

Assignments for students

Group Essay

- Write an essay on how next-generation sequencing has been used to research a specific cancer of your choice. What has the impact been? (2000 word limit)

Group Presentation

- Your group will be assigned a cancer database to analyse. Find more about this task in 2 slides.

The multi-omic methods covered

- Bulk DNA analysis
- Bulk RNA analysis
- Proteomics (e.g. western blotting)
- Single-cell and single-nucleus sequencing
- Epigenomic (e.g. ATAC-seq and ChIP-seq)

Essential criteria for assignments

- Essays and presentations must be adequately sourced, and a bibliography should be used.
- References should follow a Nature referencing style:
Author(s). Title of article. *Title of Journal* **Volume number**, Page range (Year).
- Essays and presentations should answer the question/problem.
- Essays and presentations should be formatted (Times New Roman, size 12, justified, 1.5 spacing)
- Images and figures should be sourced and labelled.
- For presentations, all members of the group need to speak.

Data Analysis Presentation Requirements

Mini-assignment:

- During lessons we will be creating quality check (QC) plots .
- In this time, you should find a dataset of your own (with your group) and generate the following QC plots:
 - Average read quality
 - Read frequency
 - Read width distribution
 - Cycle-specific GC content

I expect you to explain what each plot is showing.

Data Analysis Presentation Requirements

- Main-assignment: We will be using the TCGA dataset in Bioconductor.
 1. Filter out the low expression genes.
 2. Normalise data.
 3. Analyse the gene expression of your dataset by producing the following plots:
 - Clustering dendrogram
 - Heatmap showing the most expressed genes
 - Volcano plot

I expect you to explain during your presentation

- What each plot is showing?
- What the number 1 highly expressed gene does. What role does it have in the cancer your dataset is about?

Data Analysis Presentation Requirements

Group datasets for the main-assignment

BRCA (Breast Invasive Carcinoma): Leco, Iris and Kate

LUAD (Lung Adenocarcinoma): Andy and Elaine

COAD (Colon Adenocarcinoma): Calista and Audrey

Help

Datasets are here: <https://rtcg.github.io/RTCGA/index.html>

Information on Bioconductor here: <https://www.bioconductor.org/>

QC plot information here:

<https://www.bioconductor.org/packages/release/bioc/vignettes/Rqc/inst/doc/Rqc.html>

Note:

Please don't be worried about the data analysis presentation yet, I will teach you what you need to know!

Important dates

- **Sunday 21st July:** The group essay must be handed in by the evening.
- **Monday 22nd July:** Group presentations are held on the final day of the summit.

Plagiarism and use of AI

- Plagiarism is strictly prohibited and any students found making use of use of sources without proper citation will not be eligible for awards.
- The use of AI tools such as chatbots and virtual assistants is strictly prohibited, and any students found making use of them will not be eligible for awards.

Assignment assessment criteria

- Understanding of the subject
- Engagement with literature
- Satisfying question asked
- Originality of thought

Assessment awards

- 1 Best Group Presentation
- 1 Best Group Essay
- 1 Outstanding Student
- 1 Academic Excellence