Swiss Institute of Bioinformatics

INTRODUCTION TO SEQUENCING DATA ANALYSIS

Group work

Deepak Tanwar
Geert van Geest

November 19-21, 2025
Adapted from previous year courses

Projects

Project 1: Variant analysis
Project 2: Long read RNA-seq
Project 3: Short-read RNA-seq

Data analysis steps

Go through all the steps performed in the course:
- Quality control
- Trimming
- Alignment
- Visualization

But also:
- Perform counting for estimating gene expression

Data analysis steps (covered so far)

Go through all the steps performed in the course:

- **Quality control**
- **Trimming**
- Alignment
- Visualization

But also:

- Perform counting for estimating gene expression

Projects

Project 1: Variant analysis
Project 2: Long read RNA-seq
Project 3: Short-read RNA-seq

Important

Do not <u>only</u> **perform** the calculations, also to **evaluate** the results

Be **reproducible**!!

Bonus exercise

1. Run clumpify.sh after trimming, and the perform alignment:
   https://www.biostars.org/p/225338/
   - How many reads clumpify.sh removed?
   - Check the alignment rate difference: before and after running clumpify.sh

2. Run Qualimap on aligned files:
   http://qualimap.conesalab.org/doc_html/analysis.html
   - What extra information do you get?

In the afternoon of day 3, all groups will give a **5 minute presention**
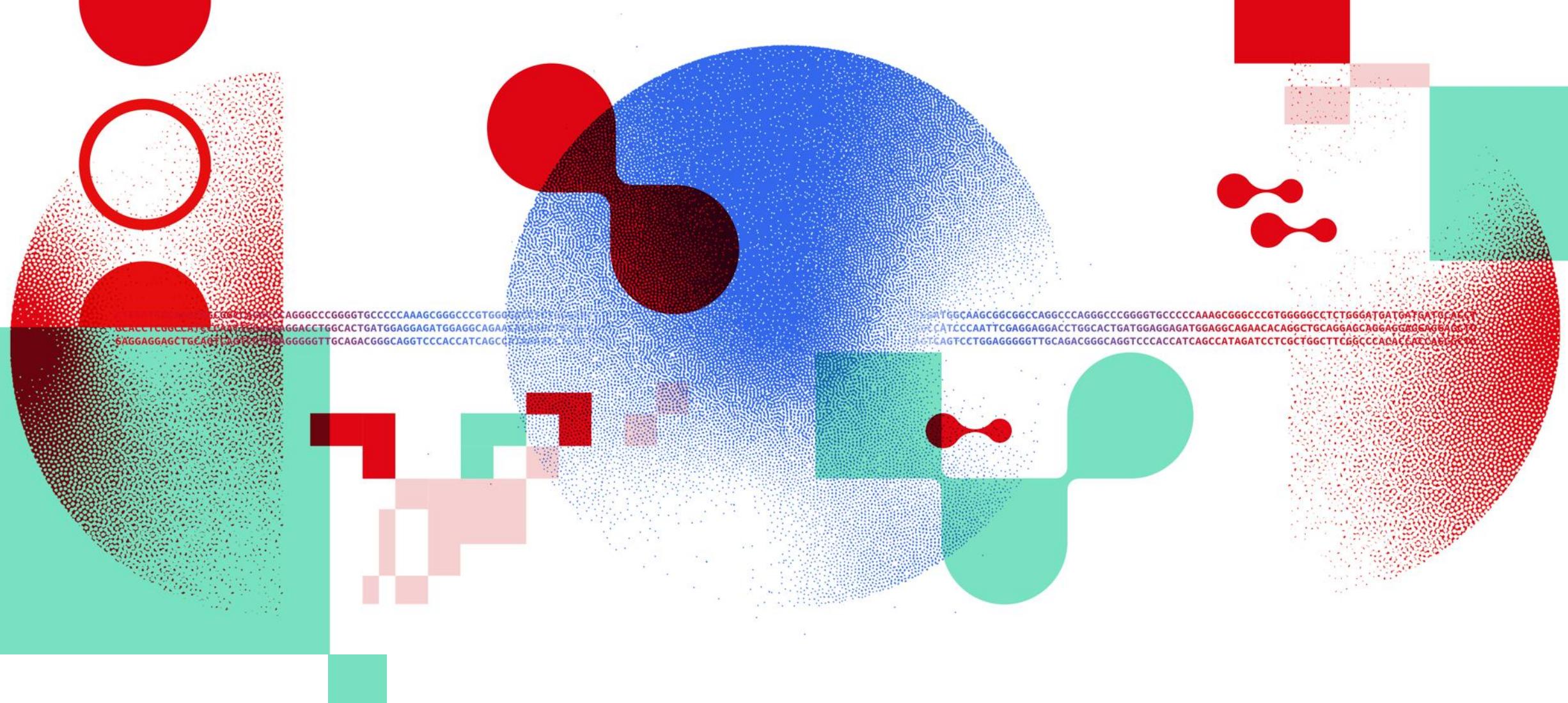
Choose your project

Project 1: Variant analysis

Project 2: Long read RNA-seq

Project 3: Short-read RNA-seq

https://forms.office.com/e/tmDa67QVpK

Thank you

DATA SCIENTISTS FOR LIFE

sib.swiss