

Chapter 5: Sequence alignment.

Answers.

1. Aligning with:

AGG WCTTY and BWCTY
 $s(i,i) = 1$, $s(i,j) = 0$, no gap penalty.

	A	G	G	W	C	T	T	Y
G	0	1	1	0	0	0	0	0
W	0	0	1	2	1	1	1	
C	0	0	1	1	3	2	2	2
T	0	0	1	1	2	4	4	3
Y	0	0	1	1	2	3	4	5

of alignments with score 5: 14

2) Same as in 1, but now a gap penalty of 1 (i.e. -1)

	A	G	G	W	C	T	T	Y
G	0	1	1	-1	-1	-1	-1	-1
W	-1	0	0	1	-1	-1	-1	-1
C	-1	-1	0	0	2	0	0	0
T	-1	-1	-1	0	0	3	2	1
Y	-1	-1	-1	-1	0	1	3	3

Assume gaps at the beginning count.

6. The size of the alphabet matters:
 you do better with proteins (20 letters)
 than with DNA (4 letters)

answer: C

7.8.

ATGCA	Series: 3	CTCCA	Series: 2.4	GTAAA	2.6
CTCCA		GTAAA		CTGCA	
ATGCA	2.9	CTCCA	3.6		
GTAAA		CTGCA			
ATGCA	3.3				
CTGCA					

Most dissimilar: CTCCA with GTAAA

Most similar: CTCCA with CTGCA

9. PAM250, just like BLOSUM, reflects
 the properties of amino acids: answer C

10. W and C are very different: W is
 aromatic, bulky, C is small, involved in
 disulfide bridges → answer B

11. All ~~5~~ amino acids (I, L, V, F, V) are hydrophobic
 (non polar) → answer A

12. $P_{\text{value}} \approx E_{\text{value}} = 10^{-4}$ answer C
 P value: probability of a random match with same score