

Федеральное государственное бюджетное образовательное  
учреждение высшего образования  
«МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ М.В.ЛОМОНОСОВА»

ФИЛИАЛ МГУ В Г. ГРОЗНОМ

УТВЕРЖДАЮ

Заместитель директора филиала – руководитель  
образовательных программ

А. С. Воронцов



20\_\_ г.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**

Наименование дисциплины:

**Информатика и биоинформатика**

Уровень высшего образования:

**Специалитет**

Специальность:

**06.05.02 Фундаментальная и прикладная биология**

Направленность (профиль)/специализация образовательной программы:

**Биотехнология**

Форма обучения:

**Очная**

Москва 2024

Рабочая программа дисциплины разработана в соответствии с самостоятельно установленным МГУ образовательным стандартом (ОС МГУ) для реализуемых основных профессиональных образовательных программ высшего образования по специальности 06.05.02 «ФУНДАМЕНТАЛЬНАЯ И ПРИКЛАДНАЯ БИОЛОГИЯ» (образовательная программа специалитета «Биотехнология»).

ОС МГУ утвержден решением Ученого совета МГУ имени М.В.Ломоносова 20.01.2022 года.

Год приема на обучение 2024.

### 1. Место дисциплины в структуре ОПОП

Дисциплина относится к обязательной части ОПОП ВО, входит в модуль «Информатика», раздел учебного плана: Базовая часть., изучается в 3 и 4 семестрах. Следует за или параллельно освоению математических дисциплин («Высшая математика», «Математические методы в биологии») и химических дисциплин («Общая и неорганическая химия», «Органическая химия», «Основы физической химии»).

Дисциплина знакомит слушателей с основными алгоритмами и приложениями геномной и структурной биоинформатики. Рассмотрены вопросы теории информации, теории графов, секвенирования нуклеиновых кислот, выравнивания последовательностей, статистического анализа последовательностей, структурной организации белков и нуклеиновых кислот, определения и предсказания типов пространственной структуры биополимеров и др.

### 2. Входные требования для освоения дисциплины

Для изучения дисциплины требуется предварительное освоение базовых курсов математики, химии («Высшая математика», «Общая и неорганическая химия»).

### 3. Планируемые результаты обучения по дисциплине (модулю), соотнесенные с требуемыми компетенциями выпускников

Компетенция	Индикатор достижения компетенции	Планируемые результаты обучения по дисциплине (модулю), соотнесенные с индикаторами достижения компетенций
<b>УК-7.</b> Способен использовать современные информационно-коммуникационные технологии в академической и профессиональной сферах.	<b>УК-7.1.</b> Использует информационно-коммуникационные технологии для сбора, обработки, хранения, распространения, отображения и использования биологической информации	<b>Знает</b> совокупность технологий, обеспечивающих сбор, обработку, хранение, распространение, отображение и использование информации <b>Умеет</b> использовать информационно-коммуникационные технологии для сбора, обработки, хранения, распространения, отображения и использования биологической информации <b>Владеет навыками</b>

		использования электронных библиотек, информационных и поисковых систем.
<b>ОПК-11.</b> Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности.	<b>ОПК-11.1.</b> Использует современные информационно-коммуникационные технологии в исследовательской деятельности в сфере биологии.	<b>Знает</b> современные проблемы биоинформатики основные алгоритмы и методы, применяемые для анализа последовательностей и структур биомакромолекул <b>Умеет</b> применять биоинформатические подходы при анализе живых систем <b>Владеет навыками</b> применения алгоритмов для анализа структуры биомакромолекул

#### 4. Объем дисциплины

Объем дисциплины - 5 з.е. (180 ак.ч), из них 120 ак.ч - контактная работа обучающихся с преподавателем на занятиях лекционного типа (лекции - 60 ак.ч) и на занятиях семинарского типа (семинары - 60 ч). Самостоятельная работа обучающихся – 60 ак.ч. Форма промежуточной аттестации – экзамены в 3 и 4 семестрах.

