Assignment 7 Solutions

Question 1
For both the TF and the polymerase matrices, determine the highest affinity binding site. You may find the highest affinity binding site from inspection or write a script. What are the scores for these sites? What sequence(s) produce these scores? Explain how you arrived at your answer.

Transcription factor scoring matrix:
Highest affinity binding site: GCTG(C/T)GCACG
Score: 48

Polymerase scoring matrix
Highest affinity binding site: GCACGGCACG
Score: 62

Question 2
Scan both promoters with both matrices using the appropriate thresholds given above. Paste the commands to do this and their output in your README.

Polymerase
Promoter 1

Command	`python3 scan_sequence.py polymerase_score_matrix.txt promoter1.txt 45`
Output	

orientation	sequence	position	score
forward	CCACGGCACG	102	59

Promoter 2

Command	`python3 scan_sequence.py polymerase_score_matrix.txt promoter2.txt 45`
Output	

orientation	sequence	position	score
forward	GCACGGCACG	186	62

TF

Promoter 1

Command	`python3 scan_sequence.py tf_score_matrix.txt promoter1.txt 40`
Output	

orientation	sequence	position	score
reverse	GCTGTGCAGG	27	43
forward	GCTATGCACG	68	45

Promoter 2

Command	`python3 scan_sequence.py tf_score_matrix.txt promoter2.txt 40`
Output	

orientation	sequence	position	score
forward	GCTGCGCACG	181	48

Knowing that the genes are differentially expressed, which promoter would you expect to be repressed by this TF? Which would be activated? Why?

I would expect transcription by promoter 2 to be repressed by this transcription factor. The high affinity binding site for the transcription factor in promoter 2 overlaps the binding site for polymerase binding at the same promoter and would likely block polymerase access to the promoter/transcription of the following gene. While there are high affinity binding sites for the tf and the polymerase in promoter 1, they do not overlap and are not close to one another.

Comments

The provided code is well commented, i.e., comments were added where it says “TODO: add comments for this code block.” Docstrings were written for functions.