Short report

2 pages max
Biological context – what will be investigate thanks to these data
Description of the dataset (samples, variables, replicates, etc.)
Biological question you will focus on
Model description allowing to take into account all of the factors that can influence your variables of interest
H0 and H1 hypothesis that you will consider to answer to your questions
Contrats matrix to use in order to test your H0 hypothesis
Description of the list of differentially expressed (DE) genes you expect to obtain
Next steps of analysis based on DE genes lists
No results of these analysis.

Report

8 pages maximum (+1 in case of a very big figure such as a heatmap)

The main part of the report are listed here + their expected content

Introduction : give some background realted to the biological question, present the experiment, introduce your question, it can introduce some hypothesis if you have some biological expectations to test

Material and methods : present the data you will use, present briefly bioinformatic and biostatistic workflow and the statistical model you apply (a chart with the analyis workflow is adviced), the theorical concept behind the functions has to be explained (not detailed, but mention what the function does and the considered hypothesis), make sure you present what you provide to the function and which output you get

Results : present the results with graphical representations, text sections allowing to describe the figure and to summarize the result obtained, pay attention to the legend and the quality of the figures, parameters used to obtained the figures need to be mentionned

Discussion : interpret and discuss the results, be critical regarding the results, the limits of the methods, biological discussion regarding the biological relevance of the results

Conclusion and perspectives : main conclusion of your analysis summarized, what would you suggest to confirm your results, what could be improved (for the experimental and analytical point of view)

If you prefer, discussion parts can be after each of the results (but clearly indicated), in this case, they should be clearly identified and the conclusion should be longer to include a part of general discussion

Oral instructions

The oral presentations take place at the end of the Big Data workshop, and you have already written (or will soon write) several reports explaining your analyses of RNAseq data (bioinformatics and statistics), as well as your analyses of ChIPseq data.

For this oral presentation, we would like you to briefly summarise the question you addressed and your results but also and above all, to take a step back from your work with us. The aim of your presentation is not to repeat the information you already presented (or will present) in your reports. We'd like you to present something new, something more integrative. For example, if you had time to continue your analyses, what would you do? And why? Have you found any consistent elements between your analyses of the RNAseq and ChIPseq data? Can you think of any complementary experiments or analyses? Have you identified any noteworthy elements in the literature that support your hypotheses?

A possible outline could be:

- 1. Intro + biological question(s) + datasets available
- 2. Results you obtained and that you find particularly relevant/important [also in your reports]
- 3. Avenue for new ideas & results stemming from your current conclusions (follow-up questions, new intersections, consistency/inconsistency, complementary investigations...).

The duration of your presentation will be 10 minutes. Ideally, more than half of this time should be devoted to the presentation of results or ideas that are not already in your reports, and we should see a great deal of diversity in the choices that each of you havemade.

It's a difficult exercise. But your studies are coming to an end and soon you will be doing nternships in research laboratories. You can count on our curiosity to ask you questions to which we don't necessarily have the answers.