SCIENTIFIC REPORT

Purpose of the visit

As stated in the project proposal for this grant, my PhD project attempts to estimate benthic biodiversity in Mediterranean marine soft bottoms (macro- and meiobenthic invertebrates), by using metabarcoding and traditional morphological approaches in parallel.

The aim of my visit to the Laboratoire d'Ecologie Alpine (LECA) was to perform the metabacording analysis on 81 samples collected on two different Mediterranean soft bottoms. In this way, I had the opportunity to learn the techniques of amplification and sequencing with the Illumina platform, as well as the bioinformatics methods applied for processing the output generated

One of the objectives was to compare the results between two different genes markers (18S and COI). However, I only worked with the 18S marker due to some amplification problems for the COI marker

Description of the work carried out during the visit

The work carried out included the learning and application of the whole set of procedures for amplifying the target region, obtaining the DNA sequences and processing the resulting information through bioinformatics pipelines.

The DNA extraction was done before the beginning of the stay, following the protocol described by Taberlet et al. 2012, but using the DNA Blood and Tissue extraction kit (QIAGEN). Thus, the following tasks were performed at the host laboratory:

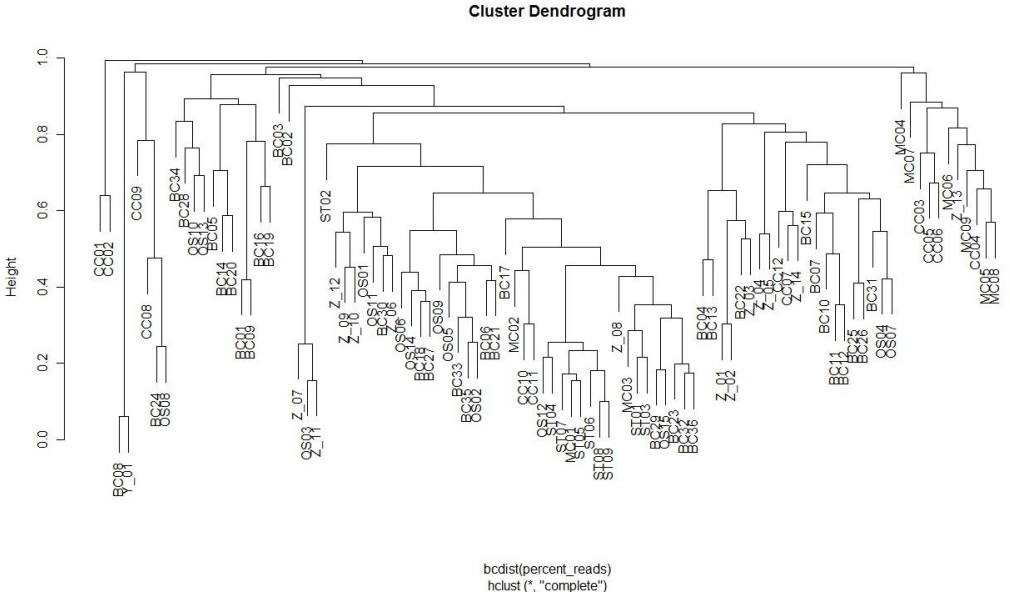
1. *Test for checking the quality of the DNA extractions and optimization of the PCR conditions*. Several PCRs were carried out on seven samples with different DNA concentrations as well as for a negative control. Different DNA taq polymerases, different 18S primers (all designed by Dr.Taberlet's team) and different PCR conditions were assayed to check which one yielded the best amplification of the DNA samples.
2. *Creation of a file with all the information for the Next Generation Sequencing (NGS) process with Illumina MiSeq*. This file was used to establish the correspondence between the each sample and the tagged primer, as well as the position of each sample at the PCR plate.
3. *DNA amplification*. A total of 288 PCR were performed (81 samples + 14 negative controls + 1 positive control, three replicates for each one, distributed on three 96-well plates) following the optimized conditions previously found.
4. *Checking the quality of DNA amplifications*. One plate was tested with QIAxcel analyser (QIAGEN) for the concentrations and the size of amplified DNA fragments.
5. *DNA purification*. The purification of each amplicon was carried out with the MinElute PCR Purification kit (QIAGEN). All the purified PCR products were mixed on a single tube and the DNA concentration was measured. This tube was then sequenced on an Illumina MiSeq platform.

6. *Data analysis*. The results from the sequencer were treated with the OBITools software (http://www.grenoble.prabi.fr/trac/OBITools), a set of Python scripts developed by the LECA team to manipulate the outputs of the next generation sequencers. Different steps were carried out for assigning the sequences to taxa. During this process, several filters are applied for cleaning the data from possible contaminations, low quality sequences, merged redundant information, etc. Once the data was clean enough, it was classified into different datasets depending on the percentage of best-identity match between the sequences and the EMBL database of the 18S marker. The subsequent analyses were carried out with the R software.

