
УТВЕРЖДЕНА

Решением Ученого совета
АНО ВО «Центральный университет»
«25» мая 2026 г.
Протокол № 2

**Рабочая программа дисциплины (модуля)
«Современное высокопроизводительное секвенирование»**

Направление подготовки: 02.04.01 Математика и компьютерные науки

Направленность (профиль) подготовки: III в биотехе

Специализация: Прикладная биоинформатика

Квалификация (степень) выпускника: магистр

Форма обучения: очная

Срок освоения программы: 2 года

Год набора: 2026

**Москва
2026**

Содержание

1. Краткая характеристика дисциплины (модуля)	3
2. Перечень планируемых результатов обучения.....	5
3. Тематический план.....	8
4. Содержание дисциплины (модуля).....	8
5. Учебно-методическое обеспечение	9
6. Материально-техническое обеспечение	9
7. Методические и оценочные материалы	11

1. Краткая характеристика дисциплины (модуля)

Рабочая программа дисциплины (модуля) «Современное высокопроизводительное секвенирование» составлена в соответствии с федеральным государственным образовательным стандартом высшего образования – магистратура по направлению 02.04.01 Математика и компьютерные науки, профиль ИИ в биотехе, утвержденный приказом Министерства науки и высшего образования Российской Федерации № 810 от 23.08.2017 года.

Изучение дисциплины (модуля) «Современное высокопроизводительное секвенирование» позволяет сформировать комплексное понимание технологического и аналитического цикла работы с геномными данными, необходимое для участия в современных научных и прикладных проектах в области геномики, персонализированной медицины и биотехнологии; владение методами NGS и навыками воспроизводимого анализа обеспечивает готовность к работе с большими данными, критической оценке результатов и эффективному использованию как локальных, так и облачных вычислительных ресурсов.

Место дисциплины (модуля) в структуре образовательной программы

Настоящая дисциплина (модуль) включена в учебный план по программе подготовки магистратуры по направлению 02.04.01 Математика и компьютерные науки, профиль ИИ в биотехе и входит в обязательную часть Блока 1 как дисциплина по выбору, подлежащая обязательному изучению на специализации «Прикладная биоинформатика».

Дисциплина (модуль) изучается на 1 курсе во 2 семестре.

Цель изучения дисциплины (модуля): формирование у обучающихся систематических знаний о принципах, технологиях и методах современного высокопроизводительного секвенирования, а также развитие практических навыков обработки, анализа и интерпретации геномных данных для решения биологических, медицинских и биотехнологических задач.

Задачи изучения дисциплины (модуля):

- освоить принципы работы основных платформ NGS и сравнить их по ключевым параметрам: длине прочтений, точности, стоимости и пропускной способности;
- изучить этапы подготовки образцов и построения библиотек для различных типов секвенирования;
- научиться выбирать оптимальную стратегию секвенирования в зависимости от биологической задачи (WGS, RNA-seq, метагеномика и др.);
- освоить форматы хранения геномных данных и методы контроля их качества;
- развить навыки биоинформатической обработки данных: от предварительной очистки до анализа генетических вариантов и визуализации результатов.

В результате освоения дисциплины (модуля) обучающийся должен:

знать:

- принципы работы современных платформ высокопроизводительного секвенирования (NGS): Illumina, PacBio, Oxford Nanopore, Ion Torrent и др;
- этапы подготовки образцов для секвенирования: фрагментация ДНК/РНК, построение библиотек, амплификация;
- основные типы секвенирования: whole genome sequencing (WGS), whole exome sequencing (WES), RNA-seq, ChIP-seq, ATAC-seq, метагеномное секвенирование и

др;

- форматы хранения данных: FASTQ, SAM/BAM, VCF, GFF/GTF и их структуру;
- принципы контроля качества данных (QC): оценка качества прочтений (read quality), дубликации, контаминации, GC-состава;
- основы биоинформатической обработки данных: выравнивание (alignment), сборка (assembly), аннотирование, анализ вариантов (SNP/INDEL);
- этические и юридические аспекты работы с геномными данными, включая конфиденциальность и биоэтику;
- возможности и ограничения разных технологий секвенирования (по длине прочтений, точности, стоимости, пропускной способности).

