

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ
ФЕДЕРАЦИИ

Федеральное государственное автономное
образовательное учреждение высшего образования
«Национальный исследовательский Нижегородский государственный университет
им. Н.И. Лобачевского»

Институт информационных технологий, математики и механики

(факультет / институт / филиал)

УТВЕРЖДЕНО

Президиумом ученого совета ННГУ

протокол от

"14" декабря 2021 г. № 4

Рабочая программа дисциплины

Введение в биоинформатику

(наименование дисциплины (модуля))

Уровень высшего образования

Бакалавриат

(бакалавриат / магистратура / специалитет)

Направление подготовки / специальность

02.03.02 Фундаментальная информатика и информационные технологии

(указывается код и наименование направления подготовки / специальности)

Направленность образовательной программы

Инженерия программного обеспечения

(указывается профиль / магистерская программа / специализация)

Форма обучения

очная

(очная / очно-заочная / заочная)

Нижний Новгород

2022 год

1. Место дисциплины в структуре ОПОП

№ варианта	Место дисциплины в учебном плане образовательной программы	Стандартный текст для автоматического заполнения в конструкторе РПД
2	Блок 1. Дисциплины (модули) Часть формируемая участниками образовательных отношений	Дисциплина <i>Б1.В.ДВ.02.02 Введение в биоинформатику</i> относится к части ООП направления подготовки 02.03.02 Фундаментальная информатика и информационные технологии, формируемой участниками образовательных отношений.

2. Планируемые результаты обучения по дисциплине, соотнесенные с планируемыми результатами освоения образовательной программы (компетенциями и индикаторами достижения компетенций)

Формируемые компетенции (код, содержание компетенции)	Планируемые результаты обучения по дисциплине (модулю), в соответствии с индикатором достижения компетенции		Наименование оценочного средства
	Индикатор достижения компетенции* (код, содержание индикатора)	Результаты обучения по дисциплине**	
ПК-3 Способен создавать и исследовать новые математические модели в естественных науках, промышленности и бизнесе, с учетом возможностей современных информационных технологий и программирования и компьютерной техники	ПК-3.1: Знает методы анализа и исследования математических моделей в области фундаментальной информатики и информационных технологий;	Знать базовые алгоритмы вычислительной математики для решения задач нелинейной динамики, условия их применимости Знать возможности современных инструментальных вычислительных средств.	<i>Собеседование,</i>
	ПК-3.2: Умеет определять ключевые свойства и ограничения системы	Уметь определять и профессионально реализовывать необходимые для решения прикладных задач нелинейной динамики вычислительные алгоритмы, анализировать полученные результаты. УМЕТЬ проводить анализ алгоритмических, методических и технологических проблем, возникающих при разработке информационных систем для решения задач	<i>Задача</i>

		биоинформатики; Уметь пользоваться вычислительными методами нелинейной динамики. Уметь пользоваться современными инструментальными вычислительными средствами.	
--	--	--	--

3. Структура и содержание дисциплины

3.1. Трудоемкость дисциплины

	очная форма обучения
Общая трудоемкость	3 ЗЕТ
Часов по учебному плану	108
в том числе	
контактная работа:	33
- занятия лекционного типа	16
- занятия лабораторного типа	16
- текущий контроль (КСР)	1
самостоятельная работа	75
Промежуточная аттестация – зачет	

3.2. Содержание дисциплины

Наименование и краткое содержание разделов и тем дисциплины	Всего (часы)	В ТОМ ЧИСЛЕ				Самостоятельная работа обучающегося, часы
		Контактная работа (работа во взаимодействии с преподавателем), часы. Из них			Всего контактных часов	
		Занятия лекционного типа	Занятия семинарского типа	Занятия лабораторного типа		
Биоинформатика: определение, цели и задачи.	21	4	2		6	15
Подходы и области применения биоинформатики. Биологические базы данных.	21	2	4		6	15
Общая характеристика живой клетки: типы биомолекул, основные компоненты, функции.	21	4	2		6	15

Компьютерные инструменты анализа биомолекул.	21	2	4		6	15
Молекулярная эволюция: вычислительные подходы	23	4	4		8	15
Текущий контроль (КСР)	1				1	
Промежуточная аттестация –зачет						
Итого	108	16	16		33	75

Текущий контроль успеваемости реализуется в формах опросов на занятиях лекционного типа.