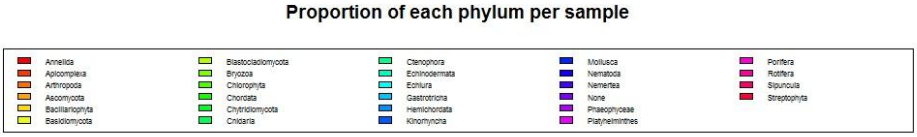
Description of the main results obtained

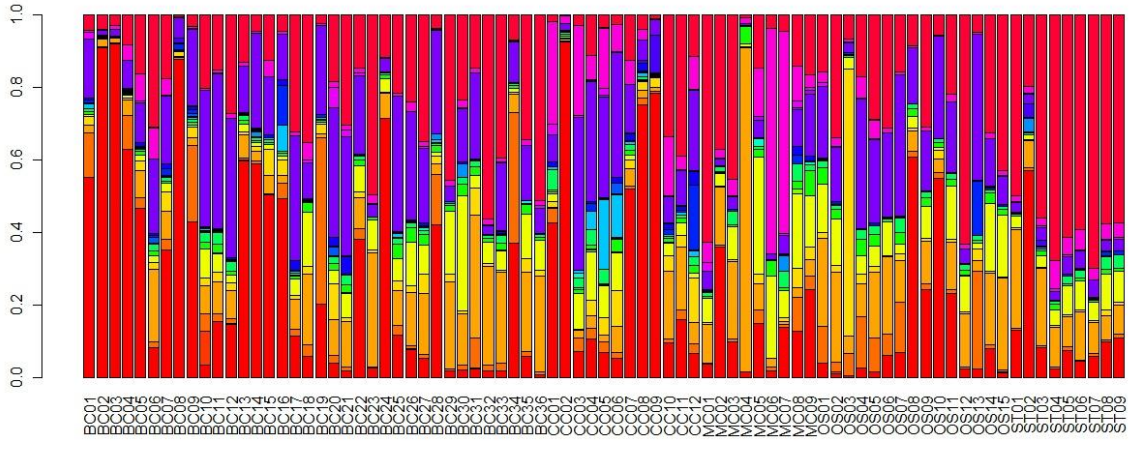
Two datasets were created for 95% and 97% best-identity cut points. For both datasets the same analyses were performed. I only present some figures for the 95% dataset.

1. Bray-Curtis distance between samples (including the negatives and the positive controls).

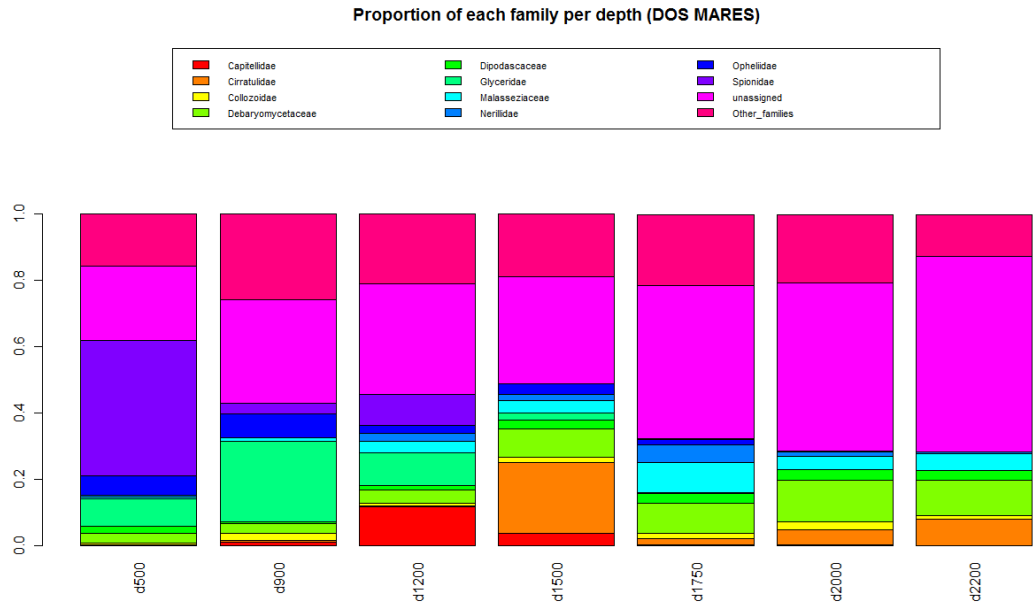


2. Number of counts and sequences per sample.
3. Proportion of different taxonomical ranks per sample.





4. Proportion of different taxonomical ranks per area.
5. Proportion of different taxonomical ranks per depth.



Future collaboration and projected publications

It is expected to publish a paper with the work developed during the stay at the Laboratoire d'Ecologie Alpine. Further analyses need to be done, as well as an appropriate discussion of the results, so the collaboration with the Dr.Taberlet's team will continue in order to obtain the positive feedback of such an experienced group in the metabarcoding field.

References:

Pierre Taberlet, Sophie M.Prud d'Homme, Etienne Campione et al. (2012) "Soil sampling and isolation of extracellular DNA from large amount of starting material suitable for metabarcoding studies" Molecular Ecology, 21.