#### 5. Форма обучения – очная

#### 6. Содержание дисциплины, структурированное по темам (разделам) с указанием отведенного на них количества академических, и виды учебных занятий

##### 6.1. Разделы дисциплины и их трудоемкость

Наименование разделов и тем дисциплины	семестр	В том числе				Самостоятельная работа, ак.часы
		Виды контактной работы, ак.часы			Всего	
		Занятия лекционного типа (Лекции)	Занятия семинарского типа (Семинары)			
Лекции						
Введение. <u>Предмет и задачи биоинформатики</u> <u>Тема 1. Элементы теории информации</u> <u>Тема 2.</u>	3	5	0	10	5	

<u>Биологические базы данных</u>					
<u>Тема 3. Элементы теории графов</u>	3	5	0	10	5
<u>Тема 4. Сравнение последовательностей</u>					
<u>Тема 5. Множественные выравнивания</u>	3	5	0	10	5
<u>Тема 6. Филогенетические деревья</u>	3	5	0	10	5
<u>Тема 7. Скрытые марковские модели</u>	3	5	0	10	5
<u>Тема 8. Методы секвенирования ДНК</u>	3	5	0	10	5
<u>Тема 9. Ресеквенирование и сборка генома</u>	3	6	0	12	6
<u>Тема 10. Приложения NGS</u>	4	4	0	8	4
<u>Тема 11. Анализ экспрессии генов по данным NGS</u>	4	4	0	8	4
<u>Тема 12. Проблемы аннотации последовательностей</u>	4	4	0	8	4
<u>Тема 13. Проблемы множественного выравнивания при большом числе замен</u>	4	4	0	8	4
<u>Тема 14. Проблемы аннотации бактериальных генов</u>	4	4	0	8	4
<u>Тема 15. Геномика: современное состояние</u>	4	4	0	8	4
<b>Семинары</b>					
<u>Тема 16. Основы структурной биологии</u>	3	0	9	18	9
<u>Тема 17. Объекты и методы структурной биоинформатики</u>	3	0	9	18	9
<u>Тема 18. Предсказание структуры белков</u>	3	0	9	18	9
<u>Тема 19. Методы исследования</u>	3	0	9	18	9

<u>конформационного пространства биополимеров</u>					
<u>Тема 20. Предсказание свойств низкомолекулярных соединений</u>	4	0	8	16	8
<u>Тема 21. Виртуальный скрининг</u>	4	0	8	16	8
<u>Тема 22. Машинное обучение и deep learning в биологии</u>	4	0	8	16	8
Всего:		60	60	180	60

## 6.2. Содержание дисциплины по разделам и темам

### *Введение*

Задачи биоинформатики. Поиск парных подобию нуклеотидных и аминокислотных последовательностей. Множественное выравнивание нуклеотидных и аминокислотных последовательностей. Проблема аннотации биологических последовательностей и поиска генов. Поиск регуляторных сигналов. Генетические сети и моделирование генетической регуляции клетки. Структурная биоинформатика. Базы данных в биоинформатике. Геномика, метаболомика и протеомика. Новые направления в биоинформатике. Лингвистика генетических текстов.

### *Тема 1. Элементы теории информации*

Понятие информации. Способы измерения информации. Теория информации. Информационная энтропия. Сжатие информации. Теорема Котельникова. Теорема Шеннона-Хартли. Шифрование информации. Хранение информации. Источники больших данных в биомедицине. Проблемы передачи больших данных.

### *Тема 2. Биологические базы данных*

Типы баз данных. Биологические БД. Типы биологических баз данных. Центры биологических баз данных. Реферативные базы данных и поиск научной литературы. Базы данных нуклеотидных последовательностей (в том числе GenBank, RefSeq). Базы данных белковых последовательностей. Базы данных трехмерных структур. Базы данных химических соединений. Геномные базы данных. Базы данных геном-фенотип. БД взаимодействий, сигнальных путей. БД результатов секвенирования. БД заболеваний и медицинской информации. БД по экспрессии генов/гистологии. Таксономические БД.

### *Тема 3. Элементы теории графов*

Определения. Примеры. Машинное представление графов. Поиск в графах. Пути и циклы. Связность. Деревья. Эйлеровы и гамильтоновы пути. Задача о мостах и задача коммивояжера. Графы де-Брёйна. Случайные графы.