уметь:

- выбирать подходящую технологию и стратегию секвенирования в зависимости от биологической задачи;
- оценивать качество исходных данных с помощью инструментов (FastQC, MultiQC и др.);
- проводить предварительную обработку данных: удаление адаптеров, фильтрация низкокачественных прочтений (Trimmomatic, Cutadapt);
- выполнять выравнивание прочтений на референсный геном (с использованием BWA, Bowtie2, STAR и др.);
- осуществлять анализ вариантов (с помощью GATK, FreeBayes, bcftools);
- интерпретировать результаты секвенирования в контексте биологических и медицинских задач;
- работать с командной строкой Linux и основными инструментами биоинформатики;
- использовать системы управления проектами и воспроизводимого анализа (Snakemake, Nextflow).

владеть:

- навыками работы с биоинформатическими пайплайнами для анализа NGS-данных;
- практическими навыками использования облачных и HPC-платформ для обработки больших объемов геномных данных;
- навыками визуализации геномных данных (IGV, R/Bioconductor, Python);
- методами интерпретации и представления результатов секвенирования в научных публикациях;
- навыками критического анализа научных статей, использующих технологии NGS;
- работой с базами данных (NCBI, ENA, 1000 Genomes, gnomAD, TCGA и др.);
- основами автоматизации и документирования биоинформатических анализов.

2. Перечень планируемых результатов обучения

Компетенции, формируемые в результате освоения дисциплины (модуля) при проведении учебных занятий в форме контактной работы обучающихся с педагогическими работниками Университета и в форме самостоятельной работы обучающихся:

Компетенция	Содержание компетенции	Индикатор компетенции	Перечень планируемых результатов обучения по дисциплине (модулю)
УК-1.	Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий	УК-1.1.	Знает основные принципы системного подхода и методы критического анализа, а также теоретические основы стратегического планирования и принятия решений
		УК-1.2.	Умеет применять методы системного анализа для выявления ключевых проблемных ситуаций, формулировать гипотезы и разрабатывать альтернативные стратегии действий на основе полученных данных
		УК-1.3.	Имеет практический опыт в проведении анализа реальных проблемных ситуаций в рамках проектов, способен вырабатывать и обосновывать стратегии действий, учитывая различные аспекты и последствия
УК-3.	Способен решать задачи профессиональной деятельности и применять методы искусственного интеллекта в прикладных и фундаментальных исследованиях в области биологии, медицины или биотехнологий	УК-3.1.	Знает принципы и методы решения профессиональных задач, включая применение профильно-специализированных знаний в области ИИ для фундаментальных и прикладных исследований в области биологии, медицины или биотехнологий
		УК-3.2.	Умеет применять математические и компьютерные методы, формулировать результаты и анализировать их последствия для научных и практических решений
		УК-3.3.	Имеет опыт участия в проектах, где методы ИИ применялись для решения исследовательских и прикладных задач в области биологии, медицины или биотехнологий
ПК-1.	Способен определять общие формы и закономерности области искусственного	ПК-1.1.	Знает основные теоретические концепции и принципы искусственного интеллекта в биотехнологии, а также методы обработки, анализа и визуализации данных в биологии, медицине или биотехнологии

	интеллекта в биологии, медицине или биотехнологии, обрабатывать и интерпретировать данные биологических исследований	ПК-1.2.	Умеет проводить систематический анализ области профессиональной деятельности, выявлять закономерности и тенденции, моделировать основные биологические процессы и интерпретировать полученные результаты с использованием методов биоинформатики и искусственного интеллекта
		ПК-1.3.	Имеет практический опыт работы в области искусственного интеллекта в биотехнологии, включая участие в научных проектах, исследованиях или практических заданиях
ПК-2.	Способен математически корректно ставить и решать естественнонаучные и прикладные задачи в области биологии, медицины или биотехнологий	ПК-2.1.	Знает основные методы математического моделирования и теоретические основы современной биологии и биоинформатики, необходимые для корректной постановки задач
		ПК-2.2.	Умеет формулировать математические модели биологических процессов, выбирать адекватные методы решения и интерпретировать результаты в профессиональном контексте
		ПК-2.3.	Имеет практический опыт постановки и решения прикладных задач с использованием методов ИИ в области биологии, медицины или биотехнологий
ПК-3.	Способен решать задачи профессиональной деятельности и применять методы искусственного интеллекта в прикладных и фундаментальных исследованиях в области биологии, медицины или биотехнологий	ПК-3.1.	Знает принципы и методы решения профессиональных задач, включая применение профильно-специализированных знаний в области ИИ для фундаментальных и прикладных исследований в области биологии, медицины или биотехнологий
		ПК-3.2.	Умеет применять математические и компьютерные методы, формулировать результаты и анализировать их последствия для научных и практических решений
		ПК-3.3.	Имеет опыт участия в проектах, где методы ИИ применялись для решения исследовательских и прикладных задач в области биологии, медицины или биотехнологий

