Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Максимов Алексей Борисович
Должность: директор департамента по образовательной политике
Дата подписания: 26.09.2023 17:17:30
Уникальный программный ключ:
8db180d1a3f02ac9e60521a5672742735885ab6

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное автономное образовательное учреждение высшего образования

«МОСКОВСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

УТВЕРЖДАЮ
Декан факультета химической технологии и биотехнологии
_____ Ю.В. Данильчук
«__» 07 2022 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ «Биоинформатика»

Направление подготовки
19.03.01 «Биотехнология»

Профиль
«Биотехнология»

Квалификация (степень) выпускника
Бакалавр

Форма обучения
Очная

Москва 2022 г.

Программа составлена в соответствии с Федеральным государственным образовательным стандартом высшего образования по направлению подготовки бакалавров **19.03.01 Биотехнология**, утвержденного приказом Министерства образования и науки Российской Федерации от 11 марта 2015 г. № 193 и основной образовательной программы высшего профессионального образования ООП ВО, разработанной в Московском политехническом университете

Программу составил:

к.б.н. /Щербинин Д.С./

Программа утверждена на заседании кафедры "ХимБиотех" «01» сентября 2021 г., протокол № 1

Заведующая кафедрой «ХимБиотех»,

профессор, д.б.н. /Т.И. Громовых/

Программа согласована:

Руководитель образовательной программы

Доцент, к.б.н. / Е.С. Горшина/

« 1 » сентября 2022 г.

1. Цели освоения дисциплины

К **основным целям** освоения дисциплины «Биоинформатика» следует отнести:

– формирование знаний о современных методах получения и работы с биологическими данными, принципами организации баз данных значимой молекулярно-биологической, биофизической и структурной информации;

– подготовка студентов к деятельности в соответствии с квалификационной характеристикой бакалавра по направлению, в том числе формирование умений по сбору и анализу необходимой информации и аннотации биологических объектов: генов, белков и др..

- формирование у студентов системного подхода к решению научных задач с применением информационных систем, способности ориентироваться в широком диапазоне специализированных научных информационных систем и их классификации с целью выбора наиболее предпочтительной исходя из рода профессиональных задач.

К **основным задачам** освоения дисциплины «Биоинформатика» следует отнести:

– освоение методологии сбора и анализа информации о биологигческих объектах, отработка практических навыков по работе с биологической информацией о последовательностях генов и белков, структурной информацией о биологических молекулах, их свойствах и функциях.

- освоение навыков анализа результатов многочисленных экспериментальных работ в области биотехнологии, молекулярной биологии, биохимии, генетики при помощи соответствующих вычислительных технологий.

2. Место дисциплины в структуре ООП бакалавриата.

Дисциплина «Биоинформатика» относится к числу профессиональных учебных дисциплин вариативной части базового цикла (Б1) основной образовательной программы бакалавриата.

«Биоинформатика» взаимосвязана логически и содержательно-методически со следующими дисциплинами и практиками ООП:

В вариативной части базового цикла (Б1.2):

- Основы генной инженерии
- Молекулярная и клеточная биотехнология
- Основы молекулярной биологии
- Математическое планирование экспериментов и прикладная статистика
- Химия биологически активных веществ
- Прикладная энзимология

3. **Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенные с планируемыми результатами освоения образовательной программы.**

В результате освоения дисциплины (модуля) у обучающихся формируются следующие компетенции и должны быть достигнуты следующие результаты обучения как этап формирования соответствующих компетенций:

Код компетенции	**В результате освоения образовательной программы обучающийся должен обладать**	**Перечень планируемых результатов обучения по дисциплине**
ОПК-1	способность осуществлять поиск, хранение, обработку и анализ информации из различных источников и баз данных, представлять ее в требуемом формате с использованием информационных, компьютерных и сетевых технологий	**знать:** • Основные открытые банки данных биологической иныормации **уметь:** • Проводить поиск по открытым архивным и курируемым базам данных **владеть:** • методами поиска информации о биологических молекулах и их функциях
ОПК-5	владение основными методами, способами и средствами получения, хранения, переработки информации, навыками работы с компьютером как средством управления информацией	**знать:** • Основы организации и формы представления биологических данных **уметь:** • Проводить аннотиротацию биологических объектов с применением данных из открытых архивных и курируемых баз данных **владеть:** • подходами для сбора необходимой биологической информации и аннотации биологических молекул
ПК-11	готовностью использовать современные информационные технологии в своей	**знать:** • Основные методы и подходы системной и структурной биоинформатики **уметь:**

	профессиональной области, в том числе базы данных и пакеты прикладных программ	• методами поиска необходимой информации по названию, ключевым словам и последовательностям биологических молекул **владеть:** • методами поиска необходимой информации по названию, ключевым словам и последовательностям биологических молекул и систематизации данных для аннотации биологических объектов

4. Структура и содержание дисциплины.

Общая трудоемкость дисциплины составляет **2** зачетные единицы, т.е. **72** академических часов (из них **36** часов – самостоятельная работа студентов).

семестр: лабораторные работы – 2 час в неделю (36 часов). форма контроля - зачет.

Структура и содержание дисциплины «Биоинформатика» по срокам и видам работы отражены в приложении 1.

Содержание разделов дисциплины

Введение

Биологические объекты в Биоинформатике. Предмет и задачи биоинформатики. Основные направления дисциплины «Биоинформатика».

История возникновения и развитие биоинформатики как отдельной науки.

Основные направления современной биоинформатики (структурная и функциональная геномика, и протеомика). Методы получения биологических данных. Работа с биологическими данными.

Методология использования подходов биоинформатики для решения фундаментальных и прикладных задач.

Классификация баз данных (автоматические, архивные, курируемые). Основные базы данных: GenBank, RefSeq, Nucleotide, SwissProt, TrEMBL, PIR, PDB, банки белковых семейств (Pfam, InterPro, SCOP), базы данных метаболических путей и белок-белковых взаимодействий (KEGG, Reactome) Направленный поиск по базам данных.

Базы данных нуклеотидных последовательностей NCBI. GeneBank, Nucleotide, Snp.

Методы получения нуклеотидных последовательностей. Секвенирование I,II и III поколения.

