

# БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

УТВЕРЖДАЮ

Ректор Белорусского  
государственного университета

А.Д.Король

27 июня 2025 г.

Регистрационный № 3523/б.



## ОСНОВЫ БИОИНФОРМАТИКИ

Учебная программа учреждения образования по учебной дисциплине для  
специальности:

**6-05-0533-10 Информатика**

Профилизация: Дискретная математика и биоинформатика

2025 г.

Учебная программа составлена на основе на основе ОСВО 6-05-0533-10-2023, учебного плана БГУ № 6-5.3-58/03 от 15.05.2023.

**СОСТАВИТЕЛЬ:**

**В.И.Белько**, заведующий кафедрой биомедицинской информатики факультета прикладной математики и информатики Белорусского государственного университета, кандидат физико-математических наук, доцент.

**РЕЦЕНЗЕНТ:**

**Н.А.Новоселова**, ведущий научный сотрудник лаборатории биоинформатики государственного научного учреждения «Объединенный институт проблем информатики Национальной академии наук Беларусь», кандидат технических наук

**РЕКОМЕНДОВАНА К УТВЕРЖДЕНИЮ:**

Кафедрой биомедицинской информатики БГУ  
(протокол № 12 от 12.06.2025);

Научно-методическим советом БГУ  
(протокол № 11 от 26.06.2025)

Заведующий кафедрой



В.И.Белько

## **ПОЯСНИТЕЛЬНАЯ ЗАПИСКА**

### **Цели и задачи учебной дисциплины.**

**Цель учебной дисциплины** – начальное знакомство студентов с основными понятиями молекулярной биологии, организацией генов и белков, технологиями получения генетических данных, включая копирование и клонирование ДНК, секвенирование геномов, рассмотрение способов представления и хранения генетических данных, организацию основных биоинформационных ресурсов – NCBI, Swiss-Prot, PDB, формирование представлений о типах биоинформационных задач, возникающих в процессе анализа биологических данных, и о вычислительных методах и алгоритмах их решения, более подробное знакомство с алгоритмами решения ряда основных задач молекулярной биологии, включая алгоритмы выравнивания нуклеотидных последовательностей, рекомбинации геномов, выделения мотивов в генетической последовательности, поиска участков генов, анализа данных экспрессии генов, определения геномных паттернов, которые в комплексе демонстрируют необходимость и эффективность применения компьютерных методов в биологии.

### **Задачи учебной дисциплины:**

1. Сформировать целостное представление о связи компьютерных наук и биологии;
2. Ознакомить с основными способами получения генетической информации и форматами ее хранения, с основными задачами биоинформатики и подходами к их решению;
3. Сформировать мотивацию к самостоятельным исследованиям в области биоинформатики.

**Место учебной дисциплины** в системе подготовки специалиста с высшим образованием.

Учебная дисциплина относится к дисциплинам профилизации профиля «Дискретная математика и биоинформатика».

Программа составлена с учетом межпредметных связей с учебными дисциплинами. Основой для изучения учебной дисциплины являются дисциплины компонента учреждения высшего образования «Методы и алгоритмы анализа данных», «Теория вероятностей и математическая статистика». Знания, полученные в учебной дисциплине, могут быть использованы при выполнении студентами курсовой и выпускной квалификационной работ, и являются необходимыми для успешной работы в сфере прикладной математики и информатики, а также для иной профессиональной деятельности по данной специальности.

### **Требования к компетенциям**

Освоение учебной дисциплины «Введение в биоинформатику» должно обеспечить формирование следующей универсальной компетенции:

УК-2. Решать стандартные задачи профессиональной деятельности на основе применения информационно-коммуникационных технологий.

