BIOM200 Module1 Homework1

Assigned Friday 9/23, due to Emily via email by 9AM Monday morning 9/26

1. We talked about many different components that make up a sequencing read. List at least 3 different sequence components that make up the read and explain their function.
2. Pick one sequencing technology that we talked about other than the Illumina HiSeq. Explain how it works and why someone would use it for their experiment.
3. Other than RNA-Seq, explain another experiment that uses high throughput sequencing. Describe the experimental setup, question that it is designed to answer, and how the sequencing data will provide the answer to that question.
4. Download the fastq files for the second replicate of lin28b\_shrna in K562 cells from the ENCODE website (remember to take the files from 8/14/15). Generate a fastqc report and perform an alignment with STAR. Copy the command that you used to perform the STAR alignment below. Write a few sentences detailing your overall workflow. What files did you download/make? Where did you put them and why? What does the output mean? Pick one figure from the fastqc report and paste it below. Explain what it tells you about your data. Report the number of input reads, uniquely mapped read %, and one other interesting output statistic from STAR. Tell me what that other interesting statistic means in terms of understanding your sequencing data.