Промежуточная аттестация проходит в традиционной форме (зачет).

4. Учебно-методическое обеспечение самостоятельной работы обучающихся

В процессе изучения дисциплины используются образовательные технологии в форме лекций, при выполнении практических работ, при самостоятельной работе и подготовке к зачету студенты имеют доступ к материалам курса, размещенным по адресу <http://www.itmm.unn.ru/pm/education/information/>, режим доступа - свободный.

Выполнение лабораторных работ на следующие темы:

1. Биоинформатика, определение и задачи. Мультидисциплинарность, фундаментальные и прикладные аспекты.
2. Мишени новых лекарств: поиск и валидация.
3. Парное выравнивание последовательностей. Алгоритмы Смита-Уотермана и Нидлмана-Вунша (локальное и глобальное выравнивание). Матрицы замен.
4. Структурная организация белка, база данных PDB.
5. Теория распознавания образов, применение в биоинформатике. Обучение с учителем.
6. Множественное выравнивание последовательностей. Матрицы замен.
7. Предсказание вторичной структуры РНК и ДНК.
8. Анализ филогений нуклеотидных и аминокислотных последовательностей, методы построения филогенетических деревьев. Понятия SNP и SAP.
9. Алгоритм BLAST. Модификации и области применения.
10. Первичная аннотация геномов. Поиск кодирующих фрагментов.
11. Предсказание пространственной структуры белка.
12. Методы сравнительной геномики, предсказание функций и регуляции генов.
13. Молекулярный докинг. Принципы и области применения.
14. Протеомные технологии. Принципы, области применения возможности и перспективы.
15. Понятие марковского процесса. Марковская цепь. Применение в биоинформатике.
16. Построение пространственных моделей белков и низкомолекулярных соединений.
17. Масс-спектрометрия. Принципы работы, типы масс-спектрометров. Идентификация пептидов и белков.
18. Динамическое программирование. Применение в задачах выравнивания последовательностей.
19. Структурно-функциональные мотивы как молекулярная основа сходства и различия белков
20. Базы данных нуклеотидных и аминокислотных последовательностей. Форматы представления данных.

22. Проект "Геном человека".
23. Компьютерное конструирование лекарств на основе структуры белка-мишени.
24. Компьютерное конструирование лекарств на основе структуры лигандов. База данных PubChem.
27. Белок-белковые взаимодействия, способы регистрации.
28. Регуляция экспрессии генов в прокариотах и эукариотов.

Образовательные материалы для самостоятельной работы студентов

1. Курс «Алгоритмы биоинформатики» Н.И.Вяххи:
<https://www.lektorium.tv/course/22933>
2. Алексеев В.Е., Таланов В.А. Графы. Модели вычислений. Структуры данных: Учебник. – Нижний Новгород: Изд-во ННГУ, 2005. 307 с. (2 экз)
3. Johnes N., Pevzner P., An introduction to bioinformatics algorithms. – MIT Press, 2004. – 454 p.
4. Гергель В.П. Высокопроизводительные вычисления для многоядерных многопроцессорных систем. Учебное пособие – Нижний Новгород; Изд-во ННГУ им. Н.И.Лобачевского, 2010. (47 экз)

Контрольные вопросы и задания для проведения текущего контроля и промежуточной аттестации по итогам освоения дисциплины приведены в п. 5.2.

