Practical data analysis courses for life science













DUKE-NUS (2016-)









EMBL Heidelberg, Germany 12 - 18 May 2019

Application closed



EMBO **Practical Course** 

EMBO EMBO **Practical Course** 

Date: 13 - 18 Nov 2022

Location: EMBL Heidelberg



#### **LOCATION & DATES**



Application closed

Date: 29 May - 3 Jun 2022

EMBL Heidelberg, Germany 30 Sep - 4 Oct 2019

Location: EMBL Heidelberg

# **Infrastructure since 2020**

- Italy: students (15) use their computer as terminals and run their analyses on miniPC (32 GB RAM 1Tb SSD, 6 cores, linux)
- Germany: computer room with 30 workstations (32 Gb RAM, 512Gb SSD, 12 cores, linux)
- Singapore: students (15-20) use their computer as terminals and run their analyses on a local server (512 Gb RAM 128 cores, linux)

# Infrastructure from 2022

- Italy: students will use their computer as terminals and run their analyses on a cloud computing infrastructure (linux).
- Germany: students will use their computer as terminals and run their analyses on a cloud computing infrastructure (linux).
- Singapore: students will use their computer as terminals and run their analyses on a local server (linux).

# **Course characteristics**

- Devoted to life scientists without any previous knowledge in scripting/programming and data analysis.
- Five days theory & hands-on covering RNAseq and scRNAseq.
  - Each day starts with "R pils" and ends with "R exercises"
  - Course starts using only a GUI and ends using only R scripts
- Topics:
  - Lecture on RNAseq and scRNAseq (experimental design and technical critical points) (T)
  - Lecture on reproducibility in bioinformatics (T)
  - Data QC (RNAseq/scRNAseq) (T&P)
  - Data reduction (RNAseq/scRNAseq) (T&P)
  - Differential expression (RNAseq) (T&P)
  - Clustering (scRNAseq) (T&P)
  - Cluster-specific markers detection (scRNAseq) (T&P)
  - Biological features characterization (RNAseq/scRNAseq) (T&P)
  - Revision exercises







Alessandri et al. GigaScience 2019



#### **Course exercises as a real experiment**

- The course exercises are build following the working path of a real experiment.
- Each section of the course is organized in the following way:



#### Course as a cooking show

	<ul><li>exercises</li><li>results</li></ul>	>	exercise1	> >	<ul><li>barcodes.tsv.gz</li><li>features.tsv.gz</li></ul>		
					matrix.mtx.gz		
•••			4SeqGUI				
File Edit ?							
<ul> <li>Tools</li> </ul>	Single Cell ►	Sparse To Dense	e				
MultiQC analysis		Execution:	🔵 sudo 🛛 🧿 docker				
S Vioplot HTML by genes		Counts table:	ero/Desktop/exercise_struct	ure/exercises/e	xercise1/matrix.mtx.gz	Browse	X Cancel
Dense To Sparse		Туре:	10xgenomics 📀				
Sparse To Dense					Execut	e 👌 Reset	😡 Close

