

**АВТНОМНАЯ НЕКОММЕРЧЕСКАЯ ОБРАЗОВАТЕЛЬНАЯ
ОРГАНИЗАЦИЯ ВЫСШЕГО ОБРАЗОВАНИЯ
«НАУЧНО-ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ «СИРИУС»
(АНОО ВО «УНИВЕРСИТЕТ «СИРИУС»)**

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Метагеномика

| | |
|---------------------------|------------------------------------------------------|
| Уровень образования: | высшее образование – программа магистратуры |
| Направление подготовки: | 06.04.01 Биология 09.04.03 Прикладная информатика |
| Направленность (профиль): | Биоинформатика |

1. Общая характеристика дисциплины

1.1. Цель: ознакомление студентов с современными методами метагеномного анализа, базирующегося на картировании и сборке геномных или транскриптомных прочтений материала микроорганизмов, полученных из среды.

1.2. Задачи: обзор предметной области и основных направлений исследований, практических задач; обзор экспериментальных методов, подходов к выделению материалов для исследования и пробоподготовки; обзор современных платформ высокопроизводительного секвенирования и их особенностей; обзор современных методов метагеномного анализа, методов оценки таксономического состава и разнообразия сообществ, их ограничений и особенности интерпретации результатов.

1.3. Общая трудоемкость: 4 з.е.

1.4. Планируемые результаты обучения:

| Формируемые компетенции (код компетенции, формулировка) | Планируемые результаты обучения по дисциплине (индикаторы достижения компетенций) |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| ПК-1. Способен применять фундаментальные математические и естественнонаучные знания для решения профессиональных задач в области биоинформатики, биоинженерии, биотехнологии и фарминдустрии | ИПК-1.1. Знает фундаментальные основы математики, биологии и других естественных наук |
| | ИПК-1.2. Применяет фундаментальные знания математики, биологии и других естественных наук для постановки и решения исследовательских и практических задач |
| | ИПК-1.3. Анализирует современные проблемы в области биоинформатики, биоинженерии, биотехнологии и фарминдустрии, формулирует гипотезы и вырабатывает подходы для решения исследовательских и практических задач |
| ПК-4. Способен комбинировать и адаптировать информационно-коммуникационные технологии для решения профессиональных задач в области биоинформатики, биоинженерии, биотехнологии и фарминдустрии с учетом требований информационной безопасности | ИПК-4.1. Знает базовые понятия информатики, информации, ее измерения, кодирования и представления в вычислительных системах, а также принципы сбора, хранения и обработки информации |
| | ИПК-4.2. Использует информационно-коммуникационные технологии для решения профессиональных задач в области биоинформатики, биоинженерии, биотехнологии и фарминдустрии |
| | ИПК-4.3. Комбинирует и адаптирует информационно-коммуникационные технологии с учетом требований информационной безопасности |
| ПК-6. Способен самостоятельно проводить расчетные работы и исследования в области биоинформатики, биоинженерии, | ИПК-6.1. Применяет классические методы решения задач, современные программные комплексы и навыки работы с высокотехнологичным лабораторным оборудованием для проведения расчетных работ и |

| | | |
|-------------------------------|---------------------------------------------------------|---------------------|
| АНОО ВО «Университет «Сириус» | Рабочая программа дисциплины (модуля) «Метагеномика» | Лист 3 Листов 11 |
|-------------------------------|---------------------------------------------------------|---------------------|

| | |
|--------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------|
| биотехнологии и фарминдустрии, применяя навыки работы с высокотехнологичным лабораторным оборудованием | исследований |
| | ИПК-6.2. Проводит расчетные работы и исследования, осуществляет обработку, анализ и интерпретацию биомедицинских и биотехнологических данных |
| | ИПК-6.3. Оформляет результаты расчетных работ и исследований в соответствии с требованиями к отчетной документации |

2. Структура и содержание дисциплины (модуля)

2.1. Объем дисциплины (модуля) и виды учебной деятельности:

| Виды учебной деятельности | 4 семестр | Всего |
|-----------------------------------------------------------------|-----------|-------|
| Контактная работа обучающихся с преподавателем, всего ч. | 52 | 52 |
| Лекционные занятия, ч. | 24 | 24 |
| Практические (семинарские) занятия, ч. | 24 | 24 |
| Лабораторные занятия, ч. | х | х |
| Промежуточная аттестация – экзамен, ч | 4 | 4 |
| Промежуточная аттестация – зачет с оценкой, ч | х | х |
| Промежуточная аттестация – зачет, ч | х | х |
| Самостоятельная работа обучающихся, всего ч. | 92 | 92 |
| Общая трудоемкость, ч. | 144 | 144 |
| Общая трудоемкость, з.е. | 4 | 4 |