NCBI. Виртуальные базы NCBI: GeneBank, Nucleotide, Pubmed, OMIM, SNP, GEO.

Идентификаторы записей в банке GeneBank. Структура записи GeneBank.

Геномная информация. Координаты гена, поиск соседних генов.

Поиск кодирующих областей гена. Сохранение последовательностей FASTA.

Поиск по базам NCBI. ENTREZ.

Парное выравнивание последовательностей. Поиск последовательностей BLAST (Basic Local Alignment Search Tool)

Работа с последовательностями. Парное выравнивание: локальное, глобальное.

Поиск маршрута выравнивания. Динамическое программирование. Оценка выравнивания: матрицы замен, штрафы за делеции.

Определения и терминология выраваниваний. Алгоритмы Смитта-Ватермана и Нидлмана-Вунша.

Пакет программ для построения выравниваний EMBOSS.

Применение локального выравнивая для поиска последовательностей. BLAST.

Основы BLAST: оценочная функция (BitScore), хэш-таблицы.

Оптимизация поиска последовательностей.

Базы данных белков SwissProt, trEMBL.

Методы получения аминокислотных последовательностей. Uniprot: SwissProt vs TrEMBL. UNIParc- история редактирования записи Uniprot.

Структура записи в Uniprot:

Описание функции белков, постранслящионные модификации, GO, метаболические пути, мутагенез, посдедовательность, белок-белковые взаимодействия, семейства белков, доменная организация, структура.

Оптимизация поиска и выдачи результатов банка Uniprot

Множественное выравнивание последовательностей. Clustal

Цели множественного выравнивания. Множественное выравнивание белков и ДНК. Как выбор последовательностей для множественного выравнивания. Множественное выравнивание: четыре группы методов:

- Прогрессивное глобальное выравнивание
- Итерационные процедуры
- Выравнивание по локальным консервативным участкам
- Статистические методы и вероятностные модели

Оценка множественного выравнивания.

Концепция прогрессивного выравнивания. Clustal. Проблемы подхода.

Локальное множественное выравнивание: DIALIGN

Анализ профилей и Марковские модели (HMM)

Изучение белковых семейств. Доменная организация. Профили выравнивания. PFAM

Родство белков. Сходство последовательностей – критерий родства белков. Гомология аминокислотных остатков.

Эволюционный домен белка – консервативный фрагмент последовательности. Поиск и описание доменов.

Работа с базами данных белковых доменов и семейств: PFAM, Интегрированная база данных InterPro

Функциональная аннотация белков. GeneOntotlogy, Kegg

Применение баз данных для аннотации новых белков по последовательности.

Белок-белковые взаимодействия.

Проект GeneOntology- иерархическая аннотация функций, процессов и локализации белков.

Метаболические и сигнальные пути. База Kegg. Поиск путей, в которых участвует целевой блок. Поиск путей, ассоциированных с заболеваниями.

Строение и функции белков.

Структурная биоинформатика. Основные методы и их применение.

Методы расшифровки и предсказания пространственных структур.

Представление пространственных структур в биоинформатике. Банки пространственных структур.

Банк PDB. Поиск по банку данных PDB и сохранение результатов.

Формат записи PDB.

Методы моделирования и оценки взаимодействий белков с лигандами и ДНК.

Молекулярный докинг. Две задачи молекулярного докинга- генерация конформаций и оценка. Подготовка структур и примеры запуска.

Молекулярная динамика. Молекулярная механика vs квантовая механика. Силовое поле. Подготовка и запуск молекулярной динамики.

Анализ профилей экспрессии генов.

Экспрессия генов. Методы оценки экспрессии генов: гибридизационные микрочипы и секвенирование транскриптома. Статистическая обработка экспрессионных данных.

Базы экспрессионных данных.

5. Образовательные технологии.

Методика преподавания дисциплины «Биоинформатика» и реализация компетентностного подхода в изложении и восприятии материала предусматривает использование следующих активных и интерактивных форм проведения групповых, индивидуальных, аудиторных занятий в сочетании с внеаудиторной работой с целью формирования и развития профессиональных навыков обучающихся:

– подготовка к выполнению лабораторных работ в компьютерных классах вуза;

– защита и индивидуальное обсуждение выполняемых этапов индивидуального лабораторного проекта по выбранным белковым мишеням;

– обсуждение и защита лабораторных проектов по дисциплине;

– подготовка, представление и обсуждение презентаций;

– организация и проведение текущего контроля выполнения зданий студентами;

Удельный вес занятий, проводимых в форме лабораторных занятий, определен главной целью образовательной программы, особенностью контингента обучающихся и содержанием дисциплины «Биоинформатика» и в целом по дисциплине составляет 100% аудиторных занятий.

6. Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины и учебно-методическое обеспечение самостоятельной работы студентов.

В процессе обучения используются следующие оценочные формы самостоятельной работы студентов, оценочные средства текущего контроля успеваемости и промежуточных аттестаций:

- Выполнение кейсовых заданий по сбору необходимой информации для генов/белков- мишеней

- Подготовка аннотации биологических объектов и защита в форме презентации.

Выполнение лабораторных работ предусматривает поэтапное выполнение кейсовых заданий, сбор результатов и подготовку краткого отчета по каждому разделу рабочей программы, включающего найденную информацию и ключевые

параметры поиска или запуска программ. В завершении работы каждый учащийся должен подготовить общую аннотацию выбранного белка в форме презентации.

Выполнение кейсовых заданий оценивается по каждому этапу программы.

Оценочные средства текущего контроля успеваемости включают контрольные вопросы и задания в форме бланкового и (или) компьютерного тестирования, для контроля освоения обучающимися разделов дисциплины, защита рефератов, курсового проекта.

Образцы экзаменационных билетов, приведены в приложении.

6.1. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине (модулю).

6.1.1. Перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы.

