

Computational Genomics: Sequences

Homework 1

Srinivas Suresh

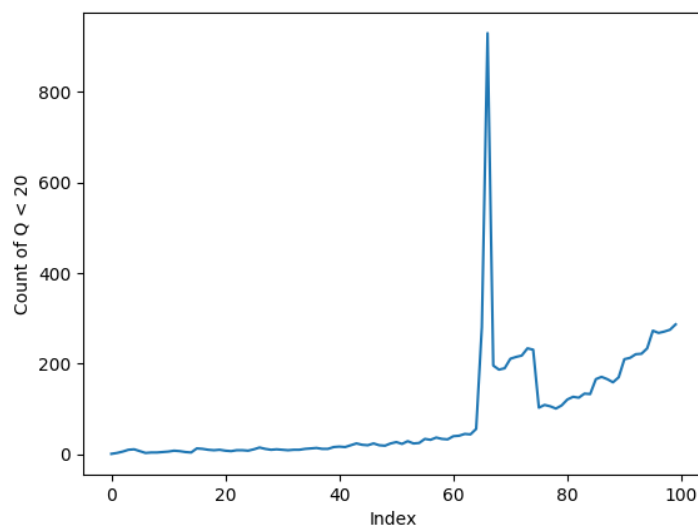
September 27, 2017

Answers

Some answers are instructions on how to run the associated program

1. `answer1.py > inputfile < outputfile`

- (a) Yes, the sequencer could not determine the 67th nucleotide quite often in every read



(b)

There appears to be an increase in bases with $Q < 20$ as the index increases. This tells me that the longer the sequence the lower the quality of bases at the end

2. Winner indicated in **Bold**

- (a) Bad Character Rule: 2 moves, 1 skip
Good Suffix Rule: 7 moves, 6 skips
- (b) **Bad Character Rule: 8 moves, 7 skips**
Good Suffix Rule: 4 moves, 3 skips
- (c) Bad Character Rule: 2 moves, 1 skip
Good Suffix Rule: 11 moves, 10 skips

3. Winner indicated in **Bold**

- (a) Bad Character Rule: 2 moves, 1 skip
Good Suffix Rule: 7 moves, 6 skips
- (b) **Bad Character Rule: 8 moves, 7 skips**
Good Suffix Rule: 8 moves, 7 skips
- (c) Bad Character Rule: 2 moves, 1 skips
Good Suffix Rule: 11 moves, 10 skips