

*Федеральное государственное автономное образовательное учреждение
высшего образования*

«Российский университет дружбы народов»

Институт биохимической технологии и нанотехнологии (ИБХТН)

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Наименование дисциплины

Введение в биоинформатику

Рекомендуется для направления подготовки/специальности

04.04.01 «Химия»

Уровень: Магистр

Москва, 2021

1. Цели освоения дисциплины

Целью изучения дисциплины «Введение в Биоинформатику» является формирование у обучающихся представления о биоинформатике, соответствующих вычислительных методах, осуществлении экспериментальных работ по молекулярной биологии, биохимии, генетике, а также разработке алгоритмов компьютерного анализа данных геномики и протеомики, программ, позволяющих предсказывать пространственную структуру биополимеров.

2. Место дисциплины в структуре основной образовательной программы образования

Дисциплина «Введение в Биоинформатику» входит в вариативную часть и является дисциплиной по выбору. Курс опирается на основные знания и представления о биохимии, молекулярной биологии, генетике. В рамках изучения дисциплины студенты используют методы прикладной математики, статистики и информатики в молекулярно-биологических исследованиях.

Подготовка по дисциплине даёт возможность получить теоретическую базу и практические навыки использования компьютерных технологий для обработки экспериментальных данных по структуре биологических макромолекул с целью получения биологически важной информации.

В таблице № 1 приведены предшествующие и последующие дисциплины, направленные на формирование компетенций дисциплины в соответствии с ОС ВО и матрицей компетенций ОП ВО.

Таблица № 1

Предшествующие и последующие дисциплины, направленные на формирование компетенций

№ п/п	Шифр и наименование компетенции	Предшествующие дисциплины	Последующие дисциплины (группы дисциплин)
Профессиональные компетенции			
1	ПК-1-н. Способен планировать работу и выбирать адекватные методы решения научно-исследовательских задач в выбранной области химии, химической технологии или смежных с химией наук	Основы фармацевтической технологии и нанотехнологии	Разработка и регистрация лекарственных препаратов; Современные принципы контроля качества лекарственных средств

3. Требования к результатам освоения дисциплины:

Процесс изучения дисциплины направлен на формирование следующих компетенций:

№ п/п	Код и наименование компетенции	Код и наименование индикатора достижения компетенции
1	ПК-1-н. Способен планировать работу и выбирать адекватные методы решения научно-исследовательских задач в выбранной области химии, химической технологии или смежных с химией наук	ПК-1-н-1. Составляет общий план исследования и детальные планы отдельных стадий; ПК-1-н-2. Выбирает экспериментальные и расчетно-теоретические методы решения поставленной задачи исходя из имеющихся материальных и временных ресурсов

В результате освоения учебной дисциплины обучающийся должен:

Знать:

- основы биоинформатики, применяемые в практике методы программирования;
- особенности разработки алгоритмов анализа биологических данных большого объема;
- последние достижения и новые разработки в области биоинформатики;
- химию и физику нуклеиновых кислот и белков;

Уметь:

- получать и грамотно использовать информацию, накопленную в базах данных по структуре геномов, белков, и другой биологической информации;
- разрабатывать новые программы, используемые для решения задач в области биоинформатики;
- модифицировать известные, и создавать специализированные и общедоступные биоинформационные сайты;
- интерпретировать различные типы биологических данных;
- использовать современное научное оборудование в профессиональной области;

Владеть:

- навыками работы с биоинформационными ресурсами;
- методами разработки программного обеспечения для управления и быстрого доступа к биологическим данным;
- методами молекулярного моделирования различных биологических объектов и изучения динамики макромолекул;
- создание и сопровождение специализированных баз данных.

4. Структура и содержание дисциплины

Общая трудоемкость дисциплины составляет 3 зачетные единицы.

Вид учебной работы	Всего часов	1 курс			
		Семестры			
		1	2	3	4
Аудиторные занятия (всего)	36			36	
Лекции	18			18	
Практические занятия (ПЗ)	18			18	
Лабораторные работы (ЛР)					
Самостоятельная работа (всего)	72			72	
Итоговая аттестация	Диф.зачет				
Общая трудоемкость, час зач. ед.	108			108	
	3			3	

5. Содержание дисциплины

Тема 1. Введение

Цели, задачи и методы, основные приложения. Основные понятия. Аминокислоты, строение и свойства. ДНК, РНК, нуклеотиды. Анализ последовательностей.

Тема 2. Базы данных

Интернет для биоинформатики. Способы представления информации о последовательностях. Основы структур баз данных: записи, поля, объекты. Форматы записи FASTA, BLAST, GenBank, PDB. Классификация баз данных (автоматические, архивные, курируемые). Основные базы данных: GenBank, EMBL, SwissProt, TrEMBL, PIR, PDB, банки белковых семейств (ProDom, PFAM, InterPro, SCOP), метаболические базы данных, генетические банки (физические карты, OMIM), специализированные банки данных. Поиск гомологичных последовательностей в базах данных.

Тема 3. Уровни структурной организации белков

Первичная структура белка. Вторичная, третичная и четвертичная структуры протеинов. Мотивы и домены. Функции белков, связь со структурой. Современные методы предсказания вторичной и третичной структуры белков на основе первичной структуры. Метод моделирования по гомологиям. Базы данных пространственных структур биополимеров.