В результате освоения дисциплины (модуля) формируются следующие компетенции:

Код компетенции	В результате освоения образовательной программы обучающийся должен обладать
ОПК-1	способность осуществлять поиск, хранение, обработку и анализ информации из различных источников и баз данных, представлять ее в требуемом формате с использованием информационных, компьютерных и сетевых технологий
ОПК-5	владение основными методами, способами и средствами получения, хранения, переработки информации, навыками работы с компьютером как средством управления информацией
ПК-11	готовностью использовать современные информационные технологии в своей профессиональной области, в том числе базы данных и пакеты прикладных программ

В процессе освоения образовательной программы данные компетенции, в том числе их отдельные компоненты, формируются поэтапно в ходе освоения

обучающимися дисциплин (модулей), практик в соответствии с учебным планом и календарным графиком учебного процесса.

6.1.2. Описание показателей и критериев оценивания компетенций, формируемых по итогам освоения дисциплины (модуля), описание шкал оценивания

Показателем оценивания компетенций на различных этапах их формирования является достижение обучающимися планируемых результатов обучения по дисциплине (модулю).

ОПК-1 - способность осуществлять поиск, хранение, обработку и анализ информации из различных источников и баз данных, представлять ее в требуемом формате с **использованием информационных, компьютерных и сетевых технологий**				
Показатель	**Критерии оценивания**			
	2	**3**	**4**	**5**
знать: Основные открытые банки данных биологической информации	Обучающийся демонстрирует полное отсутствие или недостаточное соответствие следующих знаний: Основных типов биологических данных и биоинформатических банков данных	Обучающийся демонстрирует неполное соответствие следующих знаний: Основных типов биологических данных и биоинформатических банков данных . Допускаются значительные ошибки, проявляется недостаточность знаний, по ряду показателей, обучающийся испытывает значительные затруднения при оперировании знаниями при их переносе на новые объекты.	Обучающийся демонстрирует частичное соответствие следующих знаний: Основных типов биологических данных и биоинформатических банков данных, но допускаются незначительные ошибки, неточности, затруднения при аналитических операциях.	Обучающийся демонстрирует полное соответствие следующих знаний: Основных типов биологических данных и биоинформатических банков данных, свободно оперирует приобретенными знаниями.
уметь: Проводить поиск по открытым архивным и курируемым базам данных	Обучающийся не умеет или в недостаточной степени умеет искать данные по открытым архивным и курируемым базам данных биологической информации	Обучающийся демонстрирует неполное соответствие следующих умений: производить поиск данных по открытым архивным и курируемым базам данных биологической информации . Допускаются значительные ошибки, проявляется недостаточность умений, по ряду показателей, обучающийся испытывает значительные затруднения при оперировании умениями	Обучающийся демонстрирует частичное соответствие следующих умений: производить поиск данных по открытым архивным и курируемым базам данных биологической информации Умения освоены, но допускаются незначительные ошибки, неточности, затруднения при аналитических операциях, переносе умений	Обучающийся демонстрирует полное соответствие следующих умений: производить поиск данных по открытым архивным и курируемым базам данных биологической информации . Свободно оперирует приобретенными умениями, применяет их в ситуациях повышенной сложности.

		при их переносе на новые объекты.	на новые, нестандартные объекты.	
владеть: методами поиска информации о биологических молекулах и их функциях	Обучающийся не владеет или в недостаточной степени владеет методами поиска информации о биологических молекулах и их функциях	Обучающийся владеет методами поиска информации о биологических молекулах и их функциях, допускаются значительные ошибки, проявляется недостаточность владения навыками по ряду показателей, Обучающийся испытывает значительные затруднения при применении навыков на новых объектах.	Обучающийся частично владеет методами поиска информации о биологических молекулах и их функциях, навыки освоены, но допускаются незначительные ошибки, неточности, затруднения при аналитических операциях, переносе умений на новые, нестандартные объекты.	Обучающийся в полном объеме владеет методами поиска информации о биологических молекулах и их функциях, свободно применяет полученные навыки в ситуациях повышенной сложности.
ОПК-5- владение основными методами, способами и средствами получения, хранения, переработки информации, навыками работы с компьютером как средством управления **информацией**				
Показатель	**Критерии оценивания**			
	2	**3**	**4**	**5**

Знать: Основы организации и формы представления биологических данных	Обучающийся демонстрирует полное отсутствие или недостаточное соответствие следующих знаний: Организация и формы представления биологических данных	Обучающийся демонстрирует неполное соответствие следующих знаний: Организация и формы представления биологических данных . Допускаются значительные ошибки, проявляется недостаточность знаний, по ряду показателей, обучающийся испытывает значительные затруднения при оперировании знаниями при их переносе на новые объекты.	Обучающийся демонстрирует частичное соответствие следующих знаний: Организация и формы представления биологических данных, но допускаются незначительные ошибки, неточности, затруднения при аналитических операциях.	Обучающийся демонстрирует полное соответствие следующих знаний: ООрганизация и формы представления биологических данных, свободно оперирует приобретенными знаниями.
Уметь: Проводить аннотиротацию биологических объектов с применением данных из открытых архивных и курируемых баз данных	Обучающийся не умеет или в недостаточной степени умеет проводить аннотиротацию биологических объектов с применением данных из открытых архивных и курируемых баз данных	Обучающийся демонстрирует неполное соответствие следующих умений: Проводить аннотиротацию биологических объектов с применением данных из открытых архивных и курируемых баз данных Допускаются значительные ошибки, проявляется недостаточность умений, по ряду показателей, обучающийся испытывает значительные затруднения при оперировании умениями при их переносе на новые объекты.	Обучающийся демонстрирует частичное соответствие следующих умений: Проводить аннотиротацию биологических объектов с применением данных из открытых архивных и курируемых баз данных Умения освоены, но допускаются незначительные ошибки, неточности, затруднения при аналитических операциях, переносе умений на новые, нестандартные объекты.	Обучающийся демонстрирует полное соответствие следующих умений: Проводить аннотиротацию биологических объектов с применением данных из открытых архивных и курируемых баз данных. Свободно оперирует приобретенными умениями, применяет их в ситуациях повышенной сложности.

