作业四 Python: Basics

姓名： 学号： 评分：

**请使用python3回答以下问题。**

教材《Python生物信息学数据管理》中的脚本使用的是python2，张老师课件中使用的是python3，写作业时请注意python2与python3之间的区别。

**1)《Python生物信息学数据管理》第1章、第2章所有自测题（共10题）**

**练习1.1计算所有三种组织中的ΔG值**

哪种组织释放的ATP水解能最多？使用课堂上提到的代码来解答这个问题（见PPT第2-6页）。

**练习1.2将值转化为千卡**

以kcal/mol为单位计算三种组织的ΔG值，单位换算因子是：

1 kcal/mol=4.184 kJ/mol。

**练习1.3计算pH**

溶液质子浓度为0.003162mM，求溶液pH值。

**练习1.4指数生长**

在最适合的条件下，单个大肠杆菌可以每20分钟分裂一次。保持最佳条件，求6小时后大肠杆菌数量。

**练习1.5计算单个细菌细胞体积**

已知单个大肠杆菌细胞的平均长度为2.0 μm，直径0.5 μm。假设大肠杆菌为圆柱形，求单个大肠杆菌细胞体积是多少？使用Python进行计算，用变量做参数。

**练习2.1 端粒蛋白质序列中氨基酸出现的频率**

2009年，Elizabeth H. Blackburn，Carol W. Greide和Jack W. Szostak因发现端粒酶功能而获得诺贝尔奖，端粒酶负责延伸染色体的末端。通过NCBI的蛋白质数据库检索人类亚型I的端粒的1132个残基序列，其中哪种氨基酸出现最为频繁？

**练习2.2 DNA序列中核苷酸碱基出现的频率**

更改练习2.1的程序来计算四个DNA碱基频率。用一个你知道输出结果的序列来测试程序如，“AAAACCCGGT”。这种方法可以更轻易地发现小的程序错误。

**练习2.3 一次一个残基地输出氨基酸序列**

写一个打印胰岛素第一个氨基酸的程序，然后是前2个，再然后前3个，依此类推。这个程序将需要range()和len()函数。

**练习2.4删除缩进**

在第2.2.2节中的氨基酸计数示例程序，将

print(amino\_acid, number)

更换为

print(amino\_acid, number, end = “”)

再次运行该程序并解释发生了什么。

**练习2.5 二十个命令，还是一个for循环？**

氨基酸计数的程序可以简单地不用for循环，而使用20个命令计数：

number = insulin.count("A")

print("A", number)

number = insulin.count("C")

print("C", number)

number = insulin.count("D")

...

你更倾向于选择哪种？为什么选择其中某个或者为什么不选另外一个呢？

**2) 第二大题**

针对yeast\_gene.fa文件，用Python3编写代码，输出核苷酸总数和GC含量（GC content）。

提交时请将文件名命名为：学号\_姓名\_py1

截止时间：2021年3月29日 上午10:00