### *Тема 4. Сравнение последовательностей*

Сравнение последовательностей. Точечная матрица сходства. Расстояние по Хэммингу и Левенштейну. Матрицы замен (PAM, Blosum, JTT и др.). Динамическое программирование, локальное и глобальное выравнивание, алгоритмы Нидлмана-

Вунша и Смита-Уотермана. Примеры работы программ и сайтов по парному выравниванию.

#### *Тема 5. Множественные выравнивания*

Экспресс-методы сравнения последовательностей. Программа Blast. Значимость выравнивания. Множественное выравнивание аминокислотных и нуклеотидных последовательностей, динамическое программирование и его недостатки. Визуализация выравниваний: профили. Связь между числом мутаций и числом консервативных позиций в множественном выравнивании.

#### *Тема 6. Филогенетические деревья*

Методы кластеризации (UPGMA, NJ и др.) и филогенетические деревья. Прогрессивное выравнивание - программа Clustal. Итеративное выравнивание - программа Muscle. Модели эволюции. Филогенетические деревья: переменная скорость эволюции и притяжение длинных ветвей, методы проверки. Скобочная формула.

#### *Тема 7. Скрытые марковские модели*

Марковский процесс. Задачи СММ: задача оценки, задача дешифровки, задача обучения. Алгоритм прямого-обратного хода. Оценка множественного выравнивания. Программа ClustalO. Предсказание топологии мембранных белков. Программа ТМНММ.

#### *Тема 8. Методы секвенирования ДНК*

Методы секвенирования ДНК. Проект "Геном человека". Общие принципы строения ДНК. Секвенирование по Сэнгеру. Принцип методов, основанных на терминации полимеразной реакции. Методы секвенирования второго поколения: Illumina, 454 Roche, SOLiD, полупроводниковое секвенирование. Секвенирование "третьего поколения": Oxford Nanopore, PacBio. Ошибки секвенирования. Сравнение технологий секвенирования.

#### *Тема 9. Ресеквенирование и сборка генома*

Цели, задачи, подходы. Сборка de-novo. Общий принцип методов сборки. Разнообразие форм и размеров генома. Основные проблемы, стоящие перед сборщиком: ошибки секвенирования, повторы, естественный полиморфизм, контаминация образца. Алгоритмические подходы: Overlap-consensus-layout, графы де-Брёйна. Влияние длины прочтения и покрытия на качество сборки. Оценка качества сборки. Источники ошибок в сборках. Контиги и скаффолды. Завершение сборки (автоматические, ручное). Альтернативные подходы к сборке (optical mapping, сборка по HiC).

#### *Тема 10. Приложения NGS*

Hi-C, ChipSeq, MNase, CLIP и т.п. Картирование ридов и алгоритмы, лежащие в его основе.

#### *Тема 11. Анализ экспрессии генов из данных NGS*

Понятие дифференциальной экспрессии. Особенности подсчета числа ридов, приходящихся на ген. Статистические методы обнаружения дифференциальной экспрессии. Алгоритм DEseq2. Способы представления результатов: volcano plot, MA plot. Анализ альтернативного сплайсинга. GO категории, обогащение GO категориями.

#### *Тема 12. Проблемы аннотации последовательностей*

Методы поиска генов и кодирующие потенциалы. Периодичность кодирующих последовательностей. Применение марковских моделей, нейронных сетей и теории

информации для поиска генов. Сайты для аннотации нуклеотидных последовательностей.

#### *Тема 13. Проблемы множественного выравнивания при большом числе замен*

Множественное выравнивание как NP-полная задача. Ограничение существующих методов, основанных на парном сравнении последовательностей. Новый метод поиска множественного выравнивания, позволяющий находить множественное выравнивание при числе замен больше чем 2,5 на нуклеотид. Поиск множественного выравнивания для промоторных последовательностей из геномов *A.Thaliana*, *D.melanogaster* и *H.sapiens*. Проблема поиска transcript start site. Группировка промоторов на основе множественного выравнивания. Множественное выравнивание MIR повторов. Другие приложения множественных выравниваний.

#### *Тема 14. Проблема аннотации бактериальных генов*

Недостатки существующих систем аннотации. Система аннотации бактериальных генов, основанная на филогенетической группировке. Мутации в генах типа сдвиг рамки считывания и склеек генов. Математические методы и алгоритмические подходы для поиска этих мутаций. Классы триплетной периодичности в генах. Базы данных и web-сайты для поиска склеек и сдвигов рамки считывания.

