

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ НАУЧНОЕ
УЧРЕЖДЕНИЕ «НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ИНСТИТУТ
БИОМЕДИЦИНСКОЙ ХИМИИ ИМЕНИ В.Н. ОРЕХОВИЧА»
(ИБМХ)**

ОДОБРЕНО

Решением Ученого совета

Протокол № 5 от «30» июня 2022 г.

УТВЕРЖДЕНО

Приказом директора ИБМХ

от «01» июля 2022 г. № 43

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
«МАТЕМАТИЧЕСКАЯ БИОЛОГИЯ, БИОИНФОРМАТИКА»**

Научная специальность

1.5.8 Математическая биология, биоинформатика

Москва 2022 г.

Рабочая программа дисциплины «Математическая биология, биоинформатика» разработана в соответствии с Федеральными государственными требованиями, утверждёнными приказом Министерства науки и высшего образования Российской Федерации от 20.10.2021 № 951, работниками ИБМХ.

№ п/п	Фамилия Имя Отчество	Должность
1	Поройков Владимир Васильевич	Руководитель отдела биоинформатики
2	Веселовский Александр Владимирович	Заведующий лабораторией структурной биоинформатики
3	Ильгисонис Екатерина Викторовна	Заведующая лабораторией постгеномных данных
4	Лагунин Алексей Александрович	Главный научный сотрудник лаборатории структурно-функционального конструирования лекарств
5	Соболев Борис Николаевич	Ведущий научный сотрудник лаборатории структурно-функционального конструирования лекарств
6	Иванов Сергей Михайлович	Старший научный сотрудник лаборатории структурно-функционального конструирования лекарств

ОГЛАВЛЕНИЕ

1. Цель и задачи изучения дисциплины	4
2. Объем дисциплины по видам учебной работы.....	4Ошибка! Закладка не определена
3. Содержание дисциплины	4
4. Учебно-тематический план дисциплины	6
5. Учебно-методическое обеспечение самостоятельной работы обучающихся	9
6. Оценочные средства для проведения текущего контроля успеваемости и промежуточной аттестации обучающихся	10
7. Учебно-методическое и информационное обеспечение дисциплины	14
8. Материально-техническое обеспечение дисциплины	18
9. Методические указания для обучающихся по изучению дисциплины	18
10. Методические рекомендации преподавателю по организации учебного процесса по дисциплине	19

1. Цель и задачи изучения дисциплины

Цель изучения дисциплины

Изучение дисциплины «Математическая биология, биоинформатика» предоставляет профессиональную подготовку научных и научно-педагогических кадров высшей квалификации по научной специальности 1.5.8 Математическая биология, биоинформатика.

Задачи дисциплины

1. Сформировать навыки работы с геномными данными.
2. Сформировать знания об использовании методов биоинформатики.
3. Сформировать у обучающегося умения в освоении новейших технологий и методик в сфере профессиональных интересов по научной специальности «Математическая биология, биоинформатика».
4. Сформировать навыки использования методов биоинформатики и крупнейших международных интернет ресурсов биомедицинских данных необходимых для решения фундаментальных и прикладных биомедицинских задач.

2. Объем дисциплины по видам учебной работы

Таблица 1

Виды учебной работы	Всего, час.	Объем по полугодиям (семестрам)							
		1	2	3	4	5	6	7	8
Контактная работа обучающегося с преподавателем по видам учебных занятий:	98	-	-	50	48	-	-	-	-
Лекционное занятие (Л)	70	-	-	34	36	-	-	-	-
Семинарское/практическое занятие (СПЗ)	28	-	-	16	12	-	-	-	-
Самостоятельная работа обучающегося, в том числе подготовка к промежуточной аттестации	146	-	-	90	56	-	-	-	-
Вид промежуточной аттестации: Зачет (З), Зачет с оценкой (ЗО), Экзамен (Э), Кандидатский экзамен (КЭ)	44	-	-	4 (З)	4 (З) 36 (КЭ)	-	-	-	-
Общий объем	в часах	288	-	-	144	144	-	-	-
	в зачетных единицах	8	-	-	4	4	-	-	-

3. Содержание дисциплины

1. Современные взгляды на биоинформатику. Типы биологических данных и их представления. Ресурсы EMBL-EBI и NCBI. Базы данных по нуклеотидным последовательностям. RefSeq. Геномные браузеры.
2. Информационные ресурсы по аминокислотным последовательностям: UniProt и сопряженные наборы данных, NeXtProt, PDB. Автоматизированный доступ к биоинформатическим данным.
3. Информационные ресурсы по функциональной характеристике генов и белков: Gene Ontology, KEGG, OMIM, Protein Atlas.
4. Парное выравнивание: принципы и методы. Матрицы замен. Статистическая оценка локального выравнивания. Семейство программ BLAST для поиска сходных последовательностей в базах данных (алгоритм, интерфейс).
5. Множественное выравнивание (алгоритмы, программы).

