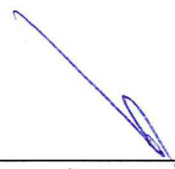


Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего образования «Новосибирский национальный исследовательский
государственный университет» (Новосибирский государственный
университет, НГУ)

Факультет естественных наук

Согласовано
Декан ФЕН
Резников В.А.



«10» октября 2020 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Методы компьютерного анализа нуклеотидных последовательностей

направление подготовки: 06.04.01 Биология

направленность (профиль): Биология

Форма обучения: очная

Разработчики:

К.б.н., Бабкин Игорь Викторович

К.б.н., Тикунов Артем Юрьевич

Руководитель программы:

Д.б.н., проф. Рубцов Н.Б.

Новосибирск, 2020

Содержание

1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы.....	3
2. Место дисциплины в структуре образовательной программы	3
3. Трудоемкость дисциплины в зачетных единицах с указанием количества академических часов, выделенных на контактную работу обучающегося с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающегося	3
4. Содержание дисциплины, структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий. Программа курса лекций.....	4
5. Перечень учебной литературы	7
6. Перечень учебно-методических материалов по самостоятельной работе обучающихся ..	7
7. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины	8
8. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине	9
9. Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине	9
10. Оценочные средства для проведения текущего контроля и промежуточной аттестации по дисциплине.....	10

1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы

Результаты освоения образовательной программы (компетенции)	В результате изучения дисциплины обучающиеся должны:		
	знать	уметь	владеть
ОПК-7. Готовность творчески применять современные компьютерные технологии при сборе, хранении, обработке, анализе и передаче биологической информации для решения профессиональных задач	Основные методы анализа последовательностей биополимеров	Выбирать адекватные методы и компьютерные программы для анализа последовательностей биополимеров на основе знаний о возможностях и ограничениях используемых методов компьютерного анализа нуклеотидных последовательностей. Находить необходимую информацию в научной литературе и базах данных	Методами выравнивания последовательностей биополимеров, методами молекулярной филогении, методами анализа данных секвенирования, методами физического картирования геномов.

2. Место дисциплины в структуре образовательной программы

Дисциплины (практики), изучение которых необходимо для освоения дисциплины
Компьютерные методы анализа нуклеотидных последовательностей:

Математика (высшая алгебра, математический анализ, математическая статистика);
Физика ;

Неорганическая химия (строение и свойства атомов, периодический закон, строение молекул, теория химической связи, стереохимия);

Физическая химия (природа химической связи в молекулах и кристаллах, химическая термодинамика);

Органическая химия (классификация и номенклатура соединений, строение молекул, изомерия);

Введение в биологию;

Молекулярная биология (структура и функции белков и нуклеиновых кислот, гены и геномы, самоорганизация живых систем).

Дисциплины (практики), для изучения которых необходимо освоение дисциплины:

Генетическая инженерия;

Молекулярная вирусология.

Методы исследования биополимеров

Выполнение научно-исследовательской задачи, выполнение ВКР

3. Трудоемкость дисциплины в зачетных единицах с указанием количества академических часов, выделенных на контактную работу обучающегося с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающегося

Трудоемкость дисциплины – 3 з.е. (108 ч)

Форма промежуточной аттестации: дифзачет

№	Вид деятельности	Семестр 5
1	Лекции, ч	30
2	Семинарские занятия, ч	14
3	Контакт. работа, ч	2
4	Занятия в контактной форме, числе, из них	46
5	аудиторных занятий, ч	44
6	в электронной форме, ч	-
7	консультаций, час.	1
8	промежуточная аттестация, ч	1
9	Самостоятельная работа, час.	62
10	Всего, ч	216

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

Лекции (30 ч)

Наименование темы и их содержание	Объем, час
Методы изучения генома, основные понятия и определения.	2
Биоинформационные подходы, позволяющие планировать эксперименты. Дизайн праймеров.	4
Анализ данных секвенирования.	4
Базы данных, извлечение и депонирование информации. Поиск гомологичных последовательностей.	4
Выравнивания последовательностей и филогенетические деревья.	6
Определение функционально важных областей.	6
Предсказание структуры и свойств биополимеров.	4

