

«ПРИНЯТО»

На заседании Ученого совета
ФИЦ Биотехнологии РАН
Протокол № 5 от 05.10.2020 г.

«УТВЕРЖДАЮ»

Директор
ФИЦ Биотехнологии РАН



А.Н. Федоров

**РАБОЧАЯ ПРОГРАММА
ПО ДИСЦИПЛИНЕ
ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ДАННЫХ В
БИОИНФОРМАТИКЕ**

Направление подготовки: 06.06.01 Биологические науки

Уровень образования: высшее образование - подготовка кадров
высшей квалификации

Квалификация выпускника: Исследователь. Преподаватель-
исследователь.

Москва

2020 г.

1. Цели и задачи изучения дисциплины

1.1. Цели дисциплины

— дать аспирантам наиболее важные представления о фундаментальных основах организации и функционирования генома, основных процессах, контролирующих экспрессию генов, современных методах исследования регуляции экспрессии генов и реализации наследственной информации, а также о базовых алгоритмах и методах машинного обучения, применяемых для исследования биологических систем.

1.2. Задачи дисциплины

— формирование базовых знаний о структуре генома высших эукариот и механизмах, ответственных за экспрессию генов в конкретной ткани и в конкретное время;

— практическое освоение студентами методов исследования структуры генома и экспрессии генов;

— формирование у аспирантов базовых знаний, в области биоинформатического анализа геномных данных и методов машинного обучения, необходимых для проведения самостоятельных научных исследований в области молекулярной генетики и вычислительной биологии.

2. Место дисциплины в структуре ООП

Интеллектуальный анализ данных в биоинформатике является дисциплиной по выбору вариативной части учебного плана основной образовательной программы.

В рамках данной дисциплины углубляются и развиваются следующие компетенции:

Универсальные компетенции

— Способность к критическому анализу и оценке современных научных достижений, генерированию новых идей при решении исследовательских и практических задач, в том числе в междисциплинарных областях (УК-1);

— Способность проектировать и осуществлять комплексные исследования, в том числе междисциплинарные, на основе целостного системного научного мировоззрения с использованием знаний в области истории и философии науки (УК-2);

— Готовность участвовать в работе российских и международных исследовательских коллективов по решению научных и научно-

образовательных задач (УК-3);

— Готовность использовать современные методы и технологии научной коммуникации на государственном и иностранном языках (УК-4);

— Способность планировать и решать задачи собственного профессионального и личностного развития (УК-5);

Общепрофессиональные компетенции

— Способность самостоятельно осуществлять научно-исследовательскую деятельность в соответствующей профессиональной области с использованием современных методов исследования и информационно-коммуникационных технологий (ОПК-1);

— Готовность к преподавательской деятельности по основным образовательным программам высшего образования (ОПК-2);

Профессиональные компетенции

— Способность к самостоятельному проведению научно-исследовательской работы и получению научных результатов, удовлетворяющих установленным требованиям к содержанию диссертаций на соискание ученой степени кандидата наук по направленности (научной специальности) (ПК-1);

— Владение представлениями о системе фундаментальных понятий и методологических аспектов биологии, форм и методов научного познания (ПК-2);

— Способность приобретать новые знания с использованием современных научных методов и владение ими на уровне, необходимом для решения задач, возникающих при выполнении профессиональных функций (ПК-3);

— Владение методами отбора материала, преподавания и основами управления процессом обучения фундаментальной биологии (ПК-5);

В результате освоения дисциплины обучающиеся должны

знать:

31(УК-1)	методы критического анализа и оценки современных научных достижений, а также методы генерирования новых идей при решении исследовательских и практических задач, в том числе в междисциплинарных областях
31(УК-2)	методы научно- исследовательской деятельности
31(УК-3)	особенности представления результатов научной деятельности в устной и письменной форме при работе в российских и международных исследовательских коллективах
32(УК-4)	стилистические особенности представления результатов научной деятельности в устной и письменной форме на государственном и иностранном языках

31(УК-5)	содержание процесса целеполагания профессионального и личностного развития, его особенности и способы реализации при решении профессиональных задач, исходя из этапов карьерного роста и требований рынка труда
31(ОПК-1)	основной круг проблем (задач), встречающихся в избранной сфере научной деятельности, и основные способы (методы, алгоритмы) их решения
32(ОПК-1)	основные источники и методы поиска научной информации
31(ОПК-2)	нормативно-правовые документы, регламентирующие организацию и содержание образовательного процесса
31(ПК-1)	современное состояние науки в области: (молекулярной биологии; биохимии; биотехнологии, микробиологии, биоинформатики)
32(ПК-1)	порядок организации, планирования и проведения научно-исследовательской работы с использованием современных научно-исследовательских, образовательных и информационных технологий
33(ПК-2)	методы поиска необходимой информации
32(ПК-3)	базовые принципы и основные приемы молекулярной биологии; биохимии; биотехнологии, бионанотехнологии, математической биологии, биоинформатики, микробиологии;
31(ПК-5)	современное состояние науки в области биологических наук