6. Мотивы. Регулярные выражения. Домены Профили. Скрытые Марковские цепи. Алгоритмы построения и поиска. Поиск мотивов *de novo* (в невыровненных последовательностях)
7. Филогения и эволюционные деревья. Концепция молекулярных часов. Алгоритмы построения филогенетических деревьев. Топология деревьев. MEGA.
8. Функциональная аннотация белков по последовательностям. Предсказание аминокислотных позиций, определяющих функциональную специфичность. Протеохемометрика.
9. Методы секвенирования следующего поколения (NGS). Сборка геномов и анализ данных NGS.
10. Метагеномика.
11. Микрочипы (microarrays) и анализ профилей экспрессии генов. Технология RNASeq. Базы транскриптомных данных.
12. Анализ генома и заболевания человека.
13. Анализ геномных вариаций и SNP в клинических исследованиях.
14. Основные биоинформатические ресурсы, предоставляющие доступ к «омиксным данным».
15. Алгоритмы и инструменты для обработки протеомных и транскриптомных данных.
16. Методы анализа белок-белковых взаимодействий. Базы данных по белок-белковым взаимодействиям, регуляторным и метаболическим путям.
17. Методы анализа молекулярных сетей. Cytoscape – программа визуализации и анализа молекулярных сетей. Понятие о степени вершины сети, «хабах» и центральности. Поиск модулей (кластеров вершин) в сети. Предсказание генов и белков, связанных с заболеваниями.
18. Язык программирования R. Синтаксис и типы данных. Основные библиотеки. Методы анализа биомедицинских данных.
19. Экспериментальные и теоретические подходы к поиску и разработке лекарств. Выбор фармакотерапевтических целевых характеристик. Биохимический и фенотипический скрининг. Оценки фармакодинамических и фармакокинетических характеристик лекарственно-подобных соединений *in silico*. Виртуальный скрининг и конструирование фармакологических веществ с требуемыми свойствами.
20. Структура белка. Методы получения трехмерной структуры белка. Визуализация белковых структур и комплексов. Банк данных пространственных структур белков PDB. Структура PDB файла.
21. Инструменты для интерактивной визуализации белковых структур. Выявления сходных 3-мерных структур белков. Визуализация белковых структур и изучение свойств белковых молекул при помощи программы PyMol.
22. Методы предсказания белковых структур по последовательностям аминокислот. Моделирование трехмерной структуры белка методом гомологического моделирования AlphaFold2.
23. 3D методы в конструировании лекарств на основе структуры лигандов. Фармакофорная модель. 3D-QSAR.
24. Анализ белок-лигандного взаимодействия методами молекулярного моделирования. Модели комплексов, силы, участвующие в связывании лигандов, термодинамика и кинетика процесса

25. Поиск и конструирование новых лигандов. Молекулярный докинг, конструирование *de novo*, оптимизация лигандов.
26. Моделирование молекулярной динамики, что из нее можно получить
27. Коллекции и корпуса текстов. Принципы работы с корпусами CHEMDNER, CHEMPROT. Корпусы DRUGPROT, GENETAG. Инструменты токенизации. Классификация токенов по принадлежности частям речи. Лемматизация. Python NLTK.
28. Представление текста для распознавания наименований биологических и химических объектов. N-граммы. Векторное представление текста. One hot encoding. Word embeddings. Python Keras, Gensim. Lingpipe 4.1.2. KNIME text classifier. KNIME text processing.
29. Предварительная обработка данных о структуре и биологической активности химических соединений в ходе формирования обучающих выборок для (Q)SAR моделирования на примере данных ChEMBL.
30. Машинное обучение. Задача классификации. Одно, двух и многоклассовая классификация. Алгоритмы классификации: деревья решений, Random Forests, наивный Байесовский классификатор, на основе сходства объектов, сходство последовательностей. Метод опорных векторов. Критерии качества классификации.
31. Регрессия, непараметрическая и параметрическая, линейная и нелинейная. Логистическая и мультиномиальная логистическая регрессия. Восстановление зависимостей по эмпирическим данным, методы регуляризации. Самосогласованная регрессия и классификация. Критерии качества зависимостей.
32. Искусственные нейронные сети. Обучение без учителя. Кластеризация. Топологические карты. Валидация результатов машинного обучения.
33. Прогнозирование метаболизма ксенобиотиков *in silico*.
34. Оценка межлекарственного взаимодействия *in silico*.
35. Анализ регуляторных сигнальных сетей с целью выявления наиболее перспективных фармакологических мишеней и оценки фармакотерапевтических и нежелательных побочных эффектов.
36. Репозиционирование лекарств.
37. Методы машинного обучения в анализе текстов.

4. Учебно-тематический план дисциплины

Таблица 2

Номер раздела, темы	Наименование разделов, тем	Количество часов					Форма контроля
		Всего	Ауд	Л	СПЗ	СР	
	3 семестр	144	50	34	16	90	4 часа (3)
Тема 1.1	Современные взгляды на биоинформатику. Типы биологических данных и их представления. Ресурсы EMBL-EBI и NCBI. Базы данных по нуклеотидным последовательностям. RefSeq. Геномные браузеры.	6	2	2	-	4	Устный опрос
Тема 1.2	Информационные ресурсы по аминокислотным последовательностям: UniProt и сопряженные наборы данных, NeXtProt, PDB. Автоматизированный доступ к биоинформатическим данным.	6	2	2	-	4	
Тема 1.3	Информационные ресурсы по функциональной характеристике генов и	6	2	2	-	4	