ПК-5.	Способен передавать результаты прикладных задач в виде профессиональных рекомендаций в области искусственного интеллекта, биоинформатики и биотехнологии	ПК-5.1.	Знает методы формулирования рекомендаций и терминологию области искусственного интеллекта и биоинформатики
		ПК-5.2.	Умеет интерпретировать результаты анализа и моделирования и формулировать рекомендации для практического применения в области биологии, медицины или биотехнологий
		ПК-5.3.	Имеет опыт разработки рекомендаций по результатам проектов в области искусственного интеллекта и биотехнологии

3. Тематический план

№ п/п	Наименование раздела дисциплины (модуля)	Трудоемкость, академические часы				ТКУ (текущий контроль успеваемости)
		Очная форма				
		Аудиторная работа		Контроль	Самостояте льная работа	
		Лекции	Семинары (практичес кие занятия)			
1	Основы NGS и подготовка данных	9	9		33	Домашнее задание Мини-проект
2	Анализ геномных и транскриптомных данных	9	9		33	Домашнее задание Мини-проект
3	Интерпретация и визуализация результатов	7	7		30	Домашнее задание Мини-проект
4	Инфраструктура и практическое применение	4	4		28	Домашнее задание Мини-проект
	Экзамен			6		
Итого:		30	30	6	124	
Объем дисциплины (модуля) (в ак. ч.)		190				
Объем дисциплины (модуля) (в зач. ед.)		5				

4. Содержание дисциплины (модуля)

№п/п	Наименование раздела дисциплины (модуля)	Содержание дисциплины (модуля) по темам
1	Основы NGS и подготовка данных	Введение в высокопроизводительное секвенирование (NGS). Технологии секвенирования: Illumina, Ion Torrent, PacBio, Oxford Nanopore. Подготовка образцов и построение библиотек. Контроль качества данных (QC). Предобработка NGS-данных.
2	Анализ геномных и транскриптомных данных	Выравнивание прочтений на референсный геном. Сборка геномов de novo. Секвенирование экзона и генома (WES, WGS). RNA-seq: анализ экспрессии генов. ChIP-seq и ATAC-seq: исследование регуляторных областей. Метагеномное секвенирование.
3	Интерпретация и визуализация результатов	Анализ вариантов (SNP, INDEL). Визуализация геномных данных.
4	Инфраструктура и практическое применение	Облачные платформы и HPC в биоинформатике.

5. Учебно-методическое обеспечение

Университет располагает полным набором лицензионного и свободно распространяемого программного обеспечения, включая продукты отечественного производства.

Каждый студент в течение всего периода обучения получает индивидуальный неограниченный доступ к электронно-библиотечной системе и электронной информационно-образовательной среде университета. Эти системы предоставляют возможность доступа к ресурсам из любой точки, где есть подключение к сети Интернет, как на территории университета, так и за его пределами.

Студентам обеспечен удаленный доступ к современным профессиональным базам данных и информационным справочным системам.

Основная литература:

1. Коничев, А. С. Молекулярная биология : учебник для вузов / А. С. Коничев, Г. А. Севастьянова, И. Л. Цветков. — 5-е изд. — Москва : Издательство Юрайт, 2026. — 422 с. — (Высшее образование). — ISBN 978-5-534-13468-1. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/586916>.

2. Костерин, О. Э. Молекулярная генетика : учебник для вузов / О. Э. Костерин, В. К. Шумный ; ответственный редактор В. К. Шумный. — Москва : Издательство Юрайт, 2025. — 683 с. — (Высшее образование). — ISBN 978-5-534-18819-6. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/568926>.

