Syllabus
Big data Workshop

Overview

This workshop aimes to present and apply bioinformatics ans statistical tools for the analysis of biological data from high-throughput sequencing. The workshop takes place over 4 weeks with alternating course time and guided exercises (teachers are in the classroom ti help and answer questions), with time for independent work (no teacher in class). During independent work time, collaboration between students is encouraged.

- Introduction to Statistics (september, 3 days) => CC1 Stats
- Bioinformatics (november, 2 weeks) => CC Bioinfo, Report BioInfo
- Statistics (december, 2 weeks) => CC2 Stats, Final report and Oral

Session 1 :
- Bioinformatics (CC and technical report) : 0.2
- Statistics (Homework and short report) : 0.2
- Final report 0.3
- Oral : 0.3

Session 2 :
Oral evaluation

To download R : https://cran.r-project.org/
To download Rstudio : https://www.rstudio.com/categories/rstudio-ide/

Introduction to Statistics

Elodie Marchadier and Christine Dillmann

Session 1.1 : variables, samples and distributions (11 september 2024 9:30-17:30) - EM

- reminder of basic statistics
- R software introduction
- probabilities calculations
- graphical representations

Session 1.2 : Statistical tests (12 september 2024 9:30-17:30) - EM

- Student test of mean conformity
- Mean homogeneity
- Power analysis

Autonomy (13 september 2024 9:30-12:30)

Session 1.3 : Principal component analysis (13 september 2024 14:00-17:30) – CD

- PCA in theory
- PCA in practice (draw an overview of a dataset, assess the quality of a experiment using replicates)

=> CC1 Stats

BioInformatics

Gaëlle Lelandais

Week 1 : Room N0-001 (bat 22 - I2BC)

	9:30-12:00	14:00-17:30
12 november 2024	Introduction and readings RNAseq data analyses *(GL)*	General presentation and visit of the sequencing platform of I2BC *(Céline Hernandez)*
13 november 2024	RNAseq data analyses *(GL)*	RNAseq data analyses *(GL)*
14 november 2024	RNAseq data analyses Autonomous work	DNAseq data analyses *(Fanny Hartmann)* Nicolaï Vavilov (IDEEV)
15 november 2024	Final gene expression matrix *(GL)*	Report writing Autonomous work 18:00 : **Deadline for bioinformatics analysis report**

Week 2 : Room N0-001 (bat 22 - I2BC)

	9:30-12:00	14:00-17:30
18 november 2024	ChIPseq data analyses *(Pierre Grognet and Benoît Moindrot)*	ChIPseq data analyses *(PG-BM)*
19 november 2024	ChIPseq data analyses *(GL)*	ChIPseq data analyses Autonomous work
20 november 2024	ChIPseq data analyses *(PG-BM)*	ChIPseq data analyses *(PG-BM)*
21 november 2024	Report writing Autonomous work	Report writing Autonomous work
22 november 2024	RNAseq report feedback *(GL)* **Deadline for ChIP-seq report**	Basics of using R and RStudio *(GL)*

Statistics

Elodie Marchadier and Christine Dillmann

Week 3 : C101 – PUIO

	9:30-12:30	14:00-17:30
25 november 2024	Session 2.1 : ANOVA 1 *(EM)*	Session 2.2 : ANOVA 2 *(CD)*
26 november 2024	Session 2.3 : Poisson and negative binomial distributions and glm *(EM)*	Session 2.4 : DiCoExpress normalization *(EM)*
27 november 2024	Session 2.5 : DiCoExpress Differential Analysis *(EM)*	Contrasts *(CD)*
28 november 2024	Session 2.6 : GO and enrichement *(CD)*	Autonomy
29 november 2024	Supervised Autonomy *(EM)* 12:00 : **deadline CC2 stats**	Autonomy Correction CD et EM

Week 4 : Room Barbara McClintock (IDEEV)

	9:30-12:30	14:00-17:30
2 december 2024	Session 2.7 : Clustering and correlations (EM)	Autonomy
3 december 2024	Supervised Autonomy *(EM)*	Autonomy
4 december 2024	Supervised Autonomy *(EM)*	Supervised Autonomy *(EM)*
5 december 2024	(Supervised ?) Autonomy *(CD ?)*	Session 2.8 : MixOmics *(EM)*
6 december 2024	Session 2.9 : Experimental design *(EM)*	Autonomy

Conferences

9 december 9:30 : Daniel Gautheret (McClintock – IDEEV)
10 december 9:30 : Mélisande Blein-Nicolas (McClintock – IDEEV)

13 december 17:00 : ***long report***

Oral : 7 january 2024 13:30-17:30 HM1323