**Exercise 1 - Sequence and Similarity Discussion Questions**

1. How could the similarity between the two passages above be quantified? What must be done prior to determining the similarity of these passages?

2. Considering amino acid residue chemical properties, explain why an Alanine substituted with a Serine is assigned a score of 1, while an Alanine substituted with a Tryptophan is assigned a score of -3 in the BLOSUM 62 substitution matrix.

3. What is the total similarity score for these two sequences? Show your calculations.

Query: MGDVEKGKKIFIMKC

Subject 1: MGEVERGKKLFIMKC

5+6+2+4+

4. If the query sequence is compared to a different sequence (given below), what is the similarity score? Show your calculations.

Query: MGDVEKGKKIFIMKC

Subject 2: MCDVWKGKSIFIMKC

5. Explain why the similarity scores calculated above are different. Consider and refer to information provided in Table 1 as part of your explanation.

6. When the query sequence is compared to itself, a similarity score of 80 is obtained. Considering this, why are the two scores you calculated above different despite having the same number of identical amino acids? Which of the two subject sequences most likely diverged evolutionarily longer ago from the query sequence?

7. Did the computationally calculated similarity scores match those that you manually calculated?