Тема 4. Сравнение последовательностей

Анализ последовательностей нуклеотидов. Строение молекулы ДНК, упаковка, комплементарность. Гены, регуляторные последовательности.

Математические основы выравнивания последовательностей символов. Матрицы аминокислотных замен, парное выравнивание и его оценка, множественное выравнивание, вычислительные ресурсы. Глобальное выравнивание: алгоритм Нидельмана-Вунша. Локальное выравнивание: алгоритм Смита-Ватермана. Другие варианты выравнивания. Статистическая значимость выравниваний. Зависимость выравнивания от параметров. Множественное выравнивание. Применение выравнивания в биоинформатике.

Тема 5. Методы определения пространственной структуры биополимеров

Структура записи PDB. Анализ структурных особенностей. Предсказание вторичной структуры. Предсказание третичной структуры белков по гомологии. Моделирование гомологов. Фолдинг и его распознавание.

Тема 6. Предсказание функции биополимеров по последовательности

Анализ гомологов и функциональные сигналы. Лидерные пептиды и трансмембранные сегменты. Сайты модификации белков (гликозилирование, фосфорилирование и т.п.).

Функциональные сайты ДНК. Гены прокариот и эукариот. Сравнительные методы предсказания генов. Поиск РНК с заданной структурой (тРНК и т.п., регуляторные участки мРНК).

Тема 7. Эволюция на уровне молекул

Эволюция молекул и организмов (горизонтальный перенос, ортологи, паралоги, деревья генов). Филогенетическое дерево. Модели эволюции. Эволюция на уровне генома. Анализ популяционных данных.

Тема 8. Актуальные проблемы биоинформатики

Аннотации генома, поиск генов, поиск сайтов репликации в геноме человека. Предсказание структуры, функции и клеточной локализации белков. Медицинская и хемоинформатика.

6. Учебно-методическое обеспечение самостоятельной работы студентов.

Самостоятельная работа студентов заключается в поиске и обработке информации по основным разделам дисциплины как в библиотечном фонде, так и в электронных базах данных. Организация и контроль выполнения самостоятельной работы студентами осуществляется через представление докладов и рефератов, подготовленных во время семестра. Доклады завершаются дискуссией по основным вопросам, затронутым в устных сообщениях. Для самостоятельной проработки различных тем и выполнения заданий, обучающиеся используют учебную и учебно-методическую литературу.

Предполагаются следующие виды самостоятельной работы - написание рефератов, составление словариков терминов, используемых в данной дисциплине, списка персоналий с указанием наиболее важных открытий названных ученых в области естественных наук, составление таблиц, схем различных процессов, создание мультимедийных презентаций.

7. Учебно-методическое и информационное обеспечение

7.1. Основная литература

Tore Samuelsson Genomics and Bioinformatics An Introduction to Programming Tools for Life Scientists From Bioinformatics Basics to Working Code - Эл. книга на платформе издательства Myilibrary. <http://lib.rudn.ru/25>

7.2. Литература для самостоятельного изучения

1. Стефанов, В. Е. Биоинформатика : учебник для академического бакалавриата / В. Е. Стефанов, А. А. Тулуб, Г. Р. Мавропуло-Столяренко. — Москва : Издательство Юрайт, 2019. — 252 с. — (Бакалавр. Академический курс). — ISBN 978-5-534-00860-9. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: <https://biblio-online.ru/bcode/433453> (дата обращения: 05.08.2019).
2. Ананько, Е.А. Введение в информационную биологию и биоинформатику : учеб. пособие : [в 5 т.]: Т. 3, гл. 3, 4: Теория генных связей. Картирование генов, контролируемых сложными признаками человека [Текст] / Ананько Е. А. и др. под ред. Н. А. Колчанова, О. В. Вишневого, Д. П. Фурман. - Новосибирск : НГУ 2015. - 297 с.
3. Леск, А.М. Введение в биоинформатику [Текст] / А.М. Леск; пер. с англ. под ред. А.А. Миронова, В.К. Швядоса. - 2-е изд. - М.: БИНОМ. Лаборатория знаний, 2015. - 318 с.
4. Стефанов, В.Е., Биоинформатика: учебник для академического бакалавриата [Текст] / В.Е. Стефанов, А.А. Тулуб, Г.Р. Мавропуло-Столяренко. - СПб.: СПГУ, 2016. - 252 с.
5. Шевелухи, В.С. Сельскохозяйственная биотехнология и биоинженерия: учебник для студентов высших учебных заведений, обучающихся по сельскохозяйственным, естественнонаучным и педагогическим специальностям [Текст] / под ред. В.С. Шевелухи. - Изд. 4-е, значит. перераб. и доп. - М.: URSS, 2015. - 700 с.
6. Белькова, Н.Л. Большой практикум по биоинженерии и биоинформатике: учеб.-метод. Пособие. в 3 ч. Ч. 2. Нуклеиновые кислоты [Текст] / Н. Л. Белькова. - Иркутск: ИГУ, 2014. - 155 с.
7. Долгих, Д. А. Введение в молекулярную биоинженерию: учебно-методический комплекс для бакалавров по дисциплине [Текст] / Д. А. Долгих [и др.]. - М.: Ин-т АйТи, 2011. - 137с.
8. Кирпичников, М.П. Методические разработки по реализации образовательных программ в области биоинженерии [Текст] / М.П. Кирпичников, К.В. Шайтан ; Моск. гос. ун-т. им. М.В.Ломоносова. - М.: МГУ, 2007. - 140 с.
9. Огурцов, А.Н. Введение в биоинформатику [Текст] / А.Н. Огурцов. - Х.: НТУ «ХПИ», 2011. - 400с.
10. Приставка, А. А. Большой практикум по биоинженерии и биоинформатике: учеб.-метод. Пособие, в 3 ч.Ч. 1. Белки [Текст] / А. А. Приставка, В. П. Саловарова. - Иркутск: ИГУ, 2013.
11. 121 с.
12. Эллиот В. Биохимия и молекулярная биология [Текст] / В. Эллиот, Д. Эллиот. - М.: МАИК, 2002. - 431 с.

