

Федеральное государственное бюджетное образовательное учреждение высшего
профессионального образования
Московский государственный университет имени М.В. Ломоносова
Механико-математический факультет



УТВЕРЖДАЮ
декан механико-
математического факультета
/А.И. Шафаревич /
«14» октября 2021г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Наименование дисциплины:

Математические модели в молекулярной биологии и эволюции

Уровень высшего образования:
магистратура

Направление подготовки / специальность:
02.04.01 "Математика и компьютерные науки" (3++)

Направленность (профиль) ОПОП:
Цифровые технологии и искусственный интеллект

Форма обучения:
очная

Рабочая программа рассмотрена и утверждена
на заседании Ученого совета Механико-математического факультета
(протокол №7, от 14 октября 2021 года)

Москва 2021

Рабочая программа дисциплины (модуля) разработана в соответствии с Федеральным государственным образовательным стандартом высшего образования (ФГОС ВО) для реализуемых основных профессиональных образовательных программ высшего образования по направлению подготовки 02.04.01 "Математика и компьютерные науки" утвержденного Приказом Министерства образования и науки РФ от 10 января 2018 г. N 13.

1. Место дисциплины (модуля) в структуре ОПОП ВО:

дисциплина относится к блоку профессиональной подготовки вариативной части ОПОП ВО.

2. Входные требования для освоения дисциплины (модуля), предварительные условия (если есть):

Курс не предполагает предварительных знаний. Для получения общей картины курса полезно посмотреть материалы на сайтах <http://logic.math.msu.ru/staff/lyubetsky/> и <http://lab6.iitp.ru/ru/pub/>.

3. Результаты обучения по дисциплине (модулю):

Планируемые результаты обучения по дисциплине (модулю)		
Содержание и код компетенции.	Индикатор (показатель) достижения компетенции	Планируемые результаты обучения по дисциплине, сопряженные с индикаторами достижения компетенций
ПК-1. Способен исследовать и разрабатывать архитектуры систем искусственного интеллекта для различных предметных областей на основе комплексов методов и инструментальных средств систем искусственного интеллекта	ПК-1.2. Выбирает комплексы методов и инструментальных средств искусственного интеллекта для решения задач в зависимости от особенностей предметной области	ПК-1.2. 3-1. Знает методы и инструментальные средства систем искусственного интеллекта, критерии их выбора и методы комплексирования в рамках создания интегрированных гибридных интеллектуальных систем различного назначения ПК-1.2. У-1. Умеет выбирать, применять и интегрировать методы и инструментальные средства систем искусственного интеллекта, критерии их выбора и методы комплексирования в рамках создания интегрированных гибридных интеллектуальных систем различного назначения
ПК-8. Способен разрабатывать и модернизировать программное и аппаратное обеспечение технологий и систем искусственного интеллекта с учетом требований информационной безопасности в различных предметных областях	ПК-8.1. Разрабатывает программное и аппаратное обеспечение технологий и систем искусственного интеллекта для решения профессиональных задач с учетом требований информационной безопасности в различных предметных областях	ПК-8.1. 3-1. Знает новые научные принципы и методы разработки программного и аппаратного обеспечения технологий и систем искусственного интеллекта для решения профессиональных задач в различных предметных областях

		ПК-8.1. У-1. Умеет разрабатывать программное и аппаратное обеспечение технологий и систем искусственного интеллекта с учетом требований информационной безопасности для решения профессиональных задач в различных предметных областях
--	--	--

4. Объем дисциплины (модуля) составляет 6 з.е., в том числе 64 академических часов, отведенных на контактную работу обучающихся с преподавателем, 152 академических часов на самостоятельную работу обучающихся.

5. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических часов и виды учебных занятий:

5.1. Структура дисциплины (модуля) по темам (разделам) с указанием отведенного на них количества академических часов и виды учебных занятий (в строгом соответствии с учебным планом)

Семестр 1

Наименование разделов и тем дисциплины (модуля), Форма промежуточной аттестации по дисциплине (модулю)	Номинальные трудозатраты обучающегося			Всего академических часов	Форма текущего контроля успеваемости* (наименование)
	Контактная работа (работа во взаимодействии с преподавателем) Виды контактной работы, академические часы		Самостоятельная работа обучающегося, академические часы		
	Занятия лекционного типа	Занятия семинарского типа			
Тема 1 Начало молекулярной биологии клетки	3		6	9	опрос

Тема 2 Модель аттенуаторной регуляции у бактерий и анализ многокомпонентной динамической системы	3		6	9	опрос
Тема 3 Начало теории эволюции и филогенетики	3		6	9	опрос
Тема 4 Эволюция регуляторного сигнала с вторичной структурой и минимизация сложного функционала (в частности, методом аннилинга)	3		6	9	опрос
Тема 5 Взаимодействия РНК-полимераз, промоторов, терминаторов и анализ соответствующей многокомпонентной динамической системы.	3		6	9	опрос
Тема 6 Моделирование движения РНК-полимераз и определение уровней экспрессии генов. Связь мутаций с геномно обусловленными болезнями человека	4		14	18	опрос
Тема 7 Влияние внешних условий среды обитания организма на структуру его генома и статистически достоверная аппроксимация данных	3		6	9	опрос
Тема 8 Эволюция геномов (видов), описываемых геномными структурами. Понятия точного и с гарантированной точностью алгоритма и времени его работы (на примерах алгоритмов оптимизации на графах). Полиномиальные низкой степени алгоритмы для задач геномного преобразования и реконструкции геномов.	4		14	18	опрос
Тема 9 Исследование генов, связанных со снижением регенеративной потенции и развитием мозга у теплокровных. Поиск генов, характерных для высокой видовой продолжительности жизни	3		6	9	опрос
Тема 10 Описание точного линейного по времени и памяти алгоритма кратчайшего преобразования циклических графов (слов) друг в друга и доказательство его точности	3		6	9	опрос

Другие виды самостоятельной работы (отсутствуют)	—	—			—
Промежуточная аттестация (экзамен)					
Итого	32		76	108	—

Семестр 2

Наименование разделов и тем дисциплины (модуля), Форма промежуточной аттестации по дисциплине (модулю)	Номинальные трудозатраты обучающегося			Всего академических часов	Форма текущего контроля успеваемости* (наименование)
	Контактная работа (работа во взаимодействии с преподавателем) Виды контактной работы, академические часы		Самостоятельная работа обучающегося, академические часы		
	Занятия лекционного типа	Занятия семинарского типа			
Тема 1 Алгоритм решения задачи кратчайшего преобразования геномных структур общего вида с	3		6	9	опрос

равным составом при равных ценах операций над ними					
Тема 2 Топологические ограничения на геном: цикличность, число хромосом в нём. Задачи кратчайшего преобразования геномов при наличии топологических ограничений. Алгоритмы их решения	3		6	9	опрос
Тема 3 Структуры с неравным генным составом и операции над ними. Алгоритм построения кратчайшего по числу операций преобразования для циклических структур без паралогов	3		6	9	опрос
Тема 4 Алгоритм построения кратчайшего по числу операций преобразования для структур общего вида без паралогов	3		6	9	опрос
Тема 5 Цены операций и задачи преобразования геномных структур с ценами. Алгоритм построения кратчайшего по суммарной цене преобразования произвольных структур с неравным составом без паралогов, если цены вставки и удаления участка хромосомы меньше цен переклеек	3		6	9	опрос
Тема 6 Алгоритм построения кратчайшего по суммарной цене преобразования произвольных структур с неравным составом без паралогов, если цены вставки и удаления участка хромосомы больше цен переклеек	4		14	18	опрос
Тема 7 Задача кратчайшего преобразования геномных структур квазиравного состава с учётом операции дупликации гена. Алгоритм её решения в простейшем случае.	3		6	9	опрос
Тема 8 Задача преобразования циклических структур с неравным составом и с паралогами. Её сведение к задаче целочисленного линейного программирования	4		14	18	опрос

Тема 9 Задача реконструкции предковых геномных структур вдоль данного филогенетического дерева. Точный квадратичный по времени алгоритм её решения при равных ценах SCJ-операций и отсутствии паралогов.	3		6	9	опрос
Тема 10 Точный кубический по времени алгоритм решения задачи реконструкции для простейших деревьев при произвольных ценах SCJ-операций и отсутствии паралогов. Эвристический алгоритм решения этой задачи на произвольном дереве, строящий локально оптимальную реконструкцию.	3		6	9	опрос
Другие виды самостоятельной работы (отсутствуют)	—	—			—
Промежуточная аттестация (экзамен)					
Итого	32		76	108	—

5.2. Содержание разделов (тем) дисциплины

Семестр 1

№ п/п	Наименование разделов (тем) дисциплины	Содержание разделов (тем) дисциплин
1.	Тема 1 Начало молекулярной биологии клетки	Основные понятия молекулярной биологии клетки. Основные функции ДНК, РНК, белков. Генетический код. Процессы транскрипции и трансляции в клетке. Альтернативный сплайсинг, его роль в эволюции.
2.	Тема 2 Модель аттенуаторной регуляции у бактерий и анализ многокомпонентной динамической системы	Экспрессия гена, её регуляция на уровне транскрипции и трансляции. Основные виды регуляции. Аттенуаторная регуляция экспрессии. Моделирование взаимодействия РНК-полимеразы, рибосомы и вторичной структуры РНК.
3.	Тема 3 Начало теории эволюции и филогенетики	Филогенетические деревья, эволюционные процессы на них. Типы деревьев. Основные методы построения филогенетических деревьев по матрице расстояний между листьями или по множественному выравниванию геномов в листьях. Согласование дерева эволюции ортологичных генов с деревом эволюции видов. Задача реконструкции предковых характеристик генома по заданным характеристикам в листьях.
4.	Тема 4 Эволюция регуляторного сигнала с вторичной структурой и минимизация сложного функционала (в частности, аннилинг)	Регуляторная область гена. Вторичная структура в регуляторной области и эволюционные события в ней. Эволюция регуляторной области как минимизация функционала. Вероятностный алгоритм аннилинга, его сходимость к минимальному значению.

