Министерство просвещения Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Ульяновский государственный педагогический университет
имени И.Н. Ульянова»
(ФГБОУ ВО «УлГПУ им. И.Н. Ульянова»)

Факультет физико-математического и технологического образования
Кафедра высшей математики

УТВЕРЖДАЮ
Проректор по учебно-
методической работе
_____ С.Н. Титов
«25» июня 2021 г.

КОМПЬЮТЕРНАЯ ГЕНОМИКА

Программа учебной дисциплины модуля информатики в биологии

основной профессиональной образовательной программы высшего образования – программы магистратуры по направлению подготовки
06.04.01 Биология

направленность (профиль) образовательной программы
Биотехнология с основами нанотехнологий

(очно-заочная форма обучения)

Составитель: Цыганов А.В., к.ф.-м.н., доцент кафедры высшей математики

Рассмотрено и одобрено на заседании ученого совета естественно-географического факультета, протокол от «22» июня 2021 г. №7

Ульяновск, 2021

ДОКУМЕНТ ПОДПИСАН
ЭЛЕКТРОННОЙ ПОДПИСЬЮ

Сертификат: 01D77C644FC1E54000000006381D0002
Владелец: Титов Сергей Николаевич
Действителен: с 19.07.2021 до 19.07.2022

Место дисциплины в структуре образовательной программы

Дисциплина «Компьютерная геномика» относится к дисциплинам части, формируемой участниками образовательных отношений Блока 1. Дисциплины (модули) модуля «Информатика в биологии» учебного плана основной профессиональной образовательной программы высшего образования – программы магистратуры по направлению подготовки 06.04.01 Биология, направленность (профиль) образовательной программы «Биотехнология с основами нанотехнологий», очно-заочной формы обучения.

Дисциплина опирается на результаты обучения, сформированные в рамках дисциплин и практик бакалавриата.

Результаты освоения дисциплины являются основой изучения дисциплин для профессиональной деятельности и прохождения государственной итоговой аттестации.

1. Перечень планируемых результатов обучения (образовательных результатов) по дисциплине

Цель изучения дисциплины: формирование у обучающихся знаний, умений и навыков в области компьютерных методов изучения геномов.

Задачи дисциплины:

– раскрыть содержание базовых понятий, предмета, методов и принципов компьютерной геномики;

– дать представление о современном биоинформационном программном обеспечении;

– обучить владению современными методами вычислительной геномики.

В результате освоения программы магистратуры обучающийся должен овладеть следующими результатами обучения по дисциплине «Компьютерная геномика» (в таблице представлено соотнесение образовательных результатов обучения по дисциплине с индикаторами достижения компетенций):

Компетенция и индикаторы ее достижения в дисциплине	Образовательные результаты дисциплины (этапы формирования дисциплины)		
	знает	умеет	владеет
ПК 4 Поиск и разработка новых эффективных путей получения биотехнологических продуктов, программного обеспечения.			
ИПК 4.1. Применяет знания стандартных и иных методик отбора, транспортировка и пробоподготовки проб согласно руководящей документации.		ОР-1 Применяет знания стандартных и иных методик отбора	
ИПК 4.2. Владеет базовыми и специализированными методами, в зависимости от типа биоматериала и поставленных задач, в		ОР-2 Владеет базовыми и специализированными методами в области генетического конструирования,	

области генетического конструирования, молекулярно-генетическими методами, методами в области клеточных технологий, согласно руководящей документации.			
ИПК 4.3. Проводит анализ современной литературы, последних достижений с целью разработки самостоятельных протоколов по созданию биотехнологических продуктов.		ОР-3 Проводит анализ современной литературы в области современной геномики	
ИПК 4.4. Умеет применять знания в области IT-технологий для решения задач анализа, прогнозирования, оптимизации лабораторных протоколов и методов исследования.		ОР-4 Умеет применять знания в области IT-технологий для решения задач анализа,	
ИПК 4.5. Владеет навыками разработки ПО, анализа полученных данных, статистической обработки, хранения и документации результатов.			ОР-5 Владеет навыками разработки ПО, анализа полученных данных, статистической обработки, хранения и документации результатов.