Рооры

7.3. Интернет-ресурсы:

- URL: <http://beta.uniprot.org> - SWISS-PROT, UniProt the protein sequence data bank, база данных UniProt
- URL: <http://cmm.info.nih.gov/modeling> - сервер Центра моделирования молекул Национального Института Здоровья НИН, США
- URL: <http://www.ebi.ac.uk/uniprot> - база данных UniProt на сервере Европейского института биоинформатики (European Bioinformatics Institute, EBI)
- URL: <http://www.expasy.org/sprot> - базы данных Swiss-Prot, TrEmbl, UniProt на сервере ExPASy (Expert Protein Analysis System) Швейцарского Института Биоинформатики SIB
- URL: <http://www.genebee.msu.su> - регулярно обновляемая копия (зеркало) базы компании GeneBio в России, на сайте Института физико-химической биологии им. А.Н. Белозерского

- URL: <http://www.genebio.com> - сайт компании GeneBio (Geneva Bioinformatics S.A.), распространяющей информацию из протеомных баз данных: SWISS-PROT, PROSITE, SWISS-2DPAGE и соответствующие программные приложения
- URL: <http://molbiol.ru> - Классическая и молекулярная биология
- URL: <http://molbiol.edu.ru> - Практическая молекулярная биология
- URL: <http://proteome.ru> - русскоязычный сайт проекта «Протеом человека»
- URL: <http://www.jcibi.ru> - Объединенный Центр вычислительной биологии и биоинформатики, русскоязычный информационный сайт с вэб-адресами и краткой характеристикой молекулярно-биологических баз данных
- URL: <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.
- URL: <http://www.rcsb.org> - Protein Data Bank, база данных PDB.

7.4. Требования к материально-техническому обеспечению

Компьютерная и мультимедийная техника, раздаточный материал, наглядный иллюстративный материал, электронные презентации.

- Microsoft Windows XP Professional Russian Upgrade Academic Open No Level, Лицензия № 15988873, дата выдачи 15.01.2003 г.
- Microsoft Office 2007 Russian Academic Open No Level Лицензия № 43178981, дата выдачи 12.12.2007 г. (Windows 7, Microsoft Office 2007, Microsoft Security Essentials)

7.5. Методические указания для обучающихся по освоению дисциплины

Студенты должны соблюдать дисциплину, вовремя приходить на занятия, предоставлять на проверку домашнюю работу, готовиться к проверочным и контрольным работам, предусмотренным курсом, проявлять активность на занятиях.

Важное место в образовательном процессе занимает самостоятельная работа студентов. Для организации самостоятельной работы студентов по курсу используются современные информационные технологии: размещенные в сетевом доступе комплексы учебных и учебнометодических материалов (программа, список рекомендуемой литературы и информационных ресурсов, задания для самоконтроля), свободный доступ к сети «Интернет» для работы с молекулярными базами данных.

В рамках самостоятельной работы студенты готовят рефераты, а также доклады к семинарским занятиям. Подготовленный реферат по выбранной теме предоставляется преподавателю на проверку. Рефераты, получившие высокую оценку, представляются другим студентам на семинарском занятии.

7.6. Материально-техническое обеспечение дисциплины

Процесс изучения дисциплины обеспечен аудиторией, оборудованной персональными компьютерами, мультимедийными средствами для демонстрации презентаций, программным обеспечением MS Office и доступом к информационно-телекоммуникационной сети «Интернет»

Москва, ул. Миклухо-Маклая, д.10, корп.2, учебная аудитория для проведения занятий лекционного типа, компьютерный класс 622.

Комплект специализированной мебели; технические средства: 20 компьютеризированных рабочих мест: Моноблок Lenovo IdeaCentre 19.5" HD+Cel J1800/4Gb/500Gb/DVDRW/DOS/kb/m/черный 1600x900 (19 шт.), Моноблок Lenovo IdeaCentre 23" C560 (57331093) i3-4160T (3.1ГГц)/4G/1Tb/DVD-SMulti/23FHD91920x1080)/NV 800M 2G/Wi-Fi/cam/White (1шт.). Обеспечен выход в интернет.

8. Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины.

8.1. Темы рефератов

1. Конструирование модифицированных и новых биологических объектов.
2. Особенности организации экспериментов с клетками и культурами клеток.
3. Внутриклеточный транспорт токсичных молекул.
4. Структурные особенности и взаимодействие макромолекул.
5. Получение искусственных белков с заданными свойствами и их синтез.
6. Культивирование клеток различного происхождения,
7. Особенности создания генно-инженерных конструкций.
8. Генетические маркеры выносливости и работоспособности человека.
9. Приоритеты компьютерного программирования в биоинженерии и биоинформатике.
10. Филогенетическое дерево для белков и организмов
11. Структурная организация хроматина и регуляция экспрессии генов (гистоновый код).
12. Сигнальные каскады: регуляция экспрессии генов, пролиферации и апоптоза.
13. Принципы работы с рекомбинантными ДНК.
14. Мутационный процесс: изменения в последовательности ДНК.
15. Секвенирование геномов. Методы сборки.
16. Использование методов биоинформатики и молекулярного моделирования для изучения механизма действия ферментов.
17. Основные базы данных о биологических последовательностях.
18. Структура записи в банках данных о геномных последовательностях нуклеотидов и белков.
19. Геномный браузер: поиск информации о геноме человека.
20. Скрытые Марковские модели.