	белков: Gene Ontology, KEGG, OMIM, Protein Atlas.						
Тема 1.4	Парное выравнивание: принципы и методы. Матрицы замен. Статистическая оценка локального выравнивания. Семейство программ BLAST для поиска сходных последовательностей в базах данных (алгоритм, интерфейс).	12	4	2	2	8	
Тема 1.5	Множественное выравнивание (алгоритмы, программы).	6	2	2	-	4	
Тема 1.6	Мотивы. Регулярные выражения. Домены Профили. Скрытые Марковские цепи. Алгоритмы построения и поиска. Поиск мотивов de novo (в невыровненных последовательностях)	12	4	2	2	8	
Тема 1.7	Филогения и эволюционные деревья. Концепция молекулярных часов. Алгоритмы построения филогенетических деревьев. Топология деревьев. MEGA.	6	2	2	-	4	
Тема 1.8	Функциональная аннотация белков по последовательностям. Предсказание аминокислотных позиций, определяющих функциональную специфичность. Протеохеометрика.	6	2	2	-	4	
Тема 1.9	Методы секвенирования следующего поколения (NGS). Сборка геномов и анализ данных NGS.	6	2	2	-	4	
Тема 1.10	Метагеномика.	6	2	2	-	4	
Тема 1.11	Микрочипы (microarrays) и анализ профилей экспрессии генов. Технология RNASeq. Базы транскриптомных данных.	6	2	2	-	4	
Тема 1.12	Анализ генома и заболевания человека.	6	2	2	-	4	
Тема 1.13	Анализ геномных вариаций и SNP в клинических исследованиях.	6	2	2	-	4	
Тема 1.14	Основные биоинформатические ресурсы, предоставляющие доступ к «омиксным данным».	6	2	2	-	4	
Тема 1.15	Алгоритмы и инструменты для обработки протеомных и транскриптомных данных.	6	2	2	-	4	
Тема 1.16	Методы анализа белок-белковых взаимодействий. Базы данных по белок-белковым взаимодействиям, регуляторным и метаболическим путям.	6	2	2	-	4	
Тема 1.17	Методы анализа молекулярных сетей. Cytoscape – программа визуализации и анализа молекулярных сетей. Понятие о степени вершины сети, «хабах» и центральности. Поиск модулей (кластеров вершин) в сети. Предсказание генов и белков, связанных с заболеваниями.	6	2	2	-	4	
Тема 1.18	Язык программирования R. Синтаксис и типы данных. Основные библиотеки. Методы анализа биомедицинских данных.	26	12	-	12	14	

	4 семестр	144	48	36	12	56	4 часа (3) 36 час – (КЭ)
Тема 2.1	Экспериментальные и теоретические подходы к поиску и разработке лекарств. Выбор фармакотерапевтических целевых характеристик. Биохимический и фенотипический скрининг. Оценки фармакодинамических и фармакокинетических характеристик лекарственных-подобных соединений <i>in silico</i> . Виртуальный скрининг и конструирование фармакологических веществ с требуемыми свойствами.	6	2	2	-	4	Устный опрос
Тема 2.2	Структура белка. Методы получения трехмерной структуры белка. Визуализация белковых структур и комплексов. Банк данных пространственных структур белков PDB. Структура PDB файла.	4	2	2	-	2	
Тема 2.3	Инструменты для интерактивной визуализации белковых структур. Выявления сходных 3-мерных структур белков. Визуализация белковых структур и изучение свойств белковых молекул при помощи программы PyMol.	5	2	2	-	3	
Тема 2.4	Методы предсказания белковых структур по последовательностям аминокислот. Моделирование трехмерной структуры белка методом гомологического моделирования AlphaFold2.	4	2	2	-	2	
Тема 2.5	3D методы в конструировании лекарств на основе структуры лигандов. Фармакофорная модель. 3D-QSAR.	4	2	2	-	2	
Тема 2.6	Анализ белок-лигандного взаимодействия методами молекулярного моделирования. Модели комплексов, силы, участвующие в связывании лигандов, термодинамика и кинетика процесса	5	2	2	-	3	
Тема 2.7	Поиск и конструирование новых лигандов. Молекулярный докинг, конструирование <i>de novo</i> , оптимизация лигандов.	4	2	2	-	2	
Тема 2.8	Моделирование молекулярной динамики, что из нее можно получить	4	2	2	-	2	
Тема 2.9	Коллекции и корпуса текстов. Принципы работы с корпусами CHEMDNER, CHEMPROT. Корпусы DRUGPROT, GENETAG. Инструменты токенизации. Классификация токенов по принадлежности частям речи. Лемматизация. Python NLTK.	5	2	2	-	3	
Тема 2.10	Представление текста для распознавания наименований биологических и химических объектов. N-граммы. Векторное представление текста. One hot encoding. Word embeddings. Python Keras, Gensim. Lingpipe 4.1.2. KNIME text classifier. KNIME text processing.	5	2	2	-	3	

Тема 2.11	Предварительная обработка данных о структуре и биологической активности химических соединений в ходе формирования обучающих выборок для (Q)SAR моделирования на примере данных ChEMBL.	4	2	2	-	2	
Тема 2.12	Машинное обучение. Задача классификации. Одно, двух и многоклассовая классификация. Алгоритмы классификации: деревья решений, Random Forests, наивный Байесовский классификатор, на основе сходства объектов, сходство последовательностей. Метод опорных векторов. Критерии качества классификации.	5	2	2	-	3	
Тема 2.13	Регрессия, непараметрическая и параметрическая, линейная и нелинейная. Логистическая и мультиномиальная логистическая регрессия. Восстановление зависимостей по эмпирическим данным, методы регуляризации. Самосоогласованная регрессия и классификация. Критерии качества зависимостей.	4	2	2	-	2	
Тема 2.14	Искусственные нейронные сети. Обучение без учителя. Кластеризация. Топологические карты. Валидация результатов машинного обучения.	4	2	2	-	2	
Тема 2.15	Прогнозирование метаболизма ксенобиотиков <i>in silico</i> .	4	2	2	-	2	
Тема 2.16	Оценка межлекарственного взаимодействия <i>in silico</i> .	4	2	2	-	2	
Тема 2.17	Анализ регуляторных сигнальных сетей с целью выявления наиболее перспективных фармакологических мишеней и оценки фармакотерапевтических и нежелательных побочных эффектов.	5	2	2	-	3	
Тема 2.18	Репозиционирование лекарств.	4	2	2	-	2	
Тема 2.19	Методы машинного обучения в анализе текстов.	24	12	-	12	12	
Общий объем		288	98	70	28	146	8 час – (3) 36 час – (КЭ)

5. Учебно-методическое обеспечение самостоятельной работы обучающихся

Цель самостоятельной работы обучающихся заключается в глубоком, полном усвоении учебного материала и в развитии навыков самообразования. Самостоятельная работа может включать: работу с текстами, литературой, учебно-методическими пособиями, нормативными материалами, в том числе материалами сети интернет, а также проработку конспектов лекций, написание докладов, рефератов, участие в работе семинаров, научных конференциях и пр.