В результате освоения учебной дисциплины студент должен:

**знать:**

- основные понятия молекулярной биологии, способы получения и хранения генетической информации;
- основные типы задач молекулярной биологии, решаемых методами биоинформатики;
- алгоритмические подходы в биоинформатике, их характеристики;
- основные геномные базы данных и биоинформационные ресурсы;

**уметь:**

- пользоваться основными биоинформационными ресурсами для изучения ДНК, РНК последовательностей, организации белков и визуального представления их структуры;
- анализировать данные генной экспрессии путем построения различных моделей кластеризации с последующей оценкой результатов;

**владеть:**

- научной терминологией данного раздела науки;
- устойчивыми навыками рационального использования методов первичного анализа биологической информации;
- базовыми навыками и умениями применения адекватного математического аппарата для решения задач биоинформатики.

### **Структура учебной дисциплины**

Дисциплина изучается в 5-м семестре. Всего на изучение учебной дисциплины «Основы биоинформатики» отведено:

– в очной форме получения высшего образования: 108 часов, в том числе 68 аудиторных часа, из них: лекции – 34 часа, лабораторные занятия – 30 часов, управляемая самостоятельная работа – 4 часа.

Трудоемкость учебной дисциплины составляет 3 зачетные единицы. Форма промежуточной аттестации – зачет.

## **СОДЕРЖАНИЕ УЧЕБНОГО МАТЕРИАЛА**

### **Раздел 1. Биоинформатика как междисциплинарная наука**

#### **Тема 1. 1 Введение.**

Понятие биоинформатики как междисциплинарной науки. Предмет исследований биоинформатики.

Основные события в истории молекулярной биологии. Основные понятия и законы молекулярной биологии. Устройство клетки, структура ДНК, генетический код, транскрипция, трансляция, репликация ДНК. Карты

и последовательности.

Способы получения генетического материала для анализа: копирование и клонирование ДНК, рестрикция и гибридизация ДНК. Полимеразная цепная реакция. Мутации генома. Источники генетических вариаций, биологические основы молекулярной эволюции, процессы адаптации и видообразования.

### **Тема 1.2. Базы данных в биологических исследованиях.**

Введение в базы данных. Типы данных в биологических базах данных. Классификация геномных баз данных. Форматы данных хранения геномных последовательностей (нуклеотидных, аминокислотных): FASTA формат, GenBank формат, EMBL формат.

GenBank – база данных нуклеотидных последовательностей.

UniProt, Swiss-Prot – базы данных информации о белках, включая аннотацию, доменной структуры белков.

### **Тема 1.3. Введение в язык программирования Python.**

Основные операции и конструкции языка Python. Библиотеки анализа биологической информации в Python. IPython – инструмент для работы с языком Python. Jupyter Notebook – графическая веб-оболочка для IPython. Организация и работа с Jupyter Notebook. Создание ноутбуков для документирования и выполнения приложений на языке Python.

### **Тема 1.4. Генетическая информация. Секвенирование и сборка геномов.**

Секвенирование геномов. Принципы секвенирования. Секвенирование путем гибридизации. Секвенирование по Сэнгеру. Секвенирование нового поколения (Next-generation sequencing). Сборка геномов из данных о сиквенсах. Проект геном человека.

## **Раздел 2. Задачи биоинформатики. Алгоритмический подход к их решению.**

### **Тема 2.1 Применение вычислительных алгоритмов в биоинформатике.**

Строки – основной тип данных. Математические алгоритмы как инструмент решения биоинформатических задач. Итерационные и рекурсивные алгоритмы. Оценка вычислительной сложности алгоритмов. Виды алгоритмов: полный перебор, метод ветвей и границ, «жадный» алгоритм, динамическое программирование, алгоритм декомпозиции, алгоритмы машинного обучения.

Основы теории графов. Реконструкция ДНК последовательности полученной путем гибридизации как задача поиска Гамильтонова пути на графе.

## **Тема 2.2 ДНК картирование. Поиск мотивов в ДНК последовательности.**

Рестрикционное картирование (restriction mapping) как инструмент анализа молекулярных данных. Постановка задачи частичного переваривания (partial digest problem). Алгоритмы решения задачи.

Регуляция гена, факторы транскрипции и ДНК регуляторные мотивы. Представление мотивов. Профили и консенсусные матрицы. Позиционная весовая матрица. Логотип последовательности. Задача поиска мотивов в ДНК последовательности. Альтернативное представление задачи как поиск медианной строки. Построение дерева поиска. Решения задачи поиска мотивов путем сканирования дерева поиска. Алгоритмы ветвей и границ для эффективного решения задачи поиска мотивов.