5. Фонд оценочных средств для промежуточной аттестации по дисциплине (модулю), включающий:

5.1. Описание шкал оценивания результатов обучения по дисциплине

Уровень сформированности компетенций (индикатора достижения компетенций)	Шкала оценивания сформированности компетенций						
	плохо	неудовлетворительно	удовлетворительно	хорошо	очень хорошо	отлично	превосходно
	Не зачтено		Зачтено				
<u>Знания</u>	Отсутствие знаний теоретического материала. Невозможность оценить полноту знаний вследствие отказа обучающегося от ответа	Уровень знаний ниже минимальных требований. Имели место грубые ошибки.	Минимально допустимый уровень знаний. Допущено много негрубых ошибок.	Уровень знаний в объеме, соответствующем программе подготовки. Допущено несколько негрубых ошибок	Уровень знаний в объеме, соответствующем программе подготовки. Допущено несколько негрубых ошибок	Уровень знаний в объеме, соответствующем программе подготовки, без ошибок.	Уровень знаний в объеме, превышающем программу подготовки.
<u>Умения</u>	Отсутствие минимальных умений.	При решении стандартных задач не	Продemonстрированы основные	Продemonстрированы все основные	Продemonстрированы все основные	Продemonстрированы все основные	Продemonстрированы

	Невозможность оценить наличие умений вследствие отказа обучающегося от ответа	продемонстрированы основные умения. Имели место грубые ошибки.	умения. Решены типовые задачи с негрубыми ошибками. Выполнены все задания, но не в полном объеме.	умения. Решены все основные задачи с негрубыми ошибками. Выполнены все задания, в полном объеме, но некоторые с недочетами.	умения. Решены все основные задачи. Выполнены все задания, в полном объеме, но некоторые с недочетами.	умения, решены все основные задачи с отдельными несущественными недочетами, выполнены все задания в полном объеме.	аны все основные умения, решены все основные задачи и. Выполнены все задания, в полном объеме без недочетов
<u>Навыки</u>	Отсутствие владения материалом. Невозможность оценить наличие навыков вследствие отказа обучающегося от ответа	При решении стандартных задач не продемонстрированы базовые навыки. Имели место грубые ошибки.	Имеется минимальный набор навыков для решения стандартных задач с некоторыми недочетами.	Продемонстрированы базовые навыки при решении стандартных задач с некоторыми недочетами	Продемонстрированы базовые навыки при решении стандартных задач без ошибок и недочетов.	Продемонстрированы навыки при решении нестандартных задач без ошибок и недочетов.	Продемонстрированы творческий подход к решению нестандартных задач.

Шкала оценки при промежуточной аттестации

Оценка		Уровень подготовки
зачтено	Превосходно	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «превосходно»
	Отлично	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «отлично», при этом хотя бы одна компетенция сформирована на уровне «отлично»
	Очень хорошо	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «очень хорошо», при этом хотя бы одна компетенция сформирована на уровне «очень хорошо»
	Хорошо	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «хорошо», при этом хотя бы одна компетенция сформирована на уровне «хорошо»

	Удовлетворительно	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «удовлетворительно», при этом хотя бы одна компетенция сформирована на уровне «удовлетворительно»
не зачтено	Неудовлетворительно	Хотя бы одна компетенция сформирована на уровне «неудовлетворительно», ни одна из компетенций не сформирована на уровне «плохо»
	Плохо	Хотя бы одна компетенция сформирована на уровне «плохо»

5.2. Типовые контрольные задания или иные материалы, необходимые для оценки результатов обучения