Контроль самостоятельной работы осуществляется на семинарских (практических) занятиях.

6. Оценочные средства для проведения текущего контроля успеваемости и промежуточной аттестации обучающихся

Вопросы для подготовки к промежуточной аттестации зачету, кандидатскому экзамену

Перечень вопросов к зачету

1. Основная догма молекулярной биологии. Матричный принцип. Процессы репликации, транскрипции, трансляции. Генетический код.
2. Структурная организация генов и геномов. Экзон-интронная структура. Хромосомы.
3. Транскрипция и её регуляция. Транскрипционные факторы. Типы регуляторных районов транскрипции.
4. Структурно-функциональные характеристики белков и нуклеиновых кислот.
5. Структура и функция РНК. Экспериментальные методы определения структуры РНК.
6. Трансляция РНК. Регуляция трансляции.
7. Структура и функция белков. Выравнивание белковых структур. Распознавание функциональных сайтов и мотивов в белках.
8. Понятие экспрессии генов. Экспериментальные методы анализа экспрессии.
9. Посттрансляционные модификации в белках. Структурно-функциональное значение. Методы экспериментального определения и компьютерного предсказания.
10. Системная биология. Генные и метаболические сети, регуляторно-сигнальные пути. Экспериментальное изучение и компьютерное моделирование.
11. Понятие об исследованиях ассоциаций (GWAS) Современное состояние и перспективы развития.
12. Секвенирование по Сенгеру, Секвенирование нового поколения (NGS) и связанные с ним подходы.
13. Протеомные исследования и масс-спектрометрия.
14. Филогения и эволюционные деревья. Подходы к изучению филогенеза, видового разнообразия и эволюционных взаимоотношений на основе геномных и протеомных исследований.
15. Микробиом человека в норме и при патологиях.
16. Основные информационные ресурсы и базы данных по молекулярной биологии. Архивные, курируемые и производные базы данных.
17. Задача сравнения генетических и белковых последовательностей. Методы выравнивания: парное и множественное, локальное и глобальное. Алгоритм глобального выравнивания Нидлмана-Вунша (Needleman-Wunsh). Алгоритм локального выравнивания Смита-Уотермана (Smith-Waterman). PSI-BLAST. Назначение и основные возможности. Алгоритм.
18. Вероятностные модели последовательностей. Марковские модели; Скрытые Марковские модели. Алгоритмы Витерби и апостериорного декодирования. Обучение марковских моделей.
19. Распознавание структурно-функциональных мотивов в генетических текстах. Понятие консенсуса, весовой матрицы. Оценка точности распознавания.

20. Сборка геномов. Графы де Брюйна и алгоритмы сборки геномов. Форматы данных секвенирования. Анализ качества данных секвенирования.
21. Геномные браузеры UCSC, NCBI Map Viewer, Ensembl.
22. Понятие гомологии. Ортологи и паралоги.
23. Биоинформатика для протеомных исследований. Вычисление массы и изоэлектрической точки белка. Протеолитические пептиды. Базы данных SWISS-2DPAGE, PeptideAtlas, Human Proteome Atlas, NeXtProt.
24. Методы анализа регуляторных и сигнальных путей с целью приоритизации фармакологических мишеней.

Перечень вопросов к кандидатскому экзамену

1. Основная догма молекулярной биологии. Матричный принцип. Процессы репликации, транскрипции, трансляции. Генетический код.
2. Структурная организация генов и геномов. Экзон-интронная структура. Хромосомы.
3. Транскрипция и её регуляция. Транскрипционные факторы. Типы регуляторных районов транскрипции. Структура и функция промотора.
4. Структурно-функциональные характеристики белков и нуклеиновых кислот.
5. Структура и функция РНК. Экспериментальные методы определения структуры РНК.
6. Трансляция РНК. Регуляция трансляции.
7. Структура и функция белков. Выравнивание белковых структур. Распознавание функциональных сайтов и мотивов в белках.
8. Функциональная геномика. Понятие экспрессии генов. Экспериментальные методы анализа экспрессии.
9. Посттрансляционные модификации в белках. Структурно-функциональное значение. Методы экспериментального определения и компьютерного предсказания.
10. Системная биология. Генные и метаболические сети, регуляторно-сигнальные пути. Экспериментальное изучение и компьютерное моделирование.
11. Биомаркеры и фармакологические мишени. Методы определения и валидации.
12. Путь от гена к лекарству *in silico*. Компьютерное конструирование лекарств.
13. Хемоинформатика. Анализ взаимосвязей «структура-свойство».
14. Персонализированная медицина. Понятие об исследованиях ассоциаций (GWAS) Современное состояние и перспективы развития.
15. Основные экспериментальные методы молекулярной биологии. Клонирование, ПЦР, секвенирование по Сенгеру, Массовое секвенирование (NGS) и связанные с ним подходы. Масс-спектрометрия.
16. Филогения и эволюционные деревья. Подходы к изучению филогенеза, видового разнообразия и эволюционных взаимоотношений на основе геномных и протеомных исследований.
17. Микробиом человека в норме и при патологиях.
18. Теория вероятностей и статистика. Случайные величины. Статистические критерии, Понятие о множественном тестировании. Байесов подход, понятие об априорной и апостериорной вероятности.

