

Федеральное государственное бюджетное образовательное учреждение высшего  
профессионального образования  
Московский государственный университет имени М.В. Ломоносова  
*Факультет биоинженерии и биоинформатики*

**УТВЕРЖДАЮ**

Декан  
факультета биоинженерии  
и биоинформатики,  
академик

\_\_\_\_\_/В.П. Скулачев /

« \_\_\_\_ » \_\_\_\_\_ 20\_\_ г.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**

**Наименование дисциплины:**

**Биоинформатика**

**Уровень высшего образования:**  
**специалитет**

**Направление подготовки (специальность):**

**06.05.01 Биоинженерия и биоинформатика**

**Форма обучения:**

**очная**

Рабочая программа рассмотрена и одобрена

*Ученым советом факультета*  
(протокол № \_\_\_\_\_, \_\_\_\_\_)

Москва 20\_\_

Рабочая программа дисциплины разработана в соответствии с самостоятельно установленным МГУ образовательным стандартом (ОС МГУ) для реализуемых основных профессиональных образовательных программ высшего образования по специальности 06.05.01 «Биоинженерия и биоинформатика» (программы специалитета) в редакции приказа МГУ от 30 декабря 2016 г.

Год (годы) приема на обучение – 2016, 2017, 2018, 2019.

© Факультет биоинженерии и биоинформатики МГУ имени М.В. Ломоносова

*Программа не может быть использована другими подразделениями университета и другими вузами без разрешения факультета.*

## **Цель и задачи дисциплины**

**Целью** курса является освоение основных концепций и методов анализа биологических данных.

### **Задачи курса:**

- дать основные представления о подходах, методах и алгоритмах анализа биологических последовательностей
- дать основные представления об анализе пространственных структур биополимеров
- дать основные представления о системной биологии
- дать основные представления о современной геномике

**1.** Место дисциплины в структуре ОПОП ВО: базовая часть, общепрофессиональный цикл, курс II – семестр 4, курс III – семестр 6, курс IV – семестр 7,8.

**2.** Входные требования для освоения дисциплины, предварительные условия (если есть): освоение дисциплины «Практическая биоинформатика», «Теория вероятностей», «Информатика», «Основы молекулярной биологии».

**3.** Планируемые результаты обучения по дисциплине:

Знать:

- основные методы анализа биологических последовательностей

Уметь:

- применять свои знания для анализа последовательностей

Владеть:

- навыками критически читать современную научную литературу по теме курса

Иметь опыт:

- применения основных методов к анализу последовательностей

**4.** Формат обучения: лекционные и семинарские занятия.

**5.** Объем дисциплины составляет 12 з.е., в том числе 184 академических часов, отведенных на контактную работу обучающихся с преподавателем, 248 академических часов на самостоятельную работу обучающихся.

### **6. Краткое содержание дисциплины (аннотация):**

Биоинформатика является одной из базовых дисциплин факультета. Биоинформатика посвящена компьютерному анализу биологических данных, в частности анализу последовательностей, в том числе геномов, анализу пост-геномных данных, анализу структур биополимеров.

Под биоинформатикой обычно понимают использование компьютеров для решения биологических задач. В настоящее время это почти исключило задачи молекулярной биологии. Причина этого в том, что за последние 20–25 лет накоплен поистине колоссальный экспериментальный материал именно о строении и функционировании биологических молекул (белков и нуклеиновых кислот), в качестве примера достаточно привести геном человека. Этот материал требует развитых компьютерных методов для своего анализа. Поэтому биоинформатика как на нашем факультете, так и в подавляющем большинстве мировых научных центров понимается как синоним вычислительной молекулярной биологии. Есть несколько основных направлений этого раздела науки, в зависимости от исследуемых объектов:

- Биоинформатика последовательностей.
- Структурная биоинформатика.
- Компьютерная геномика

С другой стороны, биоинформатику можно условно разделить на несколько направлений в зависимости от типа решаемых задач.

Наиболее известной и наиболее эффективной областью применения биоинформатики в настоящее время является анализ геномов, тесно связанный с анализом последовательностей.