2. Объем дисциплины в зачетных единицах с указанием количества академических часов, выделенных на контактную работу обучающихся с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающихся

Номер семестра	Учебные занятия						Форма промежуточной аттестации
	Всего		Лекции, час	Практические занятия, час	Лабораторные занятия, час	Самостоят. работа, час	
	Трудоемк.						
	Зач. ед.	Часы					
3	3	108	4	16	-	88	Зачет
Итого:	3	108	4	16	-	88	Зачет

3. Содержание дисциплины, структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий

3.1. Указание тем (разделов) и отведенного на них количества академических часов и видов учебных занятий:

Наименование раздела и тем	Количество часов по формам организации обучения			
	Лекционные занятия	Практические занятия	Лабораторные занятия	Самостоятельная работа
Тема 1. Введение в дисциплину	2	2		10
Тема 2. Биоинформационные базы данных	2	2		10
Тема 3. Сравнение полинуклеотидных и полипептидных последовательностей		2		18
Тема 4. Автоматическое аннотирование последовательности. Онтология.		2		10
Тема 5. Молекулярная эволюция		2		10
Тема 6. Статистика последовательностей ДНК		2		10
Тема 7. Методы вычислительной геномики		2		10
Тема 8. Биоинформационная поддержка эксперимента		2		10
Итого:	**4**	**16**		**88**

3.2. Краткое описание содержания тем (разделов) дисциплины

Краткое содержание курса

Тема 1. Введение в дисциплину.

История развития биоинформационных методов. Основные задачи и методы биоинформатики в исследовании генома и протеома. **Интерактивная форма:** групповая дискуссия.

Тема 2. Биоинформационные базы данных.

Структура биоинформационных баз данных, классификация баз по способу заполнения, основные базы данных, базы, содержащие результаты глобальных экспериментов, банки белковых семейств, метаболические базы данных генетические банки, специализированные банки данных. **Интерактивная форма:** групповая дискуссия.

Тема 2. Основные области исследований.

Анализ генетических последовательностей Вычислительная эволюционная биология. Оценка биологического разнообразия. Основные биоинформационные программы.

Биоинформатика и вычислительная биология. Структурная биоинформатика. **Интерактивная форма:** работа в микрогруппах.

Тема 3. Сравнение полинуклеотидных и полипептидных последовательностей.

Сравнение последовательностей, выравнивание двух последовательностей, глобальное и локальное выравнивание, вес выравнивания, матрицы аминокислотных замен, алгоритмы локального выравнивания, статистическая значимость выравниваний и ее зависимость от вероятностной модели последовательности, зависимость выравнивания от параметров, множественное выравнивание, последовательное выравнивание, скрытые марковские модели, поиск блоков.

Интерактивная форма: групповая дискуссия.

Тема 4. Автоматическое аннотирование последовательности. Онтология.

Поиск гомологов в базах данных, интерпретация результатов, сравнение алгоритмов, зависимость от параметров (матрицы, фильтры и т.п.). Поиск по сходству в базах данных, оценка значимости, фильтрация повторов и обработка участков малой сложности, паттерны, профили.

Интерактивная форма: работа в микрогруппах.

Тема 5. Молекулярная эволюция.

Эволюция молекул и организмов, ортологи и паралоги, горизонтальный перенос, деревья видов и деревья генов. Филогенетическое дерево как математический объект. Модели эволюции, алгоритмы построения филогенетических деревьев, матрица расстояний методы, основанные на матрице расстояний, другие методы. Алгоритмические проблемы поиска оптимального дерева, согласование деревьев. Эволюция на уровне генома. **Интерактивная форма:** работа в микрогруппах.

Тема 6. Статистика последовательностей ДНК.

Нуклеотидный состав (изохоры, GC-острова, картирование старта репликации), частые и редкие слова (вероятностные проблемы), статистика ДНК как характеристика генома.

Интерактивная форма: работа в микрогруппах.

Тема 7. Методы вычислительной геномики.

Метаболическая реконструкция, позиционный анализ, эволюция регуляторных взаимодействий, эволюция белковых семейств, их доля в геноме. **Интерактивная форма:** работа в микрогруппах.

Тема 8. Биоинформационная поддержка эксперимента.

Подбор праймеров для ПЦР. Анализ данных двумерного фореза (обработка изображений), масс-спектрометрия белков. Анализ данных по экспрессии генов (microarrays и др.), обработка изображений, кластеризация профилей экспрессии, диагностика по экспрессии генов.

Интерактивная форма: работа в микрогруппах.

4. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

Общий объем самостоятельной работы студентов по дисциплине включает аудиторную и внеаудиторную самостоятельную работу студентов в течение семестра.

Самостоятельная работа осуществляется в форме выполнения самостоятельных заданий по дисциплине и обеспечена базой заданий для самостоятельной работы.

Самостоятельная работа осуществляется в формах:

- подготовка к устным докладам (мини-выступлениям);
- подготовка к защите реферата;
- подготовка к групповому обсуждению по темам;
- подготовка стендовых докладов и постеров;
- разработка проектов.

Темы докладов

1. Основные виды биоинформационных баз данных.
2. Основные биоинформационные программы.
3. Виды выравнивания последовательностей.
4. Алгоритм Нидлмана–Вунша.
5. Алгоритмы построения филогенетических деревьев.

Тест

1. 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

2. Что из перечисленного не относится к основным типам генетических карт

а) генетические карты сцепленности генов
б) «бэндовые» схемы хромосом
в) последовательности ДНК
г) контиг

3. Контиг – это

а) набор перекрывающихся фрагментов ДНК, которые в совокупности представляют собой консенсусную область ДНК
б) локусы с варьирующим числом тандемных повторов
в) полиморфизм коротких тандемных повторов
г) короткий, секвенированный участок ДНК, локализованный в строго определенной области генома

4. Процесс спонтанного сворачивания полипептидной цепи в уникальную нативную пространственную структуру называется

а) сплайсинг
б) трансляция
в) фолдинг
г) процессинг

5. Основным инструментом биоинформатики является

а) выравнивание последовательностей
б) секвенирование
в) программирование
г) картирование генома

6. 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

7. 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-спекроскопия

г) электрофорез

8. 16. Не является этапом ПЦР

а) денатурация ДНК

б) отжиг

в) достраивание цепей ДНК

г) инициация

9. 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

10. 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 %

11. 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, разошед-

шихся в процессе видообразования

г) гомологичные гены микроорганизмов, образовавшиеся в процессе горизонтально-

го переноса

12. Трансмембранные сегменты состоят почти исключительно из гидрофобных аминокислотных остатков. Сколько остатков составляет длину транмембранных спиралей

а) 100 остатков

б) 150 остатков

в) 2-5 остатка

г) 15-30 остатков

13. Какая из перечисленных ниже программ используется для множественного выравнивания последовательностей ДНК и белков

а) ClustalW

б) BLAST

в) DALI

г) CASP

14. 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 15. 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

16. Гомологичные нуклеотидные (или аминокислотные) последовательности называют паралогичными если…

а) они появились в результате видообразования

б) они появились в результате дупликации

в) они находятся в начале гена

г) они являются уникальными

17. 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

18. 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

19. Выравнивание нуклеотидных или аминокислотных последовательностей с самым высоким весом называют

а) оптимальным

б) множественным

в) глобальным

г) структурным

20. 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

Для самостоятельной подготовки к занятиям по дисциплине рекомендуется использовать учебно-методические материалы:

1. Цыганов А.В. Вычислительная геномика. Методические указания лекционных и практических занятий для магистров 1 курса 3 семестра направления подготовки 06.04.01 «Биология». Профиль: Биотехнология с основами нанотехнологий (очная форма обучения). Ульяновск: Изд-во УлГПУ, 2016.

5. Примерные оценочные материалы для проведения текущего контроля успеваемости и промежуточной аттестации обучающихся по дисциплине

Организация и проведение аттестации магистранта

ФГОС ВО в соответствии с принципами Болонского процесса ориентированы преимущественно не на сообщение магистрантам комплекса теоретических знаний, но на выработку у студентов компетенций – динамического набора знаний, умений, навыков и личностных качеств, которые позволят выпускнику стать конкурентоспособным на рынке труда и успешно профессионально реализовываться.

В процессе оценки магистранта необходимо используются как традиционные, так и инновационные типы, виды и формы контроля. При этом постепенно традиционные средства

совершенствуются в русле компетентностного подхода, а инновационные средства адаптированы для повсеместного применения в российской вузовской практике.

Все компетенции по данной дисциплине формируются на начальном (пороговом) уровне.

Цель проведения аттестации – проверка освоения образовательной программы дисциплины-практикума через сформированность образовательных результатов. **Типы контроля:**

Текущая аттестация представлена следующими работами: устными докладами, рефератами, презентациями.