5.	Тема 5 Взаимодействия РНК-полимераз, промоторов, терминаторов и анализ соответствующей многокомпонентной динамической системы.	Моделирование процесса движения РНК-полимераз с учётом их взаимодействия с промоторами и терминаторами. Описание и анализ соответствующей многокомпонентной динамической системы.
6.	Тема 6 Моделирование движения РНК-полимераз и определение уровней экспрессии генов. Связь мутаций с геномно-обусловленными болезнями человека	Определение уровней экспрессии генов в модели взаимодействия РНК-полимераз. Выявление участков генома, ответственных за уровень экспрессии гена путём искусственных мутаций в геноме. На этой основе определение участков генома, мутации которых приводят к геномно обусловленным болезням человека.
7.	Тема 7 Влияние внешних условий среды обитания организма на структуру его генома и статистически достоверная аппроксимация данных	Зависимость структуры генома от внешних условий обитания организма. Установление зависимости характеристик генома (его размера, длины межгенных промежутков и др.) от температуры внешней среды. Методы и технологии обработки эмпирических данных для определения степени этой зависимости. Основные биоинформатические базы информации и методы извлечения данных из них.
8.	Тема 8 Эволюция геномов (видов), описываемых их геномными структурами. Понятие точного алгоритма и времени его работы на примерах алгоритмов оптимизации на графах. Применение таких алгоритмов к построению эволюционного сценария геномных преобразований	Представление генома в виде ориентированного графа. DCJ-операции (переклейки) над ним. Задача кратчайшего преобразования одного генома в другой. Алгоритмы оптимизации, их точность и время работы. Примеры алгоритмов. Представление двух геномов с равным генным составом в виде общего графа, операции над ним. Сведение задачи кратчайшего преобразования одного генома в другой к задаче приведения общего графа к финальному виду.
9.	Тема 9 Исследование генов, связанных со снижением регенеративной потенции и развитием мозга у теплокровных. Поиск генов, характерных для высокой видовой продолжительности жизни	Понятие синтении и ортологичности генов. Поиск генов, исчезновение или появление которых в процессе эволюции приводит к снижению регенеративной потенции и развитию переднего мозга у позвоночных. Алгоритм поиска генов, характерных для высокой видовой продолжительности жизни.
10.	Тема 10 Описание точного линейного по времени и памяти алгоритма кратчайшего преобразования циклических графов (слов) друг в друга и доказательство его точности	Решение задачи кратчайшего преобразования одного циклического генома в другой при равном генном составе двойными переклейками. Алгоритм приведения к финальному виду циклического общего графа, доказательство его точности и линейного времени работы. Переход от преобразования общего графа к преобразованию исходных структур друг в друга.

Семестр 2

№ п/п	Наименование разделов (тем) дисциплины	Содержание разделов (тем) дисциплин
1.	Тема 1 Алгоритм решения задачи кратчайшего преобразования геномных структур общего вида с равным составом при равных ценах операций над ними	Сведение задачи кратчайшего преобразования геномных структур общего вида к задаче приведения к финальному виду общего графа, состоящего из циклов и цепей. Алгоритм решения последней задачи, доказательство его точности и линейного времени работы. Примеры работы алгоритма. Переход от преобразования общего графа к преобразованию исходных структур друг в друга.
2.	Тема 2 Топологические ограничения на геном: цикличность, число хромосом в нём. Задачи кратчайшего преобразования геномов при наличии топологических ограничений. Алгоритмы их решения	Виды хромосом и их количество у различных организмов в живой природе. Важность топологических ограничений на структуру генома в задаче преобразования геномных структур. Возможность кратковременного нарушения ограничений, связанная с функционированием транспозонов и плазмид. Точные квадратичные по времени алгоритмы решения задач кратчайшего преобразования с топологическими ограничениями, допускающие возможность таких нарушений.
3.	Тема 3 Структуры с неравным генным составом и операции над ними. Алгоритм построения кратчайшего по числу операций преобразования для циклических структур без паралогов	DCJ-операции и операции вставки и удаления участка хромосомы для геномных структур с неравным генным составом. Общий граф двух структур и операции над ним. Сведение задачи кратчайшего преобразования циклических геномных структур к задаче приведения к финальному виду циклического общего графа. Алгоритм решения последней задачи, доказательство его точности и линейного времени работы. Примеры работы алгоритма. Переход от преобразования общего графа к преобразованию исходных структур друг в друга.
4.	Тема 4 Алгоритм построения кратчайшего по числу операций преобразования для структур общего вида без паралогов	Сведение задачи кратчайшего преобразования геномных структур общего вида с неравным генным составом к задаче приведения к финальному виду общего графа, состоящего из циклов и цепей. Общая схема алгоритма решения последней задачи, его точность и линейное время работы. Примеры работы алгоритма. Переход от преобразования общего графа к преобразованию исходных структур друг в друга.
5.	Тема 5 Цены операций и задачи преобразования геномных структур с ценами. Алгоритм построения кратчайшего по суммарной цене преобразования произвольных структур с неравным составом без паралогов, если цены вставки и удаления участка хромосомы меньше цен переклеек	Цены операций над геномом, принципы их подбора. Возможные соотношения цен и их связь со стабильностью генного состава и структуры генома в процессе эволюции. Случаи, когда цены вставки и удаления участка хромосомы следует положить меньшими цен переклеек. Особенности алгоритма решения задачи кратчайшего преобразования геномных структур в этом случае. Примеры работы алгоритма.
6.	Тема 6 Алгоритм построения кратчайшего по суммарной цене преобразования произвольных структур с неравным составом без паралогов, если цены вставки и удаления участка хромосомы больше цен переклеек	Случаи, когда цены вставки и удаления участка хромосомы следует положить большими цен переклеек. Особенности алгоритма решения задачи кратчайшего преобразования геномных структур в этом случае. Возможность аддитивной ошибки, примеры работы алгоритма.
7.	Тема 7 Задача кратчайшего преобразования геномных структур квазиравного состава с учётом операции дубликации гена. Алгоритм её решения в простейшем случае	Понятие паралогичных генов. Операции линейной и циклической дубликации гена. Общий граф двух структур квазиравного состава с паралогами, операции над ним и его финальный вид. Сведение задачи кратчайшего преобразования геномных структур квазиравного состава с паралогами к задаче приведения общего графа к финальному виду. Приближённый алгоритм решения последней задачи при наличии не более двух паралогов каждого гена в итоговой структуре и их отсутствия в исходной структуре. Переход от преобразования

		общего графа к преобразованию исходных структур друг в друга. Примеры работы алгоритма.
8.	Тема 8 Задача кратчайшего преобразования циклических структур с паралогами. Её сведение к задаче целочисленного линейного программирования	Постановка задачи кратчайшего преобразования циклических структур с паралогами, включающая поиск оптимального соответствия между паралогами. Формулировка этой задачи в терминах целочисленного линейного программирования (ЦЛП). Известные программы решения задач ЦЛП. Примеры работы алгоритма.
9.	Тема 9 Задача реконструкции предковых геномных структур на данном филогенетическом дереве. Точный квадратичный по времени алгоритм её решения при равных ценах SCJ-операций и отсутствии паралогов.	Постановка задачи реконструкции предковых геномных структур на данном филогенетическом дереве по заданным структурам в листьях. SCJ-операции над геномом: склейка и расклейка хромосомы, вставка и удаление хромосомы из одного гена. Задача реконструкции предковых геномных структур на данном филогенетическом дереве, как задача определения присутствующих в каждой вершине генов и пар склеенных краёв генов. Точный квадратичный по времени алгоритм её решения при равных ценах SCJ-операций и отсутствии паралогов. Примеры работы алгоритма.
10.	Тема 10 Точный кубический по времени алгоритм решения задачи реконструкции для простейшего дерева типа «звезда» при произвольных ценах SCJ-операций. Эвристический алгоритм решения этой задачи на произвольном дереве, строящий локально оптимальную реконструкцию.	Точный кубический алгоритм реконструкции для простейшего дерева типа «звезда», основанный на сведениях задачи к задаче построения максимального взвешенного паросочетания в графе. Случаи структур без паралогов и структур с квазиравным составом и операцией дупликации гена. Эвристический алгоритм реконструкции для произвольного дерева, строящий локально оптимальную реконструкцию. Поиск начальной расстановки с помощью квадратичного программирования.

6. Фонд оценочных средств (ФОС, оценочные и методические материалы) для оценивания результатов обучения по дисциплине (модулю).

