
УТВЕРЖДЕНА

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

**Рабочая программа дисциплины (модуля)
«Алгоритмы биоинформатики»**

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	10
6. Материально-техническое обеспечение	10
7. Методические и оценочные материалы	12

1. Краткая характеристика дисциплины (модуля)

Рабочая программа дисциплины (модуля) «Алгоритмы биоинформатики» 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) «Алгоритмы биоинформатики» позволяет подготовить специалистов, способных понимать внутренние механизмы биоинформатических инструментов, критически оценивать их применение и разрабатывать новые подходы для анализа геномных, транскриптомных и протеомных данных, что является основой для инноваций в биомедицине, биотехнологии и персонализированной медицине.

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 фундаментальных знаний о ключевых алгоритмах и структурах данных, лежащих в основе биоинформатических методов, а также развитие умений разрабатывать, адаптировать и применять алгоритмические подходы для анализа биологических последовательностей и организации вычислительных экспериментов.

Задачи изучения дисциплины (модуля):

- освоить базовые принципы молекулярной биологии и представления биологических данных в цифровой форме;
- изучить основные алгоритмы выравнивания последовательностей, сборки геномов и построения филогенетических деревьев;
- научиться применять структуры данных (графы, деревья) и вероятностные модели (НММ) для решения биоинформатических задач;
- развить навыки работы с вычислительными ресурсами и биологическими базами данных;
- сформировать способность к алгоритмическому мышлению и разработке эффективных решений для обработки больших объемов биологических данных.

В результате освоения дисциплины (модуля) обучающийся должен:

знать:

- центральную догму молекулярной биологии, структуру ДНК, РНК и белков;
- базовые структуры данных (графы, деревья), основы теории вероятностей и статистики в биологии;
- алгоритмы парного и множественного выравнивания последовательностей (Нидлмана-Вунша, Смита-Ватермана), методы скрытых марковских моделей (НММ), алгоритмы построения филогенетических деревьев и сборки геномов (теория графов де Брюйна).

уметь:

- работать с вычислительной инфраструктурой, включая использование серверов и высокопроизводительных вычислительных систем, применяемых в биоинформатике;
- выбирать, адаптировать и разрабатывать алгоритмы и структуры данных для автоматизации сбора, хранения и анализа биологических данных;
- эффективно работать с биологическими базами данных и информационными ресурсами;
- формулировать и предлагать новые алгоритмические и методические подходы для организации вычислительных экспериментов, обработки биологических данных и их последующего анализа в прикладной биоинформатике.

владеть:

- навыками анализа и интерпретации биологических данных;
- навыками применения методов современной биоинформатики для решения задач биомедицины и биотехнологии.

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	Алгоритмические основы и анализ последовательностей	8	8		21	Домашнее задание Тест
2	Графовые модели и сборка геномов	8	8	2	21	Домашнее задание Тест Контрольная работа
3	Поиск мотивов, эволюционный анализ и филогенетика	7	7		20	Домашнее задание Тест
4	Прогнозирование структуры и анализ сетей	7	7	2	20	Домашнее задание Тест Контрольная работа
	Зачет с оценкой			6		
Итого:		30	30	10	82	
Объем дисциплины (модуля) (в ак. ч.)		152				
Объем дисциплины (модуля) (в зач. ед.)		4				

4. Содержание дисциплины (модуля)

№п/п	Наименование раздела дисциплины (модуля)	Содержание дисциплины (модуля) по темам
1	Алгоритмические основы и анализ последовательностей	Асимптотический анализ и структуры данных: Оценка сложности биоинформатических алгоритмов, хэш-таблицы, деревья суффиксов. Динамическое программирование: Алгоритм Нидлмана-Вунша (глобальное выравнивание) и Смита-Ватермана (локальное выравнивание). Эвристические алгоритмы выравнивания: Базовые принципы поиска по базам данных: алгоритмы BLAST и FASTA. Множественное выравнивание последовательностей (MSA): Прогрессивные методы (алгоритм Clustal W) и профили (скрытые марковские модели).
2	Графовые модели и сборка геномов	Теория графов в биоинформатике: Поиск компонент связности, Эйлеровы и Гамильтоновы пути в контексте сборки генома. Сборка геномов de novo: Алгоритмы перекрытия-расположения-консенсуса (Overlap-Layout-Consensus) и графы де Брейна.