Владеть: подходами для сбора необходимой биологической информации и аннотации биологических молекул	Обучающийся не владеет или в недостаточной степени владеетподходами для сбора необходимой биологической информации и аннотации биологических молекул	Обучающийся владеет подходами для сбора необходимой биологической информации и аннотации биологических молекул, допускаются значительные ошибки, проявляется недостаточность владения навыками по ряду показателей, Обучающийся испытывает значительные затруднения при применении навыков на новых объектах.	Обучающийся частично владеет подходами для сбора необходимой биологической информации и аннотации биологических молекул, навыки освоены, но допускаются незначительные ошибки, неточности, затруднения при аналитических операциях, переносе умений на новые, нестандартные объекты.	Обучающийся в полном объеме владеет подходами для сбора необходимой биологической информации и аннотации биологических молекул, свободно применяет полученные навыки в ситуациях повышенной сложности.

ПК-11-готовность использовать современные информационные технологии в своей профессиональной области, в том числе базы данных и пакеты прикладных программ				
Показатель	**Критерии оценивания**			
	2	**3**	**4**	**5**
Знать: Основные методы и подходы системной и структурной биоинформатики	Обучающийся демонстрирует полное отсутствие или недостаточное соответствие следующих знаний: Основные методы и подходы системной и структурной биоинформатики	Обучающийся демонстрирует неполное соответствие следующих знаний: Основные методы и подходы системной и структурной биоинформатики . Допускаются значительные ошибки, проявляется недостаточность знаний, по ряду показателей, обучающийся испытывает значительные затруднения при оперировании знаниями при их переносе на новые объекты.	Обучающийся демонстрирует частичное соответствие следующих знаний: Основные методы и подходы системной и структурной биоинформатики, но допускаются незначительные ошибки, неточности, затруднения при аналитических операциях.	Обучающийся демонстрирует полное соответствие следующих знаний: Основные методы и подходы системной и структурной биоинформатики, свободно оперирует приобретенными знаниями.
Уметь: собирать систематизированную информацию о биологических объектах и применять биоинформатические он-лайн сервисы для решения биологических задач	Обучающийся не умеет или в недостаточной степени умеет собирать систематизированную информацию о биологических объектах и применять биоинформатические он-лайн сервисы для решения биологических задач	Обучающийся демонстрирует неполное соответствие следующих умений: собирать систематизированную информацию о биологических объектах и применять биоинформатические он-лайн сервисы для решения биологических задач Допускаются значительные ошибки, проявляется недостаточность умений, по ряду показателей, обучающийся испытывает значительные затруднения	Обучающийся демонстрирует частичное соответствие следующих умений: собирать систематизированную информацию о биологических объектах и применять биоинформатические он-лайн сервисы для решения биологических задач Умения освоены, но допускаются незначительные ошибки, неточности, затруднения	Обучающийся демонстрирует полное соответствие следующих умений: собирать систематизированную информацию о биологических объектах и применять биоинформатические он-лайн сервисы для решения биологических задач. Свободно оперирует приобретенными умениями, применяет их в ситуациях повышенной сложности.

		при оперировании умениями при их переносе на новые объекты.	при аналитических операциях, переносе умений на новые, нестандартные объекты.	
Владеть: методами поиска необходимой информации по названию, ключевым словам и последовательностям биологических молекул и систематизации данных для аннотации биологических объектов	Обучающийся не владеет или в недостаточной степени владеет методами поиска необходимой информации по названию, ключевым словам и последовательностям биологических молекул и систематизации данных для аннотации биологических объектов	Обучающийся владеет методами поиска необходимой информации по названию, ключевым словам и последовательностям биологических молекул и систематизации данных для аннотации биологических объектов , допускаются значительные ошибки, проявляется недостаточность владения навыками по ряду показателей, Обучающийся испытывает значительные затруднения при применении навыков на новых объектах.	Обучающийся частично владеет методами поиска необходимой информации по названию, ключевым словам и последовательностям биологических молекул и систематизации данных для аннотации биологических объектов , навыки освоены, но допускаются незначительные ошибки, неточности, затруднения при аналитических операциях, переносе умений на новые, нестандартные объекты.	Обучающийся в полном объеме владеет методами поиска необходимой информации по названию, ключевым словам и последовательностям биологических молекул и систематизации данных для аннотации биологических объектов , свободно применяет полученные навыки в ситуациях повышенной сложности.

Шкалы оценивания результатов промежуточной аттестации и их описание:

Форма промежуточной аттестации: зачет.

Промежуточная аттестация обучающихся в форме зачёта проводится по результатам выполнения всех видов учебной работы, предусмотренных учебным планом по данной дисциплине (модулю), при этом учитываются результаты текущего контроля успеваемости в течение семестра. Оценка степени достижения обучающимися планируемых результатов обучения по дисциплине (модулю) проводится преподавателем, ведущим занятия по дисциплине (модулю) методом экспертной оценки. По итогам промежуточной аттестации по дисциплине (модулю) выставляется оценка «зачтено» или «не зачтено».

К промежуточной аттестации допускаются только студенты, выполнившие все виды учебной работы, предусмотренные рабочей программой по дисциплине «Биоинформатика»: практические задания и защита отчетной презентации.

Шкала оценивания	**Описание**
Зачтено	Выполнены все виды учебной работы, предусмотренные учебным планом. Студент демонстрирует соответствие знаний, умений, навыков приведенным в таблицах показателей, оперирует приобретенными знаниями, умениями, навыками, применяет их в ситуациях повышенной сложности. При этом могут быть допущены незначительные ошибки, неточности, затруднения при аналитических операциях, переносе знаний и умений на новые, нестандартные ситуации.
Не зачтено	Не выполнен один или более видов учебной работы, предусмотренных учебным планом. Студент демонстрирует неполное соответствие знаний, умений, навыков приведенным в таблицах показателей, допускаются значительные ошибки, проявляется отсутствие знаний, умений, навыков по ряду показателей, студент испытывает значительные затруднения при оперировании знаниями и умениями при их переносе на новые ситуации.

Фонды оценочных средств представлены в приложении 1 к рабочей программе.

Приложение 1 к
рабочей программе

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«МОСКОВСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

(МОСКОВСКИЙ ПОЛИТЕХ)

Направление подготовки: 19.03.01 Биотехнология
ОП (профиль): «Биотехнология»
Форма обучения: очная
Вид профессиональной деятельности: научно-исследовательская

Кафедра: ___ХимБиотех_____

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

ПО ДИСЦИПЛИНЕ

__Биоинформатика____

Состав: 1. Паспорт фонда оценочных средств

2. Описание оценочных средств:

Составители:

Щербинин Д.С., к.б.н.