2.2. Структура дисциплины (модуля) по разделам (темам) и видам учебной деятельности:

| Наименования разделов (тем) дисциплины (модуля) | Лекционные занятия, ч | Практические (семинарские) занятия, ч | Лабораторные занятия, ч | Промежуточная аттестация, ч | Самостоятельная работа, ч | Всего, ч | Форма текущего контроля / промежуточной аттестации |
|-------------------------------------------------|-----------------------|---------------------------------------|-------------------------|-----------------------------|---------------------------|----------|------------------------------------------------------------------|
| Введение в метагеномику/ метатранскриптомику | 4 | 4 | | | 20 | 28 | Устный опрос, контрольные работы, проектные работы, тестирование |
| Раздел 1. Обзор экспериментальных подходов. | 4 | 4 | | | 20 | 28 | Устный опрос, контрольные работы, проектные работы, тестирование |

| | | |
|-------------------------------|---------------------------------------------------------|---------------------|
| АНОО ВО «Университет «Сириус» | Рабочая программа дисциплины (модуля) «Метагеномика» | Лист 4 Листов 11 |
|-------------------------------|---------------------------------------------------------|---------------------|

| | | | | | | | |
|----------------------------------------------------------|----|----|---|---|----|-----|------------------------------------------------------------------|
| Раздел 2. Таргетное секвенирование. | 4 | 4 | | | 20 | 28 | Устный опрос, контрольные работы, проектные работы, тестирование |
| Раздел 3. Использование референсных геномов. | 4 | 4 | | | 20 | 28 | Устный опрос, контрольные работы, проектные работы, тестирование |
| Раздел 4. Оценка разнообразия с помощью анализа k-меров. | 4 | 4 | | | 10 | 18 | Устный опрос, контрольные работы, проектные работы, тестирование |
| Раздел 5. Аннотация генов. | 4 | 4 | | | 2 | 14 | Устный опрос, контрольные работы, проектные работы, тестирование |
| Промежуточная аттестация | | | | 4 | | 4 | Экзамен |
| Итого | 24 | 24 | x | 4 | 92 | 144 | |

2.3. Содержание разделов (тем) дисциплины (модуля):

| Наименования разделов (тем) дисциплины (модуля) | Содержание разделов (тем) дисциплины (модуля) |
|----------------------------------------------------------|----------------------------------------------------------------------------------------------|
| Введение в метагеномику/метатранскриптомику | Введение в метагеномику/метатранскриптомику |
| Раздел 1. Обзор экспериментальных подходов. | Обзор современных платформ высокопроизводительного секвенирования и их особенностей |
| Раздел 2. Таргетное секвенирование. | Подходы к оценке сообщества по представленности бактериальных родов и видов |
| Раздел 3. Использование референсных геномов. | Картирование прочтений. Безреференсные подходы – сборка фрагментов de novo. Алгоритмы сборки |
| Раздел 4. Оценка разнообразия с помощью анализа k-меров. | Биннинг. Алгоритмы биннинга. Контролируемый и неконтролируемый биннинг. Кластеризация |
| Раздел 5. Аннотация генов. | Алгоритмы аннотации. Функциональный метагеномный анализ. Сравнительный анализ. |

2.4. Самостоятельная работа

Самостоятельная работа предусматривает: самостоятельное изучение теоретического материала, подготовку к ответам на семинарских заданиях, подготовку к текущему контролю и промежуточной аттестации, выполнение тестовых заданий по пройденным темам курса.

| | | |
|-------------------------------|---------------------------------------------------------|---------------------|
| АНОО ВО «Университет «Сириус» | Рабочая программа дисциплины (модуля) «Метагеномика» | Лист 5 Листов 11 |
|-------------------------------|---------------------------------------------------------|---------------------|

3. Текущий контроль и промежуточная аттестация по дисциплине (модулю). Оценочные материалы

