编程练习题

⽇期：**第10天**

注意，本次题⽬选⾃于⽣命科学学院开设课程**Python**程序设计上机练习，感谢出题⽼师的付出

# 第⼀题：附件说明

**codon.txt**： 包含所有密码⼦信息 **seq.fa**：FASTA格式，提供DNA序列，为⽂本⽂件程序分析：善⽤搜索引擎。

# 第⼆题：问题(C++)

题⽬：

1. 读取codon.txt⽂件，将所有密码⼦信息存储到字典codon\_dict中，键为密码⼦，值为对应的氨基酸缩写（终⽌密码⼦为stop）
2. 定义⼀个函数transcript，输⼊DNA序列，转录成为mRNA序列
3. 定义⼀个函数translate，输⼊DNA序列，先利⽤transcript转录为mRNA，再将mRNA翻译为氨基酸序列，输出氨基酸序列（字符串）

注意事项如下：

* 1. 起始密码⼦为AUG，从第⼀个起始密码⼦开始翻译，起始密码⼦也对应氨基酸，作为氨基酸序列第⼀位。
  2. 如果没有起始密码⼦，return空字符串
  3. 每三个碱基翻译成codon\_dict对应的氨基酸，遇到第⼀个终⽌密码⼦停⽌
  4. 读取seq.fa⽂件读取到字典seq\_dict中，键为序列名（如第⼀条名称为seq1），值为对应的

序列。

* 1. 建⽴字典protein\_dict,键为seq\_dict中序列名称，值为对应序列翻译出的氨基酸序列

（translate函数输出结果）

* 1. 创建⽂件protein.txt,将protein\_dict内容写到该⽂件中。格式如下：

程序分析：题⽬综合性较强，善于利⽤搜索引擎；翻译阶段善⽤字典，以密码⼦为键，以蛋⽩质为值，DNA转RNA同理，C++善用std::map