Group Project Instructions
IBI 1, Semester 2, 2019-20

Important deadlines:
Formative assessment: Friday, 3 April 2020.
Summative assessment poster presentation: Thursday, 7 May 2020.

Introduction

This assessment asks you to complete a task with several small sub-tasks relating to analysing nucleotide sequences. You will complete the tasks as a group and present your results in the form of a poster. Groups have been assigned randomly, and the group assignments are posted on Blackboard Learn. You will need to work together on this project despite not being together on campus.

We realise that this way of working is new to you and some of you may find it tough but we will do what we can to help this project run as smoothly as possible. You should be able to contact each other by email using your ZJUID plus '@zju.edu.cn', e.g. 3190119999@zju.edu.cn but feel free to use other means of communication that you find useful – I will be interested to hear what methods you find most useful! Please let me know (robert.young@ed.ac.uk) if you struggle with communication between your group, or any other difficulties with preparing this assessment and I will do what I can to help.

This assessment allows you to practice concepts related to programming and Python that you have encountered before and to use them in a small real-world biological application. It is also an opportunity to practice the planning and execution of a bigger project as a team, and develop your skills around online working, time management, teamwork, and presentation.

Task

1. GC content calculator

GC content, i.e. the fraction of total DNA that is comprised of Guanine and Cytosine nucleotides, varies from species to species. Guanine-Cytosine base pairs use three hydrogen bonds, whereas Adenine-Thymine base pairs use only two. This means that the stability of the DNA double helix and consequently its "melting temperature", i.e. the temperature that is required to separate both strands in the lab, varies depending on GC content. It is therefore useful to be able to compute the GC content.

Write a function that computes the GC content (in % of total bases) for any user-specified DNA sequence.

2. Complementary DNA strand calculator

Write a function that computes the complementary DNA strand (from 5' to 3') for any user-specified DNA sequence (assuming the user has also specified it 5' to 3').

3. DNA to mRNA convertor

Write a function that computes the mRNA sequence form any user-specified DNA sequence.

4. mRNA to protein

Write a function that computes a polypeptide sequence from any user-specified mRNA sequence. For simplicity, you are allowed to assume that the mRNA contains exactly one start codon and one stop codon.

5. Additional function

Write an additional function that takes a nucleotide or polypeptide sequence as input and computes something interesting and/or biologically meaningful. You can be as creative as you like, but we prefer a simple idea that works to a complex idea that is poorly implemented.

For the Formative Assessment

For the Formative Assessment, you do not need to submit any code or document. But what we would like to see is a plan. What functions will you need? What will those functions need to do? Who of you will do what when? What skills/tools will you need? What questions do you have? This formative assessment is not so much a traditional assessment, but more an opportunity for you to self-assess your skill level, to plan the assignment, and to get help.

We will open a discussion board with threads for each group on 3 April and these will remain open for one week until 10 April so that you, as a group, can talk to an instructor. Please upload your plan to your group's thread by 3 April. If cannot upload your file to Learn, please email it to me (robert.young@ed.ac.uk) and I will upload it for you. As a group you should present your ideas on your thread. This can be quite informal. It is also a good opportunity to ask questions.

For the Summative Assessment

This task is assessed in the form of a poster presentation. This has two components:

1. A poster explaining the problem(s), the approach you took, and the results. This can contain screenshots and snippets of code, but it is more important to explain the ideas behind the code than show the actual code itself on the poster.

2. An oral presentation where you explain the content of your poster. This should be about 5-8 minutes. Every member of the group has to participate in the oral presentation.

Your poster should be A1 size (594mm x 841 mm) and can be presented in either portrait or landscape format. The poster presentation should be prepared as a PowerPoint presentation with one slide and an audio recording for each group member. The final poster presentation needs to be agreed by the whole group and submitted by each individual – it must be the SAME presentation from everyone.

Marking criteria will be posted on Learn.

The deadline to submit the summative assessment is Thursday, 7 May 2020.