Практические занятия (14 ч)

Содержание практического занятия	Объем, час
Семинар по биоинформационным подходам, позволяющим планировать эксперименты.	2
Семинары по анализу данных секвенирования.	4
Семинар по базам данных, извлечению и депонированию информации.	2
Семинар по выравниванию последовательностей и филогенетическим деревьям	2
Семинар по определению функционально важных областей	2
Семинар по предсказанию структуры и свойств биополимеров	2

Самостоятельная работа студентов (62 ч)

Перечень занятий на СРС	Объем, час
Подготовка к практическим занятиям.	18
Выполнение домашних заданий	19
Изучение теоретического материала, не освещаемого на лекциях	20
Подготовка к зачету	5

Программа курса лекций

Лекция 1

Введение

Раздел 1. Методы изучения генома, основные понятия и определения.

1.1. Обзор основных методов изучения генома и анализ полученных данных.

1.2. Знакомство с базовым набором компьютерных программ молекулярной биологии и ключевыми веб-сайтами для анализа нуклеотидных последовательностей.

Лекция 2.

Раздел 2. Биоинформационные подходы, позволяющие планировать эксперименты.

2.1. Обзор биоинформационных подходов и программ, позволяющих планировать эксперимент.

2.2. Построение функциональных карт геномов. Определение возможных полиморфизмов на основании нуклеотидной последовательности. Анализ рестрикционных карт геномов. Проведение "виртуального электрофореза".

Лекция 3.

Дизайн праймеров.

2.3. Дизайн праймеров для ПЦР по нуклеотидной последовательности. Подбор праймеров для секвенирования ДНК.

Лекция 4.

Методы секвенирования

Раздел 3. Анализ данных секвенирования.

3.1. Обзор методов секвенирования.

Лекция 5.

Анализ данных секвенирования

3.2. Анализ данных секвенирования по Сэнгеру. Возможные трудности.

3.3. Анализ данных массового параллельного секвенирования.

Лекция 6.

Сборка нуклеотидных последовательностей. Построение контигов. Поиск химерных последовательностей

3.4. Сборка геномов. Построение контигов.

3.5. Поиск химерных последовательностей.

Лекция 7.

Современные биологические базы данных и их значение. Депонирование новой информации в базах данных

Раздел 4. Базы данных, извлечение и депонирование информации. Поиск гомологичных последовательностей.

4.1. Обзор современных биологических баз данных. Значение баз данных.

4.2. Национальный центр биотехнологической информации, как наиболее известная база информации по нуклеотидным и белковым последовательностям.

4.3. Пути использования баз данных.

4.4. Депонирование новой информации в базах данных.

Лекция 8.

Поиск схожих последовательностей в базах данных. Локальное выравнивание. Семейство программ серии BLAST

4.5. Поиск схожих последовательностей в базах данных. Локальное выравнивание. Семейство программ серии BLAST.

Лекция 9.

. Алгоритмы и методы глобального выравнивания последовательностей. Множественное выравнивание.

Раздел 5. Выравнивания и филогенетические деревья.

5.1. Цели и типы выравнивания последовательностей.

5.2. Алгоритмы и методы выравнивания последовательностей. Множественное выравнивание полинуклеотидных и полипептидных последовательностей, приложения для решения биологических задач.

Лекция 10.

Рекомбинационный анализ.

5.3. Рекомбинационный анализ.

Лекция 11.

Эволюционные модели и дистанции между последовательностями биополимеров.
Молекулярная филогения.

5.4. Эволюционные модели и дистанции между последовательностями биополимеров.

5.5. Молекулярная филогения: типы деревьев.

5.6. Методы кластеризации последовательностей биополимеров. Дистанционные методы построения филогенетических деревьев и методы, основанные на дискретных признаках.

Лекция 12.

Статистические методы оценки деревьев. Анализ молекулярных часов.

5.7. Статистические методы оценки деревьев.

5.8. Анализ молекулярных часов.