MEME – программное средство для поиска мотивов ДНК последовательности.

## **Тема 2.3 Перестройка генома.**

Типы перестройки генома (Genome rearrangement). Постановка задачи перестройки генома. Пример из биологии. Представление перестройки генома как последовательности инверсий геномных строк. Понятие приближенного алгоритма – алгоритма поиска приближенного решения задачи. Порядок и расположение генов в геноме. Приближенный алгоритм поиска последовательности инверсий, позволяющих трансформировать один геном в другой.

GRIMM (<http://www-cse.ucsd.edu/groups/bioinformatics/GRIMM>) – веб-сервер для расчета расстояния между геномными строками на основе минимального количества инверсий.

## **Раздел 3. Задачи биоинформатики. Алгоритмы динамического программирования.**

### **Тема 3.1. Сравнение генетических последовательностей.**

Биологические основы сравнения последовательностей. Основные операции редактирования генетической последовательности – делеция, вставка и замена. Точечные матрицы для сравнения двух последовательностей. Понятие редакционного расстояния между генетическими последовательностями – edit distance. Выравнивание как анализ схожести генетических последовательностей (строк), задача поиска наиболее длинной общей подстроки для двух строк. Пример из биологии – открытие гена кистозного фиброза.

Матрицы весов аминокислотных замен (PAM, BLOSUM). Глобальное парное выравнивание последовательностей. Алгоритм решения задачи глобального выравнивания последовательностей.

Локальное выравнивание последовательностей. Выравнивание с

учетом штрафов за внесение делеции (Affine Gap Penalties).

Множественное выравнивание последовательностей.

CLUSTAL – программное средство для множественного выравнивания нуклеотидных и аминокислотных последовательностей на основе эвристической стратегии.

### **Тема 3.2. Предсказание белок-кодирующих участков.**

Задача предсказания белок-кодирующих участков (положение гена) в генетической последовательности. Понятие экзона и интрана. Пример из биологии по изучению аденоовириуса. Два основных подхода к предсказанию белок-кодирующих участков.

Статистический подход к решению задачи предсказания белок-кодирующих участков. GENSCAN алгоритм поиска генов на основе вероятностной модели структуры гена.

Подход к решению задачи предсказания белок-кодирующих участков на основе анализа близости генетической строки и ранее предсказанного гена. Постановка задачи связывания экзонов (Exon Chaining Problem), заключающейся в поиске максимального множества неперекрывающихся экзонов. Решение Exon Chaining Problem с использованием алгоритма динамического программирования.

Glimmer и GenMark алгоритмы поиска генов в ДНК последовательности.

### **Тема 3.3. Поиск паттернов в генетической последовательности.**

Поиск повторов в генетической последовательности. Задача поиска строки в базе генетических последовательностей как задача поиска паттернов. Построение ключевых деревьев для минимизации процедур сравнения строк.

Множественный поиск строк. Построение и использование суффиксных деревьев. Алгоритм поиска точного совпадения строки.

Сокращение вычислительной сложности алгоритмов поиска в базах данных. Метрики оценки структуры последовательности нуклеотидов: GC-содержание, частота k-меров в последовательности. Оптимизация поиска путем предобработки информации в базе данных. Эвристические алгоритмы поиска совпадений на основе фильтрации.

BLAST – программный инструмент для поиска гомологов в базе данных. Оценка статистической значимости результатов поиска.

## **Раздел 4. Анализ биологических данных с использованием алгоритмов машинного обучения.**

### **Тема 4.1. ДНК Микрочипы и анализ экспрессии генов.**

Технология ДНК микрочипов. Контроль качества, нормализация.

Организация данных генной экспрессии. Анализ дифференциальной экспрессии генов.

Предсказание функций белков с использованием анализа данных генной экспрессии.