6.1. Типовые контрольные задания или иные материалы для проведения текущего контроля успеваемости, критерии и шкалы оценивания (в отсутствие утвержденных соответствующих локальных нормативных актов на факультете)

Практические задания

Семестр 1.

1. Тестирование программы моделирования взаимодействия РНК-полимеразы, рибосомы и шпилечной структуры РНК.
2. Освоение методов и программ построения филогенетических деревьев по матрице расстояний между листьями или по множественному выравниванию геномов.
3. Освоение алгоритма и программы аннелинга для установления эволюции вторичной структуры регуляторного сигнала.
4. Тестирование программы моделирования процесса движения РНК-полимераз с учётом их взаимодействия с промоторами и терминаторами. Подбор параметров модели.
5. Освоение приёмов работы с основными биоинформатическими базами данных. Обработка эмпирических данных для определения зависимости структуры генома от температуры внешней среды.
6. Освоение программ решения оптимизационных задач на графах.

7. Тестирование программы поиска генов, исчезнувших или появившихся на различных этапах эволюции.
8. Программирование алгоритма решения задачи кратчайшего преобразования одного циклического генома в другой при равном генном составе двойными переклейками.
9. Найти стационарные состояния для уравнения $x' = 2x(1-x/8) - 5x^2/(4+x^2)$. Указать, какие из них устойчивые и какие неустойчивые.
10. При каком t решение $x(t)$ уравнения Ферхюльста проходит через точку перегиба, если $x(0)=2$ и уравнение: $x' = (\ln 2)x(1-x/18)$?
11. Найти стационарные состояния для уравнения $x' = 18x^2/(2+2x) - 5x - 0.5x^2$. Указать, какие из них устойчивые и какие неустойчивые.
12. Построить бифуркационную диаграмму для уравнения $x' = x^3 - (a-1)x$. Указать устойчивые и неустойчивые ветви и точку бифуркации.
13. Провести 10 итераций для дискретной модели, заданной равенством $N_{t+1} = N_t e^{r(1-N_t/1000)}$, если $N_0=900$, $r=1.5$. Сделать вывод.
14. Для дискретной модели, заданной равенством $N_{t+1} = N_t + N_t^3 - 0.25N_t$ найти стационарные состояния, указать, какие из них устойчивые и какие неустойчивые.

15. Провести 10 итераций для матричной популяционной модели, задаваемой матрицей
$$\begin{pmatrix} 0.75 & 3 & 2 \\ 0.75 & 0 & 0 \\ 0 & 0.5 & 0 \end{pmatrix}$$
 при начальной численности возрастных групп (10, 10, 10). Сделать вывод.

16. Определить вид фазовых траекторий и нарисовать фазовый портрет линейной системы $x' = x - 2y$, $y' = 3x + 4y$.
17. Найти стационарные состояния и определить тип фазовых траекторий в окрестности каждого из них для модели химических реакций, заданной системой уравнений $x' = 2 - 2x$, $y' = 2x - y$. Нарисовать примерный фазовый портрет этой системы.
18. Найти стационарные состояния и определить тип фазовых траекторий в окрестности каждого из них для модели химических реакций, заданной системой нелинейных уравнений $x' = 6 - xy$, $y' = xy - 2y$. Нарисовать примерный фазовый портрет этой системы.
19. Найти стационарные состояния и определить тип фазовых траекторий в окрестности каждого из них для модели типа "Хищник-Жертва", заданной системой нелинейных уравнений $x' = 2x - xy$, $y' = xy - y$. Нарисовать примерный фазовый портрет этой системы.
20. Автоколебательная система химических реакций (модель "брюсселятор") задана системой уравнений $x' = 2 + x^2y - (p+1)x$, $y' = px - x^2y$, где p – параметр. Найти стационарное состояние модели и исследовать, как зависит его тип от параметра p . Построить параметрическую диаграмму, указывая типы стационарного состояния для различных областей p .
21. Модель гликолиза задана системой уравнений $x' = 1 - xy$, $y' = 2y - (x - (1+r)/(1+ry))$, где r – параметр. Найти стационарное состояние модели и исследовать, как зависит его тип от параметра r . Построить параметрическую диаграмму, указывая типы стационарного состояния для различных областей r .
22. Найти однородное стационарное состояние системы двух уравнений типа "реакция–диффузия": $x' = 1 - xy + D_x x''$, $y' = y(x - 2/(1+y)) + D_y y''$. Определить, какая переменная является активатором, а какая – ингибитором.
23. Свести задачу определения стационарного режима в данной стехиометрической модели к задаче линейного программирования. Модель задана реакциями $4a \rightarrow b + 2c$, $b + c \rightarrow 2a$, $c \rightarrow a$.
24. На кодирующем участке РНК AUGCGUGCCGCCGAACCCCGGGUCACUUUCGU GAUCGCGGGUAUUGGUUCCCGGCUGCCGCUGGCUCGCGAGCUGUGUUGUCGC UGCGUCUGUCGCUGCGCCUAAUUUCUGCACGUCUCCUGACGUCUCGG

GCGCGGUAUCGCCCCAGUUCACAACACCAAAAGCCACCCUUAGCCCGCGCUU
CAGCAUUCGUCUGAACCCGGCCUUUUUCGCCAAAGGACCCAUCUC

укажите старт-кодон, пять цистеиновых кодонов, стоп-кодон и U-богатый участок.

25. Укажите возможную последовательность состояний модели на последовательности ДНК, соответствующую её транскрипции и одновременной трансляции образующегося участка РНК. Последовательность ДНК:

CCACATGAACATACAAATACATTGGGAAGCCGGATCAACTTTCTTATTTTTCATTCATTGAAACAGATGTTAATGA

26. Нарисовать общую схему вторичной структуры РНК и указать её компоненты: спирали, внешние и внутренние петли, выпячивания, мультипетли, псевдоузлы.

27. Построить энергетически выгодную вторичную структуру в следующей последовательности РНК: AAGGUUUAUUAACACUUU.

28. Начальная вторичная структура задана в последовательности РНК подчёркиванием плеч спирали:

AUUGAAACAAAUUUCAUUGUUU. Указать возможную последовательность состояний при переходе к более мощной вторичной структуре.

29. Найти аттенуаторную вторичную структуру в последовательности РНК:

UGUUGGGGCAGGCUGCUGAGCGAAAGAAAUUCACAAAAAAGGCCUGUAUCCAACAAGAUACAGGCCUUUUUUUA

30. Найти регуляторную структуру типа Т-Вох в последовательности РНК:

AAGGAGGGUGGUACCGCGGGGCGUCUGCCUCGUCCUCCGUCAGGUGACCAGCACCCCUGAUGGAAAGGUACGCCAC

31. Найти регуляторную RFN-структуру в последовательности РНК:

AGAUCUGUCUUCAGGGCGGGGGGUGAAAUUCGCCACCGGCGGUAAAUCGGAAUGUCGC

CGCUAGCCUGAUGUGAGCCGGGCUCGCCAGAUCACGCGCGAAAUUCGCGGAUCUG

GCUCCGGAGCCGACGGUCAUAGUCCGGAUGGAAGAAGGCGGGG

32. На кольцевой ДНК

GTACSTATGTAGGTAATCCATTCATATGGCGTATTGCCATTTACTA

РНК-полимеразы могут двигаться по обеим цепям. Укажите области их возможных столкновений.

33. На ДНК

GTGTTCCCGAACAGTCAGGCAGTTCTCGCTGCCGCGCGGCGTGCGTGTGTTGTTGCGCACACCGCCGCGCCTGACTGCGCCCGGGACTGATCGAGTCCTGCTGA

укажите промотор и сайт связывания регуляторного белка.

34. Дано множество белков, заданных своими последовательностями ДНК (в скобках указан вид, из которого взят белок): ААТСАТАА(*a*), АГТГГГТТ(*a*), ААТСАТАТ(*a*), АГТГТГТТ(*b*), ААССАСАА(*b*), АГТГАГТТ(*c*), ААТСАСАА(*c*), ТААГГТГА(*c*). Проведите кластеризацию этих белков. Укажите событие возможной дупликации гена и возможного горизонтального переноса гена.

35. В наборе последовательностей ДНК: TTCGGATCGTAAGAAGATTAA, TACCGTAGAAGATAACTCTTAGGCC,

ТАССААТГГГААСССТТАГААГААТТАТАСТАТАСТАС, CCTCCTAGAAGAАТААТАСТАААТТТ укажите ультраконсервативные фрагменты.

36. У предка геном состоит из двух кольцевых хромосом (1,2,3,4); (5,6,7), а у потомка – из двух кольцевых хромосом (8,9,10); (11,12,13,14), где 1=АСАСАСАС, 2=ТАТАТАТА, 3=АГАГАГАГ, 4=АТСГАТСГ, 5=СГСГСГСГ, 6=СТСТСТСТ, 7=ТГТГТГТГ, 8=ТАТАТАТА, 9=АГАГАГАГ, 10=АСАСАСАС, 11=ССТАСТА, 12=СТСТСТСТ, 13=СГСГСГСГ, 14=ТГТГТГТГ. Какой ген предположительно возник в процессе эволюции, какой потерялся? Какие гены могли поменять механизм своей регуляции?

Семестр 2.

