

(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、表观遗传学、统计遗传学、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，要求考生能够综合运用所掌握的生物信息学理论与工具，分析并解决医学科学研究中的实际问题。

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 30%；转录组学约 30%；蛋白质组学约 10%；表观遗

传学约 10%；统计遗传学约 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. 基因组结构变异的种类；
4. 全基因组测序、靶向测序的原理及方法；
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. 识别转录调控关系的高通量实验方法技术原理、数据分析及其相关数据资源；
4. 转录因子、非编码 RNA 等参与的转录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）蛋白组学

1. 蛋白质二级结构和三级结构预测原理、数据分析、可视化方

法；

2. 蛋白质鉴定和定量的常用方法；
3. 蛋白质翻译后修饰的检测原理及数据分析方法；
4. 蛋白质-蛋白质互作的检测原理、数据资源及数据分析方法；
5. 单细胞蛋白质组学技术的检测原理及应用。

（四）表观遗传学

1. DNA 甲基化与 CpG 岛、DNA 甲基化的检测方法；
2. DNA 甲基化对转录调控的影响；
3. 核小体与组蛋白修饰；
4. DNA 甲基化与组蛋白修饰的相互影响。

（五）统计遗传学

1. 遗传多态性、等位基因的定义和等位基因频率的计算；
2. 基因重组和重组率，连锁和连锁不平衡的计算；
3. 自然选择、中性检验和哈迪-温伯格平衡；
4. 连锁分析与关联分析的常用方法和基本原理。

（六）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（分类方法：决策树、人工神经网络、支持向量机等；聚类方法：层次聚类、K 均值聚类、K 中心点聚类等）；

3. 构建分类模型的一般步骤；

4. 分类模型的评估方法。