8.2. Темы практических работ

1. Средства работы с банками данных I (Entrez) .
2. Средства работы с банками данных II (SRS).
3. Сервис GeneVee. Основные поля записи SwissProt.
4. Поиск гомологов (интерпретация результатов, сравнение алгоритмов, зависимость от параметров).
5. Построение выравниваний, реконструкция филогенетических деревьев (сравнение локальных и глобальных выравниваний, зависимость выравнивания от параметров, оценка статистической значимости).
6. Работа с банком пространственных структур PDB.
7. Структуры белков (RASMOL, SwissPDBViewer). Работа с программой визуализации макромолекул RasMol I.
8. Аннотирование последовательности (поиск белок-кодирующих областей, поиск функциональных сайтов).
9. Работа с программой визуализации макромолекул RasMol II.
10. Поиск слабых сигналов в биологических последовательностях. Интернет-ресурсы работы с полными геномами.
11. Вторичные структуры РНК.
- Предсказание структурных особенностей белков.

8.3. Вопросы для текущего контроля успеваемости

Семинар 1. Цели, задачи и методы, основные приложения

- Современное определение биоинформатики и биоинженерии.
- Специфика биоинформационных данных.
- Виды биоинформационной деятельности.
- Проект «Геном человека».
- Основные направления использования биоинформационных данных.
- Информация и энтропия.

- Молекулярно-биологические основы биоинформатики.
- Информационно компьютерные и Интернет-компоненты биоинформатики.
- Применения биоинформатики и биоинженерии.

Семинар 2. Базы данных

- Свойства информации.
- Генерация и рецепция информации.
- Особенности генетической информации.
- Биоинформационные данные, сети и базы.
- NCBI и сервисы.
- Форматы записи Fasta, Genbank, PDB.
- EMBL.
- Swiss-PDBviewer.

Семинар 3. Сравнение последовательностей

- Геномная информация.
- Генетические карты и картирование генома.
- Основные типы ДНК-маркеров, используемые при картировании генома.
- Определение нуклеотидных последовательностей, секвенирование ДНК.
- определение сиквенса клона.
- Использование EST-последовательностей.
- Методы анализа множественной экспрессии генов.
- Секвенирование белков.
- Анализ белковой экспрессии методом двумерного фореа в полиакриламидном геле.
- Глобальное выравнивание последовательностей.
- Локальное выравнивание последовательностей.
- Множественное выравнивание последовательностей.
- Мера сходства биологических последовательностей.
- Расстояния Хемминга и Левенштайна.
- операции редактирования.
- Вес операций редактирования.
- Виды штрафов за делеции.

Семинар 4. Пространственной структуры биополимеров

- Структура записи PDB.
- Анализ структурных особенностей.
- Моделирование.
- Предсказание вторичной и третичной структуры белков по гомологии.
- Предсказание параметров спирали ДНК.
- Динамическое программирование и динамические модели РНК.

Семинар 5. Эволюция на уровне молекул

- Эволюция молекул и организмов (ортологи и паралоги, горизонтальный перенос).
- Филогенетическое дерево как математический объект.
- Модели эволюции.
- Алгоритмы построения филогенетических деревьев (матрица расстояний, методы, основанные на матрице расстояний (upgma, neighbour-joining, minimal evolution, топологические инварианты и др.)).
- Алгоритмические проблемы поиска оптимального дерева, bootstrapping, согласование деревьев.
- Эволюция на уровне генома (синтения, хромосомные перестройки).
- Анализ популяционных данных.
- Snp, тандемные повторы, митохондрии и Y-хромосомы, данные по рестрикции.

- Статистика последовательностей ДНК.
- (ди)нуклеотидный состав (изохоры, gc-острова, картирование старта репликации).
- Частые и редкие слова (вероятностные проблемы).
- Статистика ДНК как характеристика генома.
- Вычислительная геномика.
- Метаболическая реконструкция (в т.ч. неортологичные замещения)
- Позиционный анализ.
- Эволюция регуляторных взаимодействий.
- Эволюция белковых семейств, их доля в геноме.

8.4. Темы докладов

1. Современные методы исследования первичной структуры белка (определение N-, C-концевых аминокислот; секвенирование).
2. Протеомика: возможности и перспективы.
3. Процессинг и фолдинг синтезированного белка, биологическое значение этих процессов.
4. Компьютерный дизайн лекарственных средств.
5. Филогенетические деревья.
6. MALDI-TOF-спектрометрия биоорганических макромолекул.
7. Гипотеза «молекулярных часов».
8. Gfr - один из самых востребованных маркеров молекулярной и клеточной биологии.
9. Структура генетического локуса.
10. Протеомика - лидер науки XXI века.
11. Филогенетический анализа геномов вирусов.
12. Компьютерный анализ генетических текстов.
13. Метаболомика и проблема антибиотикорезистентности.
14. Геномика и медицина.
15. Горизонтальный перенос генов - основной фактор эволюции прокариот.
16. Геногеография и генетический атлас народов мира.