Дополнительная литература:

1. Нефедова, Л. Н. Применение молекулярных методов исследования в генетике : учебное пособие / Л. Н. Нефедова. — Москва : ИНФРА-М, 2026. — 104 с. — (Высшее образование). — ISBN 978-5-16-019028-0. — Текст : электронный. — URL: <https://znanium.ru/catalog/product/2224072>.

6. Материально-техническое обеспечение

Университет располагает материально-технической базой, соответствующей действующим противопожарным правилам и нормам и обеспечивающей проведение всех видов дисциплинарной и междисциплинарной подготовки, практической и научно-исследовательской работ обучающихся, предусмотренных учебным планом.

Помещения, которые представляют собой учебные аудитории для проведения занятий лекционного типа, занятий семинарского (практического) типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, а также помещения для самостоятельной работы и помещения для хранения и профилактического обслуживания учебного оборудования. Помещения укомплектованы специализированной мебелью и техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Изучение дисциплины (модуля) обеспечивается в учебных аудиториях, оснащенных:

- столами и стульями;
- компьютерной техникой;
- механическими калькуляторами;
- специализированным оборудованием, включая демонстрационное оборудование.

Электронный документ

Помещения для самостоятельной работы обучающихся, в том числе приспособленные для использования инвалидами и лицами с ограниченными возможностями здоровья, оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду Университета.

Обучающимся предоставляется доступ (в том числе удаленный) к ресурсам информационно-телекоммуникационной сети «Интернет», электронным ресурсам (в том числе электронным библиотечным системам, современным профессиональным базам данных и информационным справочным системам):

№	Наименование портала (издания, курса, документа)	Ссылка
1.	Научная электронная библиотека elibrary.ru библиотека	https://elibrary.ru/defaultx.asp
2.	База данных для IT-специалистов	https://habr.com
3.	База данных ScienceDirect	https://www.sciencedirect.com
4.	Официальный сайт Министерства науки и высшего образования Российской Федерации	https://minobrnauki.gov.ru/
5.	Федеральный портал «Российское образование»	https://www.edu.ru/
6.	Информационная система "Единое окно доступа к образовательным ресурсам"	http://window.edu.ru/
7.	Единая коллекция цифровых образовательных ресурсов	http://school-collection.edu.ru/
8.	Федеральный центр информационно - образовательных ресурсов	http://fcior.edu.ru/

Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), в том числе комплект лицензионного программного обеспечения, современные профессиональные базы данных и информационные справочные системы:

Наименование ПО	Производство	Лицензионное / свободно распространяемое
Операционные системы:		
Microsoft Imagine (Windows Client, Server)	зарубежное	лицензионное
Браузеры:		
Яндекс.Браузер	отечественное	свободно распространяемое
Google Chrome	зарубежное	свободно распространяемое
Офисные приложения:		
Microsoft Imagine (Visio, OneNote)	зарубежное	лицензионное
TeXstudio	зарубежное	свободно распространяемое
Adobe Acrobat Reader	зарубежное	свободно распространяемое
Программное обеспечение для планирования и учета времени:		
Toggle app	зарубежное	свободно распространяемое
Системы управления проектами:		
Microsoft Imagine (Project)	зарубежное	лицензионное
Системы управления базами данных:		

Microsoft Imagine (SQL Server)	зарубежное	лицензионное
Системы резервного копирования (backup):		
Acronis Backup Advanced for HyperV	зарубежное	лицензионное
Справочно-правовые системы:		
КонсультантПлюс: справочно-правовая система	отечественное	лицензионное
Средства антивирусной защиты:		
Kaspersky Endpoint Security для бизнеса Стандартный Russian Edition	отечественное	лицензионное
Среды разработки:		
Visual Studio Code	зарубежное	свободно распространяемое
Bash (Unix shell)	зарубежное	свободно распространяемое
Anaconda	зарубежное	свободно распространяемое
Robotic Operating System	зарубежное	свободно распространяемое
CopelliaSim	зарубежное	свободно распространяемое
Google Colaboratory	зарубежное	свободно распространяемое
Пакеты программных средств и библиотек:		
AutoPsy	зарубежное	свободно распространяемое
Interactive Disassembler (IDA)	зарубежное	свободно распространяемое
Системы управления библиографической информацией:		
Zotero	зарубежное	свободно распространяемое
Сервисы и службы:		
Bind	зарубежное	свободно распространяемое
Docker	зарубежное	свободно распространяемое

7. Методические и оценочные материалы

Методические указания для обучающихся по освоению дисциплины (модуля)

В процессе изучения дисциплины (модуля) «Современное высокопроизводительное секвенирование» в рамках текущего контроля успеваемости используются такие виды учебной работы, как лекции, семинары, домашние задания, мини-проект, а также различные виды самостоятельной работы обучающихся по заданию преподавателя, направленные на развитие навыков профессиональной лексики, закрепление практических профессиональных компетенций, поощрение инициатив.

