

《计算生物学导论》

(Introduction to Computational Biology)

(网址 <http://www2.coe.pku.edu.cn/subpaget.asp?id=25>)

本课程是生物医学工程专业本科生专业课，同时面向生物、化学、计算机、数学、物理及工程类等对计算生物学和生物信息学感兴趣的本科生。计算生物学以现代分子生物学数据作为主要研究对象，发展理论模型和计算方法，揭示以基因组、转录组、蛋白质组以及代谢组为主的生物复杂性，以及生长、发育、遗传、进化、疾病等生命现象的根本规律。本课程主要偏重理论建模和算法的学习，包括概率统计基础、多元统计分析方法、随机过程模型、机器学习方法等，并讨论上述各种方法在实际研究中的应用(包括基因序列信息分析、基因预测、基因芯片的数据分析处理、蛋白质分类和结构预测、分子进化及系统发育树、生物网络等重要问题)。

教学大纲

第1章 绪论

1. 从人类基因组计划说起
2. 计算生物学/生物信息学——源自基因组的科学问题
3. 计算生物学/生物信息学主要研究内容和方法

第2章 计算生物学的生物学基础

1. “What is life?”——从薛定谔的思考到“双螺旋”的发现
2. 生命的演化和分类
3. 生命的分子组成
4. 遗传的分子基础

第3章 分子生物信息数据及数据库

1. 分子生物信息数据库概述
2. 生物信息数据库简介(一级数据库、二级数据库简介)
3. 关于数据库格式
4. 高通量生物信息数据的发展

第4章 计算生物学研究方法 I——概述

1. 计算生物学和生物信息学的三个层面
2. 计算生物学的数学基础(概率理论、数理统计基础、概率模型、代数基础、其他基础)
3. 计算生物学方法总览

第5章 计算生物学的基本问题

1. 基因与基因组学
2. 蛋白质与蛋白质组学
3. 生物大分子结构模拟和药物设计简介
4. 基于高通量测序的计算生物学问题
5. 其他：细胞过程模拟

第 6 章 计算生物学研究方法 II——多元统计分析

1. 多元统计分析概述
2. 回归分析方法
3. 判别分析方法
4. 聚类分析方法

第 7 章 计算生物学研究方法 III——机器学习与人工智能方法

1. 人工智能与机器学习概述
2. Markov 过程的基本概念、隐 Markov 模型 (HMM) 的基本原理、
3. 人工神经网络 (ANN) 方法简介

第 8 章 基因组学的计算生物学研究

1. 基因组学研究简介
2. 人类基因组及人类疾病的计算生物学研究
3. 基因预测的概念及其在基因组研究中的应用
4. 微生物基因预测方法简介
5. 真核生物基因预测方法简介

第 9 章 蛋白质序列与结构的计算生物学研究

1. 蛋白质序列与结构的基本知识
2. 蛋白质序列分析与信号预测
3. 蛋白质结构与蛋白质结构预测
4. 蛋白质结构预测方法简介 (基于蛋白质序列的结构预测方法、蛋白质二级结构预测、三级结构预测)

专题 1 DNA、RNA 和蛋白质序列比对及其应用

1. 序列比对的定义和意义
2. 序列比对方法概述
3. 比对搜索工具 BLAST 简介

专题 2 基因表达的计算生物学研究

1. 基因表达的研究对象和方法
2. 基因芯片研究简介
3. 芯片数据分析

专题 3 分子进化及系统发生学

1. 引言 (基因组进化的复杂性与系统发育分析、分子进化与进化树)
2. 构建系统发育树方法简介 (距离法、最大简约法、最大似然法)

参考教材:

1. Zeelebil 著, 李亦学等译, *理解生物信息学*, 科学出版社, 2012 年
2. Pevsner 著, 孙之荣等译, *生物信息学与功能基因组学*, 化学工业出版社, 2006 年
3. 郝柏林、张淑誉, *生物信息学手册* (第二版), 上海科学技术出版社, 2002 年