5.2.1 Контрольные вопросы

Вопрос	Код компетенции (согласно РПД)
1. Дайте определение биоинформатики. Цели и задачи биоинформатики. Приведите пример подходов и областей применения.	ПК-3
2. Какова общая характеристика живой клетки. Какие типы биомолекул присутствуют в клетке. Перечислите основные компоненты клетки.	ПК-3
3. Опишите строение белков и основные свойства белков.	ПК-3
4. Какие компьютерные инструменты анализа свойств белков Вы знаете.	ПК-3
5. Перечислите форматы записи первичной структуры белков.	ПК-3
6. Приведите примеры баз данных белков.	ПК-3
7. Перечислите основы структур баз данных (записи, поля, объекты), классификации баз по способу заполнения (автоматические, архивные, курируемые). Дайте характеристику основных баз данных: GenBank, EMBL, SwissProt, TrEMBL, PIR, PDB.	ПК-3
8. Приведите пример баз данных, содержащие результаты глобальных экспериментов по анализу экспрессии, протеомике, и т.п.	ПК-3
9. Какие программы расчета физико-химических свойств белков Вы знаете.	ПК-3
10. Домены: строение и свойства.	ПК-3

11. Перечислите основные требования к банкам данных белковых семейств (SCOP, Prosite, ProDom, PFAM, InterPro).	ПК-3
12. Пространственная структура белков – методы и программные средства для анализа.	ПК-3
13. Методы получения трехмерных структур белков.	ПК-3
14. Инструменты визуализации 3D структур белков.	ПК-3
15. Дайте основные характеристики базы данных 3D структур белков PDB.	ПК-3
16. Перечислите основные функции белков.	ПК-3
17. Как происходит аннотирование функций в базах данных.	ПК-3
18. База данных GO (Gene Ontology).	ПК-3
19. Сравнение последовательностей.	ПК-3
20. Выравнивание двух последовательностей.	ПК-3
21. Глобальное и локальное выравнивание, вес выравнивания, матрицы аминокислотных замен.	ПК-3
22. Дот-матрицы.	ПК-3
23. Глобальное выравнивание: алгоритм Нидельмана-Вунша.	ПК-3
24. Локальное выравнивание: алгоритм Смита-Ватермана.	ПК-3
25. Другие алгоритмы локального выравнивания.	ПК-3
26. Статистическая значимость выравниваний и ее зависимость от вероятностной модели последовательности (в т.ч. сегменты малой сложности)	ПК-3
27. Зависимость выравнивания от параметров.	ПК-3
28. Множественное выравнивание.	ПК-3
29. Динамическое программирование.	ПК-3
30. Последовательное выравнивание (Clustal).	ПК-3
31. Поиск по сходству в базах данных.	ПК-3
32. Алгоритм Smith-Waterman.	ПК-3
33. Алгоритм BLAST.	ПК-3

34. Алгоритм FASTA.	ПК-3
35. Статистическая оценка значимости (E-value, P-value).	ПК-3
36. Марковские модели (HMM, <u>H</u> idden <u>M</u> arkov <u>m</u> odels).	ПК-3
37. Позиция-специфичные матрицы расчета (Position-specific scoring matrices, PSSM). Паттерны, Sequence logos, Psi-BLAST.	ПК-3
38. Анализ структурных особенностей белков.	ПК-3
39. Моделирование белков.	ПК-3
40. Предсказание структур белков разных уровней организации.	ПК-3
41. Программное обеспечение для компьютерного анализа структур белков.	ПК-3
42. Белок-белковые взаимодействия.	ПК-3
43. Базы данных белок-белковых взаимодействий (IntAct, BIND).	ПК-3
44. Программы анализа белковых комплексов.	ПК-3
45. Компьютерный расчет областей белок-белкового контакта.	ПК-3
46. Компьютерный расчет поверхности белка, доступной для растворителя.	ПК-3
47. Нуклеиновые кислоты. Строение и свойства. Структура ДНК.	ПК-3
48. Генетический код и его свойства.	ПК-3
49. Базы данных нуклеотидных последовательностей (TrEMBL, EMBL).	ПК-3
50. Репликация ДНК и биосинтез белка.	ПК-3
51. Компьютерное предсказание белков на основе данных о нуклеотидной последовательности.	ПК-3
52. РНК. Транскриптомика.	ПК-3
53. Молекулярная эволюция.	ПК-3
54. Эволюция молекул и организмов.	ПК-3