		Алгоритмы картирования (оконтуривания): Построение и использование сжатого суффиксного массива (FM-index) и схемы Барроуза-Уилера (BWT).
3	Поиск мотивов, эволюционный анализ и филогенетика	Поиск паттернов и мотивов: Эвристический поиск регуляторных сайтов (жадные алгоритмы, EM-алгоритмы и поиск Гиббса). Расстояния между последовательностями: Модели эволюции ДНК (Джукса-Кантора, Кимуры), вычисление дистанций и бутстрэп-анализ. Кластеризация и UPGMA: Иерархические методы построения филогенетических деревьев, метод объединения соседей (Neighbor-Joining). Максимальная экономичность (Parsimony): Поиск деревьев с минимальным количеством эволюционных событий (алгоритм Фитча). Максимальное правдоподобие (Maximum Likelihood): Оценка вероятностей мутаций и построение деревьев на основе статистических моделей.
4	Прогнозирование структуры и анализ сетей	Предсказание вторичной структуры РНК: Динамическое программирование и алгоритм Зукера. Моделирование структуры белков: Скрытые Марковские модели (HMM) для профилей семейств белков (Pfam) и распознавание складок (threading). Графы взаимодействий и метаболомика: Алгоритмы поиска путей в метаболических сетях и кластеризация графов белок-белковых взаимодействий.

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. Компо, Ф. Алгоритмы биоинформатики : практическое руководство / Ф. Компо, П. Певзнер ; пер. с англ. И. Л. Люско. - Москва : ДМК Пресс, 2023. - 684 с. - ISBN 978-5-93700-175-7. - Текст : электронный. - URL: <https://znanium.ru/catalog/product/2109518>.

2. Стефанов, В. Е. Биоинформатика : учебник для вузов / В. Е. Стефанов, А. А. Тулуб, Г. Р. Мавропуло-Столяренко. — Москва : Издательство Юрайт, 2026. — 249 с. — (Высшее образование). — ISBN 978-5-534-20194-9. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/583663>.

Дополнительная литература:

1. Биотехнология. Практический курс : учебник и практикум для вузов / под редакцией А. А. Красноштановой. — Москва : Издательство Юрайт, 2026. — 174 с. — (Высшее образование). — ISBN 978-5-534-20448-3. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/589750>.

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.

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) «Алгоритмы биоинформатики» в рамках текущего контроля успеваемости используются такие виды учебной работы, как лекции, семинары, домашние задания, тесты, контрольная работа, 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.

Тест – особая форма проверки знаний. Проводится после освоения одной или нескольких тем и свидетельствует о качестве понимания основных понятий изучаемого материала. Тестовые задания составлены к ключевым понятиям, основным разделам, важным терминологическим категориям изучаемой дисциплины.

Для подготовки к тесту необходимо знать терминологический аппарат дисциплины, понимать смысл научных категорий и уметь их использовать в профессиональной лексике. Владение понятийным аппаратом, включённым в тестовые задания, позволяет преподавателю быстро проверить уровень понимания студентами важных методологических категорий.

Контрольная работа – письменная работа с набором задач, которые нужно решить за ограниченное время.

Цель контрольной работы – получить специальные знания по одной или нескольким темам дисциплины и продемонстрировать навыки их практического применения.

Бонусные баллы — это оценки, которые студенты могут получить за выполнение дополнительных заданий.

Формат бонусных баллов позволяет студентам улучшить общую оценку по дисциплине (модулю) и стимулирует углубленное изучение материала.

