《科研思维训练与实践》教学大纲

(适用于四年制生物信息学专业)

前 言

《科研思维训练与实践》课程是根据生物信息学专业本科培养目标开设的一门基础课程。旨在通过讲授生物信息学实际科研中的具体问题和前沿热点知识，培养学生的科研思维方式，掌握正确的科研方法，激发学生的创造精神，培养从事科研所需要的基本素质，为从事科研活动打下良好的基础。

本教学大纲所安排的教学内容是为四年制生物信息学专业的学生而制定。 《科研思维训练与实践》课程为专业必修课，总学时数为18，其中理论课学时为18。

# 生物信息与医学的结合——基因组医学

## 目标要求

* 1. 理解基因组学的原理
  2. 掌握基因组医学的重要应用

## 教学内容

* 1. 基因组学利用DNA测序技术解析和研究全基因组数据,揭示了人类各种疾病的遗传因素,为个体化和精准医疗奠定了基础
  2. 基因组医学的一个重要应用是遗传性疾病筛查和风险评估
  3. 基因组医学的另一个重要应用之一是药物基因组学的发展
  4. 基因组信息为复杂疾病的分子分型和精准治疗提供重要支持

## 教学学时安排

2学时

## 教学方法

多媒体辅助教学

# 从关联到因果：如何探索复杂疾病的遗传致病机制

## 目标要求

* 1. 了解复杂疾病遗传学的基本概念和目前的现状
  2. 理解如何从关联信号中精确定位因果致病变异

## 教学内容

* 1. 介绍复杂疾病遗传学的基本概念和目前的现状。
  2. 提出目前复杂疾病遗传学中最核心的科学问题之一—如何从关联信号中精确定位因果致病变异。
  3. 科学团体（主要是统计遗传学和生物信息学领域的科学家）在回答这个关键科学问题中的艰辛历程和存在的困难。
  4. 当前回答该问题新的研究思维和科研范式的展望。

## 教学学时安排

2学时

## 教学方法

多媒体辅助教学

# Reduced representation bisulfite sequencing (RRBS)实验中限制性内切酶的选择

## 目标要求

* 1. 理解DNA甲基化，Bisulfite Sequencing (BS-seq)技术，RRBS技术
  2. 掌握如何在RRBS实验中的选择限制性内切酶

## 教学内容

* 1. 简要介绍DNA甲基化，BS-seq技术，RRBS技术，以及RRBS相对于BS-seq的优势
  2. 提出问题，引导学生思考如何确定RRBS中的限制性内切酶
  3. 给出几种甲基化敏感性限制性内切酶的类型和识别序列
  4. 介绍基因组模拟酶切（simulation）的思路和具体实现方法
  5. 给出不同内切酶的模拟酶切结果，请学生们讨论选择方案
  6. 请同学们对simulation过程进行总结
  7. 用计算机模拟了哪些实验过程？
  8. 保留了哪些生物学特征？
  9. 在模拟的过程中可能出现哪些误差？
  10. Simulation的其他应用

## 教学学时安排

2学时

## 教学方法

多媒体辅助教学

# 分子进化

## 目标要求

* 1. 理解分子进化中性理论
  2. 掌握Hardy-Weinberg平衡的计算

## 教学内容

* 1. 简要介绍达尔文进化论。
  2. 介绍分子进化中性理论。
  3. 介绍Hardy-Weinberg平衡的计算方法和限制条件

## 教学学时安排

2学时

## 教学方法

多媒体辅助教学

# 蛋白质组学初探

## 目标要求

* 1. 了解蛋白质组学的意义
  2. 理解蛋白质组学的应用

## 教学内容

* 1. 蛋白质组学解析的意义：蛋白质是生命活动的最主要载体，是功能的执行者。生命现象的发生往往是通过多个蛋白质协作完成；蛋白质的参与是交织成网络的（平行或级联因果）；蛋白质执行功能是动态、多样的。
  2. 当前蛋白质组学的主要研究策略和方法