#### *Тема 15. Геномика: современное состояние*

Понятие геномики, структурная геномика, функциональная геномика, фармакогеномика, популяционная геномика, сравнительная геномика, эволюционная геномика. Структуры эукариотических и прокариотических геномов. Ортология, паралогия, синтения, COGs (clusters of orthologous groups). Геном человека. Генетические вариации. Понятие гаплотипа и гаплогруппы. Геномные проекты: геном человека, 1000 Genomes, Epigenomics Roadmap, ENCODE, 4D nucleome, TCGA, ICGC. GWAS исследования, связь генотип-фенотип. Заболевания, наследуемые по Менделю и комплексные заболевания. Базы данных dbSNP, OMIM, ClinVar. Геномные браузеры. Демонстрация браузера Ensembl.

#### *Тема 16. Основы структурной биологии*

Основы структурной биологии: рентгеноструктурный анализ биомолекул. Формат структурных файлов. Представление структуры белков и НК во внутренних координатах. Карты Рамачандрана. Другие примеры применения рентгеновского излучения в биологии.

#### *Тема 17. Объекты и методы структурной биоинформатики*

Структурная биоинформатика: объекты и методы. Уровни структурной организации белков и НК. Типы вторичной структуры белков и НК. Структура РНК и теория узлов. Разнообразие структур биомолекул. Базы данных структур. Неупорядоченные белки. Структурные выравнивания: задание поворота в пространстве (матрицы, углы Эйлера и кватернионы), среднеквадратичное отклонение, методы минимизации функции многих переменных (наискорейший спуск и сопряженные градиенты). Матрицы расстояний. Эволюция и консервативность структуры белков, классификация структур белков, системы SCOP и CATH, поиск белков со схожим типом укладки. Структурная геномика.

#### *Тема 18. Предсказание структуры белков*

Предсказание вторичной структуры и искусственные нейронные сети. Программы PSIPRED и SABLE. Гидрофобность: энтропийная природа и методы определения.

Профили гидрофобности и предсказание топологии белка. Моделирование структуры на основании гомологии: программы SwissModel и Modeller. RosettaCM? Библиотеки ротамеров. Веб-сервис ROSIE как пример универсальной платформы для моделирования. Алгоритмы распознавания пространственной укладки белка (фолда), протягивание. Радиальные функции распределения как способ оценки качества упаковки пространственных структур. Веб-сервис I-TASSER. Лейциновая застежка и предсказание суперспиралей, программа LOGICOIL. CASP. FoldIt. Прионные белки. AlphaFold.

*Тема 19. Методы исследования конформационного пространства биополимеров*  
Молекулярная динамика: общая концепция, силовые поля, возможности и ограничения. Термостат, баростат. Метод Монте-Карло. Алгоритм Метрополиса.

*Тема 20. Предсказание свойств низкомолекулярных соединений*  
Смежные вопросы хемоинформатики. Представление химических соединений: молекулярный граф, матрица смежности, линейные нотации SMILES, InChI, структурные файлы. Количественные отношения "структура-активность" (QSAR) и способы их предсказания. Обучающая и тестовая выборки. Метод наименьших квадратов и линейная регрессия. Методы проверки зависимостей: кросс-валидация, рандомизация. Молекулярные дескрипторы. Меры сходства выборок (коэффициент Жаккара, коэффициент Танимото). Предсказание биологических свойств. Веб-сервис PASS. Восприятие вкуса.

*Тема 21. Виртуальный скрининг*  
Базы данных химических соединений (PubChem, ZINC, PDBbind, DrugBank). Виртуальный скрининг: отбор по формальным признакам, фармакофорный поиск, молекулярный докинг. Веб-сервисы SwissSimilarity и PharmIt. Поверхности молекул: ван-дер-ваальсова, молекулярная и доступная растворителю. Молекулярный докинг: общая постановка задачи. Веб-сервис SwissDock. Генетический алгоритм, программа AutoDock. Оценка межмолекулярных взаимодействий, оценочные функции. Программа LigPlot, Веб-сервис Platinum. Макромолекулярный докинг, программа ZDOCK.

