

Name		ID		Section	
------	--	----	--	---------	--

1. Why the origin of replication is typically AT-rich ?
2. In a dataset of 10 sequences, a motif of length 9 is known to occur exactly once per sequence. What is the total number of candidate motifs? (assuming sequence length = 100)
3. What is the profile probability of k-mer 'ATG' given profile: “ Position 1: A=0.2 Position 2: T=0.3 Position 3: G=0.2 “ ?
4. Which of these does NOT improve motif search accuracy?
 - (a) Adding pseudocounts
 - (b) Considering reverse complements
 - (c) Increasing motif length excessively
 - (d) Using more sequences
5. Reverse complements may affect frequency counts of motifs. [True / False]
6. The consensus string is the most frequent k-mer in the dataset. [True / False]
7. The profile matrix stores raw counts only. [True / False]

8. Gibbs sampling is more likely to escape local optima than Greedy Motif Search. [True / False]
9. Which 3-mer has the highest profile probability from profile A:[0.3,0.1,0.4], C:[0.2,0.2,0.3], G:[0.2,0.4,0.1], T:[0.3,0.3,0.2]?
- (a) ATG
 - (b) TCG
 - (c) ACG
 - (d) TTG
10. How many distinct 7-mers are in a 1000 bp genome?
11. Gibbs sampling updates one motif at a time. [True / False]
12. Gibbs sampling can yield different results on the same data. [True / False]
13. If $\text{Skew}[100] = 1$ and $\text{Skew}[101] = 2$, what nucleotide is at position 101?
14. Which one is not a reason for failure in motif search algorithms?
- (a) Motif length
 - (b) No pseudocounts
 - (c) Local optimum convergence
 - (d) Uniform GC content
15. The Hamming distance can be used to compare a motif to a profile. **JUSTIFY.** [True / False]