Environmental metagenomics

Course outline and practical info

About us

Organizer:

- Carlo Pecoraro, Physalia Courses
  - info@physalia-courses.org
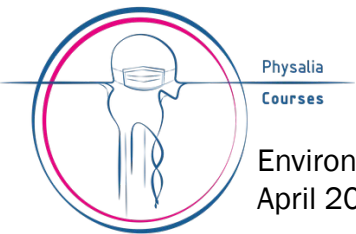
Instructors:

- Antti Karkman, University of Helsinki
  - antti.karkman@helsinki.fi
- Igor Pessi, Finnish Environment Institute
  - igor.pessi@gmail.com
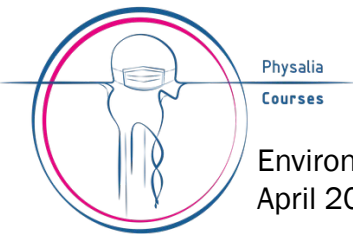
About you

- Name

- University/Institute/Company

- Research interest(s)

- Previous experience(s) with microbial ecology, metagenomics, bioinformatics, etc.

- General hopes for this course
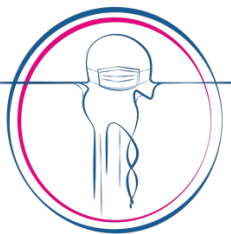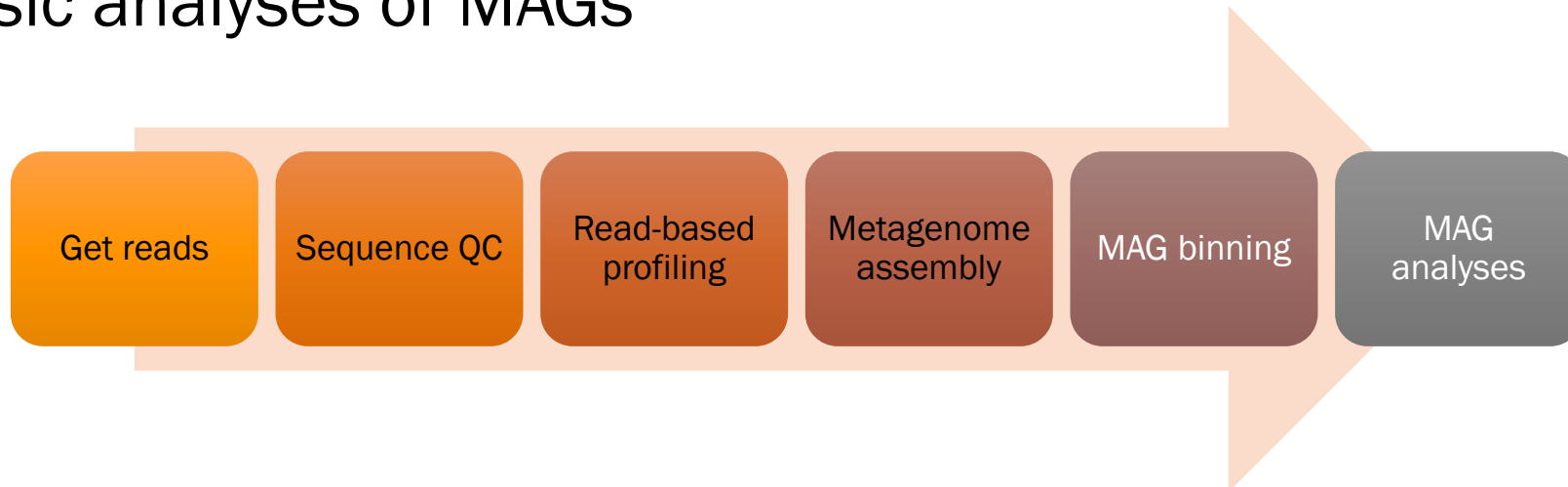
Physalia
Courses

Course outline

Day 1: QC and read-based taxonomic profiling

Day 2: Metagenome assembly

Day 3: Binning of MAGs

Day 4: Binning of MAGs (cont.)