*Тема 22. Машинное обучение и deep learning в биологии*  
Искусственные нейронные сети. Метод опорных векторов. Метод ближайших соседей. Генетический алгоритм.

## 7. Фонд оценочных средств для оценивания результатов обучения по дисциплине

### 7.1. Перечень оценочных средств

Компетенция	Результат обучения по дисциплине (модулю)	Оценочные средства
<b>УК-7.</b> Способен использовать современные информационно-коммуникационные технологии в академической и профессиональной	<b>Знает</b> совокупность технологий, обеспечивающих сбор, обработку, хранение, распространение, отображение и использование информации <b>Умеет</b> использовать информационно-коммуникационные технологии для	1. Вопросы для текущей и промежуточной аттестации (устный и письменный опрос, зачет) 2. Ситуационное кейс-задание



сферах.	сбора, обработки, хранения, распространения, отображения и использования биологической информации <b>Владеет навыками</b> использования электронных библиотек, информационных и поисковых систем.	3. Домашнее задание
<b>ОПК-11.</b> Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности.	<b>Знает</b> <ul style="list-style-type: none"> <li>современные проблемы биоинформатики</li> <li>основные алгоритмы и методы, применяемые для анализа последовательностей и структур биомакромолекул</li> </ul> <b>Умеет</b> <ul style="list-style-type: none"> <li>применять биоинформатические подходы при анализе живых систем</li> </ul> <b>Владеет навыками</b> <ul style="list-style-type: none"> <li>применения алгоритмов для анализа структуры биомакромолекул</li> </ul>	1. Вопросы для текущей и промежуточной аттестации (устный и письменный опрос, зачет) 2. Ситуационное кейс-задание 3. Домашнее задание

## 7.2. Типовые задания и иные материалы, необходимые для оценки результатов обучения

### *Примерные вопросы для проведения текущего контроля успеваемости (устный опрос и домашнее задание)*

1. Записать в шестнадцатиричной системе двоичное число.
2. Рассчитать изменение энтропии при смешении идеальных газов.
3. Рассчитать количество информации, необходимой для кодирования знаков в алфавите.
4. Найти публикацию по doi. Определить число процитированных в ней и число цитирующих её публикаций.
5. Найти последовательность белка по её коду в БД UniProt.
6. Проверить пару заданных графов на изоморфность.
7. Определить связность заданного графа.
8. Проверить, является ли последовательность нуклеотидов последовательностью де-Брёйна.
9. Посчитать долю нуклеотидов в заданной последовательности.
10. Построить разумное выравнивание и определите расстояние по Левенштейну между словами БИОИНФОРМАТИКА и КОНФОРМАЦИЯ.
11. Исходя из предположения о равновероятности замен, найти математическое ожидание числа консервативных позиций в выравнивании длиной 20, содержащем 10 замен.
12. Рассчитать число возможных выравниваний двух последовательностей длиной 3 и 5.
13. Какая аминокислотная замена (N $\leftrightarrow$ D или L $\leftrightarrow$ I) более вероятна по данным матрицы замен BLOSUM62?

14. Вероятность ошибочного прочтения нуклеотида в последовательности составила 0,32. Рассчитать качество прочтения этого нуклеотида.
15. В некоей позиции множественного выравнивания имеются набор аминокислот AVAVLLLVL. Рассчитайте высоту букв в этой позиции при построении профиля выравнивания.
16. По заданной матрице расстояний построить филогенетическое дерево методом UPGMA и представить его запись в формате Newick.
17. Во сколько раз число укоренённых деревьев с 12 листьями больше числа неукоренённых с таким же числом листьев?
18. Рассчитать кодирующий потенциал Шульмана для последовательности {atcgta}, повторенной 10 раз.
19. Какие внутренние координаты могут быть использованы для описания структуры белков? ДНК?
20. Какие свойства катодной трубки влияют на спектр испускаемого рентгеновского излучения?
21. Почему рентгеноструктурный анализ, как правило, не позволяет определить положение атомов водорода в молекулах белков?
22. Рассчитать координаты на плоскости вектора  $v_2$ , получаемому из вектора  $v_1$  путем поворота, описываемого матрицей  $M$ .
23. Для двух заданных пар точек на плоскости найти оптимальное соответствие и рассчитать среднеквадратичное отклонение при совмещении.
24. Построить профиль гидрофобности для заданной последовательности белка и определить вероятное число трансмембранных фрагментов.
25. Рассчитайте, на какое расстояние сместится атом массой  $m$  под действием силы  $F$  за время  $t$ .
26. Выберите запись в формате SMILES, соответствующую молекуле на рисунке
27. Среди веществ, чьи формулы приведены, выберите обладающие сладким вкусом.
28. Какое число акцепторов водородных связей содержится в молекуле с заданной структурной формулой?
29. Рассчитайте коэффициент Танимото по заданной библиотеке фрагментов для молекул с заданными структурными формулами.
30. Оцените температуру кипения (по шкале Цельсия) нормального алкана  $C_{40}H_{82}$ .

