

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:

Практическая биоинформатика (эволюция)

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

Цели дисциплины — обучение практическим навыкам работы за компьютером с семействами белков и соответствующими базами данных.

Задачи дисциплины:

- освоение понятия филогенетического дерева;
- освоение основных программ реконструкции филогении по множественному выравниванию биологических последовательностей;
- выработка навыков работы с основными базами данных семейств белков и их аннотаций;
- выработка навыков работы с множественным выравниванием семейства белков.

1. Место дисциплины в структуре ОПОП ВО – базовая часть, общепрофессиональный цикл, курс II, семестр 4.

2. Входные требования для освоения дисциплины, предварительные условия (если есть): освоение дисциплин «Практическая биоинформатика» и «Химические основы биологических процессов»

3. Планируемые результаты обучения по дисциплине:

Знать: биологический смысл филогенетического дерева гомологичных биологических последовательностей; способы укоренения филогенетического дерева; основные алгоритмы и программы, применяемые для реконструкции молекулярной филогении; что такое информационное содержание выравнивания; что такое позиционная весовая матрица и зачем она нужна; основные сигналы, закодированные в геномах; что такое эволюционный домен и доменная архитектура белка; что такое НММ-профиль домена и его отличие от позиционной весовой матрицы; основные статистические показатели качества решающего алгоритма при проверке на эталонных данных; какая информация хранится в базах данных KEGG и GO; чем отличаются трансмембранные белки от цитоплазматических и секретируемых.

Уметь: реконструировать филогенетическое дерево для данной совокупности последовательностей родственных белков или фрагментов ДНК; укоренять неукорененное филогенетическое дерево; создавать и использовать позиционную весовую матрицу с помощью сервиса MEME; рассчитывать информационное содержание выравнивания; использовать пакет HMMER для работы с профилями семейств белков; пользоваться базами данных KEGG, GO, Pfam, OPM.

Владеть: программами и программными пакетами: Mega, Jalview, Psi-BLAST, MEME, HMMER, Phylip; сервисами баз данных: Pfam, PROSITE, KEGG, GO, OPM.

Иметь опыт: построения филогенетических деревьев по выравниваниям биологических последовательностей; интерпретации филогенетических деревьев, включая нахождение паралогов и ортологов; построения позиционной весовой матрицы и ее использования для нахождения сигналов в геномах; построения профилей семейства белков или доменов белков и применения профилей для определения доменной архитектуры белка; исследования метаболического пути с помощью сервисов базы данных KEGG; определения терминов GO для данной совокупности белков; анализа белков на наличие трансмембранных участков.

4. Формат обучения – практические занятия.

5. Объем дисциплины составляет 2 з.е., в том числе 32 академических часов, отведенных на контактную работу обучающихся с преподавателем, 40 академических часов на самостоятельную работу обучающихся.

6. Краткое содержание дисциплины (аннотация):

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	Практические занятия	Всего	
ФИЛОГЕНЕТИЧЕСКИЕ ДЕРЕВЬЯ					
Понятие филогенетического дерева.	6	2	0	2	0
Алгоритмы и программы реконструкции филогении по множественному выравниванию.	6	1	1	2	4
Укоренение и бутстреп	6	1	1	2	4
Анализ деревьев, содержащих паралоги	6	1	1	2	4
Коллоквиум	4		4	4	0
СИГНАЛЫ В НУКЛЕОТИДНЫХ ПОСЛЕДОВАТЕЛЬНОСТЯХ					
Сигналы и мотивы. Информационное содержание выравнивания.	6	1	1	2	4
Позиционная весовая матрица мотива (PWM). Алгоритмы построения PWM и поиска мотивов. Пакет MEME.	6	1	1	2	4
СЕМЕЙСТВА БЕЛКОВЫХ ДОМЕНОВ					
Домены белков, доменная архитектура. База данных Pfam. Профиль домена. Пакет HMMER. Паттерн, база данных ProSite.	6	2	0	2	0
Анализ результатов поиска по профилю. ROC кривая	6	1	1	2	4
Коллоквиум	4		4	4	0
ФУНКЦИИ БЕЛКОВ					

Метаболические пути. «Киотская энциклопедия генов и геномов» (KEGG). Классификация ферментов (EC)	6	1	1	2	4
Функциональная аннотация генов. «Онтология генов» (GO)	6	1	1	2	4
Трансмембранные белки.	6	1	1	2	4
Выполнение зачётных заданий	2	0	2	2	0
Промежуточная аттестация — ЗАЧЁТ Студенты показывают не выполненные ранее работы или исправление ошибок в выполненных самостоятельных работах.					4
Итого	72	32		40	

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

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

См. на сайте kodomo.fbb.msu.ru/wiki

Аттестация проводится почти исключительно в форме проверки самостоятельных заданий. Приняты также мини-контрольные, предлагаемые перед занятием, на выполнение которых отводится не более 10 минут. Примерные вопросы:

1. По первому блоку. А. Для данного неукоренённого дерева нарисуйте укоренение в заданную ветвь и напишите соответствующую скобочную формулу в формате Newick. Б. Для двух данных деревьев выпишите ветви, которыми они различаются.

2. По второму блоку. А. Посчитайте информационное содержание данного выравнивания коротких нуклеотидных последовательностей. Б. Оцените число находок мотива в случайной нуклеотидной последовательности длины 1 млн пар нуклеотидов, если информационное содержание мотива равно 10. В. Как рассчитать ожидаемое число случайных встреч динуклеотида CG в геноме человека?

