

Министерство науки и высшего образования РФ
Федеральное государственное автономное образовательное учреждение
высшего образования
«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

СОГЛАСОВАНО

Заведующий кафедрой

**Кафедра геномики и
биоинформатики**

наименование кафедры

подпись, инициалы, фамилия

«___» _____ 20__ г.

институт, реализующий ОП ВО

УТВЕРЖДАЮ

Заведующий кафедрой

**Кафедра геномики и
биоинформатики**

наименование кафедры

д.б.н. Ямских И.Е.

подпись, инициалы, фамилия

«___» _____ 20__ г.

институт, реализующий дисциплину

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
ПРАКТИЧЕСКАЯ
БИОИНФОРМАТИКА**

Дисциплина Б1.В.05 Практическая биоинформатика

Направление подготовки /
специальность 06.04.01 Биология Магистерская программа
06.04.01.06 Геномика и биоинформатика

Направленность
(профиль)

Форма обучения

очная

Год набора

2020

Красноярск 2021

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

составлена в соответствии с Федеральным государственным образовательным стандартом высшего образования по укрупненной группе

060000 «БИОЛОГИЧЕСКИЕ НАУКИ»

Направление подготовки /специальность (профиль/специализация)

направление 06.04.01 Биология Магистерская программа 06.04.01.06

Геномика и биоинформатика

Программу

составили

1 Цели и задачи изучения дисциплины

1.1 Цель преподавания дисциплины

Целью изучения дисциплины является формирование у магистров знаний об организации и методах проведения исследований в области практической биоинформатики, освоение современного инструментария биоинформатики и получение навыков его практического применения.

1.2 Задачи изучения дисциплины

В задачи курса входит:

- освоение навыков работы в unix-подобных операционных системах,
- изучение основ языка программирования высокого уровня (на примере Python) и формирование навыков использования скриптовых языков программирования,
- освоение навыков статистической обработки биологических данных с использованием пакетов статистической обработки и языков программирования (на примере R),
- знакомство с основами филогенетического анализа и освоение его основных методов,
- знакомство с основами структурной биоинформатики.

1.3 Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения образовательной программы

ОПК-4: способностью самостоятельно анализировать имеющуюся информацию, выявлять фундаментальные проблемы, ставить задачу и выполнять полевые, лабораторные биологические исследования при решении конкретных задач с использованием современной аппаратуры и вычислительных средств, нести ответственность за качество работ и научную достоверность результатов	
Уровень 1	Знает лишь основные достижения и новые разработки в области биоинформатики, основные сферы применения биоинформатики.
Уровень 2	Знаком с новейшими достижениями и новыми разработками в области биоинформатики, основными сферами применения биоинформатики.
Уровень 3	Хорошо ориентируется в новейших достижениях и новых разработках в области биоинформатики, основных сферах применения биоинформатики.
Уровень 1	Умеет частично ориентироваться в современных методах биоинформатики и применять их на практике; умеет с ошибками анализировать результаты научных исследований в области геномики, биоинформатики и смежных областях; умеет выявлять только некоторые направления биоинформатики в предметной

	области, их взаимосвязь со смежными областями.
Уровень 2	Умеет хорошо ориентироваться в современных методах биоинформатики и применять их на практике; умеет анализировать результаты научных исследований в области геномики, биоинформатики и смежных областях; умеет с небольшими ошибками выявлять направления биоинформатики в предметной области, их взаимосвязь со смежными областями.
Уровень 3	Умеет отлично ориентироваться в современных методах биоинформатики и применять их на практике; умеет уверенно анализировать результаты научных исследований в области геномики, биоинформатики и смежных областях; умеет самостоятельно выявлять направления биоинформатики в предметной области, их взаимосвязь со смежными областями.
Уровень 1	Владеет лишь немногими базовыми современными компьютерными технологиями, языками программирования.
Уровень 2	Хорошо владеет современными компьютерными технологиями, языками программирования.
Уровень 3	Отлично владеет современными компьютерными технологиями, языками программирования.
ОПК-7:готовностью творчески применять современные компьютерные технологии при сборе, хранении, обработке, анализе и передаче биологической информации для решения профессиональных задач	
Уровень 1	Знает лишь отдельные современные компьютерные технологии для обработки и анализа данных геномики и протеомики.
Уровень 2	Знаком с новейшими современными компьютерными технологиями для обработки и анализа данных геномики и протеомики.
Уровень 3	Хорошо знает современные компьютерные технологии для обработки и анализа данных геномики и протеомики.
Уровень 1	Умеет частично использовать современные компьютерные технологии для обработки и анализа данных геномики и протеомики.
Уровень 2	Умеет в большинстве случаев использовать современные компьютерные технологии для обработки и анализа данных геномики и протеомики.
Уровень 3	Умеет в полном объеме использовать современные компьютерные технологии для обработки и анализа данных геномики и протеомики.
Уровень 1	Владеет лишь немногими базовыми компьютерными технологиями для обработки и анализа данных секвенирования геномов живых организмов.
Уровень 2	Хорошо владеет современными компьютерными технологиями для обработки и анализа данных секвенирования геномов живых организмов.
Уровень 3	Отлично владеет современными компьютерными технологиями для обработки и анализа данных секвенирования геномов живых организмов.
ПК-1:способностью творчески использовать в научной и производственно-технологической деятельности знания фундаментальных и прикладных разделов дисциплин (модулей), определяющих направленность (профиль) программы магистратуры	