1. Программирование алгоритма решения задачи кратчайшего преобразования геномных структур общего вида с равным составом.
2. Программирование алгоритма решения задачи кратчайшего преобразования геномных структур с топологическими ограничениями.
3. Программирование алгоритма решения задачи кратчайшего преобразования циклических геномных структур с неравным составом без паралогов.
4. Исследование зависимости результатов алгоритма построения кратчайшего преобразования геномных структур от цен операций на простых примерах. Поиск примеров, на которых алгоритм даёт аддитивную ошибку. Определение возможных способов совершенствования алгоритма для её устранения.
5. Программирование алгоритма решения задачи кратчайшего преобразования геномных структур квазиравного состава с паралогами и операцией дубликации в простейшем случае.
6. Освоение пакета целочисленного линейного программирования (ЦЛП). Его применение для решения задачи кратчайшего преобразования структур с паралогами.
7. Программирование алгоритма решения задачи реконструкции предковых геномных структур на данном филогенетическом дереве при равных ценах SCJ-операций и отсутствии паралогов.
8. Программирование алгоритма решения задачи реконструкции для простейшего дерева типа «звезда», основанный на сведении задачи к задаче построения максимального взвешенного паросочетания в графе.
9. Программирование эвристического алгоритма реконструкции для произвольного дерева, строящего локально оптимальную реконструкцию. Поиск начальной расстановки с помощью квадратичного программирования.
10. Построить глобальное выравнивание двух последовательностей: AACAGATGAC и ACTGACTGA.
11. Построить локальное выравнивание двух последовательностей: AAGTCACA и GTAACACC.
12. Построить множественное выравнивание трёх последовательностей: ACTGAC, ATATGC и TAGAC.
13. Построить дерево данного набора образцов и провести по нему поиск этих образцов в данной последовательности ДНК. Образцы: АТТС, АТАА, ААТТ, АТАТС, АТТ; последовательность: АТАТАСААТАСАТАТАААТТАТТ.
14. Построить суффиксное дерево данной последовательности ДНК и провести по нему поиск данного образца в ней. Последовательность: ACCGCTCAACCGCTCA; образец: СТСА.
15. Построить суффиксное дерево данного набора последовательностей ДНК и провести по нему поиск общего фрагмента максимальной длины в них. Последовательности: АТТААТТ, СТСТТААС, ТААТТАА.
16. Построить суффиксный массив данной последовательности ДНК и провести по нему поиск данного образца в ней. Последовательность: ACCGCTCAACCGCTCA; образец: СТСА.
17. Построить скрытую Марковскую модель и провести по ней поиск в данной последовательности ДНК участка с преобладанием в своём буквенном составе символа Т. Последовательность: АТААТТТАТТТАААТТТТА.
18. Построить скрытую Марковскую модель и провести по ней поиск в данной последовательности РНК неточного вхождения заданного образца. Последовательность: AGGGAGAGGAAAG; образец: AGAAGA.

19. Найти наиболее мощную вторичную структуру в последовательности ACACCUGGAAUUUAGU.

20. Построить корневое филогенетическое дерево по множественному выравниванию последовательностей ДНК методом UPGMA.
Выравнивание:

ACG-TATT

A-GACAT-

-CTAC-TT

CCGAT-TT

A-TGCA-T

21. Построить бескорневое филогенетическое дерево по множественному выравниванию последовательностей ДНК методом Neighbor Joining.
Выравнивание:

ACG-TATT

A-GACAT-

-CTAC-TT

CCGAT-TT

A-TGCA-T

22. Выписать задачу линейного программирования для вычисления методом максимальной экономии длин ветвей филогенетического дерева $((a,b),c),(d,e))$.

23. Выписать задачу квадратичного программирования для вычисления методом минимизации невязки длин ветвей филогенетического дерева $((a,(b,c)),(d,(e,f)))$.

24. Провести расстановку по внутренним вершинам филогенетического дерева символов, заданных в его листьях, по принципу максимальной экономии. Дерево дано в скобочной записи:

$((T,A), ((-,C), (A,C), (-,C)))$.

25. Построить оптимальное вложение данного дерева генов в данное дерево видов при отсутствии горизонтальных переносов. Указать события дупликации и потери гена, вычислить цену вложения. Дерево генов: $((a',b'),((c',d'),(e',f'))),(g',h'),f''))$; дерево видов: $((a,b),(c,d)),(e,f),(g,h)))$.

26. Построить оптимальное вложение данного дерева генов в данное дерево видов с учетом горизонтальных переносов. Указать события дупликации, потери гена, горизонтальных переносов с сохранением копии в источнике и без сохранения. Вычислить цену вложения. Дерево генов: $((a',b'),((c',d'),(e',f'))),(h',p'),e'))$; дерево видов: $((a,b),(c,d,e)),(f,g),(h,p)))$.

27. Провести бинаризацию данного небинарного дерева генов относительно данного дерева видов. Дерево генов: $((a',b',c'),(d',e',f'),(f'',g',h'))$; дерево видов: $((a,b),(c,d)),(e,f),(g,h)))$.

28. Построить филогенетическое дерево, ближайшее в среднем к следующему набору деревьев: $(a,(b,c)); (b,(a,d)); (a,(b,e)); (c,(a,d)); (c,(a,e)); (e,(a,d)); (c,(b,d)); (c,(b,e)); (d,(b,e)); (c,(d,e))$.

29. Найти строгие блоки синтении в следующем наборе линейных хромосом:

$H1=[3,-2,6,-8,-1,4,7,5,-9,11,10]$; $H2=[9,-5,-6,2,-3,11,10,-7,-4,1,8]$;

$H3 = [-10, -11, 5, -9, -8, -1, 4, 7, 3, -2, 6]$.

30. Построить по данному набору линейных хромосом PQ-дерево, представляющее множество нестрогих блоков синтении у линейных хромосом $H1 = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12]$ и $H2 = [5, 4, 2, 3, 1, 9, 6, 10, 7, 8, 12, 11]$.

32. Найти кратчайшую последовательности инверсий, переводящих линейную хромосому $H1 = [-2, 5, -3, 6, 8, 7, 9, -1, 10, 4]$ в линейную хромосому $H2 = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]$.

33. Построить общий граф двух хромосомных структур $S1$ и $S2$, где $S1 = \{(1, 2, -3); (4, -5, 6, 7)\}$ и $S2 = \{(7, -3, 5); (1, -4, -2, 6)\}$. Вычислить расстояние между данными структурами.

34. Указать кратчайшую последовательность двойных переклеек, преобразующих хромосомную структуру $S1 = \{(1, -7, -6, -3, -5, -4, -2)\}$ в хромосомную структуру $S2 = \{(1, 3, 4); (6, -2, 5, -7)\}$.

35. Указать кратчайшую последовательность переклеек, преобразующих хромосомную структуру $S1 = \{[7, 6, -4, -1, -2, 5]; [3]\}$ в хромосомную структуру $S2 = \{[-7, 1, -3, -2]; [-4, 6, 5]\}$.

36. Провести сборку кольцевой хромосомы из следующих первичных фрагментов (в скобках указано число вхождений фрагмента в хромосому, когда оно больше 1): CGAA, GAAC, AACG(3), ACGT(3), CGTT, GTTA, TTAA, TAAC(2), CGTA, GTAA, CGTC, GTCG, TCGA. Найти два возможных решения.

37. Провести поиск псевдопериодичностей на следующем участке ДНК: GATTAACSTAATAG.

Примеры контрольных работ, семестр 1

КР №1: Популяционные модели, описываемые одним дифференциальным уравнением.

Пример варианта:

1. Найти стационарные состояния для уравнения $x' = 2x(1-x/8) - 5x^2/(4+x^2)$. Указать, какие из них устойчивые и какие неустойчивые.
2. При каком t решение $x(t)$ уравнения Ферхюльста проходит через точку перегиба, если $x(0) = 2$ и уравнение: $x' = (\ln 2)x(1-x/18)$?
3. Найти стационарные состояния для уравнения $x' = 18x^2/(2+2x) - 5x - 0.5x^2$. Указать, какие из них устойчивые и какие неустойчивые.
4. Построить бифуркационную диаграмму для уравнения $x' = x^3 - (a-1)x$. Указать устойчивые и неустойчивые ветви и точку бифуркации.

КР №2: Дискретные модели роста популяции

Пример варианта:

1. Провести 10 итераций для дискретной модели, заданной равенством $N_{t+1} = N_t e^{r(1-N_t/1000)}$, если $N_0 = 900$, $r = 1.5$. Сделать вывод.
2. Для дискретной модели, заданной равенством $N_{t+1} = N_t + N_t^3 - 0.25N_t$ найти стационарные состояния, указать, какие из них устойчивые и какие неустойчивые.

3. Провести 10 итераций для матричной популяционной модели, задаваемой матрицей
$$\begin{pmatrix} 0.75 & 3 & 2 \\ 0.75 & 0 & 0 \\ 0 & 0.5 & 0 \end{pmatrix}$$
 при начальной численности возрастных групп (10, 10, 10). Сделать вывод.

