

# *Content*

1. *Bioinformatics databases*
2. *Sequence alignment and database searching (')*
3. *Phylogenic tree and multiple sequence alignment (')*
4. *Protein structure alignment (')*
5. *Protein secondary structure prediction(')*
6. *Protein tertiary structure prediction (')*
7. *Protein function prediction*
8. *HMM*

# *1. Bioinformatics databases (10')*

- 生物名词解释 ( 生物信息学, 研究方法、研究内容, 研究方向)
- 中心法则
- 有哪些重要的核酸及蛋白质数据库, 分别包含多少数据?

## 2. Sequence alignment and database searching (50')

- *PAM* 与 *BLOSUM* 矩阵的公式异同
- 提供打分方法，手动做动态规划算法（见作业）
- *BLAST* 及 *PSI-BLAST* 算法

### *3. Phylogenetic tree and multiple sequence alignment (15')*

- *UPGMA* 算法
- *CLUSTAL W* 算法
- *Henikoff-Henikoff* 权重
- *Profile-Profile* 打分公式的推导（见作业）

## 4. Protein structure alignment (10')

- *Structure superposition* 与 *Structure alignment* 的区别
- *RMSD* 与 *TM-score* 的定义及各自优缺点
- *TM-align* 算法的基本步骤

## 5. *Protein secondary structure prediction (5')*

- 机器学习方法的基本框架
- *PSI-PRED* 的基本步骤

## *6. Protein tertiary structure prediction ()*

- 蛋白质结构预测的两类方法，各自的优缺点
- 同源建模方法的基本步骤

## 7. *Protein function prediction* (不做要求)

- 蛋白质 - 配体结合位点预测方法
- *GO terms* 预测



## 8. *Hidden Markov Model*

- 前向算法
- *Baum-Welch* 算法
- *Viterbi* 算法