Уровень 1	Знает лишь отдельные механизмы сохранения информации живыми системами и реализации программ, заложенных геномами, принципы, положенные в основу методов анализа генома и протеома живых организмов.
Уровень 2	Знаком с новейшими механизмами сохранения информации живыми системами и реализации программ, заложенных геномами, принципами, положенными в основу методов анализа генома и протеома живых организмов.
Уровень 3	Хорошо ориентируется в механизмах сохранения информации живыми системами и реализации программ, заложенных геномами, принципах, положенных в основу методов анализа генома и протеома живых организмов.
Уровень 1	Умеет частично получать и грамотно использовать информацию, накопленную в базах данных по структуре геномов, белков, и другой биологической информации
Уровень 2	Умеет в большинстве случаев получать и грамотно использовать информацию, накопленную в базах данных по структуре геномов, белков, и другой биологической информации
Уровень 3	Умеет в полном объеме получать и грамотно использовать информацию, накопленную в базах данных по структуре геномов, белков, и другой биологической информации
Уровень 1	Владеет лишь немногими базовыми понятиями, положенными в основу биоинформатики, современными методами программирования, навыками работы с биоинформационными ресурсами.
Уровень 2	Хорошо владеет основными понятиями, положенными в основу биоинформатики, современными методами программирования, навыками работы с биоинформационными ресурсами.
Уровень 3	Отлично владеет основными понятиями, положенными в основу биоинформатики, современными методами программирования, навыками работы с биоинформационными ресурсами.
ПК-3: способностью применять методические основы проектирования, выполнения полевых и лабораторных биологических, экологических исследований, использовать современную аппаратуру и вычислительные комплексы (в соответствии с направленностью (профилем) программы магистратуры)	
Уровень 1	Знает лишь отдельные основы работы в unix-подобных операционных системах, основные команды для работы в режиме терминального доступа, основы программирования на Python, основные функции, типы данных, переменные, регулярные выражения этого языка программирования, основы статистической обработки результатов биологических исследований, основы обработки результатов при помощи языка программирования R, основы филогенетического анализа, методы выравнивания символьных последовательностей, методы построения филогенетических деревьев, основные форматы для работы с филогенетическими деревьями и программное обеспечение, основы структурной биоинформатики, теоретические и практические основы применения знаний о структуре белка для предсказания его свойств и функции;