Лекция 13.

Определение функционально значимых областей геномов. Предсказание функции белков
Раздел 6. Определение функционально важных областей.

6.1. Способы определения доменной структуры и консервативных участков в последовательностях биополимеров. Открытая рамка считывания.

6.2. Поиск функциональных сайтов на последовательности ДНК.

6.3. Предсказание функции белка на основании рассчитанной аминокислотной последовательности.

Лекции 14, 15

Предсказание структуры и свойств биополимеров.

Раздел 7. Предсказание структуры и свойств биополимеров.

7.1. Предсказание вторичной структуры ДНК и РНК.

7.2. Предсказание свойств и структуры белков.

5. Перечень учебной литературы

5.1 Основная литература(предоставляется электронная копия по требованию)

1. Бабкин И.В., Тикунова Н.В., Нетесов С.В. **Компьютерные методы анализа нуклеотидных последовательностей.** Новосиб гос. ун-т. Новосибирск. ИПЦ НГУ, 2017.

2. Леск А. **Введение в биоинформатику.** Изд-во «Бином», Москва, 2013.

3. Лукашов В.В. **Молекулярная эволюция и филогенетический анализ.**Изд-во «Бином», Москва, 2009.
4. Игнасимуту С. **Основы биоинформатики.** Изд-во «Регулярная и хаотичная динамика», Ижевск, 2007.
- 5.2 Дополнительная литература**(предоставляется электронная копия по требованию)
5. Сетубал Ж., Мейданис Ж. **Введение в вычислительную молекулярную биологию.** Изд-во «Регулярная и хаотичная динамика», Ижевск, 2007.
6. Pevsner J. **Bioinformatics and Functional Genomics.** Изд-во «Wiley-Blackwell», 2013.

6. Перечень учебно-методических материалов по самостоятельной работе обучающихся

Не используются.

7. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины

При освоении дисциплины используются следующие ресурсы:

- электронная информационно-образовательная среда НГУ (ЭИОС);

Взаимодействие обучающегося с преподавателем (синхронное и (или) асинхронное) осуществляется через личный кабинет студента в ЭИОС, электронную почту, социальные сети.

7.1 Современные профессиональные базы данных:

1. <http://www.ncbi.nlm.nih.gov/>
2. <http://www.ebi.ac.uk/>
3. <http://web.expasy.org/>
4. <http://greengenes.lbl.gov>,
5. <https://rdp.cme.msu.edu>,
6. www.arb-silva.de

7.2. Информационные справочные системы

- . <http://molbiol.ru/>
- . <http://evolution.genetics.washington.edu/phylip/software.html#methods>

8. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине

8.1 Перечень программного обеспечения

Mega7.0, Windows, Microsoft Office, UniproUGENE, BioEdit, BEAST, OligoCalc

8.2 Информационные справочные системы

. <http://molbiol.ru/>

. <http://evolution.genetics.washington.edu/phylip/software.html#methods>

9. Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине

Для реализации дисциплины Компьютерные методы анализа нуклеотидных последовательностей используются специальные помещения:

1. Учебные аудитории для проведения занятий лекционного типа, групповых и индивидуальных консультаций, текущего контроля, промежуточной и итоговой аттестации;
2. Компьютерные классы для проведения занятий семинарского типа.

Используемые для практикума приборы:

Системные блоки с программным обеспечением: Windows XP, Microsoft Office 2007.
Мониторы.

В качестве технического обеспечения лекционного процесса используется ноутбук, мультимедийный проектор, экран и доска.

Для демонстрации иллюстрационного материала используется программа Microsoft PowerPoint 2013.

Обеспечивается доступ студентов к сети Интернет во время семинаров.

Проведение зачета обеспечивается печатным раздаточным материалом.

Материально-техническое обеспечение образовательного процесса по дисциплине для обучающихся из числа лиц с ограниченными возможностями здоровья осуществляется согласно «Порядку организации и осуществления образовательной деятельности по образовательным программам для инвалидов и лиц с ограниченными возможностями здоровья в Новосибирском государственном университете».