Достоинства предложенной системы проведения аттестации: систематичность, непосредственно коррелирующая с требованием постоянного и непрерывного мониторинга качества обучения, а также возможность балльно-рейтинговой оценки успеваемости магистранта.

Промежуточная аттестация осуществляется в конце семестра и завершает изучение дисциплины; помогает оценить более крупные совокупности знаний и умений, формирование определенных профессиональных компетенций.

Оценочными средствами текущего оценивания являются: устные доклады, защита реферата, групповые дискуссии, стендовые доклады и презентации. Контроль усвоения материала ведется регулярно в течение всего семестра на практических занятиях.

№ п/п	**СРЕДСТВА ОЦЕНИВАНИЯ,** используемые для текущего оценивания показателя формирования компетенции	Образовательные результаты дисциплины
	Оценочные средства для текущей аттестации ОС-1 Самостоятельная работа ОС-2, ОС-3 Контрольная работа	ОР-1 Применяет знания стандартных и иных методик отбора ОР-2 Владеет базовыми и специализированными методами в области генетического конструирования, ОР-3 Проводит анализ современной литературы в области современной геномики ОР-4 Умеет применять знания в области IT-технологий для решения задач анализа, ОР-5 Владеет навыками разработки ПО, анализа полученных данных, статистической обработки, хранения и документации результатов
	Оценочные средства для промежуточной аттестации ОС-4 Зачет в устной форме	

Описание оценочных средств и необходимого оборудования (демонстрационного материала), а также процедуры и критерии оценивания индикаторов достижения компетенций на различных этапах их формирования в процессе освоения образовательной программы представлены в Фонде оценочных средств для текущего контроля успеваемости и промежуточной аттестации по дисциплине «Компьютерная геномика».

Материалы, используемые для текущего контроля успеваемости обучающихся по дисциплине

Материалы для организации текущей аттестации представлены в п.5 программы.

Материалы, используемые для промежуточного контроля успеваемости обучающихся по дисциплине

ОС-4 Зачет в устной форме
Примерный перечень вопросов к зачету

1. История развития биоинформационных методов.
2. Основные задачи и методы вычислительной геномики.
3. Структура биоинформационных баз данных. Классификация и примеры биоинформационных баз данных.
4. Сравнение последовательностей.
5. Выравнивание двух последовательностей.
6. Алгоритмы локального выравнивания.
7. Множественное выравнивание.
8. Последовательное выравнивание.
9. Поиск гомологов в базах данных, интерпретация результатов, сравнение алгоритмов.
10. Поиск по сходству в базах данных.
11. Эволюция молекул и организмов. Ортологи и паралоги.
12. Горизонтальный перенос.
13. Деревья видов и деревья генов. Филогенетическое дерево как математический объект.
14. Модели эволюции.
15. Алгоритмы построения филогенетических деревьев.
16. Алгоритмические проблемы поиска оптимального дерева, согласование деревьев.
17. Эволюция на уровне генома.
18. Нуклеотидный состав, частые и редкие слова.
19. Статистика ДНК как характеристика генома.
20. Метаболическая реконструкция.
21. Позиционный анализ.
22. Эволюция регуляторных взаимодействий.
23. Эволюция белковых семейств.
24. Подбор праймеров для ПЦР.
25. Анализ данных по экспрессии генов.

В конце изучения дисциплины подводятся итоги работы студентов на лекционных и практических занятиях путем суммирования заработанных баллов в течение семестра.

Критерии оценивания знаний обучающихся по дисциплине

Формирование балльно-рейтинговой оценки работы обучающихся

		Посещение лекций	Посещение практических занятий	Работа на практических занятиях	Зачет
3 семестр	Разбалловка по видам работ	2 х 1=2 баллов	8 х 1=8 баллов	226 балла	64 балла
	Суммарный макс. балл	2 балла max	10 баллов max	236 баллов max	300 баллов max

Критерии оценивания работы обучающегося

	Баллы (3 ЗЕ)
«отлично»	более 271
«хорошо»	211-270
«удовлетворительно»	151-210
«не удовлетворительно»	150 и менее

6. Методические указания для обучающихся по освоению дисциплины

Успешное изучение курса требует от обучающихся посещения лекций, активной работы на лабораторных занятиях, выполнения всех учебных заданий преподавателя, ознакомления с основной и дополнительной литературой.

