

ГОУ ВПО РОССИЙСКО-АРМЯНСКИЙ (СЛАВЯНСКИЙ) УНИВЕРСИТЕТ

Составлен в соответствии с
государственными требованиями к
минимуму содержания и уровню
подготовки выпускников по
указанному направлению 06.05.01.
Биоинженерия и биоинформатика и
Положением «Об УМКД РАУ».

УТВЕРЖДАЮ:

Директор А.А. Аракелян



06

2022г.

протокол № 5

Институт: Биомедицины и Фармации

Кафедра: Биоинженерии, биоинформатики и молекулярной биологии

Специальность: 06.05.01. Биоинженерия и биоинформатика

АВТОР: д.б.н. Аракелян Арсен Арташесович

УЧЕБНО-МЕТОДИЧЕСКИЙ КОМПЛЕКС

Дисциплина: Анализ данных массово-параллельных Экспериментов

ЕРЕВАН

1. Аннотация

Технологии высокопроизводительного секвенирования ДНК — секвенирования следующего и третьего поколения обеспечили бурный рост объемов биологических данных и открыли новые возможности для исследований в областях молекулярной биологии, биомедицины и биотехнологии. Широкий круг применений анализа данных секвенирования в исследовании регуляции экспрессии генов, медицинской геномике, биотехнологических приложениях требует согласования форматов, соблюдения общих стандартов представления информации, с использованием «сырых» данных секвенирования и разработкой общих конвейеров их компьютерной обработки и модельной интерпретации. Цель данной дисциплины — ознакомить студентов с основными технологиями секвенирования, их применениями, контролем качества, аннотации и визуализации геномных результатов, а также обучить их самостоятельно проводить первичную обработку экспериментов, используя соответствующие биоинформатические программы.

2. Требования к исходным уровням знаний и умений студентов

Дисциплина базируется на знаниях, приобретенных студентами при изучении теоретических и методических основ молекулярной биологии, генетики, статистики и программирования.

3. Цель и задачи дисциплины

Цель дисциплины — ознакомить студентов с основными технологиями секвенирования, их применениями, контролем качества, аннотации и визуализации геномных результатов, а также обучить их самостоятельно проводить первичную обработку экспериментов, используя соответствующие биоинформатические программы.

Задачи дисциплины:

Ознакомление с технологиями высокопроизводительного секвенирования

Ознакомление с основными применениями высокопроизводительного секвенирования

Ознакомление с основными типами геномных данных и форматов их хранения

Ознакомление с методами контроля качества экспериментов высокопроизводительного секвенирования

Практическая работа с базами данных нуклеиновых кислот

Практическая работа с программами выравнивания коротких прочтений

Практическая работа с программами манипуляции геномными интервалами

Практическая работа с программами аннотации и визуализации геномных данных

4. Требования к уровню освоения содержания дисциплины

После прохождения дисциплины студент должен:

Знать:

- Принципы и основные технологии высокопроизводительного секвенирования
- Основные применения высокопроизводительного секвенирования
- Достоинства и недостатки существующих платформ высокопроизводительного секвенирования
- Основные форматы геномных данных и интервалов
- Основные базы данных геномов и геномных данных
- Основные критерии качественной сборки геномов
- Основные стратегии и программы для де-ново секвенирования, ре-секвенирования и позиционных методов секвенирования

Уметь:

- Проводить контроль качества прочтений
- Использовать базы данных референсных геномов
- Использовать программы манипуляции и аннотации геномных интервалов
- Использовать программы выравнивания коротких прочтений
- Использовать программы сборки и аннотации геномов
- Визуализировать геномные данные
- Конвертировать и преобразовывать форматы геномных данных

Владеть:

- командной строкой UNIX/Linux для манипуляции геномными интервалами и файлами прочтений
- программами Samtools, Bedtools, IGV genome browser, FastQC
- программами Bowtie2, Blast, Picard tools, igvtools

5. Объем дисциплины и виды учебной работы по рабочему учебному плану

Виды учебной работы	Всего часов	Количество часов по семестрам							
		9 сем.							
			се м.	се м.	се м.	се м.	се м.	се м.	се м.
1	2	3	4	5	6	7	8	9	10
1. Общая трудоемкость изучения дисциплины по семестрам, в т. ч.:	72	72							
1.1. Аудиторные занятия, в т. ч.:	36	36							
1.1.1. Лекции	18	18							

[illegible]

Распределение ресурсов на фермам контроля и оценки академической успеваемости

	й оценке текущего контроля			оценки текущего контроля в итоговой оценке промежуточного контроля			в результатирую щей оценке промежуточн ого контроля	контролей и оценки итогового контроля в результатирующе й оценке итогового контроля
	M1	M2	M3	M1	M2	M3		
Вид учебной работы/контроля								
Контрольная работа						1		
Тест								
Курсовая работа								
Лабораторные работы								
Письменные домашние задания								
Эссе								
Другие формы Семинар								
Другие формы (добавить)								
Другие формы (добавить)								
Вес результативной оценки текущего контроля в итоговых оценках промежуточных контролей					0			
Вес итоговой оценки 1-го промежуточного контроля в результатирующей оценке промежуточных контролей							0	