### ***Типовые вопросы для проведения промежуточной аттестации***

1. Однобуквенные обозначение нуклеотидов и аминокислот.
2. Биоинформатика, общее представление о дисциплине
3. Информация: измерение, передача, хранение
4. Биологические базы данных (классификация и примеры)
5. Меры сходства последовательностей. Точечная матрица сходства.
6. Матрицы замен: происхождение и применение
7. Метод динамического программирования. Алгоритм Нидлмана-Вунша и его вариации
8. Экспресс-методы сравнения последовательностей. Программы FASTA и BLAST. Значимость выравнивания
9. Множественное выравнивание и алгоритмы (Clustal, Muscle). Недостатки существующих методов множественного выравнивания.
10. Методы кластеризации: UPGMA и NJ.
11. Профиль множественного выравнивания
12. Филогенетические деревья, бутстреп. Проблема переменной скорости эволюции
13. Скрытые марковские модели: основные понятия, примеры использования

14. Основные понятия теории графов. Гамильтонов обход, эйлеров обход графа. Последовательность де Брёйна
15. Секвенирование по Сэнгеру, принцип метода
16. Метод секвенирования путем синтеза с обратимой терминацией, платформа Illumina, мостиковая амплификация
17. Пиросеквенирование, принцип метода
18. Отличительные особенности методов секвенирования второго и третьего поколения
19. Сборка генома с помощью подхода OLC. Сборка с помощью графов де Брёйна
20. Влияние длины чтения и покрытия на качество сборки
21. Оценка качества сборки. Методы улучшения сборки. Методы скаффолдинга
22. Парноконцевые прочтения (Paired-end reads) - методика на примере Illumina
23. Картирование чтений. Основные понятия, связанные с картированием чтений на геном
24. Алгоритмы, используемые картировщиками: Хэш-таблица, Суффиксное дерево
25. Повторности в биологическом эксперименте.
26. Понятие случайной величины. Какое распределение имеет число ридов на ген, как случайная величина?
27. Нормализация. Какие виды нормализации применяются при анализе экспрессии генов и почему они нужны?
28. Какие параметры распределения сравнивают, чтобы узнать, изменилась ли экспрессия гена в заданном образце?
29. Если на ген А в образце 1 приходится 100 ридов, а в образце 2 - 200 ридов, значит ли это, что ген изменил экспрессию? Почему?
30. Диаграмма "vulcano plot": построение и чтение.
31. Методы для поиска периодичности в символьных последовательностях.
32. Позиционно-весовые матрицы и их использование для построения множественного выравнивания.
33. Филогенетический профиль гена. Отличие генов-ортологов от генов-паралогов. Использование филогенетического профиля гена для его аннотирования.
34. Точки разладки в генах и процессы, их создающие.
35. Сдвиг фазы триплетной периодичности гена. Как можно найти в гене потенциальный сдвиг рамки считывания?
36. Уровни структурной организации белков. Типы вторичной структуры.
37. Способы задания поворота в пространстве. Преимущества и недостатки.
38. Три варианта постановки задачи структурного выравнивания. Меры структурного сходства.
39. Классификация структуры белков. Структурная геномика.
40. Принципы рентгеноструктурного анализа. Разрешение.
41. Внутренние координаты белков и нуклеиновых кислот. Карты Рамачандрана.
42. Принцип работы простейших нейронных сетей.
43. Предсказание вторичной структуры белков.
44. Гидрофобность и её использование для предсказания топологии белков.
45. Принципы моделирования по гомологии.
46. Распознавание фолда. Протягивание. Неупорядоченные белки.
47. Молекулярное моделирование как вид вычислительного эксперимента.
48. Моделирование молекулярной динамики: алгоритмы и применение.
49. Линейные нотации (SMILES) и матрицы смежности.
50. Количественные соотношения между структурой и свойствами. Методы проверки предсказывающих функций.