19. Методы машинного обучения. Решающие деревья, случайный лес, метод поддерживающих векторов (SVM), регрессионный анализ, кластерный анализ, факторный анализ. Оценка качества обучения. Проблема переобученных моделей.
20. Основные понятия о базах данных. Структура и содержание базы данных. Таблицы, поля, запросы.
21. Основные информационные ресурсы и базы данных по молекулярной биологии. Архивные, курируемые и производные базы данных.
22. Понятия Data mining и Text mining.
23. Задача сравнения генетических и белковых последовательностей. Методы выравнивания: парное и множественное, локальное и глобальное. Алгоритм глобального выравнивания Нидлмана-Вунша (Needleman-Wunsh). Алгоритм локального выравнивания Смита-Уотермана (Smith-Waterman). PSI-BLAST. Назначение и основные возможности. Алгоритм.
24. Вероятностные модели последовательностей. Марковские модели; Скрытые Марковские модели. Алгоритмы Витерби и апостериорного декодирования. Обучение марковских моделей.
25. Распознавание структурно-функциональных мотивов в генетических текстах. Понятие консенсуса, весовой матрицы. Оценка точности распознавания.
26. Предсказание структур РНК. Алгоритм Нуссонофф, сравнительный подход и метод ковариаций.
27. Сборка геномов. Графы де Брюйна и алгоритмы сборки геномов. Форматы данных секвенирования. Анализ качества данных секвенирования.
28. Геномные браузеры UCSC, NCBI Map Viewer, Ensembl.
29. Понятие гомологии. Ортологи и паралоги.
30. Биоинформатика для протеомных исследований. Вычисление массы и изоэлектрической точки белка. Протеолитические пептиды. Базы данных SWISS-2DPAGE, PeptideAtlas, Human Proteome Atlas, NeXtProt.
31. Физико-математические модели биомакромолекул. Модели подвижности ДНК.
32. Поиск и приоритизация фармакологических мишеней. Понятие и оценки "Druggability".
33. Поиск и оптимизация базовых структур новых лекарств. Оценки "Druglikeness" и "Leadlikeness".
34. Методы конструирования лекарств на основе структуры макромолекулы-мишени.
35. Методы конструирования лекарств на основе структуры лигандов.
36. Оценка безопасности базовых структур новых лекарств *in silico*.
37. Плейотропное действие лекарственных соединений: положительные и отрицательные аспекты.
38. Моделирование пространственной структуры белка. Молекулярное моделирование, докинг, оценки энергии связывания.
39. Методы анализа регуляторных и сигнальных путей с целью приоритизации фармакологических мишеней.
40. Методы анализа количественных соотношений структура-активность и структура-свойство (QSAR/QSPR).
41. Построение фармакофоров и их применение для поиска и оптимизации структуры и свойств новых лекарств.

42. Виртуальный скрининг и критерии для отбора наиболее перспективных соединений с целью экспериментального тестирования.
43. Компьютерное конструирование лекарств на основе структуры фрагментов.

Описание критериев и шкал оценивания

В ходе текущего контроля успеваемости (устный или письменный опрос, подготовка и защита реферата, доклад, презентация, тестирование и пр.) при ответах на учебных занятиях, а также промежуточной аттестации в форме кандидатского экзамена обучающиеся оцениваются по четырёхбалльной шкале: «отлично», «хорошо», «удовлетворительно», «неудовлетворительно».

Оценка «отлично» – выставляется аспиранту, если он глубоко усвоил программный материал, исчерпывающе, последовательно, четко и логически стройно его излагает, умеет связывать теорию с практикой, свободно справляется с задачами и вопросами, не затрудняется с ответами при видоизменении заданий, умеет принять правильное решение и грамотно его обосновывать, владеет разносторонними навыками и приемами выполнения практических задач, комплексной оценкой предложенной ситуации.

Оценка «хорошо» – выставляется аспиранту, если он твердо знает программный материал, грамотно и по существу излагает его, не допуская существенных неточностей при ответе на вопрос, но недостаточно полно раскрывает междисциплинарные связи, правильно применяет теоретические положения при решении практических вопросов и задач, владеет необходимыми навыками и приемами их выполнения, комплексной оценкой предложенной ситуации.

Оценка «удовлетворительно» – выставляется аспиранту, если он имеет поверхностные знания программного материала, не усвоил его деталей, допускает неточности, оперирует недостаточно правильными формулировками, нарушает логическую последовательность в изложении программного материала, испытывает затруднения при выполнении практических задач, испытывает затруднения с комплексной оценкой предложенной ситуации, не полностью отвечает на вопросы, в том числе при помощи наводящих вопросов преподавателя.

Оценка «неудовлетворительно» – выставляется аспиранту, который не знает значительной части программного материала, допускает грубые ошибки, неуверенно, с большими затруднениями решает практические задачи или не справляется с ними самостоятельно, не владеет комплексной оценкой ситуации, неверно выбирает тактику действий.

В ходе текущего контроля успеваемости (устный или письменный опрос, подготовка и защита реферата, доклад, презентация, тестирование и пр.) при ответах на учебных занятиях, а также промежуточной аттестации в форме зачета обучающиеся оцениваются по двухбалльной шкале:

Оценка «зачтено» – выставляется аспиранту, если он продемонстрировал знания программного материала, подробно ответил на теоретические вопросы, справился с выполнением заданий и (или) ситуационных задач, предусмотренных рабочей программой дисциплины.

Оценка «не зачтено» – выставляется аспиранту, если он имеет пробелы в знаниях программного материала, не владеет теоретическим материалом и допускает грубые, принципиальные ошибки в выполнении заданий и (или) ситуационных задач, предусмотренных рабочей программой дисциплины.

Шкала оценивания (четырёхбалльная или двухбалльная), используемая в рамках текущего контроля успеваемости определяется преподавателем, исходя из целесообразности применения той или иной шкалы.

