

天津医科大学实验课教案首页

(共 2 页、第 1 页)

课程名称：生物信息学		实验名称：DNA 序列的基本信息及特征分析	
教师姓名：伊现富	职称：讲师	教学日期：2013 年 9 月 4 日 14 时 -16 时	
授课对象：生物医学工程学院 2010 级生信班（本）		实验人数：23	
实验类型（验证型、综合型、设计型、创新型）：验证型		实验分组：一人一机	
学时数：2		教材版本：生物信息学实验讲义（自编教材）	

实验目的与要求：

- 掌握使用 NCBI 查询核酸序列的方法。
- 掌握 EMBOSS 的基本使用方法。
- 掌握 ORF 的性质及其分析方法。

实验内容及学时分配：

- (5') 查询核酸序列：简单介绍 NCBI 数据库，讲解 Nucleotide 数据库的使用。
- (5') EMBOSS 简介：介绍 EMBOSS 及其参考资料，讲解 compseq 等工具的使用方法。
- (5') 序列组分分析：回顾序列组分分析的主要内容。
- (5') 开放阅读框分析：回顾 ORF 的定义、相位的概念和最长 ORF 法。
- (80') 实验操作：对人类 CD9 基因序列进行组分分析，对大肠杆菌基因组序列进行 ORF 分析。

主要仪器和实验材料：

- 实验材料：人类 CD9 基因，大肠杆菌基因组。
- 主要仪器：联网的计算机。
- 分析工具：NCBI，EMBOSS。

实验重点、难点及解决策略：

- 重点难点：NCBI 和 EMBOSS 的使用。
- 解决策略：通过演示进行学习，通过练习熟练掌握。

思考题：

- 如何使用 NCBI 查询获取核酸序列？
- EMBOSS 中进行序列组分分析的程序有哪些？
- getorf 和 ORF Finder 的分析结果有何异同？

参考资料：

- NCBI
- EMBOSS

主任签字：

年 月 日

教务处制

一、查询核酸序列 (5 分钟)

1. NCBI: Nucleotide, Gene (多种数据库)
2. Nucleotide: 查询 (ID, 基因名), 下载 (选择合适的格式)

二、EMBOSS 简介 (5 分钟)

1. EMBOSS: EMBOSS Explorer, Jembooss (开源、免费)
2. EMBOSS Explorer: NUCLEIC COMPOSITION, NUCLEIC GENE FINDING (模块化, 功能强大)

三、序列组分分析 (5 分钟)

1. 碱基组成分析: 长度, 碱基数目及其比例, GC 含量
2. 序列转换: 反向序列, 互补序列, 反向互补序列

四、开放阅读框分析 (5 分钟)

1. ORF: 在给定的阅读框架中不包含终止密码子的一串序列
2. 相位: 六相位 (+1, +2, +3, -1, -2, -3)
3. 预测方法: 最长 ORF 法 (适用于原核生物)

五、实验操作 (80 分钟)

1. 人类 CD9 基因的序列组份分析
 - 获取序列: NCBI 的 Nucleotide 数据库, AY422198, FASTA 格式
 - 打开 EMBOSS: EMBOSS Explorer
 - 碱基组分分析: compseq (注意修改默认参数)
 - 计算 GC 含量: geecee
 - 序列转换: revseq (调整参数即可分别获得反向序列、互补序列或反向互补序列)
2. 大肠杆菌基因组序列的 ORF 分析
 - 获取序列: NCBI 的 Nucleotide 数据库, U00096, FASTA 格式
 - 截取序列: extractseq, 1-3000bp (仅使用部分序列进行练习)
 - ORF 预测: getorf (注意选择合适的参数)
 - 结果分析: 和 ORF Finder 的结果进行比较