# Exercise 1

## Instructions -

- **Create a folder called "exercise1" in your WEXAC-mounted classNN folder (testing\classNN\exercise1).**
- **Create a file with your answers called "Exercise_1_2019_answers.docx"**
- **Read the following review and answer the questions below**
- **RNA sequencing: the teenage years**
  https://www.nature.com/articles/s41576-019-0150-2

## Questions

1. Why is ribosomal RNA removal needed for differential gene expression and how is it performed? (detail at least 2 methods)
2. Can RNA-seq (short reads) correctly identify and quantify multiple gene isoforms, why?
3. True or false: the detection of true differential expression between conditions may be improved by filtering genes with uniformly low reads.
4. When will you prefer to sequence long reads?
5. Why is it not recommended to use long-read sequencing methods with degraded RNA?
6. Which method retains epigenetic information?
7. When planning an experiment to identify differential gene expression (between two conditions/groups) using bulk RNA-seq, the most important factor is:
   a. Use of single- or paired-end sequencing reads
   b. Read length
   c. Number of biological replicates in each group
   d. Read depth of 10–30 million reads per sample
8. Which assumptions are made in most computational normalizations?
9. What is a spike-in and what does it attempt to solve?
10. What are the two major considerations \ trade-offs for single-cell RNA-seq experiments and how are they defined?