

Q01: Qname

- prompt: Pairwise sequence alignment 的主要目标是什么？
- model_name: ChatGPT
- model_version: ChatGPT 4o
- response:

Pairwise sequence alignment 的目标是将两个序列进行比对，通过插入 **gaps**（缺口）和匹配/不匹配操作，找到两者之间的最优匹配方式，通常用于计算序列间的相似性，识别同源区域或功能相关性。

Q02: Qname

- prompt: Multiple sequence alignment (MSA) 与 Pairwise sequence alignment 有什么主要区别？
- model_name: ChatGPT
- model_version: ChatGPT 4o
- response:

Pairwise sequence alignment 是针对两条序列进行比对，而 MSA 则是针对多条序列同时比对。

MSA 通常需要考虑所有序列之间的全局关系，难度更高，计算复杂度更大。

Q03: Qname

- prompt: 在对 MSA 进行打分时，常用的 **scoring scheme** 是什么？如何影响结果？
- model_name: ChatGPT
- model_version: ChatGPT 4o
- response:

常用的 **scoring scheme** 包括：

- 匹配分数（**match score**）：给定正分，用于奖励相同字符的匹配。
- 错配分数（**mismatch penalty**）：给定负分，用于惩罚字符之间的不匹配。
- **gap penalty**（缺口惩罚）：分为线性缺口惩罚和非线性（例如 **affine gap penalty**），用于控制缺口插入的代价。

打分方案直接影响对齐的结果。例如，高 **gap penalty** 会减少 **gaps** 的插入，可能导致对齐过于保守；而低 **gap penalty** 则可能导致过多 **gaps**。

Q04: Qname

- prompt: Iterative refined alignment 如何在多个序列对齐中改进结果?
- model_name: ChatGPT
- model_version: ChatGPT 4o
- response:

Iterative refined alignment 通过以下步骤改进对齐结果:

- a. 初始对齐: 使用简单的方法 (如 **progressive alignment**) 生成初始 MSA。
- b. 逐步优化: 反复移除某些序列, 对剩余序列重新比对, 再将移除的序列重新对齐。
- c. 更新: 迭代多次, 直到对齐得分不再显著提高或达到预设阈值。
这种方法通过动态调整和优化早期的错误, 克服了 **progressive alignment** 中误差累积的问题。

Q05: Qname

- prompt: 如何将多序列比对 (MSA) 与进化树 (**phylogenetic tree**) 预测结合, 提升生物序列分析的准确性?
- model_name: ChatGPT
- model_version: ChatGPT 4o
- response:

近年来, 许多研究尝试将 MSA 与进化树预测结合, 主要方法包括:

- 联合优化: 利用算法同时生成多序列比对和进化树, 例如通过贝叶斯方法推导最优的比对和树结构。
- 深度学习辅助: 如使用基于 MSA 的预训练模型预测序列之间的关系, 再反过来优化进化树结构。
- 循环优化: 在初步生成进化树后, 利用树的信息优化 MSA 的权重矩阵, 再更新比对结果, 最终得到更精确的结果。
这类方法在蛋白质序列进化、基因组功能预测等领域具有重要价值。