Prof. Dr. Marc-Thorsten Hütt
Prof. Dr. Felix Jonas

Final Exam – INTERMEDIATE PREVIEW 1
October 9, 2024

Note: This preview is approximately 25 percent the size of the actual final; thus you should need below 30 minutes.

Your name:

(1) Here is a small sequence: `ACACA`. Create the suffix tree.

[3 points]

(2) Explain how the posterior probability works in the context of Bayes' theorem. (*Steps: Write down Bayes' theorem. Indicate the posterior probability. Explain the conceptual difference to the likelihood.*)

[5 points]

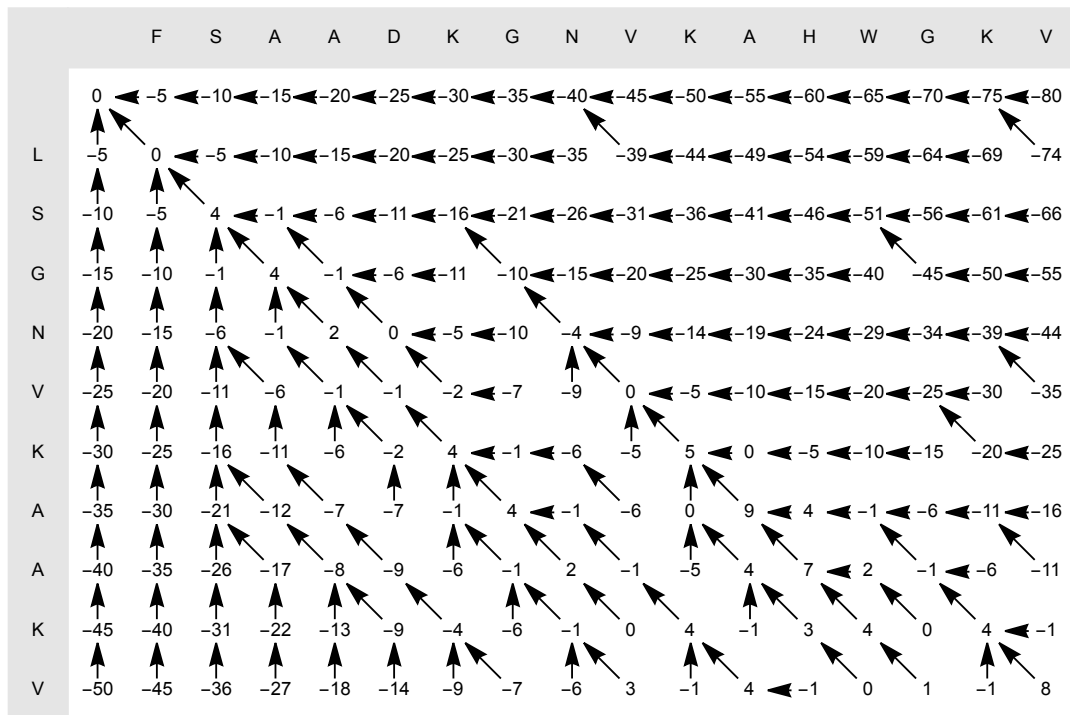**(3)** Here is a short DNA sequence: `ACTCGATACGTGCG`. Compute the observed CpG / expected CPG ratio and explain each quantity.

(4) Think about our brief discussion of the PHYLIP tool for phylogenetic analyses. Which of the matrices below is a valid distance matrix and why?

$$M_1 = \begin{pmatrix} 0. & 0.6 & 0.7 \\ 0.2 & 0. & 0.7 \\ 0.5 & 1. & 0. \end{pmatrix}, \ M_2 = \begin{pmatrix} 0. & 1.6 & 1. \\ 1.6 & 0. & 0.7 \\ 1. & 0.7 & 0. \end{pmatrix}, \ M_3 = \begin{pmatrix} 1.6 & 1. & 1.6 \\ 0.6 & 0.1 & 1.6 \\ 1.1 & 0.8 & 1.5 \end{pmatrix}$$

(5) Here is an F matrix for two sequences computed with one of the two alignment algorithms discussed:

		F	S	A	A	D	K	G	N	V	K	A	H	W	G	K	V
	0	−5	−10	−15	−20	−25	−30	−35	−40	−45	−50	−55	−60	−65	−70	−75	−80
L	−5	0	−5	−10	−15	−20	−25	−30	−35	−39	−44	−49	−54	−59	−64	−69	−74
S	−10	−5	4	−1	−6	−11	−16	−21	−26	−31	−36	−41	−46	−51	−56	−61	−66
G	−15	−10	−1	4	−1	−6	−11	−10	−15	−20	−25	−30	−35	−40	−45	−50	−55
N	−20	−15	−6	−1	2	0	−5	−10	−4	−9	−14	−19	−24	−29	−34	−39	−44
V	−25	−20	−11	−6	−1	−1	−2	−7	−9	0	−5	−10	−15	−20	−25	−30	−35
K	−30	−25	−16	−11	−6	−2	4	−1	−6	−5	5	0	−5	−10	−15	−20	−25
A	−35	−30	−21	−12	−7	−7	−1	4	−1	−6	0	9	4	−1	−6	−11	−16
A	−40	−35	−26	−17	−8	−9	−6	−1	2	−1	−5	4	7	2	−1	−6	−11
K	−45	−40	−31	−22	−13	−9	−4	−6	−1	0	4	−1	3	4	0	4	−1
V	−50	−45	−36	−27	−18	−14	−9	−7	−6	3	−1	4	−1	0	1	−1	8

(a) Which algorithm has been used (Needleman-Wunsch or Smith-Waterman)? How do you see this?

(b) Extract the corresponding optimal alignment.

[5 points]