Уровень 2	Знает на среднем уровне основы работы в unix-подобных операционных системах, основные команды для работы в режиме терминального доступа, основы программирования на Python, основные функции, типы данных, переменные, регулярные выражения этого языка программирования, основы статистической обработки результатов биологических исследований, основы обработки результатов при помощи языка программирования R, основы филогенетического анализа, методы выравнивания символьных последовательностей, методы построения филогенетических деревьев, основные форматы для работы с филогенетическими деревьями и программное обеспечение, основы структурной биоинформатики, теоретические и практические основы применения знаний о структуре белка для предсказания его свойств и функции;
Уровень 3	Хорошо знает основы работы в unix-подобных операционных системах, основные команды для работы в режиме терминального доступа, основы программирования на Python, основные функции, типы данных, переменные, регулярные выражения этого языка программирования, основы статистической обработки результатов биологических исследований, основы обработки результатов при помощи языка программирования R, основы филогенетического анализа, методы выравнивания символьных последовательностей, методы построения филогенетических деревьев, основные форматы для работы с филогенетическими деревьями и программное обеспечение, основы структурной биоинформатики, теоретические и практические основы применения знаний о структуре белка для предсказания его свойств и функции.
Уровень 1	Умеет частично работать в unix-подобных операционных системах, писать скрипты на bash и Python для решения повседневных биоинформатических задач, статистически обрабатывать данные при помощи R, строить графики, работать с большими массивами данных, использовать встроенные пакеты R для биоинформатической обработки данных, сравнивать символьные последовательности различными методами выравнивания, строить филогенетические деревья и их анализировать, работать с основными базами данных о структурах и функциях белков и других макромолекул.
Уровень 2	Умеет в большинстве случаев работать в unix-подобных операционных системах, писать скрипты на bash и Python для решения повседневных биоинформатических задач, статистически обрабатывать данные при помощи R, строить графики, работать с большими массивами данных, использовать встроенные пакеты R для биоинформатической обработки данных, сравнивать символьные последовательности различными методами выравнивания, строить филогенетические деревья и их анализировать, работать с основными базами данных о структурах и функциях белков и других макромолекул.
Уровень 3	Умеет в полном объеме работать в unix-подобных операционных системах, писать скрипты на bash и Python для решения повседневных биоинформатических задач, статистически обрабатывать данные при помощи R, строить графики, работать с

	большими массивами данных, использовать встроенные пакеты R для биоинформатической обработки данных, сравнивать символьные последовательности различными методами выравнивания, строить филогенетические деревья и их анализировать, работать с основными базами данных о структурах и функциях белков и других макромолекул.
Уровень 1	Владеет лишь немногими навыками работы unix-подобных операционных системах, навыками работы с серверами в режиме удаленного доступа, навыками программирования на скриптовых языках, статистическими методами обработки данных.
Уровень 2	Хорошо владеет навыками работы unix-подобных операционных системах, навыками работы с серверами в режиме удаленного доступа, навыками программирования на скриптовых языках, статистическими методами обработки данных.
Уровень 3	Отлично владеет навыками работы unix-подобных операционных системах, навыками работы с серверами в режиме удаленного доступа, навыками программирования на скриптовых языках, статистическими методами обработки данных.

1.4 Место дисциплины (модуля) в структуре образовательной программы

Дисциплина «Практическая биоинформатика» является обязательной и входит в вариативную часть ФГОС ВО направления 06.04.01 Биология, профиля 06.04.01.06 Геномика и биоинформатика. «Практическая биоинформатика» наряду с предшествующими курсами «Геномика» и «Основы биоинформатики» является базовой дисциплиной для подготовки магистров в области современной геномики и биоинформатики. Освоение данной дисциплины предполагает привлечение знаний, полученных как при изучении дисциплин бакалавриата – молекулярной биологии, генетики, информатики, программирования, информационно-коммуникационных технологий в естествознании, биометрии, биологической статистики, так и полученных в ходе изучения предшествующих дисциплин «Геномика» и «Основы биоинформатики», что обуславливает интеграцию полученных студентами знаний и выработку основ системного мышления. Результаты изучения дисциплины «Практическая биоинформатика» используются при изучении дисциплин вариативной части, дисциплин по выбору, при выполнении НИР и при подготовке магистерской диссертации. Курс носит междисциплинарный характер и состоит из лекционного материала и семинарских занятий. Реализуется в 3 семестре.

Молекулярная экология
Геномика

Научно-исследовательский семинар
Основы биоинформатики

Научно - исследовательская работа

Практика по получению первичных профессиональных умений и навыков

Практика по получению профессиональных умений и опыта профессиональной деятельности

Преддипломная практика

1.5 Особенности реализации дисциплины

Язык реализации дисциплины Русский.

Дисциплина (модуль) реализуется без применения ЭО и ДОТ.

2. Объем дисциплины (модуля)

Вид учебной работы	Всего, зачетных единиц (акад.час)	Семестр
		3
Общая трудоемкость дисциплины	7 (252)	7 (252)
Контактная работа с преподавателем:	2,67 (96)	2,67 (96)
занятия лекционного типа	0,89 (32)	0,89 (32)
занятия семинарского типа		
в том числе: семинары		
практические занятия	1,78 (64)	1,78 (64)
практикумы		
лабораторные работы		
другие виды контактной работы		
в том числе: групповые консультации		
индивидуальные консультации		
иная внеаудиторная контактная работа:		
групповые занятия		
индивидуальные занятия		
Самостоятельная работа обучающихся:	3,33 (120)	3,33 (120)
изучение теоретического курса (ТО)		
расчетно-графические задания, задачи (РГЗ)		
реферат, эссе (Р)		
курсовое проектирование (КП)	Нет	Нет
курсовая работа (КР)	Нет	Нет
Промежуточная аттестация (Экзамен)	1 (36)	1 (36)