Day 5: Basic analyses of MAGs

Get reads	Sequence QC	Read-based profiling	Metagenome assembly	MAG binning	MAG analyses

Physalia
Courses

Practical information: Zoom & GitHub

The course will take place in Zoom from **9 AM to 4 PM (CET)**

- Link to the Zoom room in Slack

The course page containing exercises and presentations is:
github.com/karkman/Physalia_EnvMetagenomics_2023

Please bookmark this address!

Physalia
Courses

Practical information: Amazon Cloud (AWS EC2)

We will use the Cloud Computing service from Amazon, which we will access via `ssh` (secure shell protocol)

See [here](#) for information on how to connect, but remember:

- The IP address changes every day

- Everyone is given a username, with a `home` and `shared` folders
  - List of usernames can be found in Slack
  - The `shared` folder is copy-only: do not delete, move, rename, or write
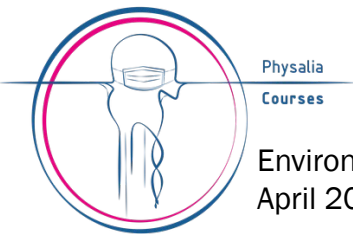
Practical information: setting up VS Code

- Download and install VS Code: code.visualstudio.com/Download

- Launch VS Code

- Go to View -> Extensions

- Search for and install the extension Remote–SSH

- See here for a step-by-step guide on how to connect to the Amazon Cloud
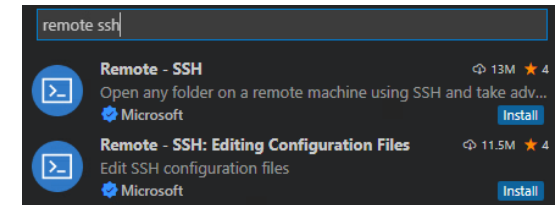
Practical information: conda

System for software management (python, R, JavaScript, C++, …)

Allows easy installation of software in dedicated environments, separated from the main environment and other conda environments

- The environments that we will use have been already set up for everyone

General conda commands

```
> conda env list
> conda activate ambiente
> conda deactivate
```
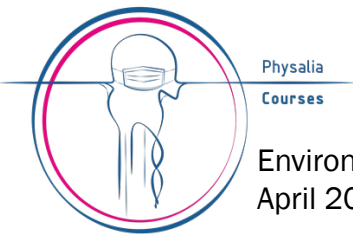
Physalia
Courses

Course data

In this edition of the course, you will get to choose between two different metagenomic datasets:

Tundra soils

vs.

WWTP

REAL (not toy) data!

Physalia
Courses

Tundra soils
Kilpisjärvi, Finnish Arctic (69°N)

Physalia
Courses

Tundra soils
Kilpisjärvi, Finnish Arctic (69°N)

Fen soils

- 1 Nanopore sample
- 4 Illumina NovaSeq samples

Learn more:

- [10.1186/s40793-022-00424-2](#)
- [10.1093/femsmc/xtac019](#)

Waste water treatment plant (WWTP)

Activated sludge from a Danish WWTP

- 1 Nanopore sample
- 3 Illumina samples

nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-021-22203-2   OPEN

Connecting structure to function with the recovery of over 1000 high-quality metagenome-assembled genomes from activated sludge using long-read sequencing

Caitlin M. Singleton[1], Francesca Petriglieri[1], Jannie M. Kristensen[1], Rasmus H. Kirkegaard[1], Thomas Y. Michaelsen[1], Martin H. Andersen[1], Zivile Kondrotaite[1], Søren M. Karst[1], Morten S. Dueholm[1], Per H. Nielsen[1] & Mads Albertsen[1]