Задача кластеризации. Основные алгоритмы кластеризации. Кластеризация данных генной экспрессии. Иерархическая кластеризация, кластеризация к-средних, кластерные алгоритмы на графах – алгоритм CAST.

Биологическая интерпретация результатов анализа. Анализ представленности функциональных групп генов — (Gene Set Enrichment Analysis). GSEA – программный инструмент, основанный на оценке перепредставленности (<http://software.broadinstitute.org/gsea/index.jsp>).

#### **Тема 4.2. Молекулярная эволюция.**

Эволюционные деревья. Реконструкция эволюционных деревьев на основе матрицы расстояний. Реконструкция деревьев на основе аддитивных матриц.

Эволюционные деревья и иерархическая кластеризация. UPGMA – вариант кластерного алгоритма для представления эволюционных деревьев.

Метод парсимонии построения филогенетических деревьев. Sankoff алгоритм. Пример: Филогенетический анализ вируса иммунодефицита человека (ВИЧ).

#### **Тема 4.3. Алгоритмы машинного обучения для поиска структур в генетических данных.**

Постановка задачи поиска CG-островков в геноме. Скрытые марковские модели (HMM) – инструмент машинного обучения. Параметры модели: количество скрытых состояний модели, вероятности переходов между состояниями, вероятностное распределение событий при условии нахождения в определённом скрытом состоянии. Алгоритм восстановления скрытых состояний модели – алгоритм Витебри. Оценка параметров скрытой марковской модели.

Использование HMM для сравнения генетических последовательностей, в частности для сравнения последовательности относительно профиля генетических строк, являющегося результатом их множественного выравнивания.

PFAM база данных белковых доменов как пример использования HMM модели.

#### **Тема 4.4. Биоинформатические ресурсы**

Специализированные базы данных и инструментарий – NCBI, EBI, KEGG, Swiss-Prot, PDB. Функциональная аннотация генов. Онтология генов.

Работа с биоинформационическими ресурсами на примере green fluorescent protein (зеленый флуоресцентный белок).

YeastMine – интегрированная среда для получения и анализа данных дрожжевых грибков (<http://yeastmine.yeastgenome.org>). Пример организации различного типа запросов для извлечения необходимой информации.

## УЧЕБНО-МЕТОДИЧЕСКАЯ КАРТА УЧЕБНОЙ ДИСЦИПЛИНЫ

Очная (дневная) форма получения высшего образования с применением дистанционных образовательных технологий  
(ДОТ)

Номер раздела, темы	Название раздела, темы	Количество аудиторных часов					Форма контроля знаний	
		Лекции	Практические занятия	Семинарские занятия	Лабораторные занятия	Иное		
1	2	3	4	5	6	7	8	9
I	<b>Биоинформатика как междисциплинарная наука</b>							
1.1	Введение. Понятие биоинформатики как междисциплинарной науки. Предмет исследований биоинформатики. Основные понятия и законы молекулярной биологии. Формат представления и хранения генетических данных	4						<i>Устный опрос</i>
1.2	Базы данных в биологических исследованиях	4			2			<i>Устный опрос Отчет о лабораторной работе</i>
1.3	Введение в язык программирования Python	4			2			<i>Отчет о лабораторной работе</i>
1.4	Генетическая информация. Секвенирование и сборка геномов							
II	<b>Задачи биоинформатики. Алгоритмический подход к их решению</b>							
2.1	Применение вычислительных алгоритмов в биоинформатике	2						<i>Устный опрос</i>

2.2	ДНК картирование. Поиск мотивов в ДНК последовательности.	2						<i>Устный опрос</i>
2.3	Перестройка генома.	2			6			<i>Отчет о лабораторной работе</i>
<b>III</b>	<b>Алгоритмы динамического программирования</b>							
3.1	Сравнение генетических последовательностей	2			6			<i>Отчет о лабораторной работе</i>
3.2	Предсказание белок-кодирующих участков	2			4			<i>Устный опрос</i> <i>Отчет о лабораторной работе</i>
3.3	Поиск паттернов в генетической последовательности	2			4			<i>Устный опрос</i>
<b>IV</b>	<b>Анализ биологических данных с использованием алгоритмов машинного обучения</b>							
4.1	ДНК. Микрочипы и анализ экспрессии генов	2			2			<i>Отчет о лабораторной работе</i>
4.2	Молекулярная эволюция	2			2		2	<i>Устный опрос</i> <i>Отчет о лабораторной работе</i>
4.3	Алгоритмы машинного обучения для поиска структур в генетических данных	2			2		2	<i>Отчет о лабораторной работе</i>
4.4	Биоинформационные ресурсы	34			30		4	