КР №3: Модели, описываемые системой двух линейных дифференциальных уравнений

Пример варианта:

1. Определить вид фазовых траекторий и нарисовать фазовый портрет линейной системы $x' = x - 2y$, $y' = 3x + 4y$.
2. Определить вид фазовых траекторий и нарисовать фазовый портрет линейной системы $x' = -2x + 2y$, $y' = x - 3y$.
3. Определить вид фазовых траекторий и нарисовать фазовый портрет линейной системы $x' = y$, $y' = 2x + y$.

КР №4: Простейшие модели химических реакций, описываемые системой двух нелинейных дифференциальных уравнений

Пример варианта:

1. Найти стационарные состояния и определить тип фазовых траекторий в окрестности каждого из них для модели химических реакций, заданной системой нелинейных уравнений $x' = 2 - 2x$, $y' = 2x - y$. Нарисовать примерный фазовый портрет этой системы.
2. Найти стационарные состояния и определить тип фазовых траекторий в окрестности каждого из них для модели химических реакций, заданной системой нелинейных уравнений $x' = 6 - xy$, $y' = xy - 2y$. Нарисовать примерный фазовый портрет этой системы.
3. Найти стационарные состояния и определить тип фазовых траекторий в окрестности каждого из них для модели химических реакций, заданной системой нелинейных уравнений $x' = 5 - 4xy$, $y' = 4xy - 2y$. Нарисовать примерный фазовый портрет этой системы.

КР №5: Модели конкуренции видов, описываемые системой двух нелинейных дифференциальных уравнений

Пример варианта:

1. Найти стационарные состояния и определить тип фазовых траекторий в окрестности каждого из них для простейшей модели типа "Хищник-Жертва", заданной системой нелинейных уравнений $x' = 2x - xy$, $y' = xy - y$. Нарисовать примерный фазовый портрет этой системы.
2. Найти стационарные состояния и определить тип фазовых траекторий в окрестности каждого из них для модели типа "Хищник-Жертва" с учётом внутривидовой конкуренции. Модель задана системой уравнений $x' = 3x - xy - x^2$, $y' = xy - y - y^2$. Нарисовать примерный фазовый портрет этой системы.
3. Найти стационарные состояния и определить тип фазовых траекторий в окрестности каждого из них для простейшей модели конкуренции равноправных видов, заданной системой нелинейных уравнений $x' = x - xy$, $y' = y - y$. Нарисовать примерный фазовый портрет этой системы.
4. Найти стационарные состояния и определить тип фазовых траекторий в окрестности каждого из них для модели конкуренции равноправных видов с учётом внутривидовой конкуренции. Модель задана системой уравнений $x' = 2x - xy - x^2$, $y' = 4y - xy - 3y^2$. Нарисовать примерный фазовый портрет этой системы.

КР №6: Колебания и предельные циклы в моделях химических реакций

Пример варианта:

1. Найти предельный цикл в колебательной системе, заданной в полярных координатах системой уравнений $r' = r(1 - 2r^2)$, $\varphi' = 2\pi$. Определить его устойчивость.
2. Найти предельные циклы в колебательной системе, заданной в полярных координатах системой уравнений $r' = r(1 + 2r^2 - 2r^4)$, $\varphi' = 2\pi$. Указать, какие из них устойчивые и какие неустойчивые.

3. Автоколебательная система химических реакций (модель "брюсселятор") задана системой уравнений $x' = 2 + x^2y - (p+1)x$, $y' = px - x^2y$, где p – параметр. Найти стационарное состояние модели и исследовать, как зависит его тип от параметра p . Построить параметрическую диаграмму, указывая типы стационарного состояния для различных областей p .

Примеры контрольных работ, семестр 2

КР №1: Реконструкция предковых состояний на филогенетическом дереве

Пример варианта:

1. Методом динамического программирования провести расстановку по внутренним вершинам филогенетического дерева символов, заданных в его листьях, по принципу максимальной экономии, используя меру сходства символов. Дерево дано в скобочной записи:
 $((T, A), ((-, C), (A, C), (-, C)))$.
2. Методом динамического программирования провести расстановку по внутренним вершинам филогенетического дерева символов, заданных в его листьях, по принципу максимальной экономии, используя меру различия символов. Дерево дано в скобочной записи:
 $((A, T), (-, T), (T, C))$.

КР №2: Согласование филогенетических деревьев между собой

Пример варианта:

1. Построить оптимальное вложение данного дерева генов в данное дерево видов при отсутствии горизонтальных переносов. Указать события дубликации и потери гена, вычислить цену вложения.
 Дерево генов: $((a', b'), ((c', d'), (e', f'))), ((g', h'), f'')$;
 дерево видов: $((a, b), (c, d)), ((e, f), (g, h)))$.
2. Построить оптимальное вложение данного дерева генов в данное дерево видов с учетом горизонтальных переносов. Указать события дубликации, потери гена, горизонтальных переносов с сохранением копии в источнике и без сохранения. Вычислить цену вложения. Дерево генов: $((a', b'), ((c', d'), (e', f'))), ((h', p'), e'')$;
 дерево видов: $((a, b), ((c, d), e)), ((f, g), (h, p)))$.
3. Провести бинаризацию данного небинарного дерева генов относительно данного дерева видов. Дерево генов: $((a', b', c'), (d', e', f'), (f'', g', h'))$; дерево видов: $((a, b), (c, d)), ((e, f), (g, h)))$.
4. Построить филогенетическое дерево, ближайшее в среднем к следующему набору деревьев: $(a, (b, c)); (b, (a, d)); (a, (b, e)); (c, (a, d)); (c, (a, e)); (e, (a, d)); (c, (b, d)); (c, (b, e)); (d, (b, e)); (c, (d, e))$.

КР №3: Блоки синтении и инверсии

Пример варианта:

1. Найти строгие блоки синтении в следующем наборе линейных хромосом:
 $H1 = [3, -2, 6, -8, -1, 4, 7, 5, -9, 11, 10]$; $H2 = [9, -5, -6, 2, -3, 11, 10, -7, -4, 1, 8]$;

$H3 = [-10, -11, 5, -9, -8, -1, 4, 7, 3, -2, 6]$.

2. Построить по данному набору линейных хромосом PQ-дерево, представляющее множество нестрогих блоков синтении у линейных хромосом $H1 = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12]$ и $H2 = [5, 4, 2, 3, 1, 9, 6, 10, 7, 8, 12, 11]$.

3. Найти кратчайшую последовательности инверсий, переводящих линейную хромосому $H1 = [-2, 5, -3, 6, 8, 7, 9, -1, 10, 4]$ в линейную хромосому $H2 = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]$.

КР №4: Хромосомные структуры и их перестройки

Пример варианта:

1. Построить общий граф двух хромосомных структур $S1$ и $S2$, где

$S1 = \{(1, 2, -3); (4, -5, 6, 7)\}$ и $S2 = \{(7, -3, 5); (1, -4, -2, 6)\}$. Вычислить расстояние между данными структурами.

2. Указать кратчайшую последовательность двойных переклеек, преобразующих хромосомную структуру $S1 = \{(1, -7, -6, -3, -5, -4, -2)\}$ в хромосомную структуру $S2 = \{(1, 3, 4); (6, -2, 5, -7)\}$.

3. Указать кратчайшую последовательность переклеек, преобразующих хромосомную структуру $S1 = \{[7, 6, -4, -1, -2, 5]; [3]\}$ в хромосомную структуру $S2 = \{[-7, 1, -3, -2]; [-4, 6, 5]\}$.

КР №5: Сборка генома и поиск в нём псевдопериодичностей

Пример варианта:

1. Используя граф де Брюйна, провести сборку кольцевой хромосомы из следующих первичных фрагментов (в скобках указано число вхождений фрагмента в хромосому, когда оно больше 1): CGAA, GAAC, AACG(3), ACGT(3), CGTT, GTTA, TТАА, ТАAC(2), CGTA, GTAA, CGTC, GTCG, TCGA. Найти два возможных решения.

2. С помощью выравнивания последовательности с ней самой провести поиск псевдопериодичностей на следующем участке ДНК: GATTAACТААТАG.

6.2. Типовые контрольные задания или иные материалы для проведения промежуточной аттестации по дисциплине, критерии и шкалы оценивания

Семестр 1

Вопросы к зачёту

1. Как выглядит ДНК, РНК, белковая цепь? В чём заключаются процессы транскрипции и трансляции в клетке? Что такое сплайсинг?
2. Каковы основные виды регуляции экспрессии гена? В чём суть аттенуаторной регуляции?
3. Какие методы построения филогенетических деревьев Вы знаете? Какие задачи связаны с филогенетическими деревьями?
4. Что такое регуляторный сигнал? Как выразить его качество? В чём суть алгоритма аннелинга?
5. Что такое промотор и терминатор? Как РНК-полимераза взаимодействует с ними?