8.5. Примеры тестов для контроля знаний

Текущий контроль проводится преподавателем в форме опросов, тестирования, наблюдения за процессом выполнения задачи на компьютере с оценкой результатов проведенного расчета

1. На конце гена α -гемоглобина имеется следующая последовательность ...ctg gcc cac aag tat cac таа
Какова аминокислотная последовательность соответствующая представленной?
2. Изобразите точечную матрицу сходства для последовательностей из генома вируса карликовости пшеницы ttttcgtgagtgccgaggagctttt против самого себя. Является ли эта последовательность настоящим палиндромом?
3. Нарисуйте все возможные укорененные деревья связывающие три таксона. Сколько таких деревьев можно создать?
4. Мутация, при которой единичная замена основания оставляет аминокислотную последовательность неизменной, называется
 - а) нонсенс-мутация;
 - б) обратная замена;
 - в) «молчащая» мутация;
 - г) мисенс-мутация.
5. Основной постулат (центральная догма) молекулярной биологии
 - а) ДНК \rightarrow РНК \rightarrow белок;
 - б) ДНК \rightarrow РНК \rightarrow белок;
 - в) ДНК \rightarrow РНК \rightarrow белок;
 - г) РНК \rightarrow ДНК \rightarrow белок.
6. Вырожденность генетического кода - это

- а) кодирование одним триплетом только одной аминокислоты;
 - б) кодирование одним триплетом одной либо нескольких аминокислот;
 - в) кодирование одной аминокислоты несколькими триплетами;
 - г) кодирование одним триплетом разных аминокислот.
7. Универсальность генетического кода - это
- а) кодирование одним триплетом одной либо нескольких аминокислот;
 - б) кодирование одной аминокислоты несколькими триплетами;
 - в) кодирование одной аминокислоты одним триплетом;
 - г) наличие единого кода для всех существ на Земле.
8. Для нахождения консервативных регионов в наборе последовательностей применяется преимущественно
- а) множественное выравнивание;
 - б) локальное выравнивание;
 - в) глобальное выравнивание;
 - г) структурное выравнивание.
9. Выравнивание - это:
- а) сравнение последовательностей нуклеотидов с «липкими концами»;
 - б) сравнение аминокислотных последовательностей белков по длине;
 - в) сравнение нуклеотидных последовательностей по длине;
 - г) сравнение последовательностей в поиске идентичных серий символов;
10. Расстояние по Левенштайну или «редакционное расстояние» между двумя строками:
- а) минимальное число «операций редактирования» для того, чтобы превратить одну строку в другую;
 - б) максимальное число «операций редактирования» для того, чтобы превратить одну строку в другую;
 - в) минимальное число замен позиций в строке для того, чтобы превратить одну строку в другую;
 - г) минимальное число вставок для того, чтобы превратить одну строку в другую.
11. PSI-BLAST- это программа, которая
- а) позволяет проводить анализ популяционно-генетических данных;
 - б) осуществляет филогенетический анализ с использованием метода парсимонии;
 - в) подбирает данные для последовательностей, аналогичных запрошенной;
 - г) проводит множественное выравнивание нуклеотидных и аминокислотных последовательностей.
12. В иерархии структур белка, домены располагаются ...
- а) после четвертичной структуры;
 - б) между вторичными и третичными структурами мономера;
 - в) подобный уровень отсутствует;
 - г) между первичными и вторичными структурами мономера.
13. Что из перечисленного не относится к основным типам генетических карт:
- а) генетические карты сцепленности генов;
 - б) «бэндовые» схемы хромосом;
 - в) последовательности ДНК;
 - г) контиг.
14. Контиг - это
- а) набор перекрывающихся фрагментов ДНК, которые в совокупности представляют собой консенсусную область ДНК;
 - б) локусы с варьирующим числом tandemных повторов;
 - в) полиморфизм коротких tandemных повторов;
 - г) короткий, секвенированный участок ДНК, локализованный в строго определенной области генома.

15. Процесс спонтанного сворачивания полипептидной цепи в уникальную *нативную* пространственную структуру называется:
- а) сплайсинг;
 - б) трансляция;
 - в) фолдинг;
 - г) процессинг.
16. Основным инструментом биоинформатики является:
- а) выравнивание последовательностей;
 - б) секвенирование;
 - в) программирование;
 - г) картирование генома.
17. Точечная матрица - это
- а) комплементарная, записанная в обратном порядке последовательность;
 - б) палиндромность последовательности;
 - в) матрицы расчета весов для замен в аминокислотных последовательностях;
 - г) простейшее изображение, которое дает представление о сходстве между двумя последовательностями.
18. Метод исследования вещества путём определения отношения массы к заряду и количества заряженных частиц, образующихся при том или ином процессе воздействия на вещество:
- а) ЯМР-спектроскопия;
 - б) масс-спектрометрия;
 - в) ИК-спектроскопия;
 - г) электрофорез.
19. Для аннотирования структуры белка не используется метод:
- а) метод выявления гомологии в последовательностях;
 - б) метод распознавания фолда;
 - в) экспериментальное определение структуры;
 - г) метод Смита-Ватермана.
20. При каком значении процента идентичных остатков в оптимальном выравнивании два белка будут, вероятно, иметь сходный паттерн фолдинга?
- а) более 25%;
 - б) более 45%;
 - в) 18%;
 - г) 10%.
21. Паралогичные гены (*paralogous genes*) это:
- а) гомологичные гены филогенетически родственных организмов;
 - б) гены, которые произошли в результате внутригеномных дупликаций в геноме данного вида;
 - в) гомологичные гены филогенетически родственных организмов, разошедшихся в процессе видообразования;
 - г) гомологичные гены микроорганизмов, образовавшиеся в процессе горизонтального переноса.
22. Трансмембранные сегменты состоят почти исключительно из гидрофобных аминокислотных остатков. Сколько остатков составляет длину трансмембранных спиралей
- а) 100 остатков;
 - б) 150 остатков;
 - в) 2-5 остатка;
 - г) 15-30 остатков.
23. Какая из перечисленных ниже программ используется для множественного выравнивания последовательностей ДНК и белков