Наименование и краткое содержание разделов и тем дисциплины,  Форма промежуточной аттестации по дисциплине	Всего (часы)	В том числе			
		Контактная работа (работа во взаимодействии с преподавателем) Виды контактной работы, часы			Самостоятельная работа обучающегося, часы (виды самостоятельной работы – эссе, реферат, контрольная работа и пр. – указываются при необходимости)
		Занятия лекционного типа	Занятия семинарского типа	Всего	
<b>2 КУРС (4 СЕМЕСТР) – КУРСОВАЯ РАБОТА</b>					
Курсовая работа	30				30
Промежуточная аттестация - зачет	6				6(количество часов, отведенных на промежуточную аттестацию)
Итого	36				36
<b>Модуль 1. Биоинформатика последовательностей. 3 КУРС (6 СЕМЕСТР)</b>					
Алгоритмы выравнивания	28	8	6	14	14
Теоретико-вероятностные аспекты биоинформатики.	24	8	4	12	12
Скрытые марковские модели.	24	8	4	12	12
Предсказание вторичной структуры РНК	24	4	6	10	14
Текущая литература по алгоритмам работы с последовательностями («журнальный клуб»).	40	4	12	16	24
Промежуточная аттестация - зачет	4				4(количество часов, отведенных на промежуточную аттестацию)
Итого	144	32	32		80
<b>Модуль 2. Методы расшифровки пространственных структур макромолекул. 4 КУРС (7 СЕМЕСТР)</b>					
Рентгеноструктурный анализ.	22	6	4	10	12
Показатели качества рентгеноструктурной	14	4	2	6	8

модели.					
Понятие о ядерном магнитном резонансе (ЯМР) и его применении в анализе структур макромолекул.	10	2	2	4	6
Модуль 3. Алгоритмы работы с данными о пространственных структурах макромолекул. 4 КУРС (7 СЕМЕСТР)					
Алгоритмы выделения вторичной структуры белков.	14	4	2	6	8
Алгоритмы выделения и классификации доменов	14	4	2	6	8
Алгоритмы пространственного совмещения и выравнивания.	14	4	2	6	8
Алгоритмы нахождения гидрофобного ядра и поверхности макромолекул.	14	4	2	6	8
Текущая литература по алгоритмам работы с пространственными структурами макромолекул («журнальный клуб»).	38	0	12	12	26
Промежуточная аттестация - экзамен	4				4 (количество часов, отведенных на промежуточную аттестацию)
Итого	144	28	28		88
Модуль 4. Функциональная аннотация. 4 КУРС (8 СЕМЕСТР)					
Gene Ontology.	6	2	2	4	2
Аннотация по гомологии. Типичные ошибки.	6	2	2	4	2
Сравнительно-геномные методы.	6	2	2	4	2
Зачетное задание: функциональная аннотация 10 генов прокариот.	8	2	2	4	4
Зачетное задание: анализ 2 бактериальных регулонов.	8	2	2	4	4
Модуль 5. Системная биология. 4 КУРС (8 СЕМЕСТР)					
Транскриптомика.	6	2	2	4	2
Белок-ДНКовые взаимодействия.	6	2	2	4	2
Эпигеномика.	6	2	2	4	2
Белок-белковые	6	2	2	4	2

взаимодействия.					
Использование негеномных данных для функциональной аннотации.	6	2	2	4	2
Зачетное задание: Статья в Википедии по методам нового поколения.	8	2	2	4	4
Зачетное задание: анализ статьи.	8	2	2	4	4
Сети и графы.	6	2	2	4	2
Модели (потокосые и кинетические).	6	2	2	4	2
Модуль 6. Эволюция геномов. 4 КУРС (8 СЕМЕСТР)					
Геномные перестройки.	6	2	2	4	2
Полногеномные дубликации.	6	2	2	4	2
Промежуточная аттестация - экзамен	4				4 (количество часов, отведенных на промежуточную аттестацию)
Итого	108	32	32		44

7. Фонд оценочных средств (ФОС) для оценивания результатов обучения по дисциплине

7.1. Типовые контрольные задания или иные материалы для проведения текущего контроля успеваемости.

Проанализировать научную статью

7.2. Типовые контрольные задания или иные материалы для проведения промежуточной аттестации

1. Геномный полипротеин полиовируса состоит из 2209 аминокислотных остатков. При поиске гомологов этого полипротеина программой BLASTP в банке данных, состоящей из белков семейства Caliciviridae, получено выравнивание, включающее остатки 1243–1368 полипротеина полиовируса, при этом BLAST сообщает следующие характеристики выравнивания:

“Score 61.2 bits(147) Expect 2e-09”

Каковы будут характеристики того же выравнивания, если на поиск по тому же банку отправить не весь полипротеин, а его фрагмент, состоящий из остатков с 1200 по 1420 включительно?

