Functional Genomics - Assignment 2

Oscar Rueda (Oscar.Rueda@cruk.cam.ac.uk)

17th November 2017

Assignment Due: 01/12/17, 23:45

Instructions
• The submission of your report must be done through the moodle site.
• The report must be a single printable PDF (no other formats are accepted)
• The name of the PDF must follow the fga2XXX.pdf format, where XXX is your CRSid (I would save my report as fga2_or249.pdf)
• Clearly state which specific question you are addressing at each stage
• Try to keep the final PDF below 1.5MB
• This course work will consist of 30% towards your overall mark for this module.

Single cell RNA-SEQ Analysis (scRNA-SEQ)

1. Research relevant literature to understand scRNA-Seq. ***In a maximum of 5 pages***, describe:

- What are the main differences between bulk RNA-SEQ and scRNA-SEQ?
    - Biological questions where scRNA-SEQ is more useful than bulk RNA-SEQ and biological questions where bulk RNA-SEQ is still more useful than scRNA-SEQ.
    - What are the main differences in the protocol?
- What is allelic dropout? Explain methods that have been proposed to overcome it.
- Choose three Bioconductor packages for the analysis of scRNA-SEQ and explain the main problems they were designed to solve and the statistical model they use. What do they have in common and what are the differences amongst them?

2. Create a pipeline for the analysis of scRNA-SEQ data using the three packages that you have selected earlier. You can get inspiration from workflows like

https://bioconductor.org/help/course-materials/2017/BioC2017/Day2/Workshops/singleCell/doc/workshop.html

https://www.bioconductor.org/help/workflows/simpleSingleCell/

Try to cover the basic steps of the analysis to answer the questions you think are more relevant. Simulate some data using the package *splatter* to illustrate your pipeline and write a (***maximum 5 pages***) tutorial, similar to the two URLs above (you can use Sweave/knitR/R markdown…)