55. Ортологи и паралоги.	ПК-3
56. Горизонтальный перенос.	ПК-3
57. Предсказание структур белков разных уровней организации.	ПК-3
58. Программное обеспечение для компьютерного анализа структур белков.	ПК-3
59. Белок-белковые взаимодействия.	ПК-3
60. Базы данных белок-белковых взаимодействий (IntAct, BIND).	ПК-3
61. Программы анализа белковых комплексов.	ПК-3
62. Компьютерный расчет областей белок-белкового контакта.	ПК-3
63. Компьютерный расчет поверхности белка, доступной для растворителя.	ПК-3
64. Нуклеиновые кислоты. Строение и свойства. Структура ДНК.	ПК-3
65. Генетический код и его свойства.	ПК-3
66. Базы данных нуклеотидных последовательностей (TrEMBL, EMBL).	ПК-3
67. Репликация ДНК и биосинтез белка.	ПК-3
68. Компьютерное предсказание белков на основе данных о нуклеотидной последовательности.	ПК-3

5.2.2. Типовые задания для оценки сформированности компетенции ПК-3

1. Какие компьютерные инструменты анализа свойств белков Вы знаете.
2. Перечислите форматы записи первичной структуры белков.
3. Приведите примеры баз данных белков.
4. Перечислите основы структур баз данных (записи, поля, объекты), классификации баз по способу заполнения (автоматические, архивные, курируемые). Дайте характеристику основных баз данных: GenBank, EMBL, SwissProt, TrEMBL, PIR, PDB.

6. Учебно-методическое и информационное обеспечение дисциплины

Рекомендуемая литература.

а) основная литература:

1. Порозов, Ю.Б. Биоинформатика. [Электронный ресурс] — Электрон. дан. — СПб. : НИУ ИТМО, 2012. — 52 с. —
Режим доступа: <http://e.lanbook.com/book/43567#authors>.
2. Асанов, М.О. Дискретная математика: графы, матроиды, алгоритмы. [Электронный ресурс] / М.О. Асанов, В.А. Баранский, В.В. Расин. — Электрон. дан. — СПб. : Лань, 2010. — 368 с. —
Режим доступа: https://e.lanbook.com/book/536?category_pk=914#authors

б) дополнительная литература:

1. Гергель В.П. Высокопроизводительные вычисления для многоядерных многопроцессорных систем. М., 2010. (47 экз)

в) программное обеспечение и Интернет-ресурсы

1. Курс «Алгоритмы биоинформатики» Н.И.Вяххи:
<https://www.lektorium.tv/course/22933>
2. <http://www.fbb.msu.ru/> - факультет биоинженерии и биоинформатики МГУ
3. <http://bioinf.spbau.ru/about> - лаборатория Алгоритмической биологии, Санкт-Петербургский академический университет РАН.
4. http://www.bionet.nsc.ru/kib/?page_id=28 - кафедра информационной биологии, НГУ
5. <http://biodynamics.ucsd.edu/> - лаборатория Биодинамики Университета Калифорнии в Сан Диего.
6. <http://bioinformatics.ru/> - российский портал по биоинформатике

7. Материально-техническое обеспечение дисциплины

Помещения представляют собой учебные аудитории для проведения учебных занятий, предусмотренных программой, оснащенные оборудованием и техническими средствами обучения.

Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечены доступом в электронную информационно-образовательную среду

Программа составлена в соответствии с требованиями ОС ННГУ 02.03.02 Фундаментальная информатика и информационные технологии.

Авторы:

в.н.с., проф. каф. прикладной

математики, д.ф.-м.н.

_____ Иванченко М.В.

Рецензент (ы) _____

Зав. каф. прикладной математики

_____ Иванченко М.В.