

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

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Транскриптомика

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

1. Общая характеристика дисциплины (модуля)

1.1. Цель: освоение современных методов исследования методов, разработанные для изучения транскриптома (то есть совокупности всех РНК-транскриптов) организма.

1.2. Задачи: формирование у студентов теоретических знаний и практических навыков, связанных с секвенированием РНК (RNA-Seq), в котором используются методы секвенирования нового поколения для получения последовательностей и оценки количества всех транскриптов.

1.3. Общая трудоемкость: 2 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, биотехнологии и фарминдустрии, применяя навыки работы с высокотехнологичным лабораторным оборудованием	ИПК-6.1. Применяет классические методы решения задач, современные программные комплексы и навыки работы с высокотехнологичным лабораторным оборудованием для проведения расчетных работ и исследований
	ИПК-6.2. Проводит расчетные работы и исследования, осуществляет обработку, анализ и интерпретацию биомедицинских и биотехнологических данных

АНОО ВО «Университет «Сириус»	Рабочая программа дисциплины (модуля) «Транскриптомика»	Лист 3 Листов 10
----------------------------------	--	---------------------

	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	3 семестр	Всего
Контактная работа обучающихся с преподавателем, всего ч.	42	42
Лекционные занятия, ч.	х	х
Практические (семинарские) занятия, ч.	х	х
Лабораторные занятия, ч.	40	40
Промежуточная аттестация – экзамен, ч	х	х
Промежуточная аттестация – зачет с оценкой, ч	х	х
Промежуточная аттестация – зачет, ч	2	2
Самостоятельная работа обучающихся, всего ч.	30	30
Общая трудоемкость, ч.	72	72
Общая трудоемкость, з.е.	2	2

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
Раздел 1. Экспрессия генов. Транскриптомика. Массовое параллельное секвенирование	х	х	2	х	2	4	письменное домашнее задание
Раздел 2. Данные массового параллельного секвенирования: формат, конвертация, процессинг.	х	х	6	х	2	8	письменное домашнее задание
Раздел 3. Задача картирования прочтений на геном. Особенности картирования транскриптомных последовательностей.	х	х	6	х	6	12	письменное домашнее задание
Раздел 4. Дифференциальная экспрессия генов.	х	х	10	х	8	18	письменное домашнее задание

АНОО ВО «Университет «Сириус»	Рабочая программа дисциплины (модуля) «Транскриптомика»	Лист 4 Листов 10
-------------------------------	--	---------------------

Многомерный анализ данных по экспрессии генов							
Раздел 5. Сборка транскриптома de novo для оценки экспрессии генов	x	x	6	x	6	12	письменное домашнее задание
Раздел 6. Функциональная аннотация транскриптомных последовательностей	x	x	10	x	6	16	письменное домашнее задание
Промежуточная аттестация				2		2	Зачет
Итого	x	x	40	2	30	72	

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

Наименования разделов (тем) дисциплины (модуля)	Содержание разделов (тем) дисциплины (модуля)
Раздел 1. Экспрессия генов. Транскриптомика. Массовое параллельное секвенирование	Введение в анализ транскриптомных данных. Высокопроизводительное секвенирование транскриптома. Подготовка библиотеки к ДНК. Одиночные, парные и цепь-специфичные прочтения. Разновидности высокопроизводительного секвенирования транскриптома (RNA-seq, miRNA-seq, и пр.) Работа с базами данных: GEO NCBI, SRA, UCSC, Ensembl.
Раздел 2. Данные массового параллельного секвенирования: формат, конвертация, процессинг.	Контроль качества и предобработка необработанных прочтений. ПЦР-дубликаты. Инструменты: sra-toolkit, FastQC, cutadapt, TrimGalore, trimmomatic, MultiQC. Компьютерное представление биологических данных; форматы представления: FASTA, FASTQ.
Раздел 3. Задача картирования прочтений на геном. Особенности картирования транскриптомных последовательностей.	Картирование прочтений на референсный геном при помощи алгоритмов STAR и Hisat2. Оценка качества выравнивания. Особенности построения и работы с различными геномными аннотациями. Оценка обогащения прочтениями различных функциональных геномных районов. Программные пакеты: samtools, Picard, RSeQC, QualiMap, MultiQC. Форматы представления данных: gtf, gff, bed, SAM, BAM.
Раздел 4. Дифференциальная экспрессия генов. Многомерный анализ данных по экспрессии генов	Квантификация выравниваний по геномным аннотациям. Различные типы нормализации данных: CPM, TPM, RPKM, TMM, DESeq. Анализ дифференциальной экспрессии генов. Визуализация и анализ качества. Поправки на множественное сравнение. Многомерный анализ экспрессии генов. Инструменты: HTSeq, featureCounts, DESeq, edgeR.
Раздел 5. Сборка транскриптома de novo для оценки экспрессии генов	Основные подходы к de novo сборке транскриптомных данных: SPAdes RNA, Trinity, Velvet. Удаление контаминаций в данных: MCSC, DeconSeq. Анализ качества сборки: BUSCO.

