости)

Доступ в Интернет. Google Drive для доступа к материалам курса. Приветствуется наличие во время занятий смартфонов/ноутбуков для участия в интерактивных упражнениях.

9. Методические указания для обучающихся по освоению дисциплины (модуля)

Студент, изучающий дисциплину, должен с одной стороны, овладеть общим понятийным аппаратом, а с другой стороны, должен научиться применять теоретические знания на практике. В результате изучения дисциплины студент должен знать основные определения дисциплины, уметь применять полученные знания для решения различных задач.

Успешное освоение курса требует:

- посещения всех занятий, предусмотренных учебным планом по дисциплине;
- ведения конспекта занятий;
- напряжённой самостоятельной работы студента.

Самостоятельная работа включает в себя:

- чтение рекомендованной литературы;
- проработку учебного материала, подготовку ответов на вопросы, предназначенных для самостоятельного изучения;
- решение задач, предлагаемых студентам на занятиях;
- подготовку к выполнению заданий текущей и промежуточной аттестации.

Показателем владения материалом служит умение без конспекта отвечать на вопросы по темам дисциплины.

Важно добиться понимания изучаемого материала, а не механического его запоминания. При затруднении изучения отдельных тем, вопросов, следует обращаться за консультациями к преподавателю.

Возможен промежуточный контроль знаний студентов в виде решения задач в соответствии с тематикой занятий.

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

по направлению: Биотехнология
профиль подготовки: Биотехнология
Физтех-школа Биологической и Медицинской Физики
кафедра биоинформатики и системной биологии
курс: 3
квалификация: бакалавр

Семестры, формы промежуточной аттестации:

6 (весенний) - Дифференцированный зачет
7 (осенний) - Экзамен

Разработчик: М.Ю. Скоблов, канд. биол. наук

1. Компетенции, формируемые в процессе изучения дисциплины

Код и наименование компетенции	Индикаторы достижения компетенции
УК-6 Способен управлять своим временем, выстраивать и реализовывать траекторию саморазвития на основе принципов образования в течение всей жизни	УК-6.1 Определяет приоритеты профессиональной деятельности и способы ее совершенствования на основе самооценки
	УК-6.2 Способен планировать самостоятельную деятельность в решении профессиональных задач; подвергать критическому анализу проделанную работу; находить и творчески использовать имеющийся опыт в соответствии с задачами саморазвития
ОПК-1 Способен изучать, анализировать, использовать биологические объекты и процессы, основываясь на математических, физических, химических, биологических законах, закономерностях и взаимосвязях	ОПК-1.1 Способен анализировать поставленную задачу, намечать пути ее решения
	ОПК-1.3 Способен определять границы применимости полученных результатов
ПК-1 Способен планировать и проводить научные эксперименты (в избранной предметной области) и (или) теоретические (аналитические и имитационные) исследования	ПК-1.1 Владеет фундаментальными понятиями, законами и теориями современной физики и биологии
	ПК-1.7 Знает основные правила поведения и работы в современной научной лаборатории

2. Показатели оценивания компетенций

В результате изучения дисциплины «Биоинформатика» обучающийся должен:

знать:

- основные геномные элементы;
- базы данных содержащих информацию о геномных элементах;
- инструменты для анализа геномных элементов;
- программы для представления данных по геномным элементам;
- различные форматы данных;
- методы с помощью которых получают биологические данные;
- методы исследования генома.

уметь:

- проводить биоинформатический анализ участка генома;
- вычленять информацию о функционировании участка генома;
- уметь собирать всю имеющуюся информацию из различных источников для анализа;
- проводить комплексный анализ различных видов данных;
- выстраивать гипотезы и планировать эксперименты для подтверждения наблюдаемых феноменов.

владеть:

- навыками освоения большого объема информации;
- культурой постановки и моделирования биологических задач.

3. Перечень типовых (примерных) вопросов, заданий, тем для подготовки к текущему контролю

Во время текущего контроля студент должен уметь ответить на следующие вопросы в 6 семестре:

1. Экспрессионный анализ. Анализ данных микрочипов Affymetrix. Устройство пробсета. База данных BioGPS.
2. Анализ данных RNAseq разных проектов. Количественный и качественный анализ. Выбор вида представления данных.
3. Данные проекта Fantom5.
4. Регуляция с помощью miRNA. CLASH data.
5. Антисмысловые взаимодействия.
6. Транскрипционные факторы.

7. Гистоновые модификации. Метилирование. Чувствительность к ДНКазам.
8. Поиск открытой рамки считывания. Определение белковых изоформ продукта. Анализ доменной структуры белка.
9. Онлайн инструменты: CDS, CDART. База данных Pfam.
10. Выявление белковых партнёров. Пост-трансляционные модификации белка.

