

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РФ
МИНИСТЕРСТВО ОБРАЗОВАНИЯ, НАУКИ, КУЛЬТУРЫ И СПОРТА РА
ГОУ ВПО РОССИЙСКО-АРМЯНСКИЙ
УНИВЕРСИТЕТ

Составлена в соответствии с федеральными
Государственными требованиями к структуре
основной профессиональной образовательной
программы послевузовского профессионального
образования (аспирантура)

УТВЕРЖДАЮ:
Проректор по науке
П.С. Аветисян
«___» _____ 2025г.

Институт: Биомедицины и фармации
Кафедра: Биоинженерии, биоинформатики и молекулярной биологии

Учебная программа подготовки аспиранта и соискателя
ДИСЦИПЛИНА: 1.5.8 Математическая биология, биоинформатика

2.1.1 Биоинформатика

Программа одобрена на заседании
кафедры

протокол № 11 от 17.04 2025г.

Утверждена Ученым Советом ИБМиФ

протокол № 19 от 21.04 2025г.

Заведующий кафедрой



д.б.н. Захарян Р.В.

Разработчик программы

д.б.н. Аракелян А.А.

Общие положения

Настоящая рабочая программа обязательной дисциплины (модуля) «**Биоинформатика**» образовательной программы послевузовского профессионального образования (ООП ППО) ориентирована на аспирантов, специализирующихся в области математической биологии, биоинформатики, биостатистики и вычислительной биологии.

Дисциплина предполагает, что обучающиеся уже владеют базовыми знаниями по молекулярной биологии, математическому анализу, программированию, линейной алгебре, теории вероятностей и статистике.

1. Цели изучения дисциплины (модуля)

Целью изучения дисциплины «Биоинформатика» является формирование у аспирантов представления о современных алгоритмах, методах и подходах в биоинформатике, необходимых для анализа, моделирования и интерпретации биологических данных с использованием математических и статистических методов. Особое внимание уделяется интеграции биологических знаний с методами обработки больших данных, машинного обучения, биостатистики и моделирования биологических процессов.

Дисциплина «Биоинформатика» входит в вариативную часть обязательных дисциплин основной образовательной программы подготовки научно-педагогических кадров в аспирантуре по направлению 06.06.01 Биологические науки (научная специальность 03.01.09 — Математическая биология, биоинформатика).

2. Компетенции обучающегося, формируемые в результате освоения дисциплины (модуля)

Аспирант должен

-Знать:

1. Основные типы биологических данных (геномные, транскриптомные, протеомные, метагеномные) и методы их получения.
2. Алгоритмы для выравнивания последовательностей, сборки геномов и аннотации.
3. Методы машинного обучения и статистики, применяемые в биологических исследованиях.
4. Основы построения биологических сетей и моделей взаимодействий.

5. Базы данных биологической информации и форматы представления данных (FASTA, GTF, BAM и др.).

- Уметь:

1. Проводить биоинформатический анализ генетических и транскриптомных данных с использованием современных программных инструментов.
2. Разрабатывать и реализовывать скрипты и пайплайны обработки биологических данных (на Python или R).
3. Применять методы кластеризации, классификации и понижения размерности в анализе биомедицинских данных.
4. Моделировать и анализировать биологические сети и динамику биосистем.
5. Использовать открытые базы данных и вычислительные ресурсы для поиска, извлечения и анализа биологической информации.

- Владеть:

1. Практическими навыками работы с биоинформатическими платформами и пакетами (Bioconductor, Galaxy, scikit-learn, nf-core и др.).
2. Навыками визуализации сложных биомедицинских данных.
3. Методами валидации и интерпретации результатов анализа.
4. Способами воспроизводимого анализа (репродуцируемость, автоматизация, версионирование).
5. Навыками критической оценки научной литературы по тематике биоинформатики и математической биологии.

3. Объем дисциплины (модуля) и количество учебных часов

Вид учебной работы	Кол-во зачетных единиц*/уч. часов
Лекции (минимальный объем теоретических знаний)	10
Семинар	18
Самостоятельная работа аспиранта	8
ИТОГО	36
Вид итогового контроля	Составляющая экзамена кандидатского минимума зачет

4. Содержание дисциплины (модуля)

4.1 Содержание лекционных занятий

№ п/п	Содержание	Кол-во уч. часов
1	Структура биологических данных, роль математических моделей и алгоритмов в биоинформатике.	1
2	Алгоритмы анализа последовательностей. Выравнивание, сборка геномов (de Bruijn, OLC), аннотация.	2
3	Статистические методы и машинное обучение Кластеризация, PCA, регрессия, классификация, модели на основе деревьев, random forest.	2
4	Анализ транскриптомики Дифференциальная экспрессия (DESeq2, EdgeR), RNA-seq пайплайн.	1
5	Моделирование биологических процессов. Построение и анализ сетей (гены–гены, белки–белки), динамическое моделирование.	1
6	Метагеномика и системная биология. Анализ микробиомов, метагеномный состав, функции.	2
7	Репродуцируемые исследования и FAIR-принципы. Git, workflow-менеджеры (Snakemake, Nextflow), публикация данных и пайплайнов.	1
Всего:		10

4.2 Практические занятия

Практические занятия не предусмотрены учебным планом

4.3 Другие виды учебной работы

Другие виды учебной работы не предусмотрены учебным планом.

