Analysis of High Throughput Sequencing RNA-Seq Data

October 23rd - 24th; October 30th - November 3rd; November 9th

Pre-course material

This will be made available two weeks ahead of the course (October 9th)

- Practice
  - Get access to a terminal, log in with the user: **training**
    (link and password will be provided in updated schedule on Oct 9th)
  - An introduction to the **unix/linux command line**. Do the 10 steps listed on:
    LinuxCommand.org.
  - An introduction to the **R language and environment**. The website requires to register, but the first lesson of the content linked is accessible for "free":
    (link will be provided in updated schedule on Oct 9th)
  - For a more in depth introduction to R, check the official one:
    "An Introduction to R"

Pre-course material (mandatory)

- Read
  - **For Day 1**
  - DNA Sequencing in the last 40 years
  - The original RNA-Seq publication
  - Best practices for RNA-seq data analysis
  - Guideline for RNA-Seq data analysis
  - **For Day 3**
  - Statistical considerations
  - Normalization Method Comparison
  - **For Day 4**
  - Differential Expression Methods Comparison
  - Study size and its effect on analysis

- Watch
  - For Day 1:
    A video on High Throughput Sequencing - Eric Chow from University California San Franscisco - Duration: **32 minutes**
  - For Day 3:
    Consideration about Study Design and statistical background - includes subtitles to get you used to my french accent - Duration: **29 minutes**

Part 1, week 1 - October 23rd - 24th (Liable to change)

Day 1 – 9:00 – 17:00 – Introduction to the Unix Command Line Interface (CLI)
09:00 - 9:30 - Welcome to the course!

Daily Learning Objectives
09:30 - 10:30 - CLI Lecture
10:30 - 10:45 - Coffee break
10:45 - 12:00 - Hands-on practical
12:00 - 13:00 - Lunch break
13:00 - 15:00 - Hands-on continued
15:00 - 15:15 - Coffee break
15:15 - 15:30 - "homework" instructions
15:30 - 15:40 - Feedback
15:40 - 17:00 - Self-study and assessment

Day 2 – 9:00 – 17:00 – Introduction to the R programming language and the R-Studio environment

09:00 - 9:30 - Day 1 Assessment review / revision session
Daily Learning Objectives
09:30 - 10:30 - RStudio environment and R programming lecture
10:30 - 10:45 - Coffee break
10:45 - 12:00 - Hands-on practical
12:00 - 13:00 - Lunch break
13:00 - 15:00 - Hands-on c'ed
15:00 - 15:15 - Coffee break
15:15 - 15:30 - "homework" instructions
15:30 - 15:40 - Feedback
15:40 - 17:00 - Self-study and assessment

Part 1, week 2 - October 30th - November 3rd - 9:00 - 17:00 every day

Day 1 - 9:00 - 13:00 - High Throughput Sequencing, 2nd generation

09:00 - 9:30 - Welcome to the course!
Daily Learning Objectives
09:30 - 10:30 - High Throughput Sequencing Lecture
10:30 - 10:45 - Coffee break
10:45 - 11:15 - Data pre-processing and analysis workflow
11:15 - 11:30 - Workshop data description
11:30 - 12:40 - Data pre-processing journal club
12:40 - 12:50 - "homework" instructions
12:50 - 13:00 - Feedback
In preparation for Day 2 - Self-study and Daily Assessment

Day 2 - 9:00 - 14:00 - Sequencing data pre-processing
09:00 - 10:00- Revision session AND VOTE for Day 5 content!
Daily Learning Objectives
10:00 - 10:30 - Sequence format and data pre-processing
10:30 - 10:45 - Coffee break
10:45 - 12:00 - Hands-on practical data pre-processing continued

12:00 - 13:00 - Lunch break
13:00 - 13:40 - Pseudo-alignment Lecture
13:40 - 13:50 - "homework" instructions
13:50 - 14:00 - Feedback
In preparation for Day 3 - Self-study and Daily Assessment

Day 3 - 9:00 - 14:00 - Statistical concepts and Quality Assessment

09:00 - 10:00 - Revision session
Daily Learning Objectives
10:00 - 10:30 - Statistical concerns - Lecture
10:30 - 10:45 - Coffee break
10:45 - 11:30 - Statistical concerns - Lecture
11:30 - 13:40 - Differential Expression analysis setup - until the Exploratory analysis and visualization (included)
12:00 - 13:00 - Lunch break
13:00 - 13:40 - Differential Expression analysis setup - until the Exploratory analysis and visualization (included) c'ed
13:40 - 13:50 - "homework" instructions
13:50 - 14:00 - Feedback
In preparation for Day 4 - Self-study and Daily Assessment

Day 4 - 12:00 - 17:00 - Differential Expression
12:00 - 13:00 - Revision session
Charlotte's extracted code
Revisited code
Revisited code as a html report
Daily Learning Objectives
13:00 - 14:30 - Lecture Differential Expression
14:30 - 14:45 - Coffee break
14:45 - 16:40 - Differential Expression analysis practical
Differential Expression analysis setup - R code
16:40 - 16:50 - "homework" instructions
16:50 - 17:00 - Feedback
In preparation for Day 5 - Self-study and Daily Assessment

Day 5 - 9:00 - 14:00 - Gene Network Inference*
09:00 - 10:00 - Revision session
- Annotated version of Charlotte's DE tutorial
10:00 - 10:30 - Hands-on to kick off network inferences.
- Data preparation package
10:30 - 10:45 - Coffee break
10:45 - 11:15 - Gene Network Inference Lecture
11:15 - 12:00 - Gene Network Inference Practical
12:00 - 13:00 - Lunch break
13:00 - 13:30 - Gene Network Inference Practical
13:30 - 14:00 - Final Feedback

* This is an exemplary topic; other possible topics are:
Day 5 extra - GSEA Lecture
Day 5 extra - Ensemble GSEA tutorial
Day 5 extra - Machine Learning Lecture
Day 5 extra - Cancer vs. Healthy Machine Learning Prediction
Day 5 extra - Cancer vs. Healthy Machine Learning Classification
Day 5 extra - Single Cell Sequencing lecture
Day 5 extra - Single cell RNA-Seq practical
Day 5 extra - Droplet single cell RNA-Seq practical

Part 2 - November 9th

Day 1 - 9:00 - 17:00

09:00 - 9:30 - Introduction to the course
9:00 - 10:30 - Interactive sessions to address analysis issues
10:30 - 10:45 - Coffee break
10:45 - 12:00 - Interactive sessions to address analysis issues c'ed
12:00 - 13:00 - Lunch break
13:00 - 14:00 - Presentation preparation
14:00 - 16:00 - Mini-symposium with talks and peer-review
16:00 - 16:45 - Feedback from the trainers on the projects
16:45 - 17:00 - Final Feedback