Раздел 6. Функциональная аннотация транскриптомных последовательностей	Функциональный анализ дифференциально экспрессирующихся генов: COG, GO, GSEA. Генные сети. Поиск гомологов, идентификация функциональных доменов, поиск открытых рамок считывания. Идентификация кластеров коэкспрессирующихся генов: GENIE3, WGCNA. Визуализация сетей коэкспрессирующихся генов.
--	--

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.

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
Раздел 1. Экспрессия генов. Транскриптомика. Массовое параллельное секвенирование	Индивидуальный/групповой проект	Комплект заданий для индивидуальных/групповых проектов
Раздел 2. Данные массового параллельного секвенирования: формат, конвертация, процессинг.	Индивидуальный/групповой проект	Комплект заданий для индивидуальных/групповых проектов
Раздел 3. Задача картирования прочтений на геном. Особенности картирования транскриптомных последовательностей.	Индивидуальный/групповой проект	Комплект заданий для индивидуальных/групповых проектов
Раздел 4. Дифференциальная экспрессия генов. Многомерный анализ данных по экспрессии генов	Индивидуальный/групповой проект	Комплект заданий для индивидуальных/групповых проектов
Раздел 5. Сборка транскриптома de novo для оценки экспрессии генов	Индивидуальный/групповой проект	Комплект заданий для индивидуальных/групповых проектов
Раздел 6. Функциональная аннотация транскриптомных последовательностей	Индивидуальный/групповой проект	Комплект заданий для индивидуальных/групповых проектов

3.2. Оценочные материалы для текущего контроля:

Текущий контроль состоит в выполнении учащимися самостоятельных заданий, представляющих собой части групповых/индивидуальных проектов по обработке транскриптомных данных. Обучающийся представляет отчет о выполнении самостоятельных заданий, содержащий постановку задачи, описание предложенного метода её решения и описание результатов анализа выбранных данных. Количество заданий варьируется в зависимости от объёма освоенного раздела.

Примерный перечень тем групповых/индивидуальных проектов:

1. Провести анализ экспериментальных данных приведенных в статье по изучению транскриптома *S. liquefaciens*, сравнить полученные результаты с таковыми в оригинальной статье, определить гены потенциально отвечающие за адаптацию бактерии *S. liquefaciens* к экстремальным условиям, стимулирующие Марсианские, а также проверить наличие генов адаптации у представителей перхлорат утилизирующих бактерий видов *Dechloromonas aromatica*, *Dechloromonas sp. HYN0024*, *Azospira oryzae*, *Sedimenticola thiotaurini*, *Pseudomonas aeruginosa PAO1*.

Предполагаемый план выполнения работы:

Осуществить оценку качества и предобработку RNA-seq данных для *S. liquefaciens* в 4 условиях культивирования;

Провести картирование ридов картирование полученных прочтений на референсный геном;

Провести анализ дифференциальной экспрессии генов;

Провести сборку транскриптома *de novo*;

Провести функциональную аннотацию и идентифицировать функциональные категории, для которых показано обогащение дифференциально экспрессируемыми генами;

Проверить наличие у представителей перхлорат-утилизирующих бактерий видов *Dechloromonas aromatica*, *Dechloromonas sp. HYN0024*, *Azospira oryzae*, *Sedimenticola thiotaurini*, *Pseudomonas aeruginosa PAO1* гомологов генов, отвечающих за адаптацию к Марсианским условиям у *Serratia sp.*

2. Провести анализ экспериментальных данных, приведенных в статье по изучению транскриптома *M. alcaliphilum*, в различных условиях культивирования: присутствие в среде культивирования *Ca* или *La* при росте на метане и при росте в условиях недостатка метана.

Выявить список генов, которые связаны с переключением клеточного метаболизма при росте в присутствии *La*.