3 Содержание дисциплины (модуля)

3.1 Разделы дисциплины и виды занятий (тематический план занятий)

№ п/п	Модули, темы (разделы) дисциплины	Занятия лекционного типа (акад. час)	Занятия семинарского типа		Самостоятельная работа, (акад. час)	Формируемые компетенции
			Семинары и/или Практические занятия (акад. час)	Лабораторные работы и/или Практикумы (акад. час)		
1	2	3	4	5	6	7
1		32	0	0	0	ПК-1
2		0	64	0	0	ОПК-7 ПК-1 ПК-3
3		0	0	0	120	ОПК-4 ОПК-7 ПК-1
Всего		32	64	0	120	

3.2 Занятия лекционного типа

№ п/п	№ раздела дисциплины	Наименование занятий	Объем в акад. часах		
			Всего	в том числе, в инновационной форме	в том числе, в электронной форме

1	1	<p>Введение в Linux. Основы работы в Linux. Основные команды. Работа в терминале. Исполняемые файлы. Ввод/вывод. Работа с файлами в сети. Работа с архивами. Поиск файлов и слов в строках. Работа на сервере. Знакомство с сервером. Обмен файлами. Запуск приложений. Контроль запускаемых программ. Многопоточные приложения. PBS Troque. Текстовый редактор vim. Bash: основные команды, ветвления и циклы. Часто используемые в биоинформатике команды.</p>	4	0	0
2	1	<p>Программирование на Python. Введение в Python. Функции в языке Python. PEP8. Синтаксис Python. Модель данных, объекты, типы и классы. Области видимости, пространства имен и классы. Работа с файлами в Python. Модули, импорт, модели. Регулярные выражения. Работа с данными в Интернет. Итераторы и генераторы .</p>	6	0	0

3	1	<p>Биологическая статистика на языке R. Рабочая среда R. Консоль, скрипт, рабочее пространство. Датафреймы, строки, столбцы. Базовый синтаксис, формулы. Проверка гипотез о различии распределений. Двухвыборочные и одновыборочные тесты. Визуализация данных. Проверка гипотез о различии распределений Дисперсионный анализ (ANOVA). Однофакторный, многофакторный дисперсионный анализ, анализ с повторными измерениями Регрессионный анализ. Регрессия с категориальными переменными. Обобщенные линейные модели на примере порядковой и логистической регрессии. Смешанные модели.</p>	8	0	0
---	---	--	---	---	---

4	1	<p>Молекулярная филогения. Введение в филогенетику. Филогенетические деревья. Листы, узлы, ветви. Поли/пара/монофилетические группы. Гомология и гомоплазия. Форматы. Выравнивание - глобальное, локальное, множественное. Методы построения филогенетических деревьев. Метод максимальной парсимонии. Методы расстояний. Метод наименьших квадратов, Фитча-Марголиаша, минимальной эволюции. Neighbour-joining. Тестирование филогении. Bootstrap. Модели нуклеотидных замен. Матрица замен. Выбор моделей. Метод максимального правдоподобия. Байесовские методы. Теорема Байеса. Использование апостериорных данных, пересчет вероятностей. MrBayes. Априорные распределения. Отдельные задачи филогении. Реконструкция предковых последовательностей. Примеры. Анализ митохондриальной ДНК. Влияние вторичных и третичных структур.</p>	8	0	0
---	---	---	---	---	---

5	1	Структурная биоинформатика Введение. Белки. Анализ белковых последовательностей. Структуры белка, их анализ и методы предсказания. База RCSB/PDB. Кинетика фолдинга белка. Стабильность белка. Белковые взаимодействия.	6	0	0
Всего			22	0	0

3.3 Занятия семинарского типа

№ п/п	№ раздела дисциплины	Наименование занятий	Объем в акад. часах		
			Всего	в том числе, в инновационной форме	в том числе, в электронной форме
1	2	Введение в Linux. 1. Основные команды. Работа в терминале. Исполняемые файлы. 2. Ввод/вывод. Работа с файлами в сети. Работа с архивами. Поиск файлов и слов в строках. 3. Знакомство с сервером. Обмен файлами. Запуск приложений. Установка приложений. Контроль запускаемых программ. Многопоточные приложения. PBS Troque. 4. Текстовый редактор vim. Bash: основные команды, ветвления и циклы. Часто используемые в биоинформатике команды.	8	0	0

2	2	Программирование на Python. 1. Функции в языке Python. 2. Модель данных, объекты, типы и классы. Области видимости, пространства имен и классы. 3. Работа с файлами в Python 4. Модули, импорт, модели. 5. Регулярные выражения 6. Работа с данными в Интернет	12	0	0
---	---	--	----	---	---

3	2	16	0	0
---	---	----	---	---

Биологическая статистика на языке R.