- а) ClustalW;
 - б) BLAST;
 - в) DALI;
 - г) CASP.
24. Модель для оценки эволюционного расстояния по нуклеотидным (либо аминокислотным) заменам в последовательности
- а) модель Тамуры-Нея;
 - б) модель Джукса-Кантора;
 - в) марковская модель;
 - г) скрытая марковская модель.
25. Основная проблема постгеномной эры:
- а) - предсказание первичной структуры белка по последовательности ДНК;
 - б) - предсказание вторичной структуры белка по последовательности ДНК;
 - в) - предсказание третичной структуры белка по последовательности ДНК;
 - г) - предсказание четвертичной структуры белка по последовательности ДНК.
26. Гомологичные нуклеотидные (или аминокислотные) последовательности называют паралоогичными если.
- а) они появились в результате видообразования;
 - б) они появились в результате дупликации;
 - в) они находятся в начале гена;
 - г) они являются уникальными.
27. Филогенетическое дерево (эволюционное дерево, дерево жизни) - дерево, отражающее эволюционные взаимосвязи между различными видами или другими сущностями, имеющими общего предка. Вершины филогенетического дерева делятся на три класса (отметить неверное):
- а) листья;
 - б) стволы;
 - в) узлы;
 - г) корень.
28. Какое из перечисленных ниже выравниваний применяется к «похожим» последовательностям приблизительно одинаковой длины и наглядно показывает разницу между этими последовательностями
- а) локальное;
 - б) множественное;
 - в) глобальное;
 - г) структурное.
29. Выравнивание нуклеотидных или аминокислотных последовательностей с самым высоким весом называют
- а) оптимальным;
 - б) множественным;
 - в) глобальным;
 - г) структурным.
30. Методы предсказания структуры белков по аминокислотной последовательности включают в себя (отметить неверное)
- а) моделирование по гомологии;
 - б) распознавание способа укладки;
 - в) предсказание новых фолдов;
 - г) отсев вырожденных мишеней.

8.6. Вопросы к зачету

1. Биоинформатика и биоинженерия: возникновение, цели, задачи, методы.
2. Базы данных: классификация, основы структур.
3. Базы данных белковых последовательностей.

4. Базы данных последовательностей нуклеиновых кислот.
5. Банки данных метаболических путей.
6. Базы, содержащие результаты глобальных экспериментов по анализу экспрессии, протеомике, и т.п.
7. Основные библиографические базы данных.
8. NCBI, ENTREZ и BLAST - назначение, инструменты, задачи
9. Выравнивание двух последовательностей, точечные матрицы.
10. Глобальное и локальное выравнивание, вес выравнивания, матрицы аминокислотных замен.
11. Локальное выравнивание, задачи, примеры.
12. Алгоритм динамического программирования. Оптимальное выравнивание последовательностей. Способы оптимизации поиска - FASTA, BLAST
13. Зависимость выравнивания от параметров, статистическая значимость выравниваний
14. Множественное выравнивание.
15. Программы для поиска множественного выравнивания. PSI-BLAST.
16. Скрытые Марковские модели (HMM).
17. Выравнивания и филогенетические отношения.
18. Кластеризационный подход к заданию филогенетических отношений
19. Кладистические методы.
20. Третичная структура белка. Фолдинг.
21. Предсказание третичной структуры белка. Моделирование гомологов. Методы, ПО, сервисы.
22. Экспериментальное определение структуры белка. Оценка качества полученной структуры.
23. Предсказание параметров спирали ДНК.
24. Динамические модели РНК.
25. Поиск РНК с заданной структурой (тРНК и т.п., регуляторные участки мРНК).
26. Молекулярный докинг: цель, задачи, подходы, применение.
27. Биоинформатика и филогенез. Молекулярные часы. Клада, OTU, ветвь, лист, корень. Ультраметрическое и неультраметрическое дерево.
28. Ортологи, паралоги, гомологи, ксенологи.
29. Горизонтальный перенос генов и его роль в эволюции геномов.
30. Алгоритмы построения филогенетических деревьев
31. Медицинская геномика, генная диагностика и генотерапия. Фармакоинформатика.
32. Компьютерная токсикология и иммуноинформатика.

8.7. Контроль знаний и компетенций студента.

Бально-рейтинговая система контроля успеваемости студентов в рамках кредитно-модульного обучения по направлению «Методы математического моделирования».