1. 如何鉴定蛋白质的序列？
2. 蛋白质鉴定面临的挑战与困难。
3. 简单介绍基于质谱的蛋白质组学的发展历史、基本原理、研究策略，讲解bottom-up策略的基本分析方法和流程。
   1. 蛋白质组学的应用：结合当前肿瘤蛋白质组学的发展，介绍蛋白质组学在分子分型、肿瘤标志物、药物靶标等领域的应用。
   2. 蛋白质组学的数据发掘
4. 蛋白质组学质谱数据的解析和分析方法。
5. 从质谱数据中寻找“未知”。
6. 搜寻蛋白质组学数据中隐藏的“暗物质”。
7. 未来的展望。

## 教学学时安排

2学时

## 教学方法

多媒体辅助教学

# 单细胞转录组数据分析实践

## 目标要求

* 1. 理解单细胞转录组测序
  2. 掌握单细胞数据分析方法

## 教学内容

* 1. 单细胞数据分析流程

1. 介绍单细胞转录组数据分析的基本流程，包括质控、预处理、聚类、分群、差异表达分析等步骤
2. 解释如何处理和规范化单细胞数据，以及如何应对数据中的噪音和批次效应
3. 学习单细胞数据分析工具Seurat，包括安装、资料检索、使用方法的学习、重要函数的讲解等
   1. 单细胞数据分析实践
4. 使用Seurat、通过实际操作进行聚类、分群、可视化和差异表达分析，深入理解分析过程中的关键步骤
5. 引导学生尝试不同的过滤标准、参数选择、可视化方法、阈值设定
6. 引导学生讨论在单细胞数据分析中的发现，促使他们思考细胞亚群的生物学意义和潜在应用

## 教学学时安排

2学时

## 教学方法

多媒体辅助教学

# 细胞状态的计算模拟以及计算发育生物学

## 目标要求

* 1. 理解发育生物学的基本概念和原理

## 教学内容

* 1. 发育生物学的基本概念以及引入“细胞状态或属性”概念：发育生物学是当今生命科学的重要组成部分，关注器官的正常发育以及发育缺陷。引入干细胞，祖细胞，前体细胞，母细胞，已经成熟细胞等概念。单细胞转录组技术以及表观组技术出来后，科学家可以在数学模型上进一步关注“细胞类型”，“细胞属性”，“细胞可能的过去状态和未来状态”等。
  2. 以造血过程为例，发育生物学的实验科学部分：讲述造血干细胞和血液免疫系统的主要细胞，以及在单细胞图谱上的呈现。
  3. 以造血过程为例，发育计算生物学的算法和实践：以几种常见的模拟方法介绍，如Monocle, Cellrank, Palantir, Dynamo等算法为例。
  4. 其它发育过程举例：肿瘤免疫中的T细胞发育；胰腺发育; 斑马鱼胚胎发育等
  5. Perturbation-seq以及Perturbation cell prediction概念以及流行算法举例：CRISPR技术以及Perturbation-seq；Celloracle预测细胞命运。此环节可能较为深奥，可供学生独立探索和思考。

## 教学学时安排

2学时

## 教学方法

多媒体辅助教学

# 新一代测序技术的应用与挑战

## 目标要求

* 1. 理解新型测序技术面临的问题
  2. 掌握常用的第二代测序技术的生物学应用

## 教学内容

* 1. 介绍常用的第二代测序技术的生物学应用。
  2. 针对不同科学问题，新型测序技术的开发与挑战。

## 教学学时安排

2学时

## 教学方法

多媒体辅助教学

# 序列分析中的人工智能应用

## 目标要求

* 1. 了解基因序列分析的原理

## 教学内容

* 1. 介绍基因序列分析的背景
  2. 介绍人工智能算法在基因序列分析中的应用

## 教学学时安排

2学时

## 教学方法

多媒体辅助教学