Лекция – систематическое, последовательное, монологическое изложение преподавателем учебного материала, как правило, теоретического характера.

В процессе лекций рекомендуется вести конспект лекций: кратко и схематично фиксировать основные идеи, выводы и обобщения лекции; выделять важные мысли, ключевые слова и термины. Необходимо отметить вопросы или материалы, которые вызывают затруднения, и попытаться найти ответы в рекомендованной литературе. Если разобраться в материале не удастся, следует сформулировать вопрос и задать его преподавателю на консультации или во время семинарского (практического) занятия.

Семинар – это форма учебной деятельности, проводимая в учебном заведении под руководством преподавателя, где студенты активно участвуют в обсуждениях, Электронный документ

практических заданиях и других формах взаимодействия.

Для успешной подготовки к семинару рекомендуется заранее ознакомиться с темой занятия и основными материалами, чтобы иметь возможность активно участвовать в обсуждении. Также полезно подготовить вопросы и идеи для обсуждения, что поможет глубже понять материал и продемонстрировать заинтересованность.

Домашнее задание – набор задач по темам недели.

При работе над домашними заданиями важно внимательно ознакомиться с требованиями и сроками выполнения. Рекомендуется разбивать задания на этапы, чтобы избежать перегрузки и лучше усвоить материал, использовать различные источники информации, включая учебники и онлайн-ресурсы, для более глубокого понимания темы.

Мини-проект — это краткосрочная практическая работа в малых группах на семинарах, направленная на применение теоретических знаний для решения конкретных задач.

В ходе мини-проекта студенты проводят анализ проблемы, разрабатывают техническое решение и представляют результаты в виде отчета, модели или презентации. Такой формат способствует развитию навыков самостоятельной работы, критического мышления и командного взаимодействия.

Самостоятельная работа – работа студентов, направленная на углубленное изучение отдельных тем и вопросов учебной дисциплины (модуля).

В процессе самостоятельной работы студенты взаимодействуют с рекомендованными материалами при минимальном участии преподавателя. Задачи студента включают работу с конспектами лекций (обработка текста), повторное изучение учебных материалов планов и тезисов ответов, изучение дополнительных тем, выполнение учебно-исследовательских заданий и другое.

Система оценивания результатов обучения по дисциплине (модулю)

Критерии получения уровня и оценивания сформированности компетенций по дисциплине (модулю) «Современное высокопроизводительное секвенирование»

Оценивание уровня учебных достижений обучающихся по дисциплине (модулю) осуществляется в виде текущего контроля успеваемости и промежуточной аттестации.

Промежуточная аттестация по дисциплине (модулю) осуществляется в форме **экзамена**, при этом проводится оценка компетенций, сформированных по дисциплине.

Для оценивания текущего контроля успеваемости и промежуточной аттестации используется десятибалльная шкала оценивания, которая соотносится с традиционной пятибалльной шкалой следующим образом, если иное не указано в силлабусе дисциплины:

Десятибалльная оценка	Пятибалльная оценка	Общая характеристика результата обучения по дисциплине (модулю)
10	Отлично	Студент полностью владеет знаниями, изложенными в рабочей программе, и глубоко осмысляет дисциплину (модуль). Он самостоятельно и логически последовательно отвечает на все вопросы, акцентируя внимание на наиболее важном. Умеет анализировать, сравнивать, классифицировать, обобщать, конкретизировать и систематизировать изученный
9	Отлично	
8	Отлично	

