

## Master project 2021-2022

### Personal Information

<b>Supervisor</b>	Javier del Campo
<b>Email</b>	<a href="mailto:jdelcampo@ibe.upf-csic.es">jdelcampo@ibe.upf-csic.es</a>
<b>Institution</b>	Institut de Biologia Evolutiva (CSIC-UPF)
<b>Website</b>	<a href="http://delcampolab.com">delcampolab.com</a>
<b>Group</b>	del Campo Lab. Microbial Ecology and Evolution

### Project

## Computational genomics

#### Project Title:

A Ribosomal Operon Reference Database

#### Keywords:

rrn, metabarcoding, eDNA, microbiome, biomonitoring

#### Summary:

The laboratory The del Campo Lab is based at the Institut de Biologia Evolutiva (UPF-CSIC) in Barcelona. The research at the del Campo Lab is focused on the study of host-associated microbes and the effect of global warming on the microbiomes of benthic and planktonic marine animals. We have a wet and dry lab, to perform experiments and bioinformatics analysis, enabling the broadest possible goals. The ongoing climate change and its effects on the environment, such as rising sea temperature, has strong impacts on free-living marine microbial communities. However, the effects of global warming have not been properly studied on host-associated microbiomes. Microbiomes (both prokaryotic and eukaryotic) associated with host organisms have a strong influence on host evolution, physiology, and ecological functions. We study how environmental changes resulting from global warming affect the composition and function of the microbiomes in key members of the marine fauna and consequently how these changes affect the hosts. Currently, our study focuses on these impacts on corals, teleost fish, and zooplankton. To tackle this novel research topic, we use a combination of molecular biology, ecophysiology, and bioinformatics. The proposed project Metabarcoding has been for many years a useful approach to study the diversity and distribution of micro and macroorganisms across environments. Furthermore, metabarcoding is currently being implemented successfully as a biomonitoring tool. This methodology is used for diagnosis of microbial pathogens, to study the health of lakes, rivers and beaches, to track the presence of invasive or endangered species, etc. However, the current metabarcoding methodologies present certain limitations, being the most significant the lack of phylogenetic resolution. The most popular metabarcoding approach is the use of short read barcodes generated using Illumina. These fragments, that are commonly not longer than 400 bp, despite providing very useful information cannot reach the level of detail that would allow us to infer from them species or strain identities (the latest in the case of microbes). We propose the use of the whole rRNA operon (rrn) as a barcode for life. Many fragments of the rrn such as the 16S and 23S in bacteria, the 18S, ITS1, ITS2 and 28S in eukaryotes, or fragments of them, are commonly used as barcodes. By using the rrn we are using a barcode that is many times longer than the current barcodes and that also includes many of them. So, it does not have only the advantage of providing more phylogenetic resolution but also allows to bring previous information generated using other rrn derived barcodes under the same phylogenetic and taxonomic framework. In order to establish the rrn as a barcode the first thing we need to generate is a reference database. As we are just starting to generate now the first rrn amplicons using third generation sequencings (Nanopore, PacBio) we still do not have access to this type of data to generate such a reference database. However, genomes and metagenomes can be sources of rrn that can be used as references after placing them in a phylogenetic tree in order to assign them an identity. We propose to use extracted rrn from publicly available genomes and metagenomes and build a phylogenetically aware reference database using R and MySQL.

#### References:

Guillou, L. et al. (2013) The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nucleic Acids Res.* 41, D597-604 del Campo, J. et al. (2018) EukRef: Phylogenetic curation of ribosomal RNA to enhance understanding of eukaryotic diversity and distribution. *PLOS Biol.* 16, e2005849 Jamy, M. et al. (2020) Long-read metabarcoding of the eukaryotic rDNA operon to phylogenetically and taxonomically resolve

environmental diversity Molecular Ecology Resources 20, 429–443

**Expected skills::**

R, HMMER, Python, MySQL, phylogeny

**Possibility of funding::**

No

**Possible continuity with PhD: :**

To be discussed

---