

Homework 8

Chandrima Bhattacharya

7 March 2019

Question 1

Read count defines the number of reads or fragments overlapping with the union of exons of a gene.

Properties of read counts are influenced by the follow.

- They are influenced by the sequencing depth. Usually, deeper sequencing gives higher expression value.
- The length of the gene and GC-bias associated with them are also responsible for difference in read counts value.
- The RNA-composition of the different individual gene abundances are responsible for affecting read counts. Highly-expressed transcripts usually can sometimes dominate reducing the reads count available for the rest of the transcripts. ##### Correction techniques for the following is defined below.
- Sequencing depth needs to be corrected when we are comparing the same gene between different samples. If we are using DESeq2, we can use one of it's function *estimateSizeFactors()* to correct the depth. It uses multiple statistical normalization steps for the following. There are other functions which are defined also for the same!