Бально-рейтинговая система основана на подсчете баллов, полученных студентом в течение семестра в строгом соответствии с количеством кредитов, предусмотренных учебным планом. Максимальное количество баллов, которое может получить обучающийся за каждый кредит – 36. Сумма баллов за все кредиты полученная за семестр составит общую итоговую оценку успеваемости. Кроме того, соответствующее количество баллов набранных в течение семестра предусматривает также буквенную индексацию полученных результатов (оценка ECTS, табл.1)

Кредитно- модульная система не допускает наличие, каких либо пропусков учебных занятий в течение всего процесса обучения. Обучающийся допускается к сдаче соответствующего кредита лишь после того, как будет ликвидирована соответствующая задолженность. Отработки задолженностей по уважительным причинам проводятся бесплатно все остальное по соответствующим расценкам, установленным положениями по ИБХТН РУДН.

Критерии оценки:

(в соответствии с действующей нормативной базой)

Соответствие систем оценок (используемых ранее оценок итоговой академической успеваемости, оценок ECTS и балльно-рейтинговой системы (БРС) оценок текущей успеваемости).

Баллы БРС	Традиционные оценки РФ	Оценки ECTS
95 - 100	5	A
86 - 94		B
69 - 85	4	C
61 - 68	3	D
51 - 60		E
31 - 50	2	FX
0 - 30		F
51-100	Зачет	Passed

Пояснение к таблице оценок:

Описание оценок ECTS

A	“Отлично” - теоретическое содержание курса освоено полностью, без пробелов, необходимые практические навыки работы с освоенным материалом сформированы, все предусмотренные программой обучения учебные задания выполнены, качество их выполнения оценено числом баллов, близким к максимальному.
B	“Очень хорошо” - теоретическое содержание курса освоено полностью, без пробелов, необходимые практические навыки работы с освоенным материалом в основном сформированы, все предусмотренные программой обучения учебные задания выполнены, качество выполнения большинства из них оценено числом баллов, близким к максимальному.
C	“Хорошо” - теоретическое содержание курса освоено полностью, без пробелов, некоторые практические навыки работы с освоенным материалом сформированы недостаточно, все предусмотренные программой обучения учебные задания выполнены, качество выполнения ни одного из них не оценено минимальным числом баллов, некоторые виды заданий выполнены с ошибками.
D	“Удовлетворительно” - теоретическое содержание курса освоено частично, но пробелы не носят существенного характера, необходимые практические навыки работы с освоенным материалом в основном сформированы, большинство предусмотренных программой обучения учебных заданий выполнено, некоторые из выполненных заданий, возможно, содержат ошибки.
E	“Посредственно” - теоретическое содержание курса освоено частично, некоторые практические навыки работы не сформированы, многие предусмотренные программой обучения учебные задания не выполнены, либо качество выполнения некоторых из них оценено числом баллов, близким к минимальному.
FX	“Условно неудовлетворительно” - теоретическое содержание курса освоено частично, необходимые практические навыки работы не сформированы, большинство предусмотренных программой обучения учебных заданий не выполнено, либо качество их выполнения оценено числом баллов, близким к минимальному; при дополнительной самостоятельной работе над материалом курса возможно повышение качества выполнения учебных заданий.
F	“Безусловно неудовлетворительно” - теоретическое содержание курса не освоено, необходимые практические навыки работы не сформированы, все выполненные учебные задания содержат грубые ошибки, дополнительная самостоятельная работа над материалом курса не приведет к какому-либо значимому повышению качества выполнения учебных заданий.

Положительными оценками, при получении которых курс засчитывается обучаемому в качестве пройденного, являются оценки А, В, С, D и Е.

Обучаемый, получивший оценку **FX** по дисциплине образовательной программы, обязан после консультации с соответствующим преподавателем в установленные учебной частью сроки успешно выполнить требуемый минимальный объем учебных работ, предусмотренных программой обучения, и представить результаты этих работ этому

преподавателю. Если качество работ будет признано удовлетворительным, то итоговая оценка FX повышается до Е и обучаемый допускается к дальнейшему обучению.

В случае, если качество учебных работ осталось неудовлетворительным, итоговая оценка снижается до F и обучаемый представляется к отчислению. В случае получения оценки F или FX обучаемый представляется к отчислению независимо от того, имеет ли он какие-либо еще задолженности по другим дисциплинам. (Приказ Ректора РУДН №996 от 27.12.2006г.)

Разработчик:

Ассистент ИБХТН

Кезимана П.

Руководитель программы/

Директор ИБХТН

A handwritten signature in blue ink, appearing to read 'Ya. Stanishewskiy', is written over a horizontal line.

Я.М. Станишевский

**ФГАОУ ВО «Российский университет дружбы народов»
Институт биохимической технологии и нанотехнологии (ИБХТН)**

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

ПО УЧЕБНОЙ ДИСЦИПЛИНЕ

Введение в биоинформатику

(наименование дисциплины)

04.04.01 Химия

(код и наименование направления подготовки)

«Биохимические технологии и нанотехнологии»

(наименование профиля подготовки)

Магистр

Квалификация (степень) выпускника

Вопросы для подготовки к зачету

По дисциплине «Введение в биоинформатику»

1. Первый вопрос из:

