Bacterial WGS training : Exercise 2

Title	**Nextflow and Micromamba**
Training dataset:	None
Questions:	• How do I load an environment with installed software with micromamba? • How do I run a nextflow pipeline?
Objectives:	• Learn how to use micromamba. • Learn how to use nextflow.
Time estimation:	15 min

‼️ Important things to remnenber:

- Use Tab to automatically complete file names and paths, so it can be easiert to write in the terminal
- Use keyboard arrows (⬆️⬇️) to move through your terminal's history, so you don't have to write the commands again.
- Try not to use spaces, accents or special characters like "Ñ" letter, when writting directory of file names.
- Basic commands you should always remember: *pwd cd ls mkdir mv rm rmdir less nano*

Go to home, just to be sure everyone is in the same folder

```
cd
pwd
#Output: /home/alumno
ls
```

Go to the exercise folder

```
cd wgs
ls
cd bacterial_wgs_training
ls
#vemos main.nf environment.yml
```

Micromamba

micromamba is a tiny version of the mamba package manager. micromamba supports a subset of all mamba or conda commands and implements a command line interface from scratch. micromamba is a package

manager that enables the installation of any type of software, in this case, bioinformatics software.

The `environment.yml` allows installing conda environments.

```
cat environment.yml
# We see all the software that we are going to use in the course
# With micromamba install -f environment.yml, an environment is created
with everything we will need for the course
```

```
micromamba env list
```

```
fastp --help
micromamba activate fastp
#Se carga el ambiente entre parentesis delante del nombre de usuario
fastp --help
```

```
micromamba deactivate
```

How do I use Nextflow?

```
micromamba activate nextflow
nextflow info
```

```
nextflow run main.nf --help
```

So, what now? In order to execute a nextflow pipeline, we need to tell it to `run` a project which contains a `main.nf` script written in groovy + the pipeline languages:

```
rm results/trace.txt
nextflow run /home/$USER/wgs/bacterial_wgs_training --help
```

Optionally, we can pass a config file, and specify the .nf script inside a project:

```
rm results/trace.txt
nextflow -C /home/$USER/wgs/bacterial_wgs_training/nextflow.config \
run /home/$USER/wgs/bacterial_wgs_training/main.nf --help
```

Finally, let's ask how to use the pipeline:

```
rm results/trace.txt
nextflow run BU-ISCIII/bacterial_wgs_training -r one_week_format --help
```

There is one big detail left. The software needed to execute the pipeline is no installed in our machine. Thankfully, we have a conda environment ready for this course, and our pipeline has already being configurated to know where to find it and how to use it. Use the right argument and go for it:

```
rm results/trace.txt
nextflow run BU-ISCIII/bacterial_wgs_training -r one_week_format -profile
conda --help
```