

## COMPUTER SCIENCE TRIPOS Part II – 2023 – Paper 9

### 2 Bioinformatics (pl219)

- (a) What is the output of the UPGMA algorithm given the distance matrix of the species  $a, b, c, d, e$  below?

$$\begin{pmatrix} & a & b & c & d & e \\ a & 0 & 16 & 21 & 31 & 20 \\ b & & 0 & 30 & 234 & 21 \\ c & & & 0 & 28 & 38 \\ d & & & & 0 & 42 \\ e & & & & & 0 \end{pmatrix}$$

[4 marks]

- (b) Compute the neighbour joining phylogeny from the four species (s1,s2,s3,s4) DNA sequences.

s1: GATAA

s2: GATAC

s3: CTTTC

s4: CTGGG

[4 marks]

- (c) Compute the Burrows-Wheeler transform on the string PARALLELISM.

[4 marks]

- (d) Discuss the concept of modularity in the Louvain algorithm.

[4 marks]

- (e) Discuss the requirement of well-stirred chemical solution for the Gillespie Algorithm.

[4 marks]