6. Как с помощью моделирования выявляют участки генома, ответственные за уровень экспрессии гена? Как такие участки могут быть связаны с геномно обусловленными болезнями человека?
7. Какие биоинформатические базы данных Вы знаете? Для чего производится статистическая обработка данных?
8. Приведите пример изображения генома в виде ориентированного графа. Произведите над ним какую-нибудь DCJ-операцию.
9. В чём состоит задача кратчайшего преобразования одного генома в другой? Приведите пример представления двух геномов виде общего графа. Как привести этот граф к финальному виду?
10. Что такое синтения и как уточняется на её основе понятие ортологичности генов у разных видов?
11. В чём суть алгоритма приведения к финальному виду циклического общего графа? Приведите пример такого преобразования. Как по общему графу двух геномных структур восстановить эти структуры?
12. Какие биологические процессы можно моделировать одним дифференциальным уравнением? Что такое стационарное состояние, как оценить его устойчивость? Что такое точка бифуркации?
13. Что описывает уравнение Ферхюльста? Какие у него стационарные состояния? Какие уточнения этого уравнения Вам известны?
14. Какие биологические процессы можно моделировать дискретной зависимостью? Как выглядит лестница Ламерея? Что такое неподвижная точка, как определить её устойчивость?
15. Каков дискретный аналог уравнения Ферхюльста? Какой тип дискретных траекторий может у него возникать?
16. Как выглядит матричная модель динамики популяции? Каков смысл элементов матрицы?
17. Какие биологические процессы можно моделировать системой двух линейных дифференциальных уравнений? Что такое фазовые траектории и фазовый портрет. Какие типы особых точек можно увидеть на нём?
18. Как найти сепаратриссы седла? Как определить направление траектории на фазовом портрете?
19. Какие биологические процессы можно моделировать системой двух нелинейных дифференциальных уравнений? Что такое стационарное состояние? Как провести линеаризацию системы в окрестности стационарного состояния?
20. Что такое фермент-субстратная реакция Михаэлиса-Ментен? Какие типы бифуркаций вы знаете? Приведите пример модельной системы для описания бифуркаций и постройте её бифуркационную диаграмму.
21. Опишите схему модели типа «Хищник-жертва». Приведите примеры систем с двумя устойчивыми стационарными состояниями.
22. Как моделируется конкуренция двух равноправных видов и как моделируется симбиоз видов? В чём разница между этими моделями?
23. Как определить области притяжения устойчивых состояний в модели конкуренции видов? При каком соотношении коэффициентов внутривидовой и межвидовой конкуренции оба вида могут устойчиво сосуществовать?
24. Что такое автоколебания. Изобразите поведение автоколебательной системы на фазовой плоскости. Как выглядят предельные циклы?
25. Чем отличается предельный цикл от центра? Чем отличаются сценарии мягкого и жесткого возбуждения автоколебаний?
26. Как выглядит система двух уравнений типа "реакция–диффузия"? Чем является её решение общем виде? Что такое гомогенное стационарное состояние этой системы? Что такое «неустойчивость Тьюринга»? Какая переменная называется активатором, а какая – ингибитором?
27. Какие базы данных можно использовать для построения метаболических моделей? Как строится матрица стехиометрических коэффициентов по метаболическому пути?

28. Что такое стехиометрическая модель системы химических реакций. Как можно найти её стационарный режим? Какие приближения используются для решения уравнений баланса? Какие условия необходимо сформулировать для постановки задачи линейного программирования?
29. Опишите основные характеристики РНК-полимеразы и рибосомы. Как моделировать процессы транскрипции и трансляции? В чём может состоять статистическая обработка результатов моделирования?
30. Как выглядят вторичные структуры РНК, каковы их основные компоненты? Опишите схему компьютерного моделирования их динамики. В чём может состоять статистическая обработка результатов моделирования?
31. Какова роль вторичных структур РНК в регуляции экспрессии генов, в транскрипции и трансляции? Каковы основные виды регуляции экспрессии? Как вычислять энергию вторичной структуры РНК? Опишите схему моделирования регуляции экспрессии генов вторичными структурами РНК.
32. Какова роль конкуренции РНК-полимераз в регуляции экспрессии генов? Опишите схему моделирования такой конкуренции. В чём может состоять статистическая обработка результатов моделирования?
33. Каковы возможные причины геномных болезней человека, связанные с уровнем экспрессии генов? Опишите схему моделирования возникновения таких болезней. В чём может состоять статистическая обработка результатов моделирования?
34. Каковы основные типы белков, взаимодействующих с ДНК, в чём их роль? Опишите схему моделирования диффузии белка вдоль ДНК. В чём может состоять статистическая обработка результатов моделирования?
35. Каков механизм связывания белков с ДНК? Какие типы промоторов вы знаете? Как ищутся сайты связывания белков и промоторы?
36. Что такое транскрипционный фактор? Какова роль транскрипционных факторов в регуляции экспрессии генов? Как можно предсказать транскрипционный фактор по филогенетическому профилю консервативного сайта или структуры РНК?
37. Что означает, что два гена являются ортологами или паралогами? В чём смысл кластеризации белков? Как она происходит?
38. Каков механизм событий дупликации и горизонтального переноса генов? Как можно обнаружить эти события?
39. По каким причинам происходят потери и возникновения генов в процессе эволюции? Как можно обнаружить такие гены? Что такое синтения генов и как она может помочь в поиске?
40. Что такое аллель, генотип и фенотип? Как понимать термин «стационарное состояние популяции»? При каких условиях популяция находится в стационарном состоянии?
41. Какие аллели определяют группу крови человека? Какие группы крови бывают у человека? По какому принципу они передаются от родителей к ребёнку?
42. Как вычислить вероятность вырождения популяции по распределению вероятностей количеств детей у одного индивида? При каких условиях популяция вырождается?
43. Что такое вариационный ряд и какие его характеристики Вы знаете?

1. Привести примеры геномных структур, общего графа двух структур, операций над структурой и соответствующих операций над общим графом.
2. Дать определения DCJ-операций. Привести пример кратчайшего по числу операций преобразования одной структуры в другую, и соответствующего приведения общего графа к финальному виду.
3. Какие топологические ограничения на структуру генома Вы знаете? В чём важность соблюдения этих ограничений?
4. Какие операции применяются к геномным структурам с переменным генным составом? Как выглядит общий граф таких структур? Привести пример кратчайшего по числу операций преобразования одной структуры в другую (для неравного состава), и соответствующего приведения общего графа к финальному виду.
5. Из каких соображений устанавливаются цены операций над геномом? Каковы возможные соотношения цен и как они связаны со стабильностью генного состава и структуры генома в процессе эволюции?
6. При каком соотношении цен операций возможна аддитивная ошибка алгоритма построения кратчайшего преобразования структур? Приведите пример.
7. Что такое паралогичные гены? Что означает, что две геномные структуры имеют квазиравный состав? Что такое операция линейной и циклической дупликации гена? Как выглядит общий граф двух структур квазиравного состава с паралогами, какие операции допускаются над ним, каков его финальный вид?
8. Что такое неориентированные геномные структуры? Опишите схему сведения задачи преобразования ориентированных структур с паралогами к задаче преобразования неориентированных структур без паралогов. Приведите пример.
9. В чём состоит задача кратчайшего преобразования циклических структур с паралогами? Опишите общую схему её сведения к задаче целочисленного линейного программирования (ЦЛП). Какие программы (пакеты) решения задач ЦЛП Вам известны?
10. Что такое SCJ-операции? В чём состоит задача реконструкции предковых геномных структур на данном филогенетическом дереве? Какова основная идея алгоритма её решения при равных ценах SCJ-операций и отсутствии паралогов? Приведите пример реконструкции.
11. Какова основная идея алгоритма реконструкции предковых геномных структур для простейшего дерева типа «звезда», при неравных ценах SCJ-операций? Приведите пример реконструкции.
12. Опишите схему эвристического алгоритма реконструкции для произвольного дерева, строящего локально оптимальную реконструкцию. Как можно искать начальную расстановку?
13. Что такое попарное выравнивание. Как определить его качество? Что такое множественное выравнивание?
14. В чём отличие глобального выравнивания от локального? В каких условиях разумно применять глобальное выравнивание и в каких локальное?
15. Что такое делеция, в чём её биологический смысл? Что такое аффинный штраф за делеции, в чём его биологический смысл?
16. Что такое дерево образцов? Когда его разумно использовать? Какие задачи решаются с помощью этого дерева?
17. Что такое суффиксное дерево? Когда его разумно использовать? Какие задачи решаются с помощью этого дерева?
18. Что такое суффиксные ссылки? Для чего они служат?
19. В каких условиях разумно использовать дерево образцов, а в каких – суффиксное дерево?
20. Что такое суффиксный массив? Какие задачи решаются с помощью суффиксных массивов?
21. Что такое скрытая Марковская модель? Какие у неё параметры? Какие задачи решаются с помощью Марковских моделей?

