

Practical in Bioinformatics

Daniel Jouvance - UPMC- UCY International School in Marine Biology
Nicosia, 29/10-05/11/2017



Vasilis J Promponas

Bioinformatics Research Laboratory
Department of Biological Sciences
University of Cyprus

web: <http://troodos.biol.ucy.ac.cy>
email: vprobon@ucy.ac.cy

Getting Started (5 minutes)

0. Please remember, no food or drinks allowed in the lab (*exception: a bottle of water*).
1. Log into your terminals using the credentials provided during the morning lecture.
2. Spend a couple of minutes to get accustomed to the computing environment.
3. Use the folder named “International School in Marine Biology” located on your Desktop to store files related to your work.

Notes:

- a. Use a USB stick (or self email) to store data files or results you may want to take back home
- b. The accounts will remain active till Monday, 06/11/2017; feel free to revisit the lab if necessary, but take a note of scheduled activities (seen on the sheet at the lab entrance).

The EBI metagenomics portal (25 minutes)

1. Point your web browser to <https://www.ebi.ac.uk/metagenomics>
2. Spend a few minutes to see what kind of information is available.
3. Use the ‘Search’ function to locate data related to a metagenomics project for the analysis of soil microorganisms (hint: Use project id **DRS001056**).
4. Open a new tab to search the NCBI Entrez resources (<https://www.ncbi.nlm.nih.gov/>) that refer to the specific metagenomic sample. Then, try to answer the following questions:
 - 4a. Which is the identifier of the project this sample was taken from?
 - 4b. When were data from this project deposited in the database?
 - 4c. Is there any relevant publication accompanying these data?
 - 4d. How many samples were sequenced under this project?
5. Go back to the EBI metagenomics portal and explore the data related to this sample. Comment on the “Quality control” and “Taxonomic Analysis” therein.

The SILVAngs portal (20 minutes)

1. Point your web browser to <https://www.arb-silva.de/ngs/Index.htm>
2. Chose to start the “demo” session (registration is required to submit your own data).
3. Choose the “Shared Projects” option, to see a pre-run example set up by the SILVAngs administrators. Note: Do not worry about deleting anything, you only have *read-only* permission.
4. Take a few minutes to explore the environment and available options. Then try to answer the following questions:
 - 4a. How many sequence files have been submitted for this toy project?
 - 4b. How many sequences are stored in the file with the longest (on average) sequences?
 - 4c. Observe the taxonomic distribution of different samples.

Short break

The MG-RAST portal (20 minutes)

1. Point your web browser to <http://metagenomics.anl.gov>
2. Spend a few minutes to see what kind of information is available.
3. Create an analysis of an existing dataset (choose a marine benthic biome studied by whole genome shotgun sequencing for this case).
4. Wait a few moments for performing the analysis, then explore the results.
5. Identify archaeal OTUs identified in the dataset.
6. Inspect the accompanying metadata.
7. Bonus question: can you identify any experimental data in MG-RAST related to [Cyprus]?

The IMG/M portal (20 minutes) – Demo only (not fully functional with Internet Explorer)