Если текущий контроль успеваемости и (или) промежуточная аттестация, предусматривает тестовые задания, то перевод результатов тестирования в четырёхбалльную шкалу осуществляется по схеме:

Оценка «Отлично» – 90-100% правильных ответов;

Оценка «Хорошо» – 80-89% правильных ответов;

Оценка «Удовлетворительно» – 71-79% правильных ответов;

Оценка «Неудовлетворительно» – 70% и менее правильных ответов.

Перевод результатов тестирования в двухбалльную шкалу:

Оценка «Зачтено» – 71-100% правильных ответов;

Оценка «Не зачтено» – 70% и менее правильных ответов.

7. Учебно-методическое и информационное обеспечение дисциплины

1. Введение в информационную биологию и биоинформатику. Под ред. Н.А. Колчанова, О.В. Вишневого, Д.П. Фурман. Том 2. Компьютерная протеомика. Новосибирск: Новосибирский гос. ун-т, 2012. 252 с.
2. Введение в информационную биологию и биоинформатику. Под ред. Н.А. Колчанова, О.В. Вишневого, Д.П. Фурман. Том 3. Теория генных сетей. Картирование генов, контролирующих сложные признаки человека. Новосибирск: Новосибирский гос. ун-т, 2015. 298 с.
3. Введение в информационную биологию и биоинформатику. Под ред. Н.А. Колчанова, О.В. Вишневого, Д.П. Фурман. Том 4. Математическое моделирование и методы биоинформатики в биологии развития. Компьютерная эволюционная биология Новосибирск: Новосибирский гос. ун-т, 2012. 336 с.
4. Введение в хемоинформатику: Ч. 1. Компьютерное представление химических структур. Казань: Из-во КФУ, 2013. 174 с.
5. Введение в хемоинформатику: Ч. 2. Химические базы данных. Казань: Из-во КФУ, 2015. 188 с.
6. Введение в хемоинформатику: Ч. 3. Моделирование «структура-свойство». Казань: Из-во КФУ, 2015. 304 с.
7. Раевский О.А. Моделирование соотношений «структура-свойство». М.: Из-во «КДУ», 2015. 288 с.
8. Компьютеры и суперкомпьютеры в биологии Под. ред. В.Д. Лахно и М.Н. Устинина. Москва-Ижевск: Институт компьютерных исследований, 2002, 528 с.
9. Финкельштейн А.В., Птицын О.Б. Физика белка. М.: Книжный дом «Университет», 2002.
10. Каменская М.А. Информационная биология. М.: Издательский центр «Академия», 2006.
11. Хёльте Х.-Д., Зиппл В., Роньян Д., Фолькерс Г. Молекулярное моделирование. Теория и практика. М.: Издательство «Бином», 2010.
12. Жимулёв И.Ф. Общая и молекулярная генетика. Учебное пособие. Новосибирск. НГУ. 2003.

13. Ризниченко Г. Ю. Лекции по математическим моделям в биологии. Издание 2-е, исправленное и дополненное. Изд-во РХД, М–Ижевск, 2011 г. 560 с.
14. Мятлев В. Д., Панченко Л. А., Ризниченко Г. Ю., Терёхин А. Т. Теория вероятностей и математическая статистика. Математические модели. Серия: Высшая математика и ее приложения к биологии. М.: Академия, 2009.
15. Фурсова П. В., Тёрлова Л. Д., Ризниченко Г. Ю. Математические модели в биологии. Москва–Ижевск: Изд-во РХД, 2008.
16. Рубин А.Б. Биофизика. М.: МГУ им. М.В. Ломоносова, 2000.
17. Шайтан К.В., Сарайкин С.С. Молекулярная динамика. М.: МГУ им. М.В. Ломоносова, 1999.
18. Урбах В.Ю. Статистический анализ в биологических и медицинских исследованиях. М.: Медицина, 1975.
19. Айвазян С.А., Бухштабер В.М., Енюков И.С., Мешалкин Л.Д. Прикладная статистика. М.: Финансы и статистика, 1989.
20. Химмельблау Д. Прикладное нелинейное программирование, М.: Мир, 1975.
21. Фомин С.В., Беркинблит М.Б. Математические проблемы в биологии. 1973.
22. Дурбин Р., Эдди Ш., Крэг А., Митчисон Г. Анализ биологических последовательностей. Издательство: Регулярная и хаотическая динамика, Институт компьютерных исследований, 2006
23. Lenghauer T. Bioinformatics. From Genomes to Drugs. Weinheim: Wiley-VCH, 2002.

Дополнительная литература

1. Раевский О.А. Свойства химических соединений и лекарств как функции их структуры. М.: Из-во «КДУ», 2013. 376 с.
2. Дружиловский Д.С., Рудик А.В., Филимонов Д.А., Лагунин А.А., Глориозова Т.А., Поройков В.В. Веб-ресурсы для прогнозирования биологической активности органических соединений. Известия Академии наук. Серия химическая, 2016, № 2, 384-393.
3. Филимонов Д.А., Лагунин А.А., Глориозова Т.А., Рудик А.В., Дружиловский Д.С., Погодин П.В., Поройков В.В. Предсказание спектров биологической активности органических соединений с помощью веб-ресурса PASS Online. Химия гетероциклических соединений, 2014, № 3, 483-499.
4. The Practice of Medicinal Chemistry, 4th Edition. C. Wermuth, Ed. Amsterdam a.o: Academic Press, 2015.
5. Livingstone D. Data Analysis for Chemists. Oxford et al.: Oxford University Press, 1995.
6. Han van de Waterbeemd. Structure-Property Correlations in Drug Design. Austin: Academic Press, 2003.
7. Cheminformatics Approaches to Virtual Screening. Alexandre Varnek and Alex Tropsha, Eds. Cambridge, UK: RSC Publishing, 2008.
8. Wild D.J., Ding Y., Sheth A.P., et al. Systems chemical biology and the Semantic Web: what they mean for the future of drug discovery research. *Drug Discovery Today*. 2012, 17 (9-10): 469-474.
9. Fernald G.H., Capriotti E., Daneshjou R., Karczewski K.J., Altman R.B. Bioinformatics challenges for personalized medicine. *Bioinformatics*, 2011, 27 (13): 1741–1748.