51. Меры сходства молекул. Молекулярные дескрипторы. Битовые строки. Коэффициент Танимото.
52. Виды виртуального скрининга. Понятие фармакофора.
53. Виды молекулярных поверхностей и их особенности.
54. Алгоритмы молекулярного докинга.

**Методические рекомендации  
для обеспечения самостоятельной работы студентов**

*Для овладения знаниями:*

- чтение текста (учебника, методического пособия, первоисточника, дополнительной литературы);
- составление плана текста;
- графическое изображение структуры текста;
- конспектирование текста;
- выписки из текста

*Для закрепления и систематизации знаний:*

- работа с конспектом лекции (обработка текста);
- повторная работа над учебным материалом (учебника, первоисточника, дополнительной литературы);
- составление плана и тезисов ответа;
- ответы на контрольные вопросы

*Для формирования умений:*

- решение задач и упражнений по образцу;
- решение вариантов задач и упражнений;
- применение ПО, разработанного для моделирования биологических систем, для решения задач.

### 7.3. Описание шкал и критериев оценивания

**Описание критериев оценивания выполнения задания**

Показатель	Баллы
Студент выполняет менее 50% задания	0-20
Задание студент выполняет все или большей частью, есть отдельные неточности, способен при направляющих вопросах исправить допущенные неточности	21-32
Задание выполнено студентом правильно, самостоятельно в полном объеме	33-40

**Шкала оценивания сформированности компетенций**

Уровень сформированности компетенции	Баллы	Оценка в 5-ти балльной шкале	Оценка на зачете
недостаточный	Менее 20	неудовлетворительно	не зачтено
базовый	20-26	удовлетворительно	зачтено
Высокий (повышенный)	27-32	хорошо	
Продвинутый (повышенный)	33-40	отлично	

<b>ШКАЛА И КРИТЕРИИ ОЦЕНИВАНИЯ результатов обучения по дисциплине (модулю)</b> (*оценка сформированности компетенций дается в соответствии со шкалой выше)				
Оценка Рез-т обучения	<b>2 (не зачтено)</b>	<b>3 (зачтено)</b>	<b>4 (зачтено)</b>	<b>5 (зачтено)</b>
Знания (приведены в п.3.)	Отсутствие знаний	Фрагментарны е знания	Общие, но не структурированные знания	Сформированные систематические знания
Умения (приведены в п.3.)	Отсутствие умений	В целом успешное, но не систематическ ое умение	В целом успешное, но содержащее отдельные пробелы умение (допускает неточности непринципиального характера)	Успешное и систематическое умение
Навыки /владения/опыт деятельности (приведены в п.3.)	Отсутствие навыков (владений, опыта деятельности)	Наличие отдельных навыков	В целом, сформированные навыки (владения), но используемые не в активной форме	Сформированные навыки (владения), применяемые при решении задач

## 8. Ресурсное обеспечение

### 8.1. Перечень основной и дополнительной учебной литературы

#### а) Основная литература:

1. Биоинформатика : учеб. для акад. бакалавриата. / В. В. Стефанов, А. А. Тулуб, Г. Р. Мавропуло-Столяренко. - М. : Юрайт, 2017
2. Молекулярная эволюция и филогенетический анализ: учеб.пособие для студентов. / В.В.Лукашов. - М. : БИНОМ. Лаб. знаний, 2009.
3. Задачи и решения по анализу биологических последовательностей / Марк Бородовский, Светлана Екишева ; пер. с англ. А. А. Чумичкина под ред. А. А. Миронова. - Ижевск : Ин-т компьютер. исслед. : R&C Dynamics, 2008.
4. Физика белка : Курс лекций с цв. стереоскоп. ил. и задачами с решениями: Учеб.пособие для студентов вузов. / А.В.Финкельштейн,О.Б.Птицын;Ин-т белка РАН. - М. : Университет, 2005.
5. Молекулярное моделирование: теория и практика. / Х.-Д. Хельтье [и др.] ; пер. с англ. А. А. Олиференко [и др.] под ред. В. А. Палюлина и Е. В. Радченко. - М. : БИНОМ. Лаб. знаний, 2013.