22. Как осуществляется подбор параметров скрытой Марковской модели?
23. Что такое вторичная структура РНК? Какие биологические функции она имеет? Что такое псевдоузел?
24. Указать на данной вторичной структуре РНК её компоненты: спирали, внешние и внутренние петли, выпячивания, мультипетли, псевдоузлы.
25. Какие вторичные структуры РНК представляют интерес? В чём суть аттенуаторной регуляции транскрипции?
26. Какие бывают филогенетические деревья? Как определить качество филогенетического дерева?
27. Какие существуют методы построения филогенетических деревьев?
28. Что отражают длины ветвей филогенетического дерева? Какие существуют методы их определения?
29. Опишите основные характеристики РНК-полимеразы и рибосомы. Как моделировать процессы транскрипции и трансляции? В чём может состоять статистическая обработка результатов моделирования?
30. Как выглядят вторичные структуры РНК, каковы их основные компоненты? Опишите схему компьютерного моделирования их динамики. В чём может состоять статистическая обработка результатов моделирования?
31. Какова роль вторичных структур РНК в регуляции экспрессии генов, в транскрипции и трансляции? Каковы основные виды регуляции экспрессии? Как вычислять энергию вторичной структуры РНК? Опишите схему моделирования регуляции экспрессии генов вторичными структурами РНК.
32. Какова роль конкуренции РНК-полимераз в регуляции экспрессии генов? Опишите схему моделирования такой конкуренции. В чём может состоять статистическая обработка результатов моделирования?
33. Каковы возможные причины геномных болезней человека, связанные с уровнем экспрессии генов? Опишите схему моделирования возникновения таких болезней. В чём может состоять статистическая обработка результатов моделирования?
34. Каковы основные типы белков, взаимодействующих с ДНК, в чём их роль? Опишите схему моделирования диффузии белка вдоль ДНК. В чём может состоять статистическая обработка результатов моделирования?
35. Каков механизм связывания белков с ДНК? Какие типы промоторов вы знаете? Как ищутся сайты связывания белков и промоторы?
36. Что такое транскрипционный фактор? Какова роль транскрипционных факторов в регуляции экспрессии генов? Как можно предсказать транскрипционный фактор по филогенетическому профилю консервативного сайта или структуры РНК?
37. Что означает, что два гена являются ортологами или паралогами? В чём смысл кластеризации белков? Как она происходит?
38. Каков механизм событий дупликации и горизонтального переноса генов? Как можно обнаружить эти события?
39. По каким причинам происходят потери и возникновения генов в процессе эволюции? Как можно обнаружить такие гены? Что такое синтения генов и как она может помочь в поиске?

Задачи, предлагающиеся на экзамене

Задача 1. Привести общий граф G из циклов и петель (циклический случай) к финальному виду операциями DM (двойная переклейка) и Rem (удаление сингулярной вершины) при одном из указанных условий на цены:

Графы:

1. $(abab)$,
2. $(aabab)$,

3. $(aabbab)$,
4. $(aabbaab)$,
5. $(aabbaabb)$.
6. цикл (abb) и b -петля,
7. два цикла (baa) ,
8. два цикла (abb) и $(aabb)$,
9. два цикла $(aabb)$,
10. цикл $(aabbaabbaabb)$.

Цены:

- 1) цены всех операций равны 1,
- 2) цена двойной переклейки равна 1, цена удаления равна 0.5,
- 3) цена двойной переклейки равна 1, цена удаления равна 1.5,
- 4) цена двойной переклейки равна 1, цена удаления равна 2.5,
- 5) цена двойной переклейки равна 1, цена удаления a -вершины равна 1, цена удаления b -вершины равна 1.5.

Задача 2. Привести общий граф G из цепей и изолированных вершин к финальному виду всеми операциями DM, SM, OM, Cut, Rem при одном из указанных условий на цены:

Графы:

1. $abbaa$.
2. $abba$.
3. $bbaab$.
4. $baabba$.
5. aab и bba .
6. $bbaabb$ и сингулярная изолированная b -вершина.
7. baa и $abba$.
8. bba и aa .
9. ab и $aabb$.
10. ab и bba .
11. две цепи aab .
12. aab , $aabb$ и сингулярная изолированная a -вершина.
13. две цепи aab и сингулярная изолированная a -вершина.

Цены:

- 1) цены всех операций равны 1,
- 2) цены DCJ-операций равны 1, цена удаления равна 0.5.

Задача 3.

Построить вторичную структуру с максимальным количеством склеек в последовательности CCCAGAGGCCGCCUCGG.

Задача 4.

Методом UPGMA построить дерево по множественному выравниванию:

a: CT-TAGTA

b: -TTAAAT-

c: C-TCATCA

d: A-TCA-CA

e: CTCCGTCG

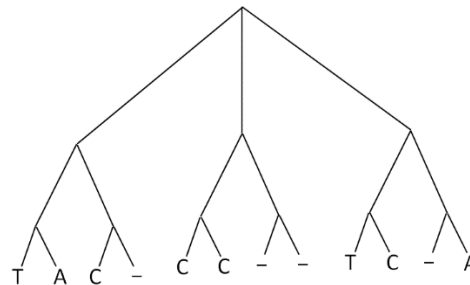
Задача 5.

3) Методом Neighbor Joining построить дерево по матрице расстояний:

	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>
<i>a</i>	-	8	3	10	9	11
<i>b</i>		-	9	12	5	13
<i>c</i>			-	11	10	12
<i>d</i>				-	13	7
<i>e</i>					-	14
<i>f</i>						-

Задача 6.

Методом динамического программирования найти предковые символы в вершинах дерева:

**Задача 7.**

Построить вложение (с горизонтальными переносами) дерева $G = (((a', b'), ((c', d'), (f', e'))), ((g', h'), d''))$ в дерево $S = (((a, b), (c, d)), ((e, f), (g, h)))$.

6) Инверсиями привести хромосому $[-2, 5, -3, 6, 8, 7, 9, -1, 10, 4]$ к каноническому виду кратчайшим образом.

Задача 8.

Бинаризовать дерево $G=((a',b',c'),(f',g',h'),(c'',d',e',f''))$ относительно дерева $S=((a,b),(c,d)),((e,f),(g,h)))$.

Задача 9.

Согласовать 10 деревьев в одно супердерево, где деревья такие:

$((2,3),1); ((1,4),2); (1,(2,5)); ((1,4),3); ((1,5),3); ((1,4),5); ((2,4),3); ((2,5),3); (4,(2,5)); ((4,5),3).$

Задача 10.

Найти строгие блоки синтении у трёх хромосом:

$H_1=[3,-2,6,-8,-1,4,7,5,-9,11,10];$

$H_2=[9,-5,-6,2,-3,11,10,-7,-4,1,8];$

$H_3=[-10,-11,5,-9,-8,-1,4,7,3,-2,6].$

Задача 11.

Найти нестрогие блоки синтении двух хромосом

$H_1=[1,2,3,4,5,6,7,8,9,10,11,12]; H_2=[5,4,2,3,1,9,6,10,7,8,12,11]$

и построить PQ-дерево.

Задача 12.

Двойными переклейками преобразовать циклическую структуру a в циклическую структуру b кратчайшим образом:

$a: \{(2,4,5,3,6,7,-1)\}. b: \{(1,3,4); (6,-2,5,-7)\}.$

Задача 13.

Преобразовать линейную структуру a в линейную структуру b кратчайшим образом с сохранением линейности:

$a: \{[7,6,-4,-1,-2,5]; [3]\}. b: \{[-7,1,-3,-2]; [-4,6,5]\}.$

Задача 14.

Провести сборку генома по графу де Брёйна (найти два решения), если секвенатор выдал следующие последовательности (в скобках указано их число, когда оно больше 1):

$cga a, gaac, aacg(3), acgt(3), cggt, gtta, ttaa, taac(2), cgta, gtaa, cgtc, gtcg, tcga.$

Задача 15.

Дана стехиометрическая матрица (в виде таблицы).

v_1	v_2	v_3	v_4	v_5	v_6	
0	1	0	1	-1	0	S_1
1	0	-1	0	-1	0	S_2
0	1	-1	-1	0	0	S_3
0	0	1	0	0	-1	S_4

Найти скорости реакций в стационарном состоянии, если $v_1 \leq 6$, и требуется максимизировать v_6 .

Задача 16.

В популяции растений, записываемой как $p^2AA+2pqAa+q^2aa$, было обнаружено 743 растения с пурпурным венчиком и 31 с розовым. Признак контролируется одним геном, розовая окраска – рецессивное проявление. Найти частоты генотипов в популяции.

Задача 17.

В популяции $(4/9)AA+(4/9)Aa+(1/9)aa$ найти вероятность появления ребёнка с рецессивным генотипом aa , если оба родителя обладают доминантным признаком.

Задача 18.

В популяции все шесть генотипов по группе крови имеют одинаковую частоту $1/6$. Один родитель имеет группу крови А, другой – группу В. Найти вероятность, что их ребёнок будет иметь группу крови 0.

Задача 19.

Найти вероятность вырождения популяции, если вначале она состояла из одного индивида и каждый индивид имеет 0 потомков с вероятностью $1/8$, одного потомка с вероятностью $3/8$, двух потомков с вероятностью $3/8$ и трёх потомков с вероятностью $1/8$.

ШКАЛА И КРИТЕРИИ ОЦЕНИВАНИЯ результатов обучения (РО) по дисциплине				
Оценка	2 (не зачтено)	3 (зачтено)	4 (зачтено)	5 (зачтено)
виды оценочных средств				
Знания	Отсутствие знаний	Фрагментарные знания	Общие, но не структурированные знания	Сформированные систематические знания

<i>(виды оценочных средств: опрос, тесты)</i>				
Умения <i>(виды оценочных средств: практические задания)</i>	Отсутствие умений	В целом успешное, но не систематическое умение	В целом успешное, но содержащее отдельные пробелы умение (допускает неточности непринципиального характера)	Успешное и систематическое умение
Навыки (владения, опыт деятельности) <i>(виды оценочных средств: выполнение и защита курсовой работы, отчет по практике, отчет по НИР и т.п.)</i>	Отсутствие навыков (владений, опыта)	Наличие отдельных навыков (наличие фрагментарного опыта)	В целом, сформированные навыки (владения), но используемые не в активной форме	Сформированные навыки (владения), применяемые при решении задач

7. Ресурсное обеспечение:

7.1. Перечень основной и дополнительной литературы

Основная литература

1. Сингер М., Берг П. Гены и геномы, тома I-II. Москва, «Мир», 1998.
2. Ней М., Кумар С. Молекулярная эволюция и филогеномика, Киев, 2004.
3. Кортэ Б., Фиген Й. Комбинаторная оптимизация. Теория и алгоритмы. М.: МЦНМО, 2015.

