Computational Biologist in building microeucaryotes reference transcriptomic database.

<u>Contacts:</u> E. Corre (corre@sb-roscoff.fr)

In the context of ELIXIR Excelerate WP6 Use Case dedicated to the « Marine metagenomic infrastructure as driver for research and industrial innovation » (1), the ABiMS bioinformatics platform of the Roscoff Marine Station offer an **18-month junior bioanalyst position** in marine microeucaryotes reference transcriptomic databases creation and analysis.

The main objective of this work package is to build a stable and sustainable infrastructure for marine metagenomics information. At the heart of this are the data resources, tools, pipelines and services offered to the marine community, providing stable and sustainable access to computational and reference data resources. These resources vary from archives or deposition databases that contain research data outputs such as DNA sequences, to highly dynamic knowledge bases that aggregate and process research data, adding layers of value through manual curation by highly qualified personnel to complex and scalable metagenomics analysis pipelines.

The bioanalyst will worked on the establishment of the marine micro-eukaryotic reference database. This database is complementary to the marine prokaryotic reference DB produced within this WP. Unlike marine prokaryotes, very little genomic data from marine protists are available so far, and the database will therefore be based, for its first version, exclusively on available trancriptomic data coming either from MMETSP (2), RCC (3) /Tara oceans project (4) and data sets of marine protists present in public databases (ENA / SRA).
Candidate will constitute the database using standardized cleaning and assembly procedure. In close connection with Genoscope CEA in France and the EBI metagenomics group in the UK, the candidate will contribute to the construction of the pipeline dedicated to automated retrieving, cleaning, assembly, evaluation and annotation of the dataset using tools operated through the Common Workflow Language (CWL).

Additional works including integration with sequence search tools in Tara dataset (sequence similarity with gene catalogs, use of occurrence data, search in k-mer space, etc…) will be considered.

Some of the following skills are expected:
- Proficiency in transcriptomic analysis.
- Knowledge on metagenome/metatranscriptomes analysis
- A track record in marine ecology is a plus
- Familiarity with biological & statistical software packages for high-throughput data analysis (e.g. gene annotation programs, microbial ecology sequencing analysis softwares, R)
- Expertise in building bioinformatics pipelines and NGS data analysis
- Good scripting/programming skills (Python, bash) proficiency in Unix/Linux and cluster expericence
- Capacity for teamwork and interest in multidisciplinary approaches
- Ability to communicate technical information effectively, both orally and in writing.

Interested candidates are encouraged to send their CV, recommendation letters, along with a letter stating their interest

Localization :
Plateforme ABiMS (Analyses and Bioinformatics for Marine Science)
Station Biologique de Roscoff
Place Georges Teissier
29680 Roscoff
France

http://abims.sb-roscoff.fr/

Starting date: first semester 2018.

(1) (https://www.elixir-europe.org/about-us/how-funded/eu-projects/excelerate/wp6)
(2) https://www.ncbi.nlm.nih.gov/pubmed/24959919
(3) http://roscoff-culture-collection.org
(4) http://oceans.taraexpeditions.org