10. Bissantz C., Kuhn B., Stahl M. A medicinal chemist's guide to molecular interactions. *Journal of Medicinal Chemistry*, 2010, 53 (14): 5061-5084. Erratum in: *Journal of Medicinal Chemistry*, 2010, 53 (16): 6241.
11. Villaamil V.M., Gallego G.A., Cainzos I.S., et al. State of the Art *in Silico* Tools for the Study of Signaling Pathways in Cancer. *International Journal of Molecular Sciences*, 2012, 13 (6): 6561–6581.
12. Grosdidier S., Totrov M., Fernández-Recio J. Computer applications for prediction of protein-protein interactions and rational drug design. *Advances and Applications in Bioinformatics and Chemistry*, 2009, 2: 101-123.
13. Huggins D.J., Sherman W., Tidor B. Rational Approaches to Improving Selectivity in Drug Design. *Journal of Medicinal Chemistry*, 2012, 55 (4): 1424–1444.
14. Briggs K., Cases M., Heard D.J., et al. Inroads to Predict *in Vivo* Toxicology—An Introduction to the eTOX Project. *International Journal of Molecular Sciences*, 2012, 13 (3): 3820–3846.
15. Lagunin A.A., Goel R.K., Gawande D.Y., et al. Chemo- and bioinformatics resources for *in silico* drug discovery from medicinal plants beyond their traditional use: a critical review. *Natural Product Reports*, 2014, 31 (11), 1585-1611.
16. Zakharov A., Lagunin A. Computational toxicology in drug discovery: opportunities and limitations. In: *Application of Computational Techniques in Pharmacy and Medicine*. Leonid Gorb, Victor Kuz'min and Eugene Muratov Eds.. Springer, 2014, p. 325-367.
17. Katsila T., Spyroulias G.A., Patrinos G.P., Matsoukas M.-T. Computational approaches in target identification and drug discovery. *Computational and Structural Biotechnology Journal*, 2016, 14, 177-184.
18. Ivanov S.M., Lagunin A.A., Poroikov V.V. (2016). *In silico* assessment of adverse drug reactions and associated mechanisms. *Drug Discovery Today*, 2016, 21 (1), 58-71.
19. Arnold J.W., Roach J., Azcarate-Peril M.A. Emerging Technologies for Gut Microbiome. *Trends in Microbiology*, 2016, 24 (11), 887-901.
20. Tycko J., Myer V.E., Hsu P.D. Methods for Optimizing CRISPR-Cas9 Genome Editing Specificity. *Molecular Cell*, 2016, 63 (3), 355-370.

Перечень ресурсов информационно-телекоммуникационной сети «Интернет»

1. Официальный сайт ИБМХ: адрес ресурса – <https://www.ibmc.msk.ru/>, на котором содержатся сведения об образовательной организации и ее подразделениях, локальные нормативные акты, сведения о реализуемых образовательных программах, их учебно-методическом и материально-техническом обеспечении, а также справочная, оперативная и иная информация. Через официальный сайт обеспечивается доступ всех участников образовательного процесса к различным сервисам и ссылкам;
2. Библиотека естественных наук Российской академии наук (БЕН РАН);
3. Электронно-библиотечная система (ЭБС) «Университетская библиотека онлайн»;
4. Электронный библиотечный абонемент Центральной научной медицинской библиотеки (ЦНМБ).
5. Электронные ресурсы издательств и баз данных.

Перечень профессиональных баз данных и информационных справочных систем

1. Электронные ресурсы издательства «Эльзевир» (Elsevier B.V.) - Freedom Collection. Freedom Collection – полнотекстовая коллекция электронных журналов издательства Elsevier B.V. по различным отраслям знаний, включающая не менее 2 500 наименований электронных журналов. Полный список электронных журналов приведен на сайте издательства: <http://info.sciencedirect.com/techsupport/journals/freedomcoll.htm>
2. Электронные ресурсы издательства Springer. Полнотекстовая коллекция электронных журналов издательства Springer по различным отраслям знаний на платформе: <http://link.springer.com>. Имеется полный доступ к статическим и динамическим справочным изданиям по любой теме.
3. Электронные ресурсы международного индекса научного цитирования Web of Science. База данных Web of Science - Политематическая реферативно-библиографическая и наукометрическая (библиометрическая) база данных компании Thomson Reuters (Scientific). Доступ по ссылке: <http://webofscience.com>
4. Электронные ресурсы международного индекса научного цитирования Scopus. База данных Scopus - Информационная поддержка научно-исследовательского процесса - доступна на платформе <https://www.scopus.com>
5. На русскоязычном сайте издательства Elsevier, в разделе, посвященном Scopus, есть детальная информация о ресурсе на русском языке: <http://elsevierscience.ru/products/scopus/>
6. В этом же разделе есть ссылка на краткое руководство пользователя на русском языке: http://elsevierscience.ru/files/pdf/Scopus_Quick_Reference_Guide_Russian_v2.pdf
7. Электронные ресурсы издательства Springer Nature: международное издательство, выпускает журналы, энциклопедии, книги и электронные продукты по целому спектру отраслей знания, включая биомедицину и науки о жизни, химию и компьютерные науки. <http://www.nature.com/siteindex/index.html>.
8. Доступ и использование базы данных Wiley Journals через Онлайн библиотеку Wiley Online Library: <http://onlinelibrary.wiley.com/>. Wiley Online Library представляет самую обширную междисциплинарную коллекцию интернет-ресурсов. Платформа обеспечивает удобный интегрированный доступ к результатам более 200 лет исследований, среди которых некоторые из наиболее авторитетных источников по биологическим наукам, медицине, физике, естественным дисциплинам.
9. Электронные ресурсы библиотеки eLIBRARY.RU: <http://elibrary.ru/contact.asp>. eLIBRARY.RU – это крупнейший российский информационный портал в области науки, технологии, медицины и образования, содержащий рефераты и полные тексты более 12 млн. научных статей и публикаций. На платформе eLIBRARY.RU доступны электронные версии более 1900 российских научно-технических журналов, в том числе более 900 журналов в открытом доступе.
10. <https://www.ncbi.nlm.nih.gov/> NCBI – Национальный центр биотехнологической информации (базы данных и компьютерные ресурсы по биоинформатике);
11. <https://genome.ucsc.edu/> – геномный браузер Университета Калифорнии Санта-Круз;
12. <https://www.ensembl.org/> – геномный браузер Ensembl;
13. <http://www.genome.jp/kegg/> KEGG – коллекция баз данных по сигнальным и регуляторным путям Университета Киото;
14. <http://www.omim.org/> OMIM – база данных менделевских наследованных признаков у людей;
15. <http://pfam.xfam.org/> PFAM – база данных семейств белков с аннотациями и