#### б) Дополнительная литература:

1. Введение в биоинформатику / А. Леск ; пер.с англ. под ред. А. А. Миронова и В. К. Швядоса. - М. : БИНОМ. Лаб. знаний, 2017.
2. Pevzner, J. Bioinformatics and functional genomics, 3rd edition
3. Математические методы для анализа последовательностей ДНК / К.Дж.Бенхэм, Б.Э.Блейсделл, К.Буркс, [и др.] ; Пер. с англ.: М.С.Гельфанд ; Ред.: М.С.Уотермен ; Ред. пер. с англ.: П.А.Певзнер . - М. : Мир, 1999.
4. Искусство молекулярной динамики / Д. К. Рапапорт ; пер. с англ. А. Н. Дьяконовой под науч. ред. Р. Г. Ефремова. - Ижевск : Ин-т компьютер. исслед., 2012.

## **8.2. Перечень лицензионного и(или) свободно распространяемого программного обеспечения, в том числе отечественного производства**

1. AutoDock
2. LigPlot
3. ZDOCK
4. Blast
5. Clustal
6. Muscle
7. Google Chrome
8. Libre Office
9. Adobe Acrobat Reader

## **8.3. Перечень ресурсов информационно-телекоммуникационной сети «Интернет»**

1. Зарубежные журналы и библиографические базы данных, доступные через Интернет <https://www.ncbi.nlm.nih.gov/>
2. Материалы на сайте «Центр развития электронных образовательных ресурсов»: <https://distant.msu.ru/course/view.php?id=1654>
3. Материалы на сайте кафедры биоинженерии: <https://intbio.org/bioinf2019-2020/>
4. Базы данных химических соединений (PubChem, ZINC, PDBbind, DrugBank)
5. Веб-сервис PASS
6. Веб-сервисы SwissSimilarity и PharmIt
7. Веб-сервис Platinum
8. Веб-сервис SwissDock

## **8.4. Описание материально-технической базы**

Для освоения дисциплины требуется свободный доступ к сети Интернет, а также:

- Аудитории для проведения лекционных и семинарских занятий, текущего контроля успеваемости и промежуточной аттестации
  - А. Помещения: аудитории для проведения лекционных/семинарских, текущего контроля и промежуточной аттестации: компьютерный класс (25 персональных компьютеров, объединенных в сеть с возможностью выхода в Интернет).
  - Б. Оборудование: наборы ученической мебели, рабочее место преподавателя, ученическая доска, компьютер, проектор, экран, доска.

## **9. Язык преподавания**

Русский.

## **10. Преподаватели**

Доктор биологических наук, профессор, руководитель группа математического анализа последовательностей ДНК и белков ФИЦ «Фундаментальные основы биотехнологии» РАН Коротков Евгений Вадимович

Кандидат физико-математических наук, старший научный сотрудник каф. биоинженерии биологического факультета МГУ Армеев Григорий Алексеевич

Кандидат биологических наук, старший научный сотрудник каф. молекулярной биологии биологического факультета МГУ Герасимов Евгений Сергеевич

Кандидат физико-математических наук, доцент каф. биоинженерии биологического факультета Новоселецкий Валерий Николаевич

Кандидат физико-математических наук, ведущий научный сотрудник каф. биоинженерии биологического факультета Шайтан Алексей Константинович.

### **11. Разработчики программы**

Доктор биологических наук, профессор, руководитель группы математического анализа последовательностей ДНК и белков ФИЦ «Фундаментальные основы биотехнологии» РАН Коротков Евгений Вадимович

Кандидат физико-математических наук, старший научный сотрудник каф. биоинженерии биологического факультета МГУ Армеев Григорий Алексеевич

Кандидат биологических наук, старший научный сотрудник каф. молекулярной биологии биологического факультета МГУ Герасимов Евгений Сергеевич

Кандидат физико-математических наук, доцент каф. биоинженерии биологического факультета Новоселецкий Валерий Николаевич.