3.1. Текущий контроль успеваемости проводится в течение семестра в следующих формах:

| Наименования разделов (тем) дисциплины (модуля) | Форма текущего контроля | Оценочные материалы |
|----------------------------------------------------------|------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------|
| Введение в метагеномику/ метатранскриптомику | Устный опрос, контрольные работы, проектные работы, тестирование | Перечень вопросов для опроса, комплект практических заданий, темы для выполнения проектов, перечень вопросов для тестирования |
| Раздел 1. Обзор экспериментальных подходов. | Устный опрос, контрольные работы, проектные работы, тестирование | Перечень вопросов для опроса, комплект практических заданий, темы для выполнения проектов, перечень вопросов для тестирования |
| Раздел 2. Таргетное секвенирование. | Устный опрос, контрольные работы, проектные работы, тестирование | Перечень вопросов для опроса, комплект практических заданий, темы для выполнения проектов, перечень вопросов для тестирования |
| Раздел 3. Использование референсных геномов. | Устный опрос, контрольные работы, проектные работы, тестирование | Перечень вопросов для опроса, комплект практических заданий, темы для выполнения проектов, перечень вопросов для тестирования |
| Раздел 4. Оценка разнообразия с помощью анализа k-меров. | Устный опрос, контрольные работы, проектные работы, тестирование | Перечень вопросов для опроса, комплект практических заданий, темы для выполнения проектов, перечень вопросов для тестирования |
| Раздел 5. Аннотация генов. | Устный опрос, контрольные работы, проектные работы, | Перечень вопросов для опроса, комплект |

| | | |
|-------------------------------|---------------------------------------------------------|---------------------|
| АНОО ВО «Университет «Сириус» | Рабочая программа дисциплины (модуля) «Метагеномика» | Лист 6 Листов 11 |
|-------------------------------|---------------------------------------------------------|---------------------|

| | | |
|--|--------------|----------------------------------------------------------------------------------------------------|
| | тестирование | практических заданий, темы для выполнения проектов, перечень вопросов для тестирования |
|--|--------------|----------------------------------------------------------------------------------------------------|

3.2. Оценочные материалы для текущего контроля

Примерный перечень вопросов для устного опроса:

1. Что обозначают термины «микробиота», «микробиом», «метагеном»?
2. Какие фракции ДНК (рНК) входят в состав метагенома? Все ли они принадлежат живым организмам?
3. Какие основные проблемы встречаются при выделении нуклеиновых кислот из объектов окружающей среды?
4. Какие гены, кроме 16S рРНК могут быть объектами таргетного секвенирования?
5. Что такое разнообразие? Какие меры разнообразия вы знаете?
6. Три поколения систем секвенирования - их особенности и производительность.
7. Алгоритмы OTE и ASV-пикинга. Особенности применения алгоритмов
8. Пайплайн обработки коротких прочтений.
9. Проблемы интерпретации абсолютного числа прочтений. Методы возврата к истинным значениям.
10. Базы данных для функциональной и таксономической аннотации метагеномов.
11. Приведите примеры использования глубокого обучения в метагеномике.
12. Каковы отличия алгоритмов выравнивая и мэппинга между собой. В каких алгоритмах чаще используются k-mer и почему?
13. В чем состоит отличие пайплайнов обработки длинных прочтений и коротких?
14. Какие существуют проблемы соотнесения данных ампликонного секвенирования и метагеномов?
15. Какие базы данных для функциональной и таксономической аннотации метагеномов вам известны?

Примерный перечень заданий для контрольных работ:

1. Оценка разнообразия микробиомов по данным таргетного секвенирования.
2. Вычисление количества копий геномов в препарате ДНК, выделенном из окружающей среды.
3. Оценка качества ампликонного секвенирования.

| | | |
|-------------------------------|---------------------------------------------------------|---------------------|
| АНОО ВО «Университет «Сириус» | Рабочая программа дисциплины (модуля) «Метагеномика» | Лист 7 Листов 11 |
|-------------------------------|---------------------------------------------------------|---------------------|

4. ОТЕ и ASV-пикинг — сравнение полученных результатов.
5. Статистический анализ результатов секвенирования — формулирование гипотез и их проверка.
6. Работа с ампликонными прочтениями различных генов (отличных от 16S rRNA).
7. Работа с базами данных, используемых для таксономической аннотации.
8. Оценка качества метагеномной сборки.
9. Биннинг: использование различных подходов и проверка качества полученных геномов.
10. Сравнение полученной таксономической аннотации полного метагенома и данных полученных ампликонным секвенированием.
11. Предобработка длинных прочтений на примере сборки отдельных геномов.
12. Работа с основными базами данных, используемых для функциональной аннотации.