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) «Алгоритмы биоинформатики»

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 *зачета с оценкой*, при этом проводится оценка компетенций, сформированных по дисциплине.

Для оценивания текущего контроля успеваемости и промежуточной аттестации
Электронный документ

используется десятибалльная шкала оценивания, которая соотносится с традиционной пятибалльной шкалой следующим образом, если иное не указано в syllabus 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 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.
9	Отлично	
8	Отлично	
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) «Алгоритмы биоинформатики» оценивается следующим образом:

Активность	Вес	Количество	Описание
Домашние задания	25%	10	Набор задач по темам недели
Тесты	25%	10	Форма проверки знаний студентов через ответы на вопросы с выбором из нескольких вариантов ответов
Контрольная работа	25%	2	Письменная работа с набором задач, которые нужно решить за ограниченное время
Зачет с оценкой	25%	1	Письменная или устная работа над заданием, направленным на проверку полученных знаний и навыков по дисциплине (модулю)

Формула расчёта итоговой оценки по дисциплине (модулю) «Алгоритмы биоинформатики»: $0, 25 \times \text{Среднее за домашние задания} + 0, 25 \times \text{Среднее за тесты} + 0, 25 \times \text{Среднее за контрольные работы} + 0, 25 \times \text{Зачет с оценкой}$.

В рамках изучения дисциплины (модуля) возможно получение бонусных баллов.

Текущий контроль успеваемости обучающихся по дисциплине (модулю)

Примерные домашние задания

Домашнее задание 1. Выравнивание последовательностей

1. Выполните **глобальное выравнивание** двух последовательностей вручную с помощью алгоритма Нидлмана–Вунша:
 - Последовательность A: GATTACA
 - Последовательность B: GCATGCU
 - Параметры: штраф за замену = -1 , штраф за гэп = -2 , премия за совпадение = $+1$.
2. Постройте матрицу динамического программирования, укажите обратный путь (traceback).
3. Запишите итоговое выравнивание и его оценку.
4. Объясните, чем локальное выравнивание (Смит–Ватерман) отличается от глобального.
5. В каких биологических задачах уместно использовать каждый из методов?

Домашнее задание 2. Сборка генома и графы де Брейна

1. Даны короткие риды длиной 5 нуклеотидов:
ATGCA, TGCAA, GCAAT, CAATT, AATTG.
2. Постройте граф де Брейна для $k=4$ (разбейте риды на k -меры, постройте узлы и рёбра).
3. Найдите Эйлеров путь в графе и восстановите исходную последовательность.
4. Объясните, почему графы де Брейна эффективны для сборки геномов, но могут приводить к ошибкам при повторях.
5. Сравните подход OLC (Overlap-Layout-Consensus) и графы де Брейна по сложности и применимости.

Домашнее задание 3. Филогенетический анализ

1. Даны четыре последовательности гена цитохрома С у разных видов (по 20 нуклеотидов).
2. Вычислите попарные расстояния по модели Джукса–Кантора.
3. Постройте филогенетическое дерево методом UPGMA.
4. Проведите бутстрэп-анализ (на примере 3 псевдореplik) — опишите, как он оценивает надёжность ветвей.
5. Объясните, в чём разница между методами максимальной экономичности и максимального правдоподобия.

Примерные задания для теста

Вопрос 1.

Какой алгоритм используется для глобального выравнивания последовательностей?

1. Смита–Ватермана
2. Нидлмана–Вунша
3. BLAST
4. Кнута–Морриса–Пратта

Правильный ответ: 2

Вопрос 2.

Что лежит в основе построения графа де Брейна?

1. Поиск перекрытий между ридом
2. Разбиение ридов на k-меры
3. Выравнивание на референс
4. Сборка по принципу "layout-consensus"

Правильный ответ: 2

Вопрос 3.

Какой метод использует скрытые марковские модели (HMM) для множественного выравнивания?