уметь:

У1(УК-1)	анализировать альтернативные варианты решения исследовательских и практических задач и оценивать потенциальные выигрыши/проигрыши реализации этих вариантов
У2(УК-1)	при решении исследовательских и практических задач генерировать новые идеи, поддающиеся операционализации исходя из наличных ресурсов и ограничений
У1(УК-2)	использовать положения и категории философии науки для анализа и оценивания различных фактов и явлений
У1(УК-3)	следовать нормам, принятым в научном общении при работе в российских и международных исследовательских коллективах с целью решения научных и научно-образовательных задач
У1(УК-4)	следовать основным нормам, принятым в научном общении на государственном и иностранном языках
У1(УК-5)	формулировать цели личностного и профессионального развития и условия их достижения, исходя из тенденций развития области профессиональной деятельности, этапов профессионального роста, индивидуально-личностных особенностей.
У1(ОПК-1)	находить (выбирать) наиболее эффективные (методы) решения основных типов проблем (задач), встречающихся в избранной сфере научной деятельности
У2(ОПК-1)	обобщать и систематизировать передовые достижения научной мысли и основные тенденции хозяйственной практики
У3(ОПК-1)	анализировать, систематизировать и усваивать передовой опыт проведения научных исследований
У4(ОПК-1)	собирать, отбирать и использовать необходимые данные и эффективно применять количественные методы их анализа
У2(ОПК-2)	осуществлять отбор и использовать оптимальные методы преподавания оценивания успеваемости обучающихся в области биологических наук
У1(ПК-1)	самостоятельно формулировать конкретные задачи научных исследований и проводить углубленную их разработку

У-1(ПК-2)	использовать современные методы сбора, анализа и обработки научной информации
У-2(ПК-2)	анализировать и систематизировать полученную информацию
У1(ПК-3)	проводить обработку результатов исследований
У1(ПК-5)	преподавать учебные предметы, курсы, дисциплины
У2(ПК-5)	разрабатывать научно- методическое обеспечение реализации курируемых учебных предметов, курсов, дисциплин

владеть:

В1(УК-1)	навыками анализа методологических проблем, возникающих при решении исследовательских и практических задач, в том числе в междисциплинарных областях
В2 (УК-1)	навыками критического анализа и оценки современных научных достижений и результатов деятельности по решению исследовательских и практических задач, в том числе в междисциплинарных областях
В1(УК-2)	навыками анализа основных мировоззренческих и методологических проблем, в т.ч. междисциплинарного характера, возникающих в науке на современном этапе ее развития
В2(УК-2)	технологиями планирования в профессиональной деятельности в сфере научных исследований
В1(УК-3)	навыками анализа основных мировоззренческих методологических проблем, в.т.ч. междисциплинарного характера, возникающих при работе по решению научных и научно-образовательных задач в российских или международных исследовательских коллективах
В4(УК-3)	различными типами коммуникаций при осуществлении работы в российских и международных коллективах по решению научных и научно-образовательных задач
В2(УК-4)	навыками критической оценки эффективности различных методов и технологий научной коммуникации на государственном и иностранном языках
В1(УК-5)	приемами и технологиями целеполагания, реализации целей и оценки результатов деятельности по решению профессиональных задач
В2(УК- 5)	способами выявления и оценки индивидуально-личностных, профессионально-значимых качеств и путями достижения более высокого уровня их развития.
В1(ОПК-1)	современными методами, инструментами и технологией научно-исследовательской и проектной деятельности в определенных областях биологии
В2(ОПК-2)	методиками и технологиями преподавания и оценивания успеваемости обучающихся (биологические науки)
В2 (ПК-1)	методами и приемами экспериментальных исследований в области (молекулярной биологии; биохимии; биотехнологии, биоинформатике, микробиологии)
В1(ПК-2)	методами работы с основными базами данных биологической информации
В1(ПК-3)	навыками использования биологических Интернет-ресурсов
В1(ПК-5)	умениями разрабатывать научно-методическое обеспечение реализации курируемых учебных предметов, курсов, дисциплин

3. Объем дисциплины и виды учебной работы:

Всего часов	Аудиторные занятия (час):	Самостоятельная работа	Всего зачетных единиц
144	36	108	4

4. Распределение аудиторных часов по темам и видам учебной работы

№ п/п	Наименование тем и разделов (с развернутым содержанием курса по каждой теме и разделу)	Аудиторные занятия (час), в том числе:
1	Введение в машинное обучение	2
2	Методы регрессионного анализа	2
3	Методы классификации	2
4	Методы кластеризации	2
5	Нейронные сети и глубокое обучение	2
6	Стандартные статистические методы для планирования эксперимента и обработки результатов	4
7	Технологии секвенирования	4
8	Контроль качества	4
9	Выравнивание коротких фрагментов сиквенса (ридов)	4
10	Аннотация генов и интерпретация результатов	2
11	Транскрипция, сплайсинг, их регуляция. Секвенирование РНК, обработка данных секвенирования РНК.	4
12	Иммунопреципитация хроматина (ChIP-seq)	2
13	Методы определения областей открытого хроматина	2
	Всего	36

5. Содержание курса

Тема 1

Введение в машинное обучение

Задачи машинного обучения, границы применимости методов для анализа биологических данных. Классификация методов машинного обучения.