Примерный перечень тем групповых и индивидуальных проектов:

1. Сравнение ампликонных данных для изучения влияния засоления на состав микробного сообщества.
2. Оценка влияния различных подходов обработки данных.
3. Оценка ризосферного эффекта на нескольких сортах злаков.
4. Оценка влияния химерных прочтений на итоговый состав сообществ.
5. Поиск «активного» и «корового» компонента в почвенных сообществах.
6. Сравнение предобработанных данных для пациентов после трансплантации кишечного микробиома.
7. Поиск оверпредставленных генов в целлюлозолитических микробных сообществах.
8. Маркеры патогенности - анализ существующих баз данных.
9. Поиск генов антибиотикоустойчивости в природных сообществах.
10. Различные подходы в поиске мобильных генетических элементов в метагеноме.

Примерный перечень тестовых заданий:

1. Какие меры разнообразия более подходят для анализа динамики малопредставленных таксонов?
 - Индекс Симпсона
 - Индекс Шеннона
 - Индекс Чао1
2. Какие меры бета-разнообразия учитывают филогенетические соотношения между таксономическими группами?

- Морисита
- Брэй-Кертис
- Юнифрак

3. Какие из нижеперечисленных систем секвенирования не используют секвенирование синтезом?

- MinION
- Illumina
- PacBio

4. Какая особенность гена 16S rDNA делает его таким удобным для анализа разнообразия прокариот?

- малая длина
- наличие консервативных и варибельных участков
- наличие этого гена у всех прокариот
- большое количество копий гена на одну клетку

5. Выберите базу данных, которая используется для таксономической аннотации ампликонов 16S rDNA

- SILVA
- GTDB
- KEGG
- UniProt

6. Какой из алгоритмов не применим для данных, полученных при помощи PacBio?

- open-reference OTU picking
- de-novo OTU picking
- ASV picking
- clustering by similarity (97% threshold)

7. Для решения каких задач корректно использовать ампликонное секвенирование?

- скрининг в сточных водах генов устойчивости к антибиотикам
- функциональная характеристика микробного сообщества горячих источников
- клинический анализ на поиск патогенной микрофлоры
- таксономический анализ биопрепаратов для переработки органических отходов

8. Какие из нижеперечисленных программ используют графы де Брюйна

- metaFlye
- metaSPADes
- Kraken2

- MetaPhlAn

9. Какая особенность геномов прокариот делает корректный метагеномный биннинг невозможным?

- малый размер геномов
- наличие мобильных генетических элементов
- высокая вариабильность генов домашнего хозяйства
- метилирование

10. Выберите базу данных, которая используется для таксономической аннотации:

- RPHAM
- GTDB
- KEGG
- MetaCyc

11. Какой из этапов пайплайна обработки специфичен для работы с длинными прочтениями?

- сборка
- бейзколлинг
- биннинг
- денойзинг

12. Для решения каких задач предпочтительно использовать полное метагеномное секвенирование?

- скрининг в сточных водах генов устойчивости к антибиотикам
- функциональная характеристика микробного сообщества горячих источников
- клинический анализ на поиск патогенной микрофлоры
- поиск генов активных в поздних стадиях деградации полиэтилена

3.3. Формой промежуточной аттестации является экзамен.

Результаты промежуточной аттестации оцениваются оценками «отлично», «хорошо», «удовлетворительно» и «неудовлетворительно».

Оценка «отлично», «хорошо», «удовлетворительно» означает успешное прохождение промежуточной аттестации.

3.4. Оценочные материалы для промежуточной аттестации

Перечень вопросов для подготовки к экзамену:

1. Введение в метагеномику/метатранскриптомику.
2. Обзор экспериментальных подходов. Обзор современных платформ высокопроизводительного секвенирования и их особенностей.

3. Таргетное секвенирование. Подходы к оценке сообщества по представленности бактериальных родов и видов.

4. Использование референсных геномов. Картирование прочтений. Безреференсные подходы – сборка фрагментов de novo. Алгоритмы сборки.

5. Оценка разнообразия с помощью анализа k-меров. Биннинг. Алгоритмы биннинга. Контролируемый и неконтролируемый биннинг. Кластеризация.

6. Аннотация генов. Алгоритмы аннотации. Функциональный метагеномный анализ. Сравнительный анализ.

7. Статистический анализ данных метагеномного картирования. Многомерная статистика и специфические метрики для экспериментов сравнения влияния факторов на бактериальные сообщества.

8. Филогенетические деревья. Методы построения, обработки и сравнения деревьев.