1. ClustalW
2. MUSCLE
3. HMMER / Pfam
4. T-Coffee

Правильный ответ: 3

Вопрос 4.

Какой индекс используется в алгоритме BWA для быстрого картирования ридов?

1. Suffix Tree
2. Hash Table
3. FM-index
4. K-mer Index

Правильный ответ: 3

Вопрос 5.

Какой метод строит филогенетическое дерево, минимизируя количество эволюционных изменений?

1. Максимальное правдоподобие

2. Метод Нейбор-Джойнинг
3. Максимальная экономичность
4. UPGMA

Правильный ответ: 3

Примерные задания для контрольной работы

Задание 1.

Выполните локальное выравнивание двух последовательностей с помощью алгоритма Смита–Ватермана:

- A: GCGAT
- B: GCAGT

Параметры: совпадение +2, несовпадение –1, гэп –2.

Постройте матрицу, найдите максимальный путь, запишите выравнивание.

Задание 2.

Объясните, как работает алгоритм BLAST. Почему он является эвристическим? Какие этапы включает (seed, extend, evaluate)? В каких случаях он может упустить значимые выравнивания?

Задание 3.

Даны k-меры: ATG, TGC, GCA, CAT, ATG.

Постройте граф де Брейна (k=3), найдите Эйлеров путь и восстановите последовательность. Объясните, как повторы в геноме могут нарушить однозначность сборки.

Задание 4.

Что такое UPGMA? Какой ключевой недостаток у этого метода построения филогенетических деревьев? Приведите пример, когда он может привести к неверной топологии.

Задание 5.

Опишите, как алгоритм Зукера используется для предсказания вторичной структуры РНК. Какие принципы динамического программирования применяются? Что такое минимальная свободная энергия (MFE)?

Задания для промежуточной аттестации по дисциплине (модулю)

№ п/п	Задание	Ответ	Компетенция
1.	Какой алгоритм решает задачу глобального выравнивания?	Нидлмана–Вунша	УК-1
2.	Что такое k-мер в графе де Брейна?	Подстрока длины k	УК-3
3.	Какой метод использует FM-index?	BWA, Bowtie2	ПК-1

4.	Что оценивает модель Джукса–Кантора?	Эволюционное расстояние между последовательностями	ПК-2
5.	Какой алгоритм подходит для поиска мотивов в ДНК?	Поиск Гиббса	ПК-3
6.	Что такое Эйлеров путь в графе?	Путь, проходящий по каждому ребру один раз	ПК-5
7.	Какой метод используется для множественного выравнивания?	ClustalW	УК-1
8.	Что такое бутстрэп в филогенетике?	Оценка устойчивости ветвей дерева	УК-3
9.	Какой алгоритм применяется для предсказания структуры РНК?	Алгоритм Зукера	ПК-1
10.	Что такое НММ в биоинформатике?	Скрытая марковская модель для профилей белков	ПК-2
11.	Какой индекс основан на преобразовании Барроуза–Уилера?	BWT → FM-index	ПК-3
12.	Что минимизирует метод максимальной экономичности?	Количество замен в эволюции	ПК-5
13.	Какой метод строит дерево по попарным расстояниям?	UPGMA, Neighbor-Joining	УК-1
14.	Что такое суффиксное дерево?	Структура данных для поиска подстрок	УК-3
15.	Какой алгоритм использует "жадный" подход к выравниванию?	FASTA	ПК-1
16.	Что такое профильное выравнивание?	Выравнивание с использованием НММ или профилей	ПК-2
17.	Какой граф используется в сборке геномов?	Граф де Брейна	ПК-3
18.	Что такое парсимония в филогенетике?	Принцип минимального числа	ПК-5

		изменений	
19.	Какой метод оценивает вероятность дерева по модели эволюции?	Максимальное правдоподобие	УК-1
20.	Какой алгоритм используется для поиска путей в метаболических сетях?	Алгоритм Дейкстры, поиск в глубину	УК-3