Дополнительная литература

1. Коничев А.С., Севастьянова Г.А. Молекулярная биология. Москва, Издательский центр «Академия», 2008.
2. Хаубольд Б., Вие Т. Введение в вычислительную биологию: эволюционный подход. М. Ижевск: НИЦ «Регулярная и хаотическая динамика», 2011.
3. Омельченко А.В. Теория графов. М.: МЦНМО, 2018.
4. Ризниченко Г.Ю. Лекции по математическим моделям в биологии. М. Ижевск: НИЦ «Регулярная и хаотическая динамика», 2011.
5. Новожилов А.С., Платонов А.П., Братусь А.С. Динамические системы и модели в биологии. М.: Физматлит, 2010.

Дополнительная специальная литература

- Lyubetsky V., Pirogov S., Rubanov L., Seliverstov A. Modeling classic attenuation regulation of gene expression in bacteria, 2007, Journal of Bioinformatics and Computational Biology, vol 5, no 1, p. 155-180.
- L.I. Rubanov, A.G. Zaisky, G.A. Shilovsky, A.V. Seliverstov, O.A. Zverkov, V.A. Lyubetsky. Screening for mouse genes lost in mammals with long lifespans. BioData Mining, Nov 9 2019, Vol. 12, Art. 20.
- В.А. Любецкий, Е.А. Жижина, Л.И. Рубанов. Гиббсовский подход в задаче эволюции регуляторного сигнала экспрессии гена. Проблемы передачи информации, 2008, том 43, вып. 4, стр. 52–71, eLIBRARY: 11755654, Mi: ppi1289.
- Lyubetsky V.A. et al. Modeling RNA polymerase competition: the effect of σ -subunit knockout and heat shock on gene transcription level, Biology Direct, 2011, 6:3, <http://www.biology-direct.com/content/6/1/3>.
- Lyubetsky et al. Optimal growth temperature and intergenic distances in bacteria, archaea, and plastids of rhodophytic branch. BioMed Research International, 2020, Art. 3465380.
- V.A. Lyubetsky, R.A. Gershgorin, K.Yu. Gorbunov. Chromosome structures: reduction of certain problems with unequal gene content and gene paralogs to integer linear programming. BMC Bioinformatics, Dec 6 2017, Vol. 18, Art. 537, 18 pp.
- Korotkova D.D., Lyubetsky V.A., Ivanova A.S., Rubanov L.I., Seliverstov A.V., Zverkov O.A., Martynova N.Yu., Nesterenko A.M., Tereshina M.B., Peshkin L., Zaisky A.G. Bioinformatics screening of genes specific for well-regenerating vertebrates reveals c-answr, a regulator of brain development and regeneration // Cell Reports, 2019, V. 29, № 4, P. 1027–1040.e6. DOI: 10.1016/j.celrep.2019.09.038.
- Rubanov L.I., Zaisky A.G., Shilovsky G.A., Seliverstov A.V., Zverkov O.A., Lyubetsky V.A. Screening for mouse genes lost in mammals with long lifespans // BioData Mining, 2019, V. 12, ID 20. DOI: 10.1186/s13040-019-0208-x.
- K.Yu. Gorbunov, V.A. Lyubetsky. Linear time additively exact algorithm for transformation of chain-cycle graphs for arbitrary costs of deletions and insertions. Mathematics, Nov 10 2020, Vol. 8, No. 11, Art. 2001, 30 pp.
- K.Yu. Gorbunov, V.A. Lyubetsky. Multiplicatively exact algorithms for transformation and reconstruction of directed path-cycle graphs with repeated edges. Mathematics, Oct 14 2021, Vol. 9, No. 20, Art. 2576. DOI: 10.3390/math9202576, ISTINA: 402488632, IITP: 8274.

7.2. Перечень лицензионного программного обеспечения, в том числе отечественного производства

При реализации дисциплины может быть использовано следующее программное обеспечение:

1. Программный продукт Red Hat Enterprise Linux Server for HPC Compute Node for Power, LE, Self-support 4 шт. №5540331
2. Программный продукт Red Hat Enterprise Linux Server for HPC Head Node for Power, LE, Standard 1 шт. №5540332
3. Операционная система SUSE Linux Enterprise Server 11 SP4 for x86_64 16 шт.
4. Операционная система Red Hat Enterprise Linux Server 5.0 for x86_64 14шт.
5. Операционная система SUSE Linux Enterprise Server 10 SP3 for ppc64 7 шт.
6. Операционная система Ubuntu 18.04.

7. Программное обеспечение для виртуализации Oracle VM VirtualBox
8. Операционная система ALT Linux MATE Starterkit 9 лицензия GPL
9. Программный продукт JetBrains IntelliJ IDEA Community Edition Free Educational Licenses
10. Программный продукт JetBrains PyCharm Community Edition Free Educational Licenses
11. Программный продукт JetBrains CLion Community Edition Free Educational Licenses
12. Программный продукт UPPAAL (<http://www.uppaal.org/>) академическая лицензия
13. Программный продукт Java 8 (64-bit) Oracle Corporation
14. Программный продукт Java SE Development Kit 8 (64-bit) Oracle Corporation
15. Программный продукт NetBeans IDE 8.2 NetBeans.org
16. Программный продукт Dev-C++ Bloodshed Software
17. Программный продукт CodeBlocks The Code::Blocks Team
18. Программный продукт Free Pascal 3.0.0 Free Pascal Team
19. Программный продукт Python 3.5.1 (64-bit) Python Software Foundation
20. Программный продукт R for Windows 3.2.2 R Core Team
21. Программный продукт Haskell Platform 7.10.3 Haskell.org
22. Операционная система Microsoft Windows 7 корпоративная академическая лицензия
23. Операционная система Microsoft Windows 10 Education академическая лицензия
24. Программный продукт Microsoft Project Professional 2013 академическая лицензия
25. Программный продукт Microsoft Visio Professional 2013 академическая лицензия
26. Программный продукт Microsoft Visual Studio Professional 2013 - RUS [Русский (Россия)] академическая лицензия

7.3. Перечень профессиональных баз данных и информационных справочных систем

1. <http://www.edu.ru> – портал Министерства образования и науки РФ
2. <http://www.ict.edu.ru> – система федеральных образовательных порталов «ИКТ в образовании»
3. <http://www.openet.ru> - Российский портал открытого образования
4. <http://www.mon.gov.ru> - Министерство образования и науки Российской Федерации
5. <http://www.fasi.gov.ru> - Федеральное агентство по науке и инновациям

7.4. Перечень ресурсов информационно-телекоммуникационной сети «Интернет»

1. Math-Net.Ru [Электронный ресурс] : общероссийский математический портал / Математический институт им. В. А. Стеклова РАН ; Российская академия наук, Отделение математических наук. - М. : [б. и.], 2010. - Загл. с титул. экрана. - Б. ц.
URL: <http://www.mathnet.ru>
2. Университетская библиотека Online [Электронный ресурс] : электронная библиотечная система / ООО "Директ-Медиа" . - М. : [б. и.], 2001. - Загл. с титул. экрана. - Б. ц. URL: www.biblioclub.ru

3. Универсальные базы данных East View [Электронный ресурс] : информационный ресурс / East View Information Services. - М. : [б. и.], 2012. - Загл. с титул. экрана. - Б. ц.

URL: www.ebiblioteka.ru

4. Научная электронная библиотека eLIBRARY.RU [Электронный ресурс] : информационный портал / ООО "РУНЭБ" ; Санкт-Петербургский государственный университет. - М. : [б. и.], 2005. - Загл. с титул. экрана. - Б. ц.

URL: www.eLibrary.ru

5. Публикации сотрудников лаборатории математических методов и моделей в биоинформатике: <http://lab6.iitp.ru/ru/pub/>

6. Материалы спецкурсов и спецсеминаров кафедры математической логики и теории алгоритмов: <http://logic.math.msu.ru/staff/lyubetsky/mb/>

7.5. Описание материально-технического обеспечения.

Механико-математический факультет, ответственный за реализацию данной Программы, располагает соответствующей материально-технической базой, включая современную вычислительную технику, объединенную в локальную вычислительную сеть, имеющую выход в Интернет. Используются специализированные компьютерные классы, оснащенные современным оборудованием. Материальная база факультета соответствует действующим санитарно-техническим нормам и обеспечивает проведение всех видов занятий (лабораторной, практической, дисциплинарной и междисциплинарной подготовки) и научно-исследовательской работы обучающихся, предусмотренных учебным планом.

8. Соответствие результатов обучения по данному элементу ОПОП результатам освоения ОПОП указано в Общей характеристике ОПОП.

9. Разработчик (разработчики) программы.

д.ф.-м.н., профессор Любецкий Василий Александрович