

公共选修课课程介绍

课程名称	转录组大数据处理与Linux 实战		总学分：2 总学时：34			其中	理论	34	
							实验	0	
开课院部	生命科学与技术学院	教研室	微生物与合成生物学	教学对象	不限				
教材名称	生物信息学	主编	陈 铭	年 版	2022第4版	出版社名称/自编	科学出版社		

课程简介（课程的背景、教学目的、主要内容等，不少于 600 字，）:

课程背景

转录组学作为揭示基因表达调控和生命活动规律的重要手段，在生命科学与医学研究中发挥着核心作用。本课程以“数据驱动与实践结合”为特色，系统介绍转录组大数据的产生、处理与分析方法，重点围绕 RNA-Seq 数据的质控、比对、定量、差异分析、功能注释与可视化展开。通过案例驱动教学，学生将深入理解从原始测序数据到生物学解释的完整流程，掌握主流分析工具的使用和结果解读。

课程同时强调 Linux 系统在大数据处理中的重要性，带领学生从基础命令行入门，逐步过渡到脚本化、自动化的数据处理方法，提升高通量生物信息数据的处理效率。课程采用项目式教学，结合真实转录组数据案例，帮助学生在动手实践中理解数据背后的生物学意义。

通过本课程的学习，学生不仅能够掌握转录组学数据分析的核心技能，还能培养独立解决问题和团队协作的能力，为未来在科研、医药研发和大数据分析等领域的发展奠定坚实基础。

主要内容

本课程结合中国药科大学的学科特色，以新药研发为核心背景，系统介绍生物信息学的基本概念、研究方向、数据分析方法及其在医药领域的应用。课程内容涵盖生物信息学基础、大数据管理、Linux 操作系统、生物信息学数据库、测序技术、生物数据分析方法以及数据可视化等多个方面。通过理论讲解与实践操作相结合的方式，使学生能够在短时间内掌握生物信息学的核心技能，并能够在实际科研中运用所学知识解决生物医学问题。

教学目的

总目标:

本课程面向生命科学、医学、药学及相关交叉学科的学生，旨在系统讲解生物信息学的基本理论、核心方法和数据分析技能。通过理论讲授与实践操作结合，培养学生掌握计算工具、

生物数据处理和可视化分析的能力，并能独立完成转录组数据分析等核心任务。同时，课程融入国家生物信息战略，增强学生的科研诚信意识、数据安全意识和科技自立自强精神，为未来科研与行业应用奠定基础。

课程子目标：

- 1. 掌握生物信息学基本技术（基础知识与技能）
 - 1) 了解生物信息学的基本概念、学科定位及其在生命科学中的重要性。
 - 2) 熟悉转录组学，并掌握常用数据分析方法。
 - 3) 掌握计算机基础知识，包括计算机系统原理、Linux 操作。
- 2. 提升生物数据分析能力（方法与实践）
 - 1) 掌握生物数据的处理与存储方式，包括数据库检索、序列数据格式、云计算环境的使用。
 - 2) 能够独立完成混合转录组（bulk RNA-seq）数据的分析流程，包括数据的获取、质量控制、修建、比对以及定量等上游操作流程。
- 3. 强化科研思维与问题解决能力（综合应用）
 - 1) 通过案例分析和实际操作，培养学生的数据驱动科研能力，提高数据建模、分析和决策能力。
 - 2) 能够运用多组学数据整合方法，解决复杂生命科学问题，并理解人工智能在生物信息学中的应用。
 - 3) 能够独立分析实验数据，并对结果进行科学解释。
- 4. 增强国家科技自立自强意识与数据安全伦理意识（思政目标）
 - 1) 结合我国生物信息学发展历程，引导学生认识生物数据主权的重要性，增强科技自立自强意识。
 - 2) 通过国家生物信息战略、数据安全法规案例，让学生理解数据合规、伦理道德和科研诚信的核心价值。
 - 3) 引导学生关注精准医学、新药研发等领域的社会价值，培养其责任感和科研使命感，促进科技报国精神的形成。

讲授提纲（每一章节的名称）

章节	转录组大数据处理与 Linux 实战	学时
第一章	绪论	3

	<ul style="list-style-type: none">• 生物信息学与转录组学的基本概念• 转录组大数据的特点与挑战• 本课程的定位、学习目标与案例导向教学思路	
第二章	数据科学基础与 Linux 入门	9
	<ul style="list-style-type: none">• 数据科学在生命科学中的角色与价值• Linux 基础命令行操作（文件管理、权限、软件安装）• Shell 脚本初步与批量处理思维• 课程亮点：多种方式领略 Linux 与云计算的魅力（本地电脑、本地服务器、虚拟机、云平台、Docker 等）	
第三章	二代测序原理	2
	<ul style="list-style-type: none">• 测序技术的发展历程• 二代测序技术发展原理详解• RNA-Seq 实验流程与数据格式	
第四章	二代测序（转录组）上游分析实践	9
	<ul style="list-style-type: none">• 数据质控与过滤（FastQC、Trimmomatic 等工具）• 序列比对与转录组组装（HISAT2、StringTie）• 表达量定量（FeatureCounts、Salmon、Kallisto）• 批量处理与自动化分析流程（Pipeline 思维）	
第五章	R 语言初识	3
	<ul style="list-style-type: none">• R 语言环境• R 语言的功能展示• R 在生物信息可视化中的优势与常用包介绍	
第六章	转录组下游分析与图形可视化展示	3
	<ul style="list-style-type: none">• 差异表达分析（DESeq2）• 功能注释与通路富集分析• 可视化输出：火山图、热图、富集图等结果展示	
第七章	国内生物信息中心发展与数据安全	2
	<ul style="list-style-type: none">• 国内主要生物信息数据中心与数据库• 公共数据库与数据共享现状	

结课	• 生物信息数据安全与伦理问题	3
	随堂检测	

考核方式或评分标准（笔试、论文、实际操作考察等）：

序号	成绩类别	考核方式	考核要求	评价权重（%）
1	过程性考核 （平时成绩）	考勤、课堂表现、随堂任务	课堂提问、讨论、小组案例分析作业、编程作业	70%
2	期末成绩	30-40 分钟随堂检测	对重点知识点进行考察	30%

任课教师简介（不少于 50 字）：

夏云，工学博士，中国药科大学生命科学与技术学院副教授、独立 PI，2024 年高层次人才引进至中国药科大学。主要研究方向为生物医学大数据整合、新药发现、肿瘤免疫与生物标志物识别。团队致力于诊断与预后模型开发、肿瘤与免疫分析数据库构建，并已申请多项国内外专利。科研成果发表于 Nature Immunology (Cover)、Briefings in Bioinformatics、Thyroid、iMeta 等国际期刊。曾就职于华中科技大学科学技术发展院，具备全国专利代理师资格，从事生物医药领域的专利挖掘与成果转化工作。