- Бионформатика: определение, основные понятия, цели и задачи.
- Взаимосвязи бионформатики с другими дисциплинами биологии
- Основы молекулярной биологии
- Методы секвенирования
- Биополимеры – белки, нуклеиновые кислоты.
- Последовательности аминокислот и нуклеотидов как основная информационная составляющая биоинформатики.
- Форматы файлов, используемых в биоинформатике.
- Запись аминокислотных последовательностей.
- Запись нуклеотидных последовательностей.
- База данных NCBI - GenBank.
- Источники биологической информации и базы данных
- Классификация и типы баз данных.
- Всеобъемлющие, универсальные и комбинированные базы данных.
- Базы данных по конкретным организмам.
- Базы данных по типам молекул.
- Проблемы баз данных
- База нуклеотидных последовательностей EMBL.
- База данных по белкам SwissProt.
- База структурной информации о белках PDB – Protein Data Bank.
- Базы данных по малым молекулам и лекарственным препаратам.
- Базы данных по метаболомике и её приложениям.
- Токсикологические базы данных
- Встроенные инструменты для работы с базами данных в Интернете.
- Интернет-инструменты для работы с информацией из биологических
- BLAST
- Выравнивание, локальное, глобальное.
- Множественное выравнивание.
- Филогенетический анализ и визуализации

2. Второй и третий вопрос: решение конкретных задач с помощью методов по:

- Выравнивание

- Филогенетический анализ

Контрольная работа № 1

По дисциплине «Введение в биоинформатику»

1. Биоинформатика: определение, основные понятия, цели и задачи.
2. Взаимосвязи биоинформатики с другими дисциплинами биологии
3. Основы молекулярной биологии
4. Методы секвенирования
5. Биополимеры – белки, нуклеиновые кислоты.
6. Последовательности аминокислот и нуклеотидов как основная информационная составляющая биоинформатики.
7. Форматы файлов, используемых в биоинформатике.
8. Запись аминокислотных последовательностей.
9. Запись нуклеотидных последовательностей.

Контрольная работа № 2

По дисциплине «Введение в биоинформатику»

1. База данных NCBI - GenBank.
 2. Источники биологической информации и базы данных
 3. Классификация и типы баз данных.
 4. Всеобъемлющие, универсальные и комбинированные базы данных.
 5. Базы данных по конкретным организмам.
 6. Базы данных по типам молекул.
 7. Проблемы баз данных
 8. База нуклеотидных последовательностей EMBL.
 9. База данных по белкам SwissProt.
 10. База структурной информации о белках PDB – Protein Data Bank.
 11. Базы данных по малым молекулам и лекарственным препаратам.
 12. Базы данных по метаболомике и её приложениям.
- Токсикологические базы данных

Контрольная работа № 3

По дисциплине «Введение в биоинформатику»

1. Встроенные инструменты для работы с базами данных в Интернете.
2. Интернет-инструменты для работы с информацией из биологических
3. BLAST
4. Выравнивание, локальное, глобальное.
5. Множественное выравнивание.
6. Филогенетический анализ и визуализации

Лабораторные (семинарские) работы

По дисциплине «Введение в биоинформатику»

1. Ознакомление с популярными БД – GeneBank, EMBL, Swiss-Prot, UniProt и другие
2. Поисковая система BLAST, MUSCLE, CLUSTAL и другие
3. Имеется сиквенс некоторого гена. Дана его нуклеотидная последовательность.
Определить, что это за ген и подобрать праймеры, чтобы можно было определить присутствие этого гена в трансформированном организме.
4. Какова нормальная функция белка, кодируемого этим геном? Имеется ли гомолог этого гена в геноме какого-то организма? Если да, какова функция этого гена?
Какая хромосома содержит участок, гомологичный региону?
5. Алгоритмы сборки геномных последовательностей из фрагментов
6. Построение выравнивания двух коротких искусственных последовательностей
7. Построение выравнивания двух реальных последовательностей
8. Редактирование выравнивания
9. Выравнивание последовательностей со схожей функцией (возможных гомологов)
10. Выравнивание последовательностей, содержащих участки гомологии
11. Найти последовательность по фрагменту
12. С помощью программы BLAST проведите поиск по банку данных конкретный ген/протеин
13. Сравнить множественное выравнивание, построенное программой Clustal
14. Составить паттерн по множественному выравниванию
15. Найдите в БД документ, содержащий информацию о белке
16. Филогенетические анализы PhyML

Разработчик:

Ассистент ИБХТН

П. Кезимана

**ФГАОУ ВО «Российский университет дружбы народов»
Институт биохимической технологии и нанотехнологии (ИБХТН)**

САМОСТОЯТЕЛЬНАЯ УЧЕБНАЯ РАБОТА СТУДЕНТА

ПО УЧЕБНОЙ ДИСЦИПЛИНЕ

Введение в биоинформатику

(наименование дисциплины)

ТЕМЫ РЕФЕРАТОВ

1. Структура и функции генов и белков.
2. Открытия и достижения в молекулярной биологии, генетике связаны с возникновением биоинформатика?
3. Характеристики генома организм.
4. Основные задачи биоинформатики
5. Методы распознавания функциональных участков в нуклеотидных последовательностях
6. Методы сравнения первичных структур молекул биополимеров
7. Проблемы филогении геномных последовательностей.
8. Предсказание функций генов.
9. Сравнение геномов
10. Фармакогеномика и фармакогенетика
11. Генетические основы индивидуальной чувствительности к лекарствам.
12. Генетические различия рецепторов лекарств.
13. Клиническое значение фармакодинамических полиморфизмов генов

Разработчик:

Ассистент ИБХТН

П. Кезимана