Вес итоговой оценки 2-го промежуточного контроля в результатирующей оценке промежуточных контролей							0	
Вес итоговой оценки 3-го промежуточного контроля в результатирующей оценке промежуточных контролей т.д.							1	
Вес результирующей оценки промежуточных контролей в результатирующей оценке итогового контроля								1
Экзамен/зачет (оценка итогового контроля)								0
	$\Sigma = 1$	$\Sigma = 1$	$\Sigma = 1$	$\Sigma = 1$	$\Sigma = 1$	$\Sigma = 1$	$\Sigma = 1$	$\Sigma = 1$

7. Содержание дисциплины:

7.1. Тематический план (Разделы дисциплины и виды занятий) по учебному плану:

Разделы и темы дисциплины	Всего ак. часов	Лекц ии, ак. часов	Прак т. занят ия, ак. часов	Семина ры, ак. часов	Лабор , ак. часов	Контр ольна я работа
1	3=4+5+6 +7+8	4	5	6	7	8
Тема 1. Введение в молекулярную	4	2	2			

биологию и секвенирование.						
Тема 2. Основные задачи и применения высокопроизводительных методов секвенирования.	8	4	4			
Тема 3. Работа с данными высокопроизводительного секвенирования: прочтения и контроль качества.	8	4	4			
Тема 4. Выравнивание прочтений.	4	2	2			
Тема 5. Статистика выравнивания на геном и транскриптом.	4	2	2			
Тема 6. Геномная визуализация.	4	2	2			
Тема 7. Взаимопревращение форматов данных.	4	2	2			
Всего	36	18	18			

7.2. Содержание разделов и тем дисциплины:

Тема 1. Введение в молекулярную биологию и секвенирование. Основные концепции молекулярной биологии. Геномы эукариот и прокариот. Компоненты клетки. Полимеразная цепная реакция (ПЦР). Принцип секвенирования. Секвенирование по Сэнгеру. Обзор технологий высокопроизводительных методов секвенирования.

Тема 2. Основные задачи и применения высокопроизводительных методов секвенирования. Приготовление типичной библиотеки для высокопроизводительного секвенирования. Основные применения высокопроизводительного секвенирования. Методы de novo: сборка геномов и транскриптомов. Методы ресеквенирования в применении к геномам и транскриптомам. Позиционные методы: MNase-seq, DNase-seq, ATAC-seq. Позиционные методы: ChIP-seq. Методы определения 3D-структуры ДНК: 3C, 4C, 5C, Hi-C.

Тема 3. Работа с данными высокопроизводительного секвенирования: прочтения и контроль качества. Референсные геномы. Форматы FASTA и FASTQ. Геномные интервалы. Формат BED. Аннотация геномов. Форматы GTF и GFF. Контроль качества сырых прочтений: FastQC.

Тема 4. Выравнивание прочтений. Концепция и алгоритмы выравнивания последовательностей. BLAST и его варианты. Форматы SAM, BAM и CRAM. Использование samtools.

Тема 5. Статистика выравнивания на геном и транскриптом. Сложность библиотеки. Дубликаты прочтений и их типы. Мульти-выравнивания и выравниваемость. Выравнивание на транскриптом. Ните-специфичность, загрязнение рибосомной РНК, особенности покрытия

Тема 6. Геномная визуализация. Визуализация покрытия, интервалов, и прочтений. Samtools tview. Геномные браузеры UCSC и IGV. Индексация и визуализация BAM. Форматы для отображения покрытия. Учет дубликатов. Отображение вариаций в геноме. Важные опции в IGV. Использование дополнительных треков и аннотаций. Браузеры ExAc и CTEH.

Тема 7. Взаимопревращение форматов данных. Форматы для покрытия. Форматы для интервалов. Форматы для аннотации. Утилиты UCSC.

8. Учебно-методическое обеспечение дисциплины

8.1. Рекомендуемая литература:

1. Д.В. Ребриков. NGS: высокопроизводительное секвенирование. БИНОМ. Лаборатория знаний 2014
2. István Albert. The Biostar Handbook: 2nd Edition.
<https://www.biostarhandbook.com/index.html>

Адреса электронных ресурсов

1. Cell Biology by the Numbers: <http://book.bionumbers.org/>
2. An Introduction to Molecular Biology:
https://en.wikibooks.org/wiki/An_Introduction_to_Molecular_Biology
3. <http://seqanswers.com/forums/index.php>
4. <https://www.biostars.org/>

8.2. Материально-техническое обеспечение дисциплины

Компьютер, проектор, мультимедийные средства, программное обеспечение и т.д.