Предполагаемый план выполнения работы:

Осуществить оценку качества и предобработку RNA-seq данные для *M. alcaliphilum*, полученных в 4 условиях культивирования, провести их препроцессинг;

Провести картирование прочтений на референсный геном;

Провести анализ дифференциальной экспрессии генов;

Провести сборку транскриптома de novo;

Провести функциональную аннотацию и идентифицировать функциональные категории, для которых показано обогащение дифференциально экспрессируемыми генами;

Выявить кластеры коэкспрессирующихся генов, связанные с адаптацией клеток к росту в условиях ограниченного количества метана и используя эти гены, выявить именно те, которые связаны с адаптацией к росту в присутствии La. Провести сравнение адаптации клетки к ограничению количества метана для условий культивирования в присутствии La и в присутствии Ca.

3.3. Формой промежуточной аттестации является зачет.

Результаты промежуточной аттестации оцениваются как «зачтено» и «не зачтено». Оценка «зачтено» означает успешное прохождение промежуточной аттестации.

3.4. Оценочные материалы для промежуточной аттестации

Перечень вопросов для подготовки к зачету:

1. Основные понятия транскриптомики (ген, транскрипция, трансляция, геном, транскриптом).

2. Последовательности ДНК, РНК, белков. Форматы представления данных.

3. Базы данных геномной информации (UCSC, Ensembl, NCBI).

4. Форматы геномных данных: FASTQ, FASTA, gtf, gff, bed.

5. Выравнивание последовательностей. Сходство последовательностей. Гомология.

6. Функции генов. Функциональная аннотация генов. Генная онтология.

7. Полногеномные данные по экспрессии генов. Базы полногеномных данных по экспрессии генов.

8. Принципы высокопроизводительного секвенирования. Платформы для высокопроизводительного секвенирования. Особенности различных платформ.

9. Высокопроизводительное секвенирование транскриптома. Подготовка библиотеки к ДНК. Одиночные, парные и цепь-специфичные прочтения. Разновидности высокопроизводительного секвенирования транскриптома (RNA-seq, miRNA-seq, и пр.).

10. Формат данных высокопроизводительного секвенирования. Конвертация данных. Основные информационные ресурсы по транскриптомным данным.

11. Ошибки секвенирования и качество прочтений. Программы и алгоритмы оценки качества прочтений. Коррекция ошибок.
12. Основные задачи анализа транскриптома.
13. Форматы записи данных картирования (SAM/BAM).
14. Особенности картирования транскриптомных последовательностей. Программы для картирования транскриптомных данных (HISAT2, STAR).
15. Оценка качества картирования транскриптомов.
16. Покрывание (глубина секвенирования). Контроль качества секвенирования на основании аннотации генома.
17. Оценка уровня экспрессии генов. Подсчет прочтений. Методы нормирования.
18. Сравнение уровней экспрессии генов на уровне транскриптома. Особенности статистического анализа.
19. Основные программы для анализа дифференциальной экспрессии генов. Выбор программы в зависимости от дизайна биологического эксперимента.
20. Оценка вариаций изоформ генов.
21. Многомерный анализ данных по экспрессии генов: кластеризация, анализ главных компонент, корреляция экспрессии генов.
22. Задача сборки последовательностей de novo. Жадный алгоритм сборки контигов.
23. Задача сборки последовательностей de novo
24. Идентификация коэкспрессирующихся кластеров генов. Генные сети.
25. Особенности сборки транскриптомных последовательностей. Программы для сборки транскриптомов.
26. Оценка качества сборки транскриптомных данных.
27. Оценка экспрессии генов по данным de novo сборки транскриптома.
28. Аннотация транскриптомных последовательностей. Выравнивание на геном. Поиск гомологов, идентификация доменов.
29. Аннотация транскриптомных последовательностей. Поиск открытых рамок считывания.

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

4.1. Перечень основной литературы:

1. Ребриков, Д.В., NGS высокопроизводительное секвенирование / Д.В. Ребриков, Д.О. Коростин, Е.С. Шубина, В.В. Ильинский— 2-е изд.— Москва: БИНОМ. Лаборатория знаний, 2015. — 232 с.

4.2. Перечень дополнительной литературы:

1. Korpelainen E., Tuimala J., Somervuo P., Huss M., Wong G. RNA-seq data analysis: A practical approach— Boca Raton: CRS Press, 2014. - 322 p.

2. Введение в информационную биологию и биоинформатику: учеб. пособие: в 5 т. / Отв. ред. Н.А. Колчанов, О.В. Вишнеvский, Д.П. Фурман; Новосибир. гос. ун-т.—Новосибирск: РИЦ НГУ, 2012.