10. Оценочные средства для проведения текущего контроля и промежуточной аттестации по дисциплине

Перечень результатов обучения по дисциплине Методы компьютерного анализа нуклеотидных последовательностей и индикаторов их достижения представлен в виде знаний, умений и владений в разделе 1.

10.1 Порядок проведения текущего контроля и промежуточной аттестации по дисциплине

Текущий контроль успеваемости:

Формой текущего контроля при прохождении дисциплины Методы компьютерного анализа нуклеотидных последовательностей является контроль посещаемости практических занятий и проверочные вопросы на лекциях.

Для того чтобы быть допущенным к зачету, студент должен:

- в ходе обучения посетить не менее 70% лекционных занятий;
- выполнить поставленные задачи.

Промежуточная аттестация:

Итоговую оценку за семестр студент может получить на устном зачете в конце семестра в виде оценки.

Описание критериев и шкал оценивания индикаторов достижения результатов обучения по дисциплине Методы компьютерного анализа нуклеотидных последовательностей

Таблица 10.1

Код компетенции	Результат обучения по дисциплине	Оценочное средство
ОПК-7. Готовность творчески применять современные компьютерные технологии при сборе, хранении, обработке, анализе и передаче биологической информации для решения профессиональных задач	Умение выбирать адекватные методы и компьютерные программы для анализа последовательностей биополимеров на основе знаний о возможностях и ограничениях используемых методов компьютерного анализа нуклеотидных последовательностей. Умение находить необходимую информацию в научной литературе и базах данных.	дифференцированный зачет, практические занятия
	Знание основных методов анализа последовательностей биополимеров.	дифференцированный зачет, практические занятия
	Владение методами выравнивания последовательностей биополимеров, методами молекулярной филогении, методами анализа данных секвенирования, методами физического картирования геномов..	дифференцированный зачет, практические занятия

Таблица 10.2

Критерии оценивания результатов обучения	Шкала оценивания
<p><u>Зачет:</u></p> <ul style="list-style-type: none"> – знание основных методов анализа последовательностей биополимеров, – полнота их изложения, – самостоятельность, осмысленность, структурированность, логичность и аргументированность изложения материала, – точность и корректность применения терминов и понятий, – наличие исчерпывающих ответов на дополнительные вопросы. <p>При изложении ответа на вопрос(ы) экзаменационного билета обучающийся мог допустить не принципиальные неточности.</p>	<i>Отлично</i>
<p><u>Зачет:</u></p> <ul style="list-style-type: none"> – знание основных методов анализа последовательностей биополимеров, – полнота их изложения, – самостоятельность, осмысленность, структурированность, логичность и аргументированность изложения материала, – точность и корректность применения терминов и понятий при наличии незначительных ошибок, <p>При изложении ответа на вопросы экзаменационного билета обучающийся мог допустить небольшие неточности.</p>	<i>Хорошо</i>
<p><u>Зачет:</u></p> <ul style="list-style-type: none"> - неполное знание основных методов анализа последовательностей биополимеров, – частичное их понимание и неполное изложение, – наличие ошибок в логике и аргументации – наличие неполных и/или содержащих существенные ошибки ответов на дополнительные вопросы. 	<i>Удовлетворительно</i>
<p><u>Зачет:</u></p> <ul style="list-style-type: none"> – фрагментарное и недостаточное знание основных методов анализа последовательностей биополимеров, – непонимание причинно-следственных связей, – отсутствие осмысленности, структурированности, логичности и аргументированности в изложении материала, – грубые ошибки в применении терминов и понятий, – отсутствие ответов на дополнительные вопросы. 	<i>Неудовлетворительно</i>

Типовые контрольные задания и иные материалы, необходимые для оценки результатов обучения

Примеры тестовых вопросов :