Запись **лекции** – одна из форм активной самостоятельной работы обучающихся, требующая навыков и умения кратко, схематично, последовательно и логично фиксировать основные положения, выводы, обобщения, формулировки. В конце лекции преподаватель оставляет время (5 минут) для того, чтобы обучающиеся имели возможность задать уточняющие вопросы по изучаемому материалу. Из-за недостаточного количества аудиторных часов некоторые темы не удается осветить в полном объеме, поэтому преподаватель, по своему усмотрению, некоторые вопросы выносит на самостоятельную работу студентов, рекомендуя ту или иную литературу. Кроме этого, для лучшего освоения материала и систематизации знаний по дисциплине, необходимо постоянно разбирать материалы лекций по конспектам и учебным пособиям. В случае необходимости обращаться к преподавателю за консультацией.

Подготовка к семинару (практическому занятию).

Большая часть семинарских (практических) занятий предусматривает изучение материала учебного пособия, хрестоматии, дополнительной литературы (в том числе и материалов периодической печати), подготовку рефератов и сообщений по предложенным вопросам.

Подготовка к практическому занятию, должна основываться на изучении источников и новейших исследований отечественных и зарубежных. Кроме того, практическое занятие может включать и мероприятия по контролю знаний по дисциплине в целом.

При подготовке к практическому занятию обучающийся должен изучить все вопросы, предлагаемые по данной теме, но ответить развернуто может по одному из вопросов, наиболее интересному на его взгляд. При этом обучающийся должен иметь конспект лекций и сделанные конспекты вопросов, рекомендованные для практического занятия.

Подготовка к **устному докладу.**

Доклады делаются по каждой теме с целью проверки теоретических знаний обучающегося, его способности самостоятельно приобретать новые знания, работать с информационными ресурсами и извлекать нужную информацию.

Доклады заслушиваются в начале лабораторного занятия после изучения соответствующей темы. Продолжительность доклада не должна превышать 5 минут. Тему доклада студент выбирает по желанию из предложенного списка.

При подготовке доклада магистрант должен изучить теоретический материал, используя основную и дополнительную литературу, обязательно составить план доклада (перечень рассматриваемых им вопросов, отражающих структуру и последовательность материала), подготовить раздаточный материал или презентацию. План доклада необходимо предварительно согласовать с преподавателем.

Выступление должно строиться свободно, убедительно и аргументировано. Преподаватель следит, чтобы выступление не сводилось к простому воспроизведению текста, не допускается простое чтение составленного конспекта доклада. Выступающий также должен быть готовым к вопросам аудитории и дискуссии.

7. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины

Основная литература

1. Системная компьютерная биология: Монография / Колчанов В.Б., Гончаров С., Лихошвай В.А. - Новосибирск :СО РАН, 2008. - 769 с. ISBN 978-5-7692-0871-3. - Текст : электронный. - URL: https://znanium.com/catalog/product/924675

2. Курносов, М. Г. Вычислительные методы, алгоритмы и аппаратурно-программный инструментарий параллельного моделирования природных процессов: Монография / Курносов М.Г., Хорошевский В.Г. - Новосибирск :СО РАН, 2012. - 355 с. ISBN 978-5-7692-1237-6. - Текст : электронный. - URL: https://znanium.com/catalog/product/924904

Дополнительная литература

1. Гупал, В. М. Математика и загадочный генетический код : монография / В.М. Гупал. — 2-е изд. — Москва : РИОР : ИНФРА-М, 2017. — 288 с. — (Научная мысль). — https://doi.org/10.12737/6032. - ISBN 978-5-369-01276-5. - Текст : электронный. - URL: https://znanium.com/catalog/product/883158

2. Володченкова, Л. А. Биоинформатика: учебное пособие : [16+] / Л. А. Володченкова ; Омский государственный университет им. Ф. М. Достоевского. – Омск : Омский государственный университет им. Ф.М. Достоевского, 2018. – 44 с. : ил. – ISBN 978-5-7779-2214-4 – URL: https://biblioclub.ru/index.php?page=book&id=563147

Интернет-ресурсы

1. https://compbiol.ru – compbiol.ru https://biomolecula.ru – «Биомолекула»
2. https://www.ncbi.nlm.nih.gov – National Center for Biotechnology Information