Десятибалльная оценка	Пятибалльная оценка	Общая характеристика результата обучения по дисциплине (модулю)
		материал, выделяя ключевые моменты и устанавливая причинно-следственные связи. Четко формулирует ответы, уверенно интерпретирует результаты анализов и других исследований, а также решает сложные задачи. Студент хорошо знаком с методами исследования, необходимыми для практической деятельности, и умеет связывать теоретические аспекты дисциплины (модуля) с практическими задачами.
7	Хорошо	Студент обладает знаниями предмета почти в полном объеме рабочей программы и самостоятельно, логически последовательно и всесторонне отвечает на все вопросы, акцентируя внимание на наиболее значимых моментах. Он умеет анализировать, сравнивать, классифицировать, обобщать, конкретизировать и систематизировать изученный материал, выделяя его ключевые аспекты и устанавливая причинно-следственные связи. Формулирует свои ответы, уверенно интерпретирует результаты анализов и других исследований, а также решает сложные ситуационные задачи. Студент хорошо знаком с методами исследования, необходимыми для практической деятельности, и умеет связывать теоретические аспекты предмета с практическими задачами.
6	Хорошо	
5	Удовлетворительно	Студент обладает базовыми знаниями по дисциплине (модулю), но испытывает трудности при самостоятельных ответах и использует неточные формулировки. В ходе ответов он допускает ошибки, касающиеся сути вопросов. Студент способен решать только самые простые задачи и владеет лишь минимальным набором методов исследования.
4	Удовлетворительно	
3	Не сдан	Студент не овладел обязательным минимумом знаний по предмету и не может ответить на вопросы, даже если преподаватель задает дополнительные наводящие вопросы.
2	Не сдан	
1	Не сдан	

Дисциплина (модуль) «Современное высокопроизводительное секвенирование» оценивается следующим образом:

Активность	Вес	Количество	Описание
Домашние задания	40%	10	Набор задач по темам недели
Мини-проект	30%	1	Краткосрочная практическая работа в малых группах на занятиях, направленная на решение конкретных задач
Экзамен	30%	1	Письменная или устная работа над заданием, направленным на проверку полученных знаний и навыков по дисциплине (модулю)

Формула расчёта итоговой оценки по дисциплине (модулю) «Современное высокопроизводительное секвенирование»: $0,4 \times \text{Среднее за домашние задания} + 0,3 \times \text{Мини-проект} + 0,3 \times \text{Экзамен}$.

Текущий контроль успеваемости обучающихся по дисциплине (модулю)

Примерные домашние задания

Домашнее задание 1. Основы NGS и подготовка данных

1. Сравните технологии секвенирования Illumina, Oxford Nanopore и PacBio по следующим параметрам:
 - длина прочтений,
 - точность,
 - принцип детекции,
 - стоимость,
 - типичные области применения.
2. Опишите этапы построения библиотеки для WGS на платформе Illumina.
3. Скачайте FASTQ-файл из базы данных SRA (например, SRR123456).
4. Используя FastQC, проведите контроль качества: оцените распределение качества по циклам, GC-состав, уровень дубликации.
5. С помощью Trimmomatic или Cutadapt выполните предобработку: удалите адаптеры, обрежьте низкокачественные основания ($Q < 20$), отфильтруйте короткие риды (< 36 п.н.).
6. Оформите отчёт (2–3 стр.) с графиками и выводами о качестве данных до и после обработки.

Домашнее задание 2. Анализ геномных и транскриптомных данных

1. Выровняйте предобработанные риды из ДЗ1 на референсный геном (например, GRCh38) с помощью BWA (для WGS) или STAR (для RNA-seq).
2. Преобразуйте SAM-файл в BAM, отсортируйте и проиндексируйте с помощью samtools.
3. Для RNA-seq: подсчитайте количество ридов на гены с помощью featureCounts.
4. Для WGS: используя bcftools, выполните вызов SNP и коротких инделов.
5. Опишите, в каких случаях используется de novo-сборка, а в каких — выравнивание на референс.
6. Оформите результаты в виде краткого отчёта с командами и пояснениями.

Домашнее задание 3. Визуализация и интерпретация результатов

1. Загрузите VCF-файл с вызванными вариантами (можно использовать результат из ДЗ2 или из 1000 Genomes).
2. Используя IGV (Integrative Genomics Viewer), визуализируйте участок хромосомы с несколькими SNP.
3. Найдите вариант в известном гене (например, BRCA1), определите его тип (missense, intronic и др.).
4. Используя R (библиотека ggplot2) или Python (matplotlib), постройте:
 - гистограмму распределения качества вариантов,

- график плотности SNP по хромосоме.
5. Оформите отчёт с изображениями и интерпретацией: какие участки генома наиболее вариабельны, какие фильтры можно применить.