1. Приведите пример базы данных нуклеотидных последовательностей.
2. Чем отличаются первичные и вторичные базы данных?
3. Что такое химерные нуклеотидные последовательности, как они возникают?
4. Перечислите основные текстовые форматы записи нуклеотидных последовательностей.
5. В каком направлении принято записывать нуклеотидные последовательности?
6. С помощью какого бесплатного ПО вы смогли бы построить рестрикционную карту?
7. С помощью какого ПО вы смогли бы построить физическую карту плазмиды?
8. Назовите on-line ресурс поиска гомологичных последовательностей в базах данных с помощью их локального выравнивания.
9. Назовите основные параметры, выдаваемые программой BLAST при поиске гомологичных последовательностей.
10. Назовите матрицы сравнения аминокислот, используемые для выравнивания аминокислотных последовательностей.
11. Что такое выравнивание последовательностей?
12. Цель построения выравниваний последовательностей и типы выравниваний.
13. Приведите два примера алгоритмов глобального выравнивания последовательностей.
14. Какие существуют методологические подходы к подбору олигонуклеотидных праймеров?
15. С помощью какого ПО вы смогли бы рассчитать температуру отжига праймеров?
16. Назовите три стадии ПЦР.
17. Какие методы секвенирования ДНК существуют, и какой принцип используется в автоматических секвенаторах (не NGS)?
18. С помощью какого ПО вы можете анализировать данные, генерируемые автоматическим секвенатором?
19. Как называется формат файлов, которые генерирует автоматический секвенатор AppliedBiosystems. Как называют такие данные?
20. Назовите методы проведения филогенетического анализа.
21. Назовите метод оценки статистической достоверности филогенетического дерева.
22. Что такое кладограмма?
23. Назовите ПО для построения хронограмм и анализа скорости молекулярной эволюции.
24. Назовите on-line ресурс для предсказания вторичной структуры ДНК и РНК.

Вопросы для подготовки к дифференцированному зачету

1. Методы изучения генома, основные понятия и определения.
2. Основные методы изучения генома и анализ полученных данных.
3. Базовый набор компьютерных программ молекулярной биологии и ключевые веб-сайты для анализа нуклеотидных последовательностей.
4. Биоинформационные подходы к планированию экспериментов.
5. Построение функциональных карт геномов.
6. Определение возможных полиморфизмов на основании нуклеотидной последовательности. Анализ рестрикционных карт геномов. Проведение "виртуального электрофореза".
7. Дизайн праймеров для ПЦР по нуклеотидной последовательности. Подбор праймеров для секвенирования ДНК.
8. Методы секвенирования. Анализ данных секвенирования.
9. Анализ данных секвенирования по Сэнгеру.
10. Анализ данных массового параллельного секвенирования.
11. Сборка нуклеотидных последовательностей. Построение контигов. Поиск химерных последовательностей
12. Сборка геномов. Принципы построения контигов.
13. Поиск химерных последовательностей.
14. Современные биологические базы данных и их значение. Депонирование новой информации в базах данных
15. Базы данных, извлечение и депонирование информации. Поиск гомологичных последовательностей.
16. Современные биологические базы данных. Значение баз данных. Пути использования баз данных.
17. Депонирование новой информации в базах данных.
18. Поиск схожих последовательностей в базах данных. Локальное выравнивание. Семейство программ серии BLAST
19. Алгоритмы и методы глобального выравнивания последовательностей. Множественное выравнивание.
20. Выравнивания и филогенетические деревья.
21. Алгоритмы и методы выравнивания последовательностей. Множественное выравнивание полинуклеотидных и полипептидных последовательностей, приложения для решения биологических задач.
22. Рекомбинационный анализ.
23. Эволюционные модели и дистанции между последовательностями биополимеров. Молекулярная филогения.
24. Методы кластеризации последовательностей биополимеров. Дистанционные методы построения филогенетических деревьев и методы, основанные на дискретных признаках.
25. Статистические методы оценки деревьев. Анализ молекулярных часов.
26. Определение функционально значимых областей геномов. Предсказание функции белков
27. Способы определения доменной структуры и консервативных участков в последовательностях биополимеров. Открытая рамка считывания.
28. Поиск функциональных сайтов на последовательности ДНК.

29. Предсказание структуры и свойств биополимеров. Предсказание вторичной структуры ДНК и РНК.
30. Предсказание свойств и структуры белков.