2. Предположим, что в результате поиска программой BLAST найдено выравнивание, не содержащее гэпов. Если повторить поиск, увеличив вдвое штрафы за открытие и удлинение гэпа и не меня остальных параметров, то как изменятся вес, вес в битах и E-value того же выравнивания?

3. По данному дереву определите веса объектов A–D методом Гернштейна – Сонхаммера – Чотья:

4. В следующих вопросах  $d$  — штраф за гэп,  $r$  — штраф за несовпадение букв; вес совпадения букв всюду равен 1. Требуется обвести номера всех правильных ответов.

1. Наибольшее общее слово — это оптимальное выравнивание:

1.1. Глобальное при  $d = 0$ ,  $r = 0$

1.2. Локальное при  $d = 0$ ,  $r = 0$

1.3. Глобальное при  $d = \infty$ ,  $r = 0$

- 1.4.Локальное при  $d = \infty, r = 0$
- 1.5.Глобальное при  $d = 0, r = \infty$
- 1.6.Локальное при  $d = 0, r = \infty$
- 1.7.Глобальное при  $d = \infty, r = \infty$
- 1.8.Локальное при  $d = \infty, r = \infty$
- 2.Наибольшая общая подпоследовательность — это оптимальное выравнивание:
  - 2.1.Глобальное при  $d = 0, r = 0$
  - 2.2.Локальное при  $d = 0, r = 0$
  - 2.3.Глобальное при  $d = \infty, r = 0$
  - 2.4.Локальное при  $d = \infty, r = 0$
  - 2.5.Глобальное при  $d = 0, r = \infty$
  - 2.6.Локальное при  $d = 0, r = \infty$
  - 2.7.Глобальное при  $d = \infty, r = \infty$
  - 2.8.Локальное при  $d = \infty, r = \infty$

**Шкала и критерии оценивания результатов обучения по дисциплине.**

Результаты обучения	«Неудовлетворительно»	«Удовлетворительно»	«Хорошо»	«Отлично»
Знания: основные методы анализа биологических последовательностей	Знания отсутствуют	Фрагментарные знания	Общие, но не структурированные знания	Сформированные систематические знания
Умения: применять свои знания для анализа последовательностей	Умения отсутствуют	В целом успешное, но не систематическое умение	В целом успешное, но содержащее отдельные пробелы умение (допускает неточности не принципиального характера)	Успешное и систематическое умение
Владения: навыками критически читать современную научную литературу по теме курса	Навыки владения критически читать современную научную литературу по теме курса отсутствуют	Наличие отдельных навыков (наличие фрагментарного опыта)	В целом, сформированные навыки (владения), но используемые не в активной форме	Сформированные навыки (владения), применяемые при решении задач

8. Ресурсное обеспечение:

- Перечень основной и дополнительной литературы
- Основная литература:
  1. Дурбин и др. Анализ биологических последовательностей. - Москва: НИЦ "Регулярная и хаотическая динамика", 2006
  2. Бородовский, Екишева. Задачи и решения по анализу биологических последовательностей. - Москва: НИЦ "Регулярная и хаотическая динамика", 2008
  3. Леск. Введение в биоинформатику. - Москва: Бином. Лаборатория знаний, 2013

4. Бландел, Джонсон. Кристаллография белка. - Москва: Мир, 1979
  5. Eugene V Koonin and Michael Y Galperin. Sequence - Evolution – Function. Computational Approaches in Comparative Genomics. - Boston: Kluwer Academic, 2003
  6. Ж. Сетубал, Ж. Мейданис. Введение в вычислительную молекулярную биологию. - Ижевск: Регулярная и хаотическая динамика., 2007
  7. С. Игнасимуту. Основы биоинформатики. - Ижевск: Регулярная и хаотическая динамика., 2007
  8. М. Бородовский, С. Екишева. Задачи и решения по анализу биологических последовательностей. - Ижевск: Регулярная и хаотическая динамика., 2008
  9. Бернхард Хаубольд, Томас Вие. Введение в вычислительную биологию. - Ижевск: Регулярная и хаотическая динамика., 2011
  10. Jonathan Pevsner. Bioinformatics and Functional Genomics. - Hoboken: Wiley-Blackwell, 2009
- Перечень лицензионного программного обеспечения (при необходимости)
  - Перечень профессиональных баз данных и информационных справочных систем
  - Перечень ресурсов информационно-телекоммуникационной сети «Интернет» (при необходимости): [www.ncbi.nlm.nih.gov/pubmed](http://www.ncbi.nlm.nih.gov/pubmed)
  - Описание материально-технического обеспечения: аудитория, проектор, компьютерный класс