3. По третьему блоку. А. Чем отличаются входные данные для построения НММ-профиля от входных данных для построения PWM? Б. Что является единицей хранения в БД Pfam?

4. По четвертому блоку. А. Для чего в клетке используется аденозинтрифосфат (АТФ)? Б. Какие клеточные функции выполняют мембранные белки?

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

Каждый блок завершается коллоквиумом. Вопросы коллоквиумов публично доступны на веб-ресурсе kodomo.fbb.msu.ru: <http://kodomo.fbb.msu.ru/wiki/2017/4/colloquium1> , <https://kodomo.fbb.msu.ru/wiki/2017/4/colloq3>, <http://kodomo.cmm.msu.ru/wiki/2017/4/colloquium> .

Шкала и критерии оценивания результатов обучения по дисциплине.

Результаты обучения	«Неудовлетворительно»	«Удовлетворительно»	«Хорошо»	«Отлично»
Знания: <i>биологический смысл</i>	Знания отсутствуют	Фрагментарные знания	Общие, но не структурированные знания	Сформированные систематически

<p>филогенетического дерева гомологичных биологических последовательностей; способы укоренения филогенетического дерева; основные алгоритмы и программы, применяемые для реконструкции молекулярной филогении; что такое информационное содержание выравнивания; что такое позиционная весовая матрица и зачем она нужна; основные сигналы, закодированные в геномах; что такое эволюционный домен и доменная архитектура белка; что такое НММ-профиль домена и его отличие от позиционной весовой матрицы; основные статистические показатели качества решающего алгоритма при проверке на эталонных данных; какая информация хранится в базах данных KEGG и GO; чем отличаются трансмембранные белки от цитоплазматических и секреторируемых</p>				е знания
<p>Умения: реконструировать</p>	<p>Умения отсутствуют</p>	<p>В целом успешное, но не</p>	<p>В целом успешное, но</p>	<p>Успешное и систематическое</p>

<p>филогенетическое дерево для данной совокупности последовательностей ей родственных белков или фрагментов ДНК; укоренять неукорененное филогенетическое дерево; создавать и использовать позиционную весовую матрицу с помощью сервиса MEME; рассчитывать информационное содержание выравнивания; использовать пакет HMMER для работы с профилями семейств белков; пользоваться базами данных KEGG, GO, Pfam, OPM.</p>		<p>систематическое умение</p>	<p>содержащее отдельные пробелы умение (допускает неточности неприципиального характера)</p>	<p>е умение</p>
<p>Владения: программами и программными пакетами: Mega, Jalview, Psi-BLAST, MEME, HMMER, Phylip; сервисами баз данных: Pfam, PROSITE, KEGG, GO, OPM.</p>	<p>Навыки владения отсутствуют</p>	<p>Наличие отдельных навыков (наличие фрагментарного опыта)</p>	<p>В целом, сформированные навыки (владения), но используемые не в активной форме</p>	<p>Сформированные навыки (владения), применяемые при решении задач</p>

8. Ресурсное обеспечение:

- Прежде всего, Интернет ресурсы, см. ниже.
- Перечень основной и дополнительной литературы:
Дополнительно к Интернет ресурсам можно (но не обязательно) использовать литературу:
 1. А. Леск. Введение в биоинформатику. – Москва: "Бином. Лаборатория знаний", 2013
 2. Б. Хаубольд, Т. Вие. Введение в вычислительную биологию. Эволюционный подход. – Москва: НИЦ "Регулярная и хаотическая динамика", 2011. 455 с. + CD-ROM
 3. Р. Дурбин, Ш. Эдди, А. Крог, Г. Митчисон. Анализ биологических последовательностей. – Москва: НИЦ "РХД", 2006. 480 с.

- Перечень лицензионного программного обеспечения (при необходимости): Все используемые биоинформатические программы и базы молекулярных данных являются свободно распространяемыми. Выполненные самостоятельные работы предоставляются студентом на учебной веб-странице. Для этого достаточен любой свободно распространяемый текстовый редактор (WordPad, NotePad++ и др.). Полезно иметь лицензионный Microsoft Office или любой бесплатный его аналог.
- Перечень профессиональных баз данных и информационных справочных систем: *все необходимые базы данных и справочные системы свободно доступны через Интернет, см. следующий пункт*
- Перечень ресурсов информационно-телекоммуникационной сети «Интернет» (при необходимости):
 - Базы данных Uniprot, Refseq (protein), Pfam, Prosite, KEGG, GO, OPM, Pubmed.
 - Справочные материалы по пакету EMBOSS, bash, HTML и CSS, программам MEGA, Jmol, JSmol, Jalview, доступные в Интернет.
 - Справочные материалы на веб-ресурсе kodom.fbb.msu.ru.

*(Примечание: без международных Интернет ресурсов освоение программы курса **НЕВОЗМОЖНО**).*
- Описание материально-технического обеспечения:
 - Помещения: компьютерные классы
 - Оборудование: сервер для обеспечения учебного процесса, компьютеры для компьютерных классов, в количестве не менее чем число студентов одного курса, два принтера, два проектора, оборудование для выхода в Интернет.
 - Иные материалы: программное обеспечение для сервера и компьютеров компьютерных классов