## **ИНФОРМАЦИОННО-МЕТОДИЧЕСКАЯ ЧАСТЬ**

### **Основная литература**

1. Н. Ю. Часовских. Биоинформатика. Учебник – ГЭОТАР-Медиа, 2021. – 352 с.
2. Ф. Компо, П. Певзнер Алгоритмы биоинформатики, ДМК Пресс, 2023. – 682 с.
3. Юмагулов, М. Г. Введение в нелинейную динамику: теория, приложения, модели: Учебное пособие для вузов. - Санкт-Петербург: Лань, 2024. – 36 с.

### **Дополнительная литература**

1. Леск А. Введение в биоинформатику – Бином. Лаборатория знаний, 2015. – 318 с.
2. Финкельштейн А.В., Птицын О.Б. “Физика белка. Курс лекций”. – Москва: Университет, 2002. – 365 с.
3. Н. Ю. Часовских Биоинформатика. Учебник – ГЭОТАР-Медиа 2020. – 352 с.
4. Сетабул Ж., Мейданис Ж. Введение в вычислительную биологию. – Москва-Ижевск: «Регулярная и хаотическая динамика», Институт компьютерных исследований, 2007. – 420 с..
5. Игнасимуту С. Основы биоинформатики. – Ижевск: НИЦ. «Регулярная и хаотическая динамика», Институт компьютерных исследований, 2007. – 320 с.

### **Перечень рекомендуемых средств диагностики и методика формирования итоговой отметки**

Объектом диагностики компетенций студентов являются знания, умения, полученные ими в результате изучения учебной дисциплины. Выявление учебных достижений студентов осуществляется с помощью мероприятий текущего контроля и промежуточной аттестации.

Для диагностики компетенций в рамках учебной дисциплины рекомендуется использовать следующие формы:

1. Устная форма: дискуссия, коллоквиум, собеседование.
2. Письменная форма: контрольные работы.
3. Устно-письменная форма: отчеты по домашним упражнениям с их устной защитой.

В качестве рекомендуемых технических средств диагностики используется Образовательная платформа Moodle (<https://edufpmi.bsu.by>).

Формой промежуточной аттестации по дисциплине «Введение в биоинформатику» учебными планами предусмотрен зачет.

Рейтинговая система предусматривает использование весовых коэффициентов в ходе проведения контрольных мероприятий текущей аттестации.

Примерные весовые коэффициенты, определяющие вклад текущей аттестации в отметку при прохождении промежуточной аттестации:

Формирование отметки за текущую аттестацию:

- отчет по домашним упражнениям с их устной защитой – 80 %;
- контрольные работы – 10 %;
- коллоквиум – 10 %.

Объектом диагностики компетенций студентов являются знания, умения, полученные ими в результате изучения учебной дисциплины. Выявление учебных достижений студентов осуществляется с помощью мероприятий текущего контроля и промежуточной аттестации.

Для диагностики компетенций в рамках учебной дисциплины рекомендуется использовать следующие формы:

4. Устная форма: дискуссия, коллоквиум, собеседование.
5. Письменная форма: контрольные работы.
6. Устно-письменная форма: отчеты по домашним упражнениям с их устной защитой.

В качестве рекомендуемых технических средств диагностики используется Образовательная платформа Moodle (<https://edufpmi.bsu.by>).

Формой промежуточной аттестации по дисциплине «Введение в биоинформатику» учебными планами предусмотрен зачет.

Рейтинговая система предусматривает использование весовых коэффициентов в ходе проведения контрольных мероприятий текущей аттестации.