1. Консоль, скрипт, рабочее пространство. Дата фреймы, строки, столбцы. Базовый синтаксис, формулы. Расчет базовых описательных статистик, выделение подгрупп, вывод данных в PDF/HTML.

2. Проверка гипотез о различии распределений (двухвыборочные и одновыборочные тесты): Базовый анализ номинативных данных. Таблицы сопряженности. Хи-квадрат и точный критерий Фишера. V-Крамера и другие меры ассоциации. Визуализация номинативных данных.

3. Базовый корреляционный анализ. Порядковые и метрические переменные. Множественный линейный регрессионный анализ.

4. Проверка гипотез о различии распределений (двухвыборочные и одновыборочные тесты): Пирсон, Спирмен, Кендалл. Базовое сравнение групп (категориальная и некатегориальная переменная). Тест Стьюдента, Вилкоксона-Манна-Уитни.

Визуализация данных корреляционного анализа и сравнения групп.

5. Дисперсионный анализ (ANOVA). Однофакторный дисперсионный анализ, многофакторный дисперсионный анализ.

6. Дисперсионный анализ с повторными измерениями. Тест Краскела-Уоллиса и Фридмана. Визуализация результатов.

7. Регрессия с категориальными

4	2	<p>Молекулярная филогения.</p> <p>1. Введение в филогенетику. Филогенетические деревья. Листы, узлы, ветви. Поли/пара/монофилетические группы. Внешняя группа. Гомология и гомоплазия. Формат Newick.</p> <p>2. Выравнивание - глобальное, локальное, множественное. ClustalW, MUSCLE, MAFFT. Форматы (FASTA, NEXUS). Выбор маркеров. Филогенетический сигнал, насыщение. DNA barcoding.</p> <p>3. Метод максимальной парсимонии. Поиск дерева минимальной длины. Метод Фитча. Branch-and-bound. Эвристики - NNI, SPR, TBR.</p> <p>4. Методы расстояний. Метод наименьших квадратов, Фитча-Марголиаша, минимальной эволюции. Neighbour-joining. Тестирование филогении. Bootstrap.</p> <p>5. Модели нуклеотидных замен. Матрица замен. Модель как марковская цепь 1-го порядка. Неравномерность скоростей замен, гамма-распределение. Выбор моделей. Информационный критерий Akaike, BIC. ModelTest. Indel treatment. GapCoder.</p> <p>6. Метод максимального правдоподобия. Понятие функции правдоподобия. RAxML, PhyML. Выбор модели. ModelTest. Использование нескольких моделей (partitioning). Partition Finder.</p> <p>7. Байесовские методы.</p>	16	0	0
---	---	---	----	---	---

5	2	<p>Структурная биоинформатика.</p> <p>1. Основы рентгеноструктурного анализа(РСА) макромолекул. Индикаторы качества модели макромолекулы, построенной по данным РСА.</p> <p>2. Поверхность макромолекулы, алгоритмы её вычисления.</p> <p>3. Гидрофобное ядро молекулы белка, алгоритмы его нахождения.</p> <p>4. Структурные домены белков, алгоритмы их нахождения.</p> <p>5. Пространственное выравнивание структур белков. Структурные классификации доменов SCOP и CATH.</p> <p>6. Молекулярная динамика, основные принципы и ограничения.</p>	12	0	0
Всего			64	0	0

3.4 Лабораторные занятия

№ п/п	№ раздела дисциплины	Наименование занятий	Объем в акад. часах		
			Всего	в том числе, в инновационной форме	в том числе, в электронной форме
Всего					

4 Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине (модулю)

	Авторы, составители	Заглавие	Издательство, год
Л1.1	Колесниченко Д. Н.	Самоучитель Linux. Установка, настройка, использование: [самоучитель]	Санкт-Петербург: Наука и техника, 2009

Л1.2	Кузьмин Д. А., Удалова Ю. В.	Разработка компонентов системного программного обеспечения. Процессы в Linux: учеб.-метод. пособие для студентов спец. 010501, 090102, 230100	Красноярск: СФУ, 2012
------	---------------------------------	---	--------------------------

5 Фонд оценочных средств для проведения промежуточной аттестации

Оценочные средства находятся в приложении к рабочим программам дисциплин.

