

BIOSC 1540 - Computational Biology

Quiz 01

Jan 28, 2025

20 points

Please read the following instructions carefully before beginning your assessment.

- **Time limit:** You have 15 minutes to complete and turn in this assessment.
- **Closed note:** You may not use any notes or additional resources during this assessment.
- **No digital devices:** The use of digital devices, including calculators, is not allowed.

I agree to follow the above instructions. I affirm that all work on this assessment will be my own and that I will not give or receive any unauthorized assistance. To have your assessment graded, you must write your name, sign, and provide your student ID below.

Name

Signature

Student ID

Problem 1

What is the main reason adapters are added to DNA fragments during library preparation for sequencing?

(3 points)

- (A) To prevent contamination of samples.
- (B) To protect DNA from degradation.
- (C) To facilitate binding to the sequencing platform.
- (D) To label fragments with fluorescent markers.

Problem 2

The A260/A280 ratio in UV spectrophotometry helps determine DNA purity, with a ratio close to 0.5 indicating high-quality DNA.

(1 point)

- (A) True
- (B) False

Problem 3

In FastQC quality plots, a high percentage of N base calls near the end of reads often indicates what? Your answer should be at most one sentence.

(2 points)

Problem 4

During Illumina sequencing, why are clusters formed?

(3 points)

- (A) To remove non-specific DNA fragments.
- (B) To amplify the signal from DNA clusters.
- (C) To fragment long DNA molecules.
- (D) To attach fluorescent dyes to nucleotides.

Problem 5

Describe one advantage and one limitation of using Sanger sequencing compared to next-generation sequencing methods. Your answer should be at most two sentences.

(2 points)

Problem 6

Which of the following challenges is most likely to result in fragmented contigs during genome assembly?

(3 points)

- (A) High sequencing depth.
- (B) Repetitive DNA sequences.
- (C) Use of single-ended reads.
- (D) Longer read lengths.

Problem 7

The length of the shortest contig covering 50% of the assembly is called the _____.

Higher values of this metric typically indicate _____ assembly quality.

(3 points)

Problem 8

Explain the main principle of the greedy algorithm for genome assembly. Why is it called “greedy”?

Your answer should be at most two sentences.

(2 points)

Problem 9

Using the following reads: ATCGA, TCGTT, and CGATT, construct a De Bruijn graph with k (overlap) of 3.
(1 point)