9. Эволюционное расстояние. Модели нуклеотидных и аминокислотных замен. Проверка эволюционных гипотез. Гомология последовательностей. Ортологи и паралоги. Филогенетический анализ на основе подходов без выравниваний. Анализ частот встречаемости k-меров.

4. Учебно-методическое и информационное обеспечение дисциплины (модуля)

4.1. Перечень основной литературы:

1. Принципы и методы биохимии и молекулярной биологии / К. Уилсон и Дж. Уолкер. — М.: БИНОМ. Лаборатория знаний, 2015. — 848 с. — ISBN 978-5-9963-1895-7.

2. Нельсон Дэвид, Кокс Майкл. Основы биохимии Ленинджера. В 3 т. — М.: Лаборатория знаний, 2017. — Т. 1. — ISBN 978-5-00101-014-2.

4.2. Перечень дополнительной литературы:

1. Егорова Т.А. Основы биотехнологии / Т.А. Егорова, С.М. Клунова, Е.А. Живухина. – М.: Издат. центр Академия, 2003. – 208 с.

2. Глик Б. Молекулярная биотехнология. Принципы и применение / Б. Глик, Дж. Пастернак. – М.: Мир, 2002. – 589 с.

3. Егоров Н.С. Основы учения об антибиотиках / Н.С. Егоров. – М.: Наука, 2004. – 525 с.

4.3. Перечень современных профессиональных баз данных и ресурсов информационно-телекоммуникационной сети «Интернет»:

1. SEQanswers [Электронный ресурс] <http://seqanswers.com/> - информационный ресурс (дискуссионный форум) о методах высокопроизводительного секвенирования.

| | | |
|-------------------------------|---------------------------------------------------------|----------------------|
| АНОО ВО «Университет «Сириус» | Рабочая программа дисциплины (модуля) «Метагеномика» | Лист 11 Листов 11 |
|-------------------------------|---------------------------------------------------------|----------------------|

2. Курс «Введение в биоинформатику. Метагеномика» [Электронный ресурс]. – <https://ru.coursera.org/learn/bioinformatics>

3. Введение в биоинформатику: метагеномика [Электронный ресурс]. – <https://stepik.org/course/2557/>

5. Материально-техническое и программное обеспечение дисциплины (модуля)

5.1. Материально-техническое обеспечение:

| <i>Вид аудитории</i> | <i>Технические средства и оборудование</i> |
|-----------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <i>Учебная аудитория для проведения лекционных занятий</i> | Альфа 5.3 - учебная аудитория для проведения учебных занятий, предусмотренных программой магистратуры. Доска магнитно-маркерная поворотная BoardSYS Twist 100x160 ПО-15Ф 1 шт. Флипчарт 70*100 на роликах 1 шт. Стол-кафедра 1 шт. Стол аудиторный 1 шт. Столы-трансформеры Summa GA ученические 40 шт. Стулья на колесах ученические 40 шт. Ноутбук HP 1 шт. Интерактивная панель NexTouch Nextpanel 86” 1 шт. Радиосистема Arthur Forty U-9700C PSC (UHF) в комплекте. Акустическая система Behringer B215D 2 шт. Веб-камера 4К с технологией искусственного интеллекта JazzTel JT-Vintage-4K 1 шт. Комплект электронных презентаций. |
| <i>Учебная аудитория для проведения практических занятий – Компьютерный класс</i> | Бета 4.1 – учебная аудитория для проведения практических занятий (компьютерный класс). Доска магнитно-маркерная поворотная BoardSYS Twist 100x160 ПО-15Ф 1 шт. Флипчарт 70*100 на роликах 1 шт. Стол преподавателя аудиторный 1 шт. Столы и стулья ученические 42 шт. Компьютеры Lenovo ThinkCentre M920s SFF в комплекте с мониторами PУАМА 27” и периферией – 42 шт. Интерактивная панель NexTouch Nextpanel 86” 1 шт. Радиосистема Arthur Forty U-9700C PSC (UHF) в комплекте. Акустическая система Behringer B215D 2 шт. Веб-камера 4К с технологией искусственного интеллекта JazzTel JT-Vintage-4K 1 шт. Комплект электронных презентаций. |

5.2. Перечень лицензионного и свободно распространяемого программного обеспечения, в том числе российского производства:

- пакет библиотек для Python (Anaconda);
- инструмент для сборки Haskell (Stack);
- компилятор C++ (clang).

Кроме того, для работы будут использованы программы Gaia, MG rast, Phymm, Galaxy.