Примерное описание задания и критерии оценивания к мини-проекту

Студент выполняет мини-исследование, охватывающее полный цикл анализа NGS-данных — от выбора технологии до интерпретации результатов.

Темы (на выбор):

- Анализ SNP в экзомных данных пациента с наследственным заболеванием.
- Сравнение экспрессии генов в опухолевой и нормальной ткани (RNA-seq).
- Метагеномный анализ образца кишечной микробиоты.
- Сборка генома бактерии de novo с длинными ридами (Nanopore).
- Анализ открытых хроматиновых участков (ATAC-seq) в клеточной линии.

Этапы проекта:

1. Формулировка биологической задачи и выбор подходящей технологии секвенирования.
2. Загрузка реальных данных (из SRA, ENA, TCGA).
3. Контроль качества (FastQC), предобработка (Trimmomatic).
4. Основной анализ: выравнивание, подсчёт, вызов вариантов, сборка или анализ регуляторных областей.
5. Визуализация (IGV, R/Python) и биологическая интерпретация.
6. Оформление: Jupyter-ноутбук, R Markdown или PDF-отчёт (8–12 стр.) + презентация (5–7 слайдов).

Критерии оценивания мини-проекта:

1. Полнота и корректность анализа
— Все этапы выполнены правильно: QC, предобработка, выравнивание, анализ, визуализация.
2. Обоснованность выбора методов и технологии
— Обоснован выбор платформы, инструментов, стратегии в зависимости от задачи.
3. Качество визуализации и интерпретации
— Графики информативны, подписаны, сопровождаются биологической или медицинской интерпретацией.
4. Воспроизводимость и документирование
— Чёткие команды, указание версий программ, структурированный код, README.
5. Критическое мышление и выводы
— Обсуждение ограничений, возможных артефактов, альтернативных объяснений, этических аспектов.

Задания для промежуточной аттестации по дисциплине (модулю)

№ п/п	Задание	Ответ	Компетенция
1.	Какая технология NGS основана на измерении изменения pH при встраивании нуклеотидов?	Ion Torrent	УК-1

2.	Какой метод используется для контроля качества FASTQ-данных?	FastQC	УК-3
3.	Какой формат используется для хранения выровненных ридов?	SAM/BAM	ПК-1
4.	Что означает аббревиатура WGS?	Полное секвенирование генома	ПК-2
5.	Какой инструмент подходит для выравнивания ридов при RNA-seq?	STAR	ПК-3
6.	Какой формат хранит информацию о генетических вариантах?	VCF	ПК-5
7.	Что такое de novo-сборка?	Сборка генома без использования референсной последовательности	УК-1
8.	Какой платформой характерно секвенирование сверхдлинных ридов?	Oxford Nanopore	УК-3
9.	Какой инструмент используется для удаления адаптеров из ридов?	Cutadapt, Trimmomatic	ПК-1
10.	Что показывает график per-base sequence quality в FastQC?	Качество прочтений по позициям в риде	ПК-2
11.	Какой тип секвенирования используется для анализа экспрессии генов?	RNA-seq	ПК-3
12.	Что позволяет выявить ATAC-seq?	Открытые участки хроматина (регуляторные области)	ПК-5
13.	Какой инструмент используется для вызова SNP?	GATK, bcftools, FreeBayes	УК-1
14.	Что такое метагеномное секвенирование?	Секвенирование ДНК из смешанного образца (например, микробиом)	УК-3
15.	Какой формат используется для аннотации генов?	GFF/GTF	ПК-1
16.	Какой ресурс подходит для хранения больших объёмов геномных данных?	Облачные платформы (AWS, Google Cloud), HPC	ПК-2
17.	Что такое IGV?	Инструмент для	ПК-3

		визуализации геномных данных	
18.	Какой тип данных получают при ChIP-seq?	Информацию о связывании белков с ДНК	ПК-5
19.	Что означает дубликация ридов в данных NGS?	Многократное секвенирование одной и той же молекулы ДНК	УК-1
20.	Зачем использовать Snakemake или Nextflow?	Для автоматизации и воспроизводимости пайплайнов	УК-3