Тема 2

Методы регрессионного анализа

Линейная регрессия. Оценка регрессионных моделей. Чувствительность, специфичность и аккуратность моделей. Более сложные регрессионные модели.

Тема 3

Методы классификации

Задача классификации и основные алгоритмы. Оценка моделей классификации. Современные методы классификации, используемые для биологических данных.

Тема 4

Методы кластеризации

Методы оценки похожести сложных объектов. Модели и алгоритмы кластеризации. Оценки качества моделей.

Тема 5

Нейронные сети и глубокое обучение

Модель нейрона. Нейронные сети и другие способы выявления нелинейных свойств. Глубокое обучение и задачи распознавания образов.

Тема 6

Стандартные статистические методы для планирования эксперимента и обработки результатов

Воспроизводимость исследований. Выборка. Контроль. Схема эксперимента случай-контроль. Статистическая значимость. Проблема множественного тестирования.

Тема 7

Технологии секвенирования

Обзор методов, основанных на секвенировании нового поколения. Основные этапы.

Тема 8

Контроль качества

Методы биоинформатического контроля качества прочтений. FastQC. Обработка прочтений пед выравниванием, обрезка адатперов, тримминг.

Тема 9

Выравнивание коротких фрагментов сиквенса (ридов)
Алгоритмы выравнивания, оптимизация.

Тема 10

Аннотация генов и интерпретация результатов.

Этапы обработки данных секвенирования РНК. Введение, картирование и подсчёт транскриптомных ридов, проверка качества, нормализация.

Тема 11

Транскрипция, сплайсинг, их регуляция. Секвенирование РНК, обработка данных секвенирования РНК.

Процессинг РНК; донорные и акцепторные сайты сплайсинга; регуляция сплайсинга; альтернативный сплайсинг; энхансеры сплайсинга. Выравнивание прочтений РНК на геном с учетом сплайсинга. Определение новых транскриптов. Дифференциальная экспрессия.

Тема 12

Иммунопреципитация хроматина (ChIP-seq)

Модификации гистоновых белков. Определение районов модификации гистонов. ChIP-Seq, особенности обработки при анализе данных секвенирования.

Тема 13

Методы определения областей открытого хроматина.

Метилирование ДНК и бисульфитное секвенирование. Метилирование ДНК. Полногеномные методы анализа метилирования ДНК. Бисульфитная конверсия. BS-seq и его варианты. Обработка результатов бисульфитного секвенирования.

6. Самостоятельная работа

В процессе освоения дисциплины предусмотрено самостоятельное изучение отдельных вопросов лекционного курса с целью подготовки к устному опросу по теме.

7. Учебно-методическое обеспечение дисциплины

7.1. Основная литература

1. Козлов Н.Н. Математический анализ генетического кода — М. : БИНОМ. Лаборатория знаний, 2017. — 215 с. : ил., [8] с. цв. вкл. — (Математическое моделирование).
2. Кребс Дж., Голдштейн Э., Килпатрик С. Гены по Льюису; пер. 10-го англ. изд. — М. : Лаборатория знаний, 2017. — 919 с. : цв. ил.
3. Альбертс Б., Брей Д., Хопкин К. и др. Основы молекулярной биологии клетки; пер. с англ. — 2-е изд., испр. — М. : Лаборатория знаний, 2018. — 768 с. : ил.
4. Кассимерис Л. [и др.] Клетки по Льюису; пер. 2-го англ. изд. — М. : Лаборатория знаний, 2016.— 1056 с. : цв. ил.
5. Фрешни Р.Я. Культура животных клеток : практическое руководство; пер. 5-го англ. изд. — М. : БИНОМ. Лаборатория знаний, 2017. — 691 с. : ил., [24] с. цв. вкл.

7.2. Дополнительная литература

1. Джералд М. Великая биология. От происхождения жизни до эпигенетики. 250 основных вех в истории биологии; пер. с англ. А. А. Синюшина.—М. : Лаборатория знаний, 2018.—540 с. : ил.
2. Тейлор Д., Грин Н., Стаут У. / Биология: в 3-х томах (комплект) Т. 1; под ред. Р. Сопера ; пер. 3-го англ. изд. — 9-е изд. — М. : Лаборатория знаний, 2018. — 454 с. : ил.

