

809 生物信息学基础考试大纲

一、考试性质

生物信息学基础是全国硕士研究生入学考试南京医科大学招收学术型生物医学工程（生物信息学方向）与流行病学（生物信息学方向）专业的研究生而设置的自主命题考试科目，其目的是考查考生是否具备生物信息学研究方向的基本理论知识与基本技能，评价的标准是高等学校生物信息学专业的优秀本科毕业生能达到的及格或及格以上水平，确保本专业的研究生招生质量。

二、考查目标

生物信息学基础的考试范围为基因组学、转录组学、蛋白质组学与蛋白质信息学、表观遗传学、机器学习算法与建模等生物信息学分析必备的基础知识与基本理论，能够运用所掌握的生物信息学理论知识和工具分析和解决临床应用问题。

三、考试形式和试卷结构

1. 试卷满分及考试时间

本试卷满分为 150 分，考试时间为 180 分钟。

2. 答题方式

闭卷、笔试

3. 试卷内容结构

基因组学约 35%；转录组学约 35%；蛋白质组学与蛋白质信息学约 10%；表观遗传学约 10%；机器学习算法与建模约 10%。

4.试卷题型结构

填空题、简答题、综合应用题

四、考查内容

（一）基因组学

1. 常见的序列比对方法和工具；
2. 基因功能注释与富集分析的基本原理与常用软件；
3. 常见基因组变异的类型及基本分析方法；
4. 全基因组测序、全外显子组测序和靶向测序的基本原理及应用；
5. 测序相关数据的格式与数据特征；
6. 基因组测序数据的常见分析方法与应用。

（二）转录组学

1. 表达谱芯片检测原理、常用分析方法与数据资源；
2. RNA-seq 的基本原理、常见分析方法与应用；
3. 转录因子结合位点及靶基因的识别，高通量转录调控研究技术的基本原理、数据分析方法及相关数据资源；
4. 转录因子、非编码 RNA 等介导的转录调控关系及调控网络的构建原理与分析方法；
5. 单细胞转录组测序的基本原理、常见分析方法与应用；
6. 空间转录组技术的基本原理、常见分析方法与应用。

（三）蛋白质组学与蛋白质信息学

1. 蛋白质序列与结构分析，蛋白质二级结构和三级结构预测

的基本原理及常用方法；

2. 蛋白质鉴定和定量的基本原理、常用方法及数据分析；
3. 蛋白质翻译后修饰的主要类型、检测原理及数据分析方法；
4. 蛋白质-蛋白质相互作用的研究方法、常用数据资源及网络分析方法；
5. 单细胞蛋白质组学技术的基本原理及应用。

（四）表观遗传学

1. DNA 甲基化与 CpG 岛，DNA 甲基化的基本原理及常用检测方法；
2. DNA 甲基化对基因转录调控的影响；
3. 核小体与染色质结构，常见组蛋白修饰及其生物学功能和检测方法；
4. 染色质重塑、染色质可及性及其检测方法；
5. DNA 甲基化、组蛋白修饰与染色质状态之间的相互作用及其对基因表达的调控。

（五）机器学习算法与建模

1. 基本概念：有监督学习、无监督学习、半监督学习、训练集、验证集、K 倍交叉证实、过拟合等；
2. 常用机器学习方法的基本原理，包括分类方法（决策树、随机森林、支持向量机、人工神经网络等）和聚类方法（层次聚类、K 均值聚类、K 中心点聚类等）；
3. 机器学习模型构建的一般步骤，包括特征选择、模型训练、

验证与测试；

4. 模型的常用评价指标与性能评估方法。