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

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

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

5.1. Материально-техническое обеспечение:

Вид аудитории	Технические средства и оборудование
<i>Учебная аудитория для проведения промежуточной аттестации</i>	Альфа 5.2 - 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>Лаборатория для проведения практикума</i>	10.09 - проектная лаборатория для проведения научных исследований, предусмотренных программой магистратуры. Основное оборудование: И1204 Дозатор 1-кан. 5-50 мкл механический Transferpette S Transferpette; И1207 Дозатор 1-кан. 0,5-10 мкл механический Transferpette S Transferpette; И1209 Дозатор 1-кан. 5-50 мкл механический Transferpette S Transferpette; O3783 Миницентрифуга-вортекс FV-2400 Biosan; O3802 ТрансиллюминаторTCP-20 Vilber; O3803 Термостат твердотельный DKT-100 MIULAB; O3804 Термостат твердотельный DKT-100 MIULAB; O3805 Термостат твердотельный DKT-100 MIULAB; O3809 Центрифуга настольная без охлаждения Mini-4K MIULAB; O3810 Центрифуга настольная без охлаждения Mini-4K MIULAB; O3817 Бокс биологической безопасности класс ПБМБ-II-"Ламинар-

	<p>С."-1,2 SAVVY SL Ламинарные Системы; O3818 Морозильник низкотемпературный DW-HL528S Zhongke Meiling Cryogenics Company Limited; O3819 Спектрофотометр NanoDrop OneC Thermo Fisher Scientific; O3823 Амплификатор MiniAmp Plus Thermo Fisher Scientific; O3824 Амплификатор MiniAmp Plus Thermo Fisher Scientific; O3833 Центрифуга настольная без охлаждения Microfuge 20 Beckman Coulter; O3835 Морозильник фармацевтический DW-FL450 Meling; O3839 Центрифуга настольная с охлаждением SL 16R Thermo Fisher Scientific; O3841 pH-метр ST3100-F OHAUS; O3842 Весы аналитические Pioneer New PX224 OHAUS; O3844 ПЦР-бокс UVC/T-M-AR Biosan; O3847 Амплификатор CFX96 C1000 Bio-Rad; O3878 Холодильник бытовой STN 185 STINOL; O3887 Система гель-документирующая E-Box-CX5.TS Vilber; O3890 Дозатор 1-кан. 500-5000 мкл механический Блэк Thermo Fisher Scientific; O3891 Дозатор 1-кан. 2-20 мкл механический Блэк Thermo Fisher Scientific; O3907 Камера для горизонтального электрофореза MINIE-135 Miulab; O3943 Дозатор 1-кан. 2-20 мкл механический Блэк Thermo Fisher Scientific; O3947 Дозатор 1-кан. 20-200 мкл механический Блэк Thermo Fisher Scientific; O3949 Дозатор 1-кан. 20-200 мкл механический Блэк Thermo Fisher Scientific; O3952 Дозатор 1-кан. 100-1000 мкл механический Блэк Thermo Fisher Scientific; O3953 Дозатор 1-кан. 100-1000 мкл механический Блэк Thermo Fisher Scientific; O4054 Мешалка магнитная с нагревом MSH-300 Biosan; O4175 Дозатор 8-кан. 30-300 мкл механический Research Plus Eppendorf; O4215 Термогигрометр TH-14 RGK; O4288 Дозатор 1-кан. 0,5-10 мкл механический Transferpette S Transferpette; O4289 Дозатор 1-кан. 0,5-10 мкл механический Transferpette S Transferpette; O4292 Дозатор 1-кан. 5-50 мкл механический Transferpette S Transferpette; O4346 Система генерации липосом Dolomite Dolomite Microfluidics; O4485 Вортекс V-1 plus BioSan; O4487 Дозатор 1-кан. 100-1000 мкл механический Assist BioSan; O4488 Дозатор 1-кан. 100-1000 мкл механический Assist BioSan; O4489 Дозатор 1-кан. 20-200 мкл механический Assist BioSan; O4670 Дозатор 1-кан. 20-200 мкл механический MicroPette DLAB Scientific; O4676 Гомогенизатор 15D-SET-A Witeg; O4704 Дозатор 1-кан. 20-200 мкл механический Assist BioSan; O708 Система для электрофореза на микрочипе MCE-202 MultiNA Shimadzu; O7150 Гомогенизатор Q125 Qsonica</p>
--	---

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