6 Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины (модуля)

6.1. Основная литература			
	Авторы, составители	Заглавие	Издательство, год
Л1.1	Колисниченко Д. Н.	Linux. От новичка к профессионалу: наиболее полное руководство	Санкт-Петербург: БХВ-Петербург, 2012
Л1.2	Лав Р., Сивченко О.	Linux. Системное программирование	Санкт-Петербург: Питер, 2015
6.2. Дополнительная литература			
	Авторы, составители	Заглавие	Издательство, год
Л2.1	Игнасимуту С.	Основы биоинформатики: перевод с английского	МоскваМосква: [R&C Dynamics] Регулярная и хаотическая динамика [РХД], 2007
Л2.2	Глик Б., Пастернак Д., Янковский Н. К.	Молекулярная биотехнология: принципы и применение: перевод с английского	Москва: Мир, 2002
Л2.3	Леск А., Миронов А. А., Швядас В. К.	Введение в биоинформатику: учеб. пособие: пер. с англ.	Москва: БИНОМ, Лаборатория знаний, 2009
Л2.4	Хаубольд Б., Вие Т., Чудов С. В., Артамонова И. И.	Введение в вычислительную биологию. Эволюционный подход	Москва: Регулярная и хаотическая динамика, 2011
6.3. Методические разработки			
	Авторы, составители	Заглавие	Издательство, год

ЛЗ.1	Колесниченко Д. Н.	Самоучитель Linux. Установка, настройка, использование: [самоучитель]	Санкт-Петербург: Наука и техника, 2009
ЛЗ.2	Кузьмин Д. А., Удалова Ю. В.	Разработка компонентов системного программного обеспечения. Процессы в Linux: учеб.-метод. пособие для студентов спец. 010501, 090102, 230100	Красноярск: СФУ, 2012

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

Э1	American Physical Society [Электронный ресурс]	http://publish.aps.org http://bik.sfu-kras.ru/nb/american-physical-society
Э2	Annual Reviews Science Collection [Электронный ресурс]	http://www.annualreviews.org http://bik.sfu-kras.ru/nb/annual-reviews-science-collection
Э3	arXiv [Электронный ресурс]	http://arxiv.org
Э4	Cambridge University Press [Электронный ресурс]	http://www.journals.cambridge.org
Э5	Coghlan, A. A little book of R for bioinformatics [Электронный ресурс] / A. Coghlan. – Электрон. текстовые дан. – Cambridge: [б.и.], 2013.	http://www.info.univ-angers.fr/~gh/wstat/Introduction_R/a-little-book-of-r-for-bioinformatics
Э6	DOAJ [Электронный ресурс]	http://drf.lib.hokudai.ac.jp
Э7	EBSCO Publishing [Электронный ресурс]	http://search.ebscohost.com
Э8	Elsevier [Электронный ресурс]	http://www.sciencedirect.com
Э9	Oxford Russia Fund eContent library [Электронный ресурс]	http://lib.mylibrary.com http://bik.sfu-kras.ru/nb/oxford-russia-fund-econtent-library
Э10	Scopus [Электронный ресурс]	http://bik.sfu-kras.ru/nb/scopus
Э11	Springer [Электронный ресурс]	http://www.springerlink.com
Э12	Web of Science [Электронный ресурс]	http://isiknowledge.com
Э13	Колчанов, Н.А. Электронно-лекционный курс «Введение в информационную биологию» [Электронный ресурс] / Н.А. Колчанов, С.А. Лашин. – Электрон. текстовые	
Э14	Лашин, С.А. Электронно-лекционный курс «Информационные технологии и языки программирования» [Электронный ресурс] / С.А. Лашин. – Электрон. текстовые дан. – Новосибирск: [б.и.], 2012	http://kib.nsu.ru/?page_id=2837
Э15	Онлайн-курс «Bioinformatics Methods» [Электронный ресурс]	https://www.coursera.org/specializations/bioinformatics

Э16	Онлайн-курс «Bioinformatics: Life Sciences on Your Computer» [Электронный ресурс]	https://www.coursera.org/course/bioinform
Э17	Онлайн-курс «Computational Molecular Evolution» [Электронный ресурс]	https://ru.coursera.org/course/molevol
Э18	Специализация из 7-ми онлайн-курсов «Bioinformatics» [Электронный ресурс]	https://www.coursera.org/specializations/bioinformatics
Э19	Центр Исследования Генома (Genomics Resource Centre) [Электронный ресурс]	http://www.rockefeller.edu/genomics

8 Методические указания для обучающихся по освоению дисциплины (модуля)

Самостоятельная работа студента по дисциплине «Практическая биоинформатика» предусматривает изучение теоретического материала и выполнение самостоятельных работ с использованием основной и дополнительной литературы – 3,3 з.е. (120 часов).

