Genomic 2019 - Homework 2

Dehe Wang

October 30, 2019

Homework

Question 1: Chain termination method

Please read the document "SangerGel.pdf" and answer: What is the template sequence corresponding to the Gel? (5')

Question 2: Four-color fluorescence method

Please read the document "FulGel.pdf" and answer: What is the template sequence corresponding to the Gel? (5') (A: Green, C: Blue, G: Black, T: Red)

Question 3: Assemble

Please assemble the sequence in the file "Assemble.fa" into a complete sequence (10') and determine whether it is a circular DNA. (5') (Sequence length=49)

Question 4: K-mer depth

1. Please calculate the kmer-depth of the sequence in the "GenomeKmerDepth.fa" file and answer, what is the maximum kmer-depth (5') and what is the corresponding kmer (5'). (k=3)

2. Please read the document "GenomeRepeatKmerDepth.pdf" and find out the pictures correspond with random genome (5'), large fragment repeat genome (5') and SINE enrichment genome (5') separately.

Question 5: Rfam database

Please use Rfam web server (`http://rfam.xfam.org`) to identify RNA sequence in the "Rfam.fa" file and answer, what is the Rfam family accession for given sequence (5') and and take a screenshot for RNA Secondary structure page as pdf format (5').

Question 6: Sequencing technology

Please read first section of the second part of "Genomics" (Page 67-93) and answer:

1. Please describe the "four breakthrough" in sequencing technology. (10')

2. Please describe the SBC (5') and SBS (5') method and explain why the SBC method cannot be promoted on a large scale (10').

3. Please explain what is bridge-PCR in your own words according to Illumina's introduction (`https://www.illumina.cn/science/technology/next-generation-sequencing`) (10').

Please answer the first five questions and save them in the "StudentID_Answer.tsv" file. Please write the answers for question six in the given "StudentID_Answer.txt". Please save the screenshot as "StudentID_ScreenShot.pdf" and package the 3 files as "StudentID_HW2.zip". You should answer in the format as in the given example. And you should submit your homework to Course website before Nov. 06, 2019.

Grading

The score S of this homework is calculated by the following formula:

$$S = (S_0 - P) * 0.97^d$$

where S_0 is the total score from all questions, P is a penalty and is obtained from the following table, and d is the late time (in day) calculated by the e-mail event stamp.

Table 1: Penalty list

Contents	Score
File naming error	-5' / file
Packaging error	-5'