3. Тейлор Д., Грин Н., Стаут У. / Биология: в 3-х томах (комплект) Т. 2; под ред. Р. Сопера ; пер. 3-го англ. изд. — 9-е изд. — М. : Лаборатория знаний, 2018. — 454 с. : ил.
4. Тейлор Д., Грин Н., Стаут У. / Биология: в 3-х томах (комплект) Т. 3; под ред. Р. Сопера ; пер. 3-го англ. изд. — 9-е изд. — М. : Лаборатория знаний, 2018. — 454 с. : ил.
5. Разин С.В., Быстрицкий А.А. Хроматин: упакованный геном — 3-е изд. — М. : БИНОМ. Лаборатория знаний, 2017. — 172 с. : ил., [16] с. цв. вкл.
6. Шмид Р. Наглядная биотехнология и генетическая инженерия; пер. с нем. — М.: БИНОМ. Лаборатория знаний, 2017. — 325 с.: ил.

7.3. Электронные ресурсы

1. <http://www.benran.ru/> - Библиотека по естественным наукам Российской академии наук
2. <https://apps.webofknowledge.com/> - Научно-библиографическая база данных Web of Science.
3. <http://www.scopus.com/> - Научно-библиографическая база данных Scopus.
4. <http://elibrary.ru/> - Научная электронная библиотека НЭБ.
5. <http://www.rsl.ru/> - Электронная библиотека РГБ.
6. <http://www.diss.rsl.ru/> - Электронная библиотека диссертаций РГБ.
7. <http://www.sciencedirect.com/> - Журналы издательства Elsevier.
8. <http://link.springer.com/> - Журналы издательства Springer.
 - a) <http://www.springerprotocols.com> - SpringerProtocols
 - b) <http://www.springermaterials.com> - SpringerMaterials
 - c) <http://www.springerimages.com> - SpringerImages
 - d) <http://www.zentralblatt-math.org/zblmath/en> - Zentralblatt MATH
9. <http://link.springer.com/> - Архивные материалы на платформе Springer.
 - a) Журналы (Journals) 1832-1996 и 2002-2011 гг.
 - b) Журналы (Journals) 1997-2001 гг.
 - c) Книги (Books) 2005-2010 гг., включая книжные серии и справочники.
 - d) Книжные серии (Book Series) 1902-1996 гг.
 - e) Книжные серии (Book Series) 2005-2010 гг.
 - f) Электронные справочники (E-References) 2005-2010 гг.
10. <http://xlink.rsc.org?genre=journal&eissn=1364-548X&date=1996> - Chemical Communications (Cambridge)
11. <http://xlink.rsc.org?genre=journal&eissn=1460-4744&date=1972> - Chemical Society Reviews
12. <http://xlink.rsc.org?genre=journal&eissn=1477-9234&date=2003> - Dalton

Transactions

13. <http://xlink.rsc.org?genre=journal&eissn=1364-5501&date=1991> - Journal of Materials Chemistry
14. <http://xlink.rsc.org?genre=journal&eissn=2050-7496&date=2012> - Journal of Materials Chemistry A
15. <http://xlink.rsc.org?genre=journal&eissn=2050-7518&date=2013> - Journal of Materials Chemistry B
16. <http://xlink.rsc.org?genre=journal&eissn=2050-7534&date=2013> - Journal of Materials Chemistry C
17. <http://xlink.rsc.org?genre=journal&eissn=1463-9084&date=1999> - Physical Chemistry Chemical Physics
18. <http://pubs.rsc.org/en/journals/journalissues/ob#!recentarticles&all> - Organic & Biomolecular Chemistry
19. <http://journals.cambridge.org/> - Журналы издательства Cambridge University Press.
20. <http://www.oxfordjournals.org/en/> - Журналы издательства Oxford University Press.
21. <http://onlinelibrary.wiley.com/> - Журналы издательства Wiley.
22. <http://pubs.acs.org/> - American Chemical Society.
23. <http://www.nature.com/> - Журнал «Nature» (и другие журналы группы Nature).
24. www.sciencemag.org - Журнал «Science».
25. <http://www.fips.ru/> — Патентная база данных РФ (РОСПАТЕНТ).
26. <http://www.uspto.gov/> - Патентная база данных США (USPATFULL).
27. <http://arxiv.org> - arXiv.org/ - международный архив электронных научных статей.
28. <http://www.ccdc.cam.ac.uk/> - Кэмбриджская база структурных данных органических и металлоорганических соединений

7.4. Лицензионное программное обеспечение

1. Office Professional
2. Project Professional
3. Visio Professional
4. Windows
5. Exchange Server Standard CAL - Device CAL

8. Составители программы:

к.б.н. Ю.А. Медведева