Примерные весовые коэффициенты, определяющие вклад текущей аттестации в отметку при прохождении промежуточной аттестации:

Формирование отметки за текущую аттестацию:

- отчет по домашним упражнениям с их устной защитой – 80 %;
- контрольные работы – 10 %;
- коллоквиум – 10 %.

### **Примерный перечень заданий для управляемой самостоятельной работы студентов**

Управляемая самостоятельная работа предлагается в виде заданий проектного типа.

Тема 3.1 Сравнение генетических последовательностей (2ч.).

Примером задания может быть самостоятельная реализация студентом одного из известных алгоритмов локального или глобального выравнивания белковых последовательностей на языке программирования Python.

Форма контроля – собеседование.

Тема 4.2 Молекулярная эволюция (2ч.).

Примером задания может быть построение филогенетических деревьев для отдельных белков по результатам работы различных алгоритмов

множественного выравнивания с применением различных алгоритмов построения филогенетических деревьев, сравнение полученных результатов.

Форма контроля – собеседование.

### **Примерная тематика лабораторных занятий**

Занятие 1. Введение в язык программирования Python.

Занятие 2. Работа с различными структурами данных в языке программирования Python.

Занятие 3. Поиск мотивов в ДНК последовательности.

Занятие 4. Алгоритмы поиска подстроки в строке.

Занятие 5. Поиск биологической информации с помощью биоинформационических баз данных с использованием языка программирования Python.

Занятие 6. Реализация алгоритма глобального выравнивания последовательностей Нидлмана-Вунша.

Занятие 7. Реализация алгоритма локального выравнивания последовательностей Ватермана-Смита.

Занятие 8. Предсказание белок-кодирующих участков.

Занятие 9. Применение скрытых марковских моделей для предсказания экзонов.

Занятие 10. Алгоритмы машинного обучения для поиска структур в генетических данных.

Занятие 11. Метод опорных векторов.

Занятие 12. Алгоритмы машинного обучения для поиска структур в генетических данных.

Занятие 13. Методы кластеризации.

Занятие 14. Алгоритмы машинного обучения для поиска структур в генетических данных.

Занятие 15. Нейронные сети.

### **Рекомендуемая тематика контрольных работ и коллоквиума:**

1) Контрольная работа № 1 «Секвенирование. Методы секвенирования».

2) Контрольная работа № 2 «Выравнивание последовательностей».

3) Контрольная работа № 3 «Кластеризация генов».

4) Коллоквиум «Выравнивание генетических последовательностей. Методы».

### **Примерные варианты контрольных работ**

1. Сравнить два указанных метода секвенирования, назвать преимущества и недостатки обоих, где и почему лучше их использовать.

2. Рассчитать редакционное расстояние между заданными аминокислотными последовательностями, используя различные виды матриц весов замен.
3. Выполнить аннотацию генома клеток одинакового типа, затем выполнить построение графа  $k$ -ближайших соседей и кластеризацию генов, применив алгоритм обнаружения сообществ Лейдена.

### **Описание инновационных подходов и методов к преподаванию учебной дисциплины**

При организации образовательного процесса большинства практических занятий используется *практико-ориентированный подход*, который предполагает:

- освоение содержания образования через решения практических задач;
- приобретение навыков эффективного выполнения разных видов профессиональной деятельности.

Также при организации образовательного процесса используются *методы группового обучения, проектного обучения и учебной дискуссии*. Студентам предлагается выполнить часть домашних заданий в форме проекта в группах до 4 человек. Задания предполагают предварительное обсуждение в форме мозгового штурма.

Выполнение проекта предусматривает самостоятельную работу с научными и техническими источниками по теме курса, самостоятельный поиск и выбор способа решения задачи, составление плана и разделение задач между участниками группы.

В конце курса предусмотрена устная защита проекта с критическим анализом идей, сгенерированных в ходе мозгового штурма, и ретроспективой выполненной работы.

