

생물정보학 및 실습 2 실습 보고서 1

2021-20471 박명규

Num	Title	XP	Cost	Due Date	Questions	Solutions
1	Compute the Number of Times a Pattern Appears in a Text	1	1	12월 25, 2022	🗨	🗨
2	Find the Most Frequent Words in a String	1	1	12월 25, 2022	🗨	🗨
3	Find the Reverse Complement of a String	1	1	12월 25, 2022	🗨	🗨

2022년 9월 7일 첫 TA session에서 실습이 어떻게 진행되는지를 익히기 위해 Rosalind에서 연습 문제를 위에서 부터 세 개 해결하였음.

Q1

주어진 문자열에서 특정 패턴이 몇 번 나타나는지 확인하는 문제.

해당 작업을 동일한 문자열에서 패턴을 바꿔가며 반복한다면 아래 Q2처럼 dictionary를 구성하는 것을 선호하지 만, 한 번만 하면 되므로 Rosalind 문제에서 주어진 pseudocode를 Python code화하여 문제를 해결하였다.

Q2

주어진 문자열에서 가장 빈번하게 나타나는 길이 k의 패턴을 찾는 문제.

먼저 주어진 문자열로부터 k개 길이의 모든 substring을 구해서 key로 하고 그 개수를 세어 value로 하는 dictionary를 구성.

만들어진 dictionary에서 가장 value가 큰 key들의 list를 반환하는 식으로 문제를 해결함.

Q3

주어진 염기서열 문자열의 상보적인 염기서열(reversed)을 찾는 문제.

입력의 편의를 위해 각 염기 A, C, G, T를 변수로 선언해두고 상보적인 염기를 value로 하는 dictionary를 구성.

주어진 문자열의 각 character 마다 dictionary에서 상보적인 염기를 찾도록 반복한 후 list를 만들고 reverse하였 다. 마지막에는 문자열로 만들기 위해 ".join 함수 사용함.