Самостоятельное изучение теоретического материала предполагает работу с учебной литературой, научными монографиями, оригинальными научными статьями, диссертациями. Итогом работы является презентация.

На самостоятельное изучение выносятся вопросы в соответствии с тематикой лекций. Организация самостоятельной работы производится в соответствии с графиком учебного процесса и самостоятельной работы.

9 Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю) (при необходимости)

9.1 Перечень необходимого программного обеспечения

9.1.1	Современные биоинформатические исследования требуют умения решать поставленные задачи с использованием самого разнообразного программного обеспечения, от пользовательских скриптов, размещенных в репозиториях, до дорогостоящего проприетарного ПО, такого как CLCbio. Философия современного биоинформатического сообщества заключается в том, что любую задачу можно решить несколькими способами: с использованием бесплатно распространяемого ПО, при помощи онлайн-сервисов (пайплайнов) и проприетарного ПО, или самостоятельно создать новый программный продукт для решения конкретной пользовательской задачи. В рамках данного курса используется только свободно распространяемое ПО: BLAST, FastQC, Trimmomatic, ABySS, MaSuRCA, SPAdes, Bowtie2, BWA, Samtools, GATK, SSPACE , MAKER , Trinity, Trinotate, Blast2GO, QUAST, UGENE, MEGA, BioEdit.
-------	--

9.2 Перечень необходимых информационных справочных систем

9.2.1	Биоинформатика – та область знаний, в которой ресурсы Интернет используются практически для решения любой задачи. Вся биоинформатика основана на создании баз данных, наполнении их результатами научных работ исследователями со всего мира, открытости доступа к этим данным и сравнении новых результатов с уже опубликованными.
9.2.2	В рамках освоения дисциплины используется одна из крупнейших информационных систем в области биологии медицины, биофизики Национального центра биотехнологической информации (National Center for Biotechnology Information (NCBI)), США (www.NCBI.nlm.nih.gov).
9.2.3	БД NCBI являются достаточно сложным инструментарием с разнообразным функционалом.
9.2.4	Ниже приведено краткое описание основных БД NCBI, которые могут быть полезны при освоении тем дисциплины.
9.2.5	БД Nucleotide (http://www.NCBI.nlm.nih.gov/sites/Entrez?db=nucleotide) объединяет данные последовательностей нуклеиновых кислот из нескольких исходных БД, в том числе GenBank, RefSeq и др. Данные могут быть найдены по регистрационному номеру, имени автора, наименованию организма, генома/белка, а также ряду других параметров.
9.2.6	БД Protein (http://www.NCBI.nlm.nih.gov/sites/Entrez?db=protein) является коллекцией аминокислотных последовательностей из нескольких источников, в том числе из GenBank, RefSeq и TrEMBL, а также SwissProt, PIR, PRF и PDB.
9.2.7	БД Structure (http://www.NCBI.nlm.nih.gov/Structure/index.shtml) организуют доступ к результатам молекулярного моделирования макромолекул и связанным с ними БД: трехмерных биомолекулярных структур полученных с помощью рентгеновской кристаллографии и ЯМР-спектроскопии; БД химических структур небольших органических молекул; к информации об их биологической активности и т. д.
9.2.8	БД Gene (http://www.NCBI.nlm.nih.gov/sites/Entrez?db=gene) представляет собой инструмент для просмотра данных из широкого спектра геномов. Каждая запись – это один из генов определенного организма. Минимальный набор данных в гене запись включает уникальный идентификатор, т. н. Gene-ID.
9.2.9	БД dbMHC (http://www.NCBI.nlm.nih.gov/gv/mhc/main.cgi?cmd=init) предоставляет открытую платформу, где научное сообщество может размещать, просматривать и редактировать данные MajorHistocompatibilityComplex (MHC) для человека. БД dbMHC полностью интегрирована с другими ресурсами NCBI, а также с Международной рабочей группой гистосовместимости (IHWG).
9.2.10	DbSNP (http://www.NCBI.nlm.nih.gov/SNP/) – БД одиночных нуклеотидных полиморфизмов, полиморфных повторяющихся элементов, включающая как гибридные данные, так и полученные только экспериментальным путем.
9.2.11	БД ReferenceSequence (RefSeq) (http://www.NCBI.nlm.nih.gov/RefSeq/), содержащая последовательности, в том числе геномных ДНК, белков и т. д., является основой для проведения функциональных исследований, геномной идентификации, сравнительного анализа и т. п. В частности, релиз от 11.07.2012 включал в себя описания 16 393 342 белков и 17 605 организмов.