Во время текущего контроля студент должен уметь ответить на следующие вопросы в 7 семестре:

1. Предсказание функции белка на основе его взаимодействий с белками-партнерами.
2. Анализ данных базы «IntAct Molecular Interaction Database».
3. Анализ клеточных путей. База данных neXtProt.
4. Данные проекта «The human protein atlas». PeptideAtlas. Uniprot
5. Сравнительно-геномный анализ локуса.
6. Консервативность гена, эволюция локуса.
7. Ортологи, паралоги, гомологи. Анализ генов паралогов в геноме человека.
8. Особенности анализа экспрессии генов в опухолях и при исследовании хронических заболеваний человека База данных GEO.
9. Датасеты, их устройство. Депонирование данных.
10. Инструменты для анализа данных. GEO2R.

Во время занятий могут проходить интерактивные обсуждения в чатах курса, что будет являться домашним заданием. Успешное выполнение всех заданий по курсу и выполнение контрольных срезов знаний дает преимущество на дифференцированном зачете и экзамене.

4. Перечень типовых (примерных) вопросов и тем для проведения промежуточной аттестации обучающихся

6 семестр:

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, TATA-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

7 семестр:

1. Сервер NCBI. Основные базы данных расположенные на нём: GenBank, Genome, Gene, dbSNP,
2. Базы данных dbEST, UniGene, PubMed, OMIM, Protein, и другие. Запросы онлайн.
3. Геномные браузеры: MapView, USCS, ensemble.
4. Устройство и пользование браузером USCS. Создание индивидуальных сессий. Обмен информацией в сессиях. Сборки геномов.
5. Трэки, управление и настройки трэков. Трэк гены. Сравнение различных аннотаций генов человека.
6. Трэки мРНК и EST. Совместный анализ трэков гены, мРНК и EST.
7. Анализ структуры локуса генома человека.
8. Кодирующие и некодирующие РНК. lncRNA, miRNA.
9. Псевдогены. Повторяющиеся последовательности. Семейство Alu повторов.

Билеты для дифференцированного зачета в 6 семестре:

Билет №1.

1. Повторяющиеся последовательности в ДНК. Тандемные повторы: микросателлиты, минисателлиты и сателлиты. Болезни экспансии тринуклеотидных повторов.
2. Функциональные элементы генома

Билет №8.

1. Полиморфизм ДНК.
2. Основные ресурсы для анализа функциональных элементов генома

Билеты для экзамена в 7 семестре:

Билет №3.

1. Антисмысловые взаимодействия.
2. Данные проекта Fantom5

Билет №5.

1. Трэки мРНК и EST. Совместный анализ трэков гены, мРНК и EST.
2. Регуляция с помощью miRNA. CLASH data.

Критерии оценивания

Оценка отлично (10 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 отлично (9 баллов) - выставляется студенту, показавшему всесторонние, систематизированные, глубокие знания учебной программы дисциплины и умение уверенно применять их на практике при решении конкретных задач, свободное и правильное обоснование принятых решений.

Оценка отлично (8 баллов) - выставляется студенту, показавшему всесторонние, систематизированные, глубокие знания учебной программы дисциплины и умение уверенно применять их на практике при решении конкретных задач, правильное обоснование принятых решений, с некоторыми недочетами.

Оценка хорошо (7 баллов) - выставляется студенту, если он твердо знает материал, грамотно и по существу излагает его, умеет применять полученные знания на практике, но недостаточно грамотно обосновывает полученные результаты.

Оценка хорошо (6 баллов) - выставляется студенту, если он твердо знает материал, грамотно и по существу излагает его, умеет применять полученные знания на практике, но допускает в ответе или в решении задач некоторые неточности.

Оценка хорошо (5 баллов) - выставляется студенту, если он в основном знает материал, грамотно и по существу излагает его, умеет применять полученные знания на практике, но допускает в ответе или в решении задач достаточно большое количество неточностей.

Оценка удовлетворительно (4 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 удовлетворительно (3 балла) - выставляется студенту, показавшему фрагментарный, разрозненный характер знаний, допускающему ошибки в формулировках базовых понятий, нарушения логической последовательности в изложении программного материала, слабо владеет основными разделами учебной программы, необходимыми для дальнейшего обучения и с трудом применяет полученные знания даже 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.

Оценка неудовлетворительно (1 балл) - выставляется студенту, который не знает основного содержания учебной программы дисциплины, допускает грубейшие ошибки в формулировках базовых понятий дисциплины и вообще не имеет навыков решения типовых практических задач.

5. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности

При проведении устного дифференцированного зачета обучающемуся предоставляется 30 минут на подготовку. Опрос обучающегося по билету на устном дифференцированном зачете не должен превышать одного астрономического часа.

При проведении устного экзамена обучающемуся предоставляется 30 минут на подготовку. Опрос обучающегося по билету на устном экзамене не должен превышать одного астрономического часа.