- множественным выравниванием последовательностей сгенерированными с использованием скрытых марковских моделей;
16. <http://www.uniprot.org/> UniProt – база данных с информацией о белках;
 17. <http://www.ebi.ac.uk/> – базы данных European Bioinformatics Institute (EBI);
 18. <https://www.r-project.org/> - основной сайт R;
 19. <https://cran.r-project.org/> - архив пакетов для статистического анализа данных в R;
 20. <http://www.bioconductor.org/> Bioconductor – архив пакетов R, предназначенных для анализа молекулярно-биологических данных.

8. Материально-техническое обеспечение дисциплины

Таблица 3

№ п/п	Наименование оборудованных учебных аудиторий	Перечень специализированной мебели, технических средств обучения
1	Учебные аудитории для проведения занятий лекционного и семинарского типов, групповых и индивидуальных консультаций, текущего контроля успеваемости и промежуточной аттестации	Персональные компьютеры с доступом в Интернет.
2	Помещения для самостоятельной работы (Компьютерный класс)	Компьютерная техника с возможностью подключения к сети "Интернет" и обеспечением доступа к электронной информационно-образовательной среде ИБМХ.

Программное обеспечение

- MICROSOFT WINDOWS;
- MICROSOFT OFFICE;
- Adobe Reader;

9. Методические указания для обучающихся по изучению дисциплины

Преподавание дисциплины осуществляется в соответствии с Федеральными государственными требованиями.

Основными формами получения и закрепления знаний по данной дисциплине являются занятия лекционного и семинарского типа, самостоятельная работа обучающегося, в том числе под руководством преподавателя, прохождение контроля.

Изучение дисциплины согласно учебному плану предполагает самостоятельную работу обучающихся. Самостоятельная работа включает в себя изучение литературы, её конспектирование, подготовку к семинарским (практическим) занятиям, текущему контролю успеваемости и промежуточной аттестации.

Текущий контроль успеваемости по дисциплине и промежуточная аттестация осуществляются в соответствии с Порядком организации и проведения текущего контроля успеваемости и Порядком проведения промежуточной аттестации обучающихся, устанавливающим формы проведения промежуточной аттестации, ее периодичность и систему оценок.

Наличие в Институте электронной информационно-образовательной среды, а также электронных образовательных ресурсов позволяет изучать дисциплину инвалидам и лицам с ОВЗ.

Особенности изучения дисциплины инвалидами и лицами с ОВЗ определены в Положении об организации получения образования для инвалидов и лиц с ограниченными возможностями здоровья.

10. Методические рекомендации преподавателю по организации учебного процесса по дисциплине

Преподавание дисциплины осуществляется в соответствии с Федеральными государственными требованиями.

При изучении дисциплины рекомендуется использовать следующий набор средств и способов обучения:

- рекомендуемую литературу;
- задания для подготовки к семинарам (практическим занятиям) – вопросы для обсуждения и др.;
- задания для текущего контроля успеваемости (задания для самостоятельной работы обучающихся);
- вопросы и задания для подготовки к промежуточной аттестации по итогам изучения дисциплины .

При проведении занятий лекционного и семинарского типа, в том числе в форме вебинаров и on-line курсов необходимо строго придерживаться учебно-тематического плана дисциплины, приведенного в разделе 4 данного документа. Необходимо уделить внимание рассмотрению вопросов и заданий, включенных в оценочные задания, при необходимости, решить аналогичные задачи с объяснением алгоритма решения.

Следует обратить внимание обучающихся на то, что для успешной подготовки к текущему контролю успеваемости и промежуточной аттестации нужно изучить литературу, список которой приведен в разделе 7 данной рабочей программы дисциплины и иные источники, рекомендованные в подразделах «Перечень ресурсов информационно-телекоммуникационной сети «Интернет» и «Перечень профессиональных баз данных и информационных справочных систем», необходимых для изучения дисциплины .

Текущий контроль успеваемости и промежуточная аттестация осуществляются в соответствии с Порядком организации и проведения текущего контроля успеваемости и Порядком проведения промежуточной аттестации обучающихся, устанавливающим формы проведения промежуточной аттестации, ее периодичность и систему оценок, с которыми необходимо ознакомить обучающихся на первом занятии.