9.2.1 2	БД Genomic Biology представляет собой объединение нескольких ресурсов и инструментов геномной биологии, в том числе геномных карт для Fruitfly, Human, Malariaparasite, Mouse, Rat, Retroviruses, Zebrafish и т. д., которые дополнительно содержат ссылки на интернет-ресурсы и БД, касающиеся рассматриваемых видов.
9.2.1 3	В БД UniGene (http://www.NCBI.nlm.nih.gov/unigene/) полноразмерные mRNA последовательности организованы в уникальные кластеры, представляющие известные или предполагаемые гены. Для кластеров доступна информация по картированию, экспрессии и другие ресурсы.
9.2.1 4	HomoloGene (http://www.NCBI.nlm.nih.gov/homologene) – инструмент для автоматизированного выявления гомологов среди аннотированных генов, который сравнивает нуклеотидные последовательности между парами организмов в целях выявления предполагаемых ортологов.
9.2.1 5	Basic Local Alignment Search Tool (http://www.NCBI.nlm.nih.gov/BLAST/) - основной метод поиска гомологичных последовательностей на основе локального выравнивания.
9.2.1 6	Public repository Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) - публичная электронная библиотека данных экспрессии генов «Омнибус Экспрессии Генов»
9.2.1 7	GenBank (http://www.NCBI.nlm.nih.gov/genbank/index.html) – БД, содержащая доступные последовательности нуклеотидов для более чем 260 000 организмов, вся информация в генетическом банке данных сопровождается библиографическими ссылками и биологическими аннотациями. GenBank автоматически интегрирует информацию о геноме и БД белковых последовательностей для изучения, учитывая таксономию, геном, белковую структуру и другую информацию.
9.2.1 8	Для представления последовательностей в GenBank предложено два инструмента:
9.2.1 9	• BankIt – интернет-представление одной или нескольких последовательностей;
9.2.2 0	• Sequin – интернет-представление для длинных последовательностей, полных геномов, результатов популяционных и филогенетических исследований.
9.2.2 1	Объединяющим фактором и при этом крайне удобным инструментом поиска в NCBI является поисковая система Search NCBI databases (http://www.NCBI.nlm.nih.gov/sites/gquery). Она обеспечивает одновременный доступ как к нуклеотидным и белковым последовательностям (GenBank, EMBL, DDBJ, PIR-International, PRF, Swiss-Prot и PDB, GenPept, RPF), 3-мерным структурам и популяционным данным, так и к библиографическим БД (PubMed, PubMed Central и т. д.). Доступ к поисковой системе Search NCBI databases может быть легко получен с помощью прямого интернет-адреса (http://www.NCBI.nlm.nih.gov/gquery/) либо посредством использования стартовой страницы NCBI (http://www.NCBI.nlm.nih.gov/). На этой странице приведен полный перечень инструментария и БД NCBI и существует возможность получить доступ к любой из перечисленных БД.

9.2.2 2	Крайне полезным инструментом, который сохраняет информацию о пользователе, используется для более точной настройки поисковых запросов в NCBI (http://www.NCBI.nlm.nih.gov/index.html) и т. д., является сервис «My NCBI» (http://www.NCBI.nlm.nih.gov/sites/My NCBI/). Этот инструмент позволяет сохранять результаты поиска, выбирать форматы отображения, фильтрации, настраивать автоматический поиск и отправлять его результаты по электронной почте. Пользователи «My NCBI» могут сохранять свои БД, построенные на основе поисковых запросов в NCBI, и управлять политикой общественного доступа.
------------	--

10 Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине (модулю)

Аудиторный класс, наличие проектора для демонстрации наглядных пособий и экрана. Компьютерный класс, лицензионное программное обеспечение, Internet.