Москва, 2021 год

Таблица 1

ПОКАЗАТЕЛЬ УРОВНЯ СФОРМИРОВАННОСТИ КОМПЕТЕНЦИЙ

Биоинформатика					
ФГОС ВО **19.03.01 Биотехнология**					
В процессе освоения данной дисциплины студент формирует и демонстрирует следующие **общекультурные компетенции**:					
КОМПЕТЕНЦИИ		**Перечень компонентов**	**Технология формирования компетенций**	**Форма оценочного средства****	**Степени уровней освоения компетенций**
ИН-ДЕКС	**ФОРМУЛИРОВ-КА**				
ОПК-1	способность осуществлять поиск, хранение, обработку и анализ информации из различных источников и баз данных, представлять ее в требуемом формате с *использованием информационных, компьютерных и сетевых технологий*	**знать:** • Основные открытые банки данных биологической информации **уметь:** • Проводить поиск по открытым архивным и курируемым базам данных **владеть:** • методами поиска информации о биологических молекулах и их функциях	Лабораторные работы, самостоятельная работа,	К-З, ДС УО	**Базовый уровень** - способен искать и систематизировать информацию из разных рассмотренных на занятиях баз данных для биологических объектов **Повышенный уровень** - способен самостоятельно разбираться в структуре новых баз данных. Способен искать и систематизировать информацию из новых баз данных для биологических объектов

ОПК-5	владение основными методами, способами и средствами получения, хранения, переработки информации, навыками работы с компьютером как средством управления ***информацией***	**знать:** • Основы организации и формы представления биологических данных **уметь:** • Проводить аннотацию биологических объектов с применением данных из открытых архивных и курируемых баз данных **владеть:** • подходами для сбора необходимой биологической информации и аннотации биологических молекул	Лабораторные работы, самостоятельная работа,	К-З, ДС УО	**Базовый уровень:** Умеет проводить аннотацию новых объектов с использованием рассмотренных на занятиях подходов. **Повышенный уровень:** Способен проводить более детализированную аннотацию объектов с использованием современных биоинформатических подходов.

ПК-11	готовностью использовать современные информационные технологии в своей *профессиональной области, в том числе базы данных и пакеты прикладных программ*	**знать:** • Основные методы и подходы системной и структурной биоинформатики **уметь:** • методами поиска необходимой информации по названию, ключевым словам и последовательностям биологических молекул **владеть:** • методами поиска необходимой информации по названию, ключевым словам и последовательностям биологических молекул и систематизации данных для аннотации биологических объектов	лекция, самостоятельная работа, семинарские занятия	К-З, ДС УО	**Базовый уровень** Способен применять рассмотренные подходы на новых объектах **Повышенный уровень** Умеет оптимизировать рассмотренные на занятиях подходы для получения более точных результатов Способен проводить автоматизированную аннотацию объектов.

**- Сокращения форм оценочных средств см. в приложении 2 к РП.

Приложение 2
к рабочей программе

Перечень оценочных средств по дисциплине «Биоинформатика»

№ ОС	Наименование оценочного средства	Краткая характеристика оценочного средства	Представление оценочного средства в ФОС
1	Кейс-задача (К-З)	Проблемное задание, в котором обучающемуся предлагают осмыслить реальную профессионально- ориентированную ситуацию, необходимую для решения данной проблемы.	Задания для решения кейс-задачи
2	Коллоквиум (К)	Средство контроля усвоения учебного материала темы, раздела или разделов дисциплины, организованное как учебное занятие в виде собеседования педагогического работника с обучающимися.	Вопросы по темам/разделам дисциплины
3	Доклад, сообщение (ДС)	Продукт самостоятельной работы студента, представляющий собой публичное выступление по представлению полученных результатов решения определенной учебно- практической, учебно-исследовательской или научной темы	Темы докладов, сообщений

Примеры тем для проведения устного опроса-собеседования/коллоквиума
(Вопросы по темам/разделам дисциплины**)**

1. Современные возможности и перспективы биоинформатики.
2. Базовые направления биоинформатики
3. Структурная и функциональная геномика
4. Фармакогеномика
5. Клиническая, функциональная, структурная протеомика,
6. Клиническая биоинформатика. Примеры применения биоинформатических подходов.
7. Методология использования подходов биоинформатики для решения фундаментальных и прикладных задач.
8. Классификация баз данных (автоматические, архивные, курируемые).
9. Нуклеотидные базы данных NCBI. Архитектура записи, виртуальные банки данных, Entrez. Внешние ссылки.
10. Белковые базы данных. Ubiprot/SwissProt, Ubiprot/TrEMBL. Структура и курирование записей. Ссылки на внешние базы.
11. Банки белковых семейств (Pfam, InterPro, SCOP)
12. Базы данных метаболических путей, белок-белковых взаимодействий. KEGG, Reactome.
13. Алгоритмы выравнивания последовательностей белков и нуклеотидов. Их применение
14. Матрицы аминокислотных и нуклеотидных замен. Их получение и применение. Парное выравнивание.
15. Глобальное выравнивание, алгоритм Нидлмана-Вунша. Программная реализация- Needle, Stretcher (EMBOSS).
16. Локальное выравнивание, алгоритм Смита-Ватермана. Программная реализация- Water, Matcher (EMBOSS).
17. Множественные выравнивания. ClustalW
18. Алгоритм BLAST. Score, Bit-Score, e-value
19. Предсказание свойств белков методами биоинформатики.
20. Структуры белков (первичная, вторичная, третичная, четвертичная).
21. Метода получения третичной структуры биологических молекул. Базы данных пространственных структур. PDB
22. Современные методы предсказания вторичной и третичной структуры белков.
23. Моделирование трехмерной структуры белков. Основные подходы, алгоритмы и их недостатки
24. Программы визуализации и анализа пространственных структур биологических молекул.
25. Предсказание и анализ генов прокариот и эукариот
26. Доменная организация белков.
27. Моделирование взаимодействия лигандов с белками методами докинга и молекулярной динамики.
28. Сервисы для работы с геномами.
29. Секвенирование I,II и III поколения.