4.4 Самостоятельная работа аспиранта

№ п/п	Виды самостоятельной работы	Кол-во уч. часов
1	Анализ RNA-seq данных. Построение пайплайна: QC → alignment → quantification → DE analysis	1
2	Решение задач по биостатистике. t-тест, ANOVA, PCA, кластеризация	2

3	Разработка Python/R скрипта. Фильтрация FASTA-файла, визуализация кластеров	1
4	Мини-проект. Сравнение методов классификации (SVM vs Random Forest) на данных TCGA	2
5	Реферирование статей. Краткий обзор современных методов single-cell анализа или network biology	2
Всего:		

5 Перечень контрольных мероприятий и вопросы к экзаменам кандидатского минимума

Контрольные мероприятия по дисциплине

	<i>Вид контроля</i>	<i>Описание</i>	<i>Форма представления</i>
1	<i>Входной контроль</i>	<i>Оценка базовых знаний по молекулярной биологии, программированию и статистике</i>	<i>Тест</i>
2	<i>Мини-проект</i>	<i>Анализ биологических данных (RNA-seq, variant calling, clustering и др.)</i>	<i>Отчет/реферат</i>
3	<i>Презентация</i>	<i>Научная презентация по одной из тем дисциплины или собственной теме в рамках биоинформатики</i>	<i>Слайды и устное выступление</i>
4.	<i>Реферат</i>	<i>Краткий обзор научной статьи по математическим или алгоритмическим аспектам биоинформатики</i>	<i>Реферат</i>

Перечень вопросов к экзаменам кандидатского минимума:

Раздел 1. Алгоритмы и структура данных в биоинформатике

Назовите и охарактеризуйте основные алгоритмы выравнивания последовательностей.

В чем состоит отличие алгоритмов de Bruijn и OLC для сборки геномов?

Какие структуры данных применяются при обработке больших объемов биологических данных?

Опишите основные принципы аннотации геномов.

Что такое формат FASTA и каковы его особенности?

Раздел 2. Статистика и машинное обучение в биологии

Каковы цели и этапы применения методов кластеризации в биоинформатике?

Какие методы понижения размерности применяются в анализе омикс-данных?

Объясните суть методов регрессии и классификации.

В чем разница между SVM и Random Forest при анализе биомедицинских данных?

Как оценивать точность и переобучение моделей в биоинформатике?

Раздел 3. Анализ биологических данных

Опишите полный пайплайн анализа данных RNA-seq.

Что такое дифференциальная экспрессия генов? Какие методы используются?

Как проводится анализ и визуализация биологических сетей?

Как организована работа с данными в биоинформатических workflow-системах (Snakemake, Nextflow)?

Что такое репродуцируемость и FAIR-принципы в анализе данных?

6 Образовательные технологии

В процессе обучения применяются следующие образовательные технологии:

1. Сопровождение лекций показом визуального материала.
2. Проведение лекций с использованием интерактивных методов обучения.

7 Учебно-методическое и информационное обеспечение дисциплины (модуля)

Учебно-методические и библиотечно-информационные ресурсы обеспечивают учебный процесс и гарантируют качественное освоение аспирантом образовательной программы. Университет располагает обширной библиотекой, включающей научно-экономическую литературу, научные журналы и труды научно-практических конференций по основополагающим проблемам науки и практики управления.

7.1. Основная литература:

Compeau, P., Pevzner, P. (2014). *Bioinformatics Algorithms: An Active Learning Approach*

Setubal, J., Meidanis, J. (1997). *Introduction to Computational Molecular Biology*

Lesk, A. (2019). *Introduction to Bioinformatics*. Oxford University Press.

Alberts, B. et al. (2015). *Molecular Biology of the Cell*, 6th ed. Garland Science.

Martinez, W., Martinez, A. (2019). *Statistical Methods in Bioinformatics with R*

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.

7.2. Дополнительная литература

Научные статьи и обзоры из базы данных NCBI Pubmed.

7.3. Интернет-ресурсы

<https://bioconductor.org/> – инструменты и пакеты на R

<https://www.ensembl.org> – геномная аннотация

<https://www.ncbi.nlm.nih.gov> – база данных NCBI

<https://galaxyproject.org> – визуальный интерфейс для биоинформатических анализов

<https://www.scikit-learn.org> – машинное обучение на Python

<https://nextflow.io> – система управления вычислительными пайплайнами

<https://www.snakemake.readthedocs.io> – репродуцируемые биоинформатические пайплайны

8 Материально-техническое обеспечение

Электронные материалы (электронные учебники, учебные пособия, курсы и краткие конспекты лекций, презентации PPT и т.п.).