

Internship mid-term report

Genetic diversity limits of a marine microbiome

Student: Sergio Gozalo Miranda

Host institution: CSIC-ICM

Host institution tutor: Ramiro Logares Haurie (ramiro.logares@icm.csic.es)

Start/End: 06/01/2021-07/04/2021

Microbial marine communities are made up of thousands of organisms that contribute with genes that are fundamental for the functionality of the ecosystem. The first challenge of marine metagenomics was to construct a list of taxa with respective abundances starting from sequencing information. During the last decade, the taxonomic diversity values have been narrowed down and, hence, the first challenge has been completed. The next challenge is to find the total diversity of genes, that has been studied but is less known and depends on the techniques used, so it cannot be said that this challenge has been completed. To solve this challenge and approach this marine genome unknown diversity, this study is going to analyse 50 different metagenomes from the same place obtained within 2 days and, in consequence, try to reach the limits of the marine microbial genetic diversity in this location.

Sample obtention was performed in a marine observatory from the Mediterranean Sea (SOLA) during 2 days with different protocols, using different filters and water volumes and filter size-fractions. A total of 50 metagenomes were obtained that were analysed with two different bioinformatics pipelines, MGnify (<https://www.ebi.ac.uk/metagenomics/>) and a custom pipeline from ICM. The results from the different pipelines will be analysed using R software (R- Development-Core-Team 2008) and the package vegan (Oksanen et al. 2008). These results will be used to try to find