Комбинация методов предполагает

- ориентацию на генерирование идей, реализацию групповых студенческих проектов, развитие предпринимательской культуры;
- способ организации учебной деятельности студентов, развивающий актуальные для учебной и профессиональной деятельности навыки планирования, самоорганизации, сотрудничества и предполагающий создание собственного продукта;
- приобретение навыков для решения исследовательских, творческих, социальных, предпринимательских и коммуникационных задач.
- появление нового уровня понимания изучаемой темы, применение знаний (теорий, концепций) при решении проблем, определение способов их решения.

## **Методические рекомендации по организации самостоятельной работы обучающихся**

Для организации самостоятельной работы студентов по учебной дисциплине следует использовать современные информационные технологии: разместить в сетевом доступе комплекс учебных и учебно-методических материалов (учебно-программные материалы, учебное издание для теоретического изучения дисциплины, презентации лекций, методические указания к практическим занятиям, электронные версии домашних заданий, материалы текущего контроля и промежуточной аттестации, позволяющие определить соответствие учебной деятельности обучающихся требованиям образовательного стандарта высшего образования и учебно-программной документации, в том числе вопросы для подготовки к зачету, задания, вопросы для самоконтроля, список рекомендуемой литературы, информационных ресурсов и др.).

Управляемая самостоятельная работа (консультационно-методическая поддержка и контроль) дисциплины обеспечивается средствами образовательного портала БГУ LMS Moodle.

### **Примерный перечень вопросов к зачету**

1. Репликация ДНК, полимеразная цепная реакция.
2. Какие основные базы данных используются в биоинформатике и какую информацию они содержат?
3. Геном, секвенирование генома.
4. Перечислите основные методы секвенирования геномов, назовите их достоинства и недостатки.
5. Граф перекрытий в задаче реконструкции строк.
6. Граф де Брюйна. Перечислите способы его построения.
7. Задача поиска мотивов в ДНК.
8. Глобальное и локальное выравнивания последовательностей. Известные программные средства для решения задачи выравниваний.
9. Редакционное расстояние. Алгоритм Нидлмана-Вунша.
10. Два основных подхода к предсказанию белок-кодирующих участков генома.
11. Алгоритм реконструкции филогении по расстоянию.
12. Задача кластеризации данных генной экспрессии.
13. Кластеризация k-средних.
14. Алгоритм Ллойда.
15. Задача поиска CG-островов.
16. Использование скрытых Марковских моделей в задаче выравнивания последовательностей.
17. Построение префиксного дерева.

18. Построение суффиксного дерева.
19. Алгоритм Ахо-Корасик.
20. Реализация бинарного поиска.
21. Процесс сборки генома.
22. Задача поиска медианной строки.
23. Задача реконструкции эволюционного дерева по признакам.
24. Иерархическая кластеризация.
25. Жадный алгоритм множественного выравнивания.
26. Устройство клетки, структура ДНК, РНК, ген и геном. Мутации генома.
27. Форматы хранения биологических данных FASTA, FASTQ, PDB, CIF.
28. Проект геном человека.
29. Задача поиска самой длинной общей подпоследовательности.

## ПРОТОКОЛ СОГЛАСОВАНИЯ УЧЕБНОЙ ПРОГРАММЫ УО

Название учебной дисциплины, с которой требуется согласование	Название кафедры	Предложения об изменениях в содержании учебной программы учреждения высшего образования по учебной дисциплине	Решение, принятое кафедрой, разработавшей учебную программу (с указанием даты и номера протокола)
Учебная дисциплина не требует согласования			

Заведующий кафедрой биомедицинской информатики

кандидат физико-математических наук, доцент



В.И.Белько

12.06.2025

**ДОПОЛНЕНИЯ И ИЗМЕНЕНИЯ К УЧЕБНОЙ ПРОГРАММЕ УО**  
на \_\_\_\_\_ / \_\_\_\_\_ учебный год

№ п/п	Дополнения и изменения	Основание

Учебная программа пересмотрена и одобрена на заседании кафедры  
(протокол № \_\_\_\_\_ от \_\_\_\_\_ 202\_ г.)

Заведующий кафедрой

УТВЕРЖДАЮ  
Декан факультета