30. GEO (Gene Expression Omnibus) - база экспрессионных данных.
31. Поиск и обработка результатов экспериментов в базах данных GEO.
32. Исследование коэкспрессии генов. R.
33. Построение и изучение генных сетей. Cytoscape
34. Программное обеспечение для высокопроизводительных вычислений.

Примеры кейсовых задач:

Работа предполагает использование белковых мишеней из списка приказа Минздрав и Минпромторг России №1605/308н от 18.05.2016

Работа с базами NCBI (GeneBank, Nucleotide, RefSeq, OMIM, SNP)

Составьте запрос для ENTREZ для поиска гена, кодирующего выбранный белок, по которому искомый ген выпадает первым в списке. Запрос не должен содержать идентификаторов!

Найдите и сохраните а) полную последовательность гена, б) кодирующую последовательность и в)последовательность кодируемого белка в отдельных текстовых файлах в FASTA-формате

Определите и выпишите координаты всего исследуемого гена и координаты экзонов

Сохраните нуклеотидные последовательности 5 ортологов

Найдите и опишите встречаемые несинонимичные snp (общее количество)

Выпишите гены и их координаты, расположенные слева и справа от исследуемого гена по той же цепи в геноме.

Определите, взаимодейтвуют ли соседние гены (или их продукты) с исследуемым (или его продуктом)?

Приведите список заболеваний, ассоциированных с исследуемым геном (раздел Phenotype и база OMIM)

Определите, есть ли паралоги исследуемого гена?

Парные выравнивания, BLAST

Проведите глобальное парное выравнивание кодирующей области (собранные экзоны) с полной последовательностью выбранного гена.

Подберите лучшие параметры для используемых алгоритмов и сервисов

Представьте результаты предоставить в виде текстового файла, содержащего параметры запуска, полученным выравниванием и кратким описанием работы (название гена, используемое приложение, описание предварительной подготовки последовательностей, проверка

По кодирующей области исследуемого гена найдите ортологи в 5 разных организмах

Для отобранных ортологов сохраните парные выравнивания с описанием результатов в отдельные файлы. В файле должно содержаться:

1.Название гена- Query (Ваш)

2. DEFINITION (развернутое название найденной последовательности)

3. Sequence ID, Длина

4. Score, Identities, Gaps, Strand

5. Само выравнивание

Оптимизируйте поиск таким образом, чтобы в результатах не было последовательностей из ДНК человека и последовательностей низкой достоверности. Общие результаты поиска необходимо сохранить в отдельном файле:

1. Название гена- Query (Ваш ген, буквенное обозначение)

2. RID (действительно 1 ,5 дня)

3. Ссылка на RID

4. Search Summary

Сервисы для выполнения заданий:

http://www.ebi.ac.uk/Tools/psa/

http://Blast.ncbi.nlm.nih.gov/

http://www.bioinformatics.org/sms/rev_comp.html

http://arep.med.harvard.edu/labgc/adnan/projects/Utilities/revcomp.html

Работа с базами Uniprot

Найдите рецензированную запись о вашем выбранном белке в базе Uniprot и сохраните ее ID.

Определите количество известных изоформ исследуемого белка и опишите их отличия

Приведите описание функции белка. (на английском и русском языке)

Выпишите ключевые участки в аминокислотной последовательности и опишите, что в них происходит (пример- сайты связывания лигандов, субстратов, координация ионов и т.д.)

Приведите список функций белков и биологических процессов в классификации Gene Ontology

Определите локализацию белка и топологически важные участки (пример- мембранные участки, внеклеточные и т.д.)

Выпишите остатки, для которых получены экспериментальные данные мутагенеза и природные варианты с описанием замены и фенотипических проявлений.

Определите, есть ли зарегистрированные лекарства и соединения, проявляющие ингибирующую активность по отношению к исследуемому белку? (ChEMBL, DrugBank. Выпишите ID соединения, название лекарств и общее количество соединений с ингибирующей активностью)

Определите, происходит ли процессинг белка после трансляции (Пример- отрезание сигнальных пептидов, внесение разрывов)? Если да, то выпишите тип и место процессинга белковой цепи.

Определите, какие происходят постранслящионные модификации в белке(пример-фосфорилирование, гликозилирование и т.д.). Укажите а.о, тип модификации и метод ее верификации (статья, анализ последовательности, гомология..)

Определите, с какими белками взаимодействует исследуемый белок.

Определите, есть ли пространственные структуры для исследуемого белка? Укажите, для каких участков разрешено пространственное расположение

Сохранить последовательность исследуемого белка и таких же белков из других организмов в Fasta-формате.

Множественное выравнивание, доменная организация белков

Постройте множественное выравнивание вашего белка с аналогичными из 5 других организмов. Сохраните результаты выравнивания в отдельный файл. Сохраните филогенетическое дерево.

Определите доменную организацию исследуемого белка и приведите схематическую структуру с соответствующими профилями выравнивания (Logo-формат).

В построенном выравнивании найдите и отметьте домены белков.

Оптимизируйте параметры выравнивания (при необходимости) для улучшения выравнивание доменов.

- Можно искать в Pfam, Smart, CDD…

Сервисы для выполнения заданий:

http://www.ebi.ac.uk/Tools/msa/

http://pfam.xfam.org/

Работа с базой метаболических и сигнальных путей KEGG

В базе KEGG найдите известные пути, в которых принимает участие исследуемый белок. Приведите их список.

Сохраните карты найденных путей, связанных с заболеванием.

Приведите описание путей. (Для чего они нужны, что происходит, какие ключевые компоненты …)

Выделить ключевые компоненты и реакции на сохраненных картах.

Сервисы для выполнения заданий:

http://www.genome.jp/kegg/

Работа с пространственными структурами белков. Банк PDB. PyMol

Выпишите известные структуры для исследуемого белка

Отметьте лучшие структуры на ваш взгляд. Обоснуйте соответствующий выбор

Опишите состав выбранных PBD- файлов: состав цепей, наличие лигандов и дополнительных низкомолекулярных соединений.

Определите фактическую длину разрешенной последовательности исследуемого белка в PDB (а.о. начала, а.о конец, разрывы цепи)

Подготовьте иллюстрацию структуры исследуемого белка с выделением важных участков и структур в программе Pymol (домены, модификации а.о., контактирующие остатки в комплексе и т.д.)

Сервисы для выполнения заданий:

http://www.wwpdb.org/

http://rcsb.org/

PyMol

Контрольные вопросы для проведения аттестации по итогам освоения дисциплины (зачет):

1. Биоинформатика: объекты исследования и методы.
2. Основные направления биоинформатики. Примеры практического применения
3. Базы данных биологической информации. Классификация, примеры.
4. Базы данных белковых последовательностей.
5. Базы данных последовательностей нуклеиновых кислот.
6. Банки данных метаболических путей.
7. Базы, содержащие результаты экспериментов по анализу экспрессии.
8. Основные библиографические базы данных.
9. Реализация поиска по базам NCBI: ENTREZ и BLAST
10. Выравнивание двух последовательностей. Матрицы замен.
11. Глобальное и локальное выравнивание, вес выравнивания, матрицы аминокислотных замен.
12. Локальное выравнивание, задачи, примеры.
13. Применение динамического программирования при поиске оптимальное выравнивание последовательностей.
14. Зависимость выравнивания от параметров. Оценка качества выравниваний.
15. Множественное выравнивание. Clustal
16. Алгоритм BLAST. Score, Bit-Score, e-value
17. Скрытые Марковские модели (HMM).
18. Выравнивания и филогения.
19. Третичная структура белка. Фолдинг.
20. Предсказание третичной структуры белка. Моделирование гомологов. Методы, ПО, сервисы.
21. Экспериментальное определение структуры белка. Оценка качества полученной структуры.
22. Ортологи, паралоги, гомологи. Определения, характеристики, поиск.
23. Алгоритмы построения филогенетических деревьев
24. Расшифровка нуклеотидных последовательностей. Секвенирование I, II и III поколения

Примеры зачетно-экзаменационных заданий

Выполнение заданий осуществляется с применением открытых биоинформатических сервисов и Баз данных:

NCBI https://www.ncbi.nlm.nih.gov/

UNIPROT http://www.uniprot.org/

PFAM *http://pfam.xfam.org/*

PDB to FASTA *http://www.meilerlab.org/index.php/servers/show?s_id=3*

PDB *http://www.rcsb.org/pdb/home/home.do*

BLAST *https://blast.ncbi.nlm.nih.gov/Blast.cgi*

MSA *http://www.ebi.ac.uk/Tools/msa/*

1. Для работы вам предлагается PDB- файл “1.pdb”. В файле содержится информация о пространственной структуре некоторого домена белка.
- Определите полное название белка и организм? Выпишите Uniprot ID данного белка.
- Какой домен белка содержится в структуре?
- Какой тип четвертичной структуры представлен в PDB-файле?

2. Вам дан файл последовательности одного из доменов белка (2.fasta).
- Определите полное название белка и организм? Выпишите Uniprot ID данного белка.
- Последовательность какого домена содержится в файле?
- Постройте множественное выравнивание этих белков из 5 организмов (включая человека).

3. В файле 3.fasta находится последовательность одного из экзонов гена А.
На той же хромосоме, на прямой цепи, справа от гена А находится ген Б.
- Приведите название и uniprot-идентификатор белка, кодируемого геном А.
- Определите белок, кодируемый геном Б. Приведите идентификатор Uniprot и опишите функцию данного белка.

4. Для работы вам предлагается PDB- файл “4.pdb”.
- Определите полное название белка и организм? Выпишите Uniprot ID данного белка.
- С какими заболеваниями (в базе KEGG) ассоциирован данный белок?
5. В файле 5.fasta находится последовательность одного из экзонов гена А.
- Приведите название и uniprot-идентификатор белка, кодируемого геном А.
- Приведите каталитическую реакцию, в которой принимает участие этот белок и назовите кофактор.
- Для каких доменов белка известна пространственная структура
6. Вам дан файл части последовательности одного мышиного белка (6.fasta).
- Определите полное название белка? Выпишите Uniprot ID гомологичного белка в организме человека.
- Отметьте в Pymol известные домены данного белка

7. Для белка Интерлейкин-6 найдите ассоциированные заболевания в OMIM и KEGG Desease.

- Укажите ID: Uniprot, KO, OMIM, Gene (NCBI)

- Найдите пути, в которых данный белок принимает участие, и которые ассоциированы с заболеваниями, связанными с артритом. С какими белками Интерлейкин в этих путях непосредственно взаимодействует?

7. Учебно-методическое и информационное обеспечение дисциплины.

Основная литература:

1. Володченкова, Л.А. Биоинформатика : [16+] / Л.А. Володченкова ; Министерство образования и науки РФ, Омский государственный университет им. Ф. М. Достоевского. – Омск : ОмГУ им. Ф.М. Достоевского, 2018. – 44 с. : ил. – Режим доступа: по подписке. –URL: http://biblioclub.ru/index.php?page=book&id=563147

Дополнительная литература:

2. C. A. Floudas etc. Optimization in Computational Chemistry and Molecular Biology.Local and Global Approaches [Электронный ресурс]. – Springer, 2000. – 342 p. ISBN: 978-1-4757-3218-4 (Online)

3. Stephen A. Krawetz etc. Introduction to Bioinformatics. A Theoretical and Practical Approach [Электронный ресурс]. – Springer, 2003. – 746 p. ISBN: 978-1-59259-335-4 (Online)

4. Jonathan M. Keith. Bioinformatics. Structure, Function and Applications [Электронный ресурс]. – Springer, 2008. – 502 p. ISBN: 978-1-60327-429-6 (Online)

5. Хёльтье Х.-Д., Зиппль В. Роньян Д., Фолькерс Г. Молекулярное моделирование: теория и практика [Электронный ресурс]. М.: Бином. Лаб. знаний, 2015. - 322 с. ISBN 978-5-9963-2401-9

8. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины (модуля)

6. http:// uniprot.org/ – SWISS-PROT|UniProt the protein sequence data bank, база данных UniProt

7. http://www.expasy.org/sprot/ – базы данных Swiss-Prot, TrEmbl, UniProt на сервере ExPASy (Expert Protein Analysis System) Швейцарского Института Биоинформатики SIB

8. http://www.ebi.ac.uk/uniprot/ – база данных UniProt на сервере Европейского института биоинформатики (European Bioinformatics Institute, EBI)

9. http://www.jcbi.ru/ – Объединенный Центр вычислительной биологии и биоинформатики, русскоязычный информационный сайт с Web-адресами и краткой характеристикой молекулярно-биологических баз данных

10. http://www.ebi.ac.uk/services – Биоинформатические сервисы, представленные на портале европейского института биоинформатики

11. http://bioinformatics.ru/en/Resources/Books/ - Список учебной литературы

12. http://www.reactome.org/ - база данных метаболических путей

13. http://www.kegg.jp/kegg/ - база данных метаболических путей

14. http://www.rcsb.org/ – Protein Data Bank, база данных PDB.
15. http://www.ncbi.nlm.nih.gov/ (http://www.pubmed.com/) – сервер Национального центра биотехнологической информации США (NCBI): базы данных GenBank, NCBI Protein Database, UniGene, HomoloGene и др.
16. http://www.genebee.msu.su/ – регулярно обновляемая копия (зеркало) базы компании GeneBio в России, на сайте Института физико-химической биологии им. А.Н. Белозерского
17. http://www1.fips.ru/wps/wcm/connect/content_ru/ru - РОСПАТЕНТ
18. http://patft.uspto.gov/ - United States Patent and Trademark Office Бесплатная патентная база.
19. www.febs.org/ - Официальный сайт Федерации европейских биохимических обществ.
20. www.chemport.org/ - Научные издания в области биохимии, химии и смежных наук.
21. www.molbiol.ru/ - Учебники, научные монографии, обзоры, лабораторные практикумы в свободном доступе на сайте практической молекулярной биологии.
22. www.chem.qmul.ac.uk/ - биохимическая классификация и номенклатура. Свободный доступ на сайте Международного союза биохимии и молекулярной биологии.
23. www.scopus.com (Scopus) – единая реферативная и наукометрическая база данных (индекс цитирования) (доступ в библиотеке МАМИ)
24. www.scincedirect.com/ (Архивные коллекции журналов издательства Elsevier) –архивные коллекции различных тематик, в том числе Biochemistry, Engineering and Technology.
25. http://cyberleninka.ru/article/c/biotehnologiya - научная электронная библиотека «КИБЕРЛЕНИНКА»
26. http://www.springerprotocols.com/ - доступ к базе данных SpringerLink
27. http://grebennikon.ru/ - электронная бибоиотека Grebennicon
28. http://login.webofknowledge.com/ - ресурсы на платформе Web of Knowledge

9. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения и информационных справочных систем (при необходимости)

Программное обеспечение:

1. Microsoft Windows 8/7/Vista/XP/ / OpenSuse/ CentOS
2. Microsoft Offcie (Word, Excel, Outlook, PowerPoint)2003/2007/2010/2013
3. Веб-сервис сайта http://www.clustal.org/
4. Веб-сервис сайта https://blast.ncbi.nlm.nih.gov/Blast.cgi
5. Веб-сервис сайта http://www.ebi.ac.uk/Tools/emboss/
6. ClustalW
7. Pymol

При осуществлении образовательного процесса по дисциплине используются следующие информационно-телекоммуникационные технологии:
Лекции с использованием презентаций в PowerPoint и доступом к биоинформатическим сервисам и базам данных
Практические занятия с пакетами программ для работы с последовательностями (EMBOSS); ClustalW, ClustalX (веб-сервис сайта http://www.clustal.org/); Pymol, и базами данных NCBI и Uniprot- GenBank, EMBL Nucleotide, Swiss-Prot, и др.;
Самостоятельная работа студентов с базами данных:

www.ncbi.nlm.nih.gov/PubMed - Свободный доступ в базу научных данных в области биомедицинских наук MedLine

http://www.kegg.jp/kegg/ - база данных метаболических путей

http://www.uniprot.org/ - Сервер протеомной информации

5. https://www.ncbi.nlm.nih.gov/ Портал национального центра биотехнологической информации

10. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине

Для реализации рабочей программы необходимы: компьютерные классы с рабочими местами, обеспечивающими выход в Интернет, стандартный набор учебной мебели, специализированное программное обеспечение.

Реализация учебной программы должна обеспечиваться доступом каждого студента к информационным ресурсам – библиотечному фонду и сетевым ресурсам Интернет.

Структура и содержание дисциплины «Биоинформатика»
Направление подготовки
19.03.01 «Биотехнология»

Раздел	Семестр	Неделя семестра	Виды учебной работы, включая самостоятельную работу студентов и трудоемкость в часах					Виды самостоятельной работы студентов					Формы аттестации	
			Л	П/С	Лаб.	СРС	КСР	КР	КП	РГР	Реф.	К/Р	Э	З
1. Введение. Предмет, задачи и объекты биоинформатики. Основные направления дисциплины «Биоинформатика». Геномика и протеомика Лабораторная работа 1	6	1			2	0	+							
2. Базы данных нуклеотидных последовательностей NCBI. GeneBank, Nucleotide, Snp. Лабораторная работа 2	6	2			6	6						+		
3. Парное выравнивание последовательностей. Поиск последовательностей BLAST (Basic Local Alignment Search Tool) Лабораторная работа 3	6	5			6	6	+					+		
4. Базы данных белков Uniprot. SwissProt, trEMBL.	6	8			6	6	+					+		
5. Множественное выравнивание последовательностей. Clustal	6	11			4	4	+					+		
6. Изучение белковых семейств. Доменная организация. Профили выравнивания. PFAM	6	13			4	4	+							
7. Функциональная аннотация белков.	6	15			4	6	+					+		

GeneOntotlogy, Kegg														
8. Строение и функции белков. Работа с белковыми структурами. Методы расшифровки и предсказания пространственных структур. Банк PDB.	6	17			4	4	+					+		
Форма аттестации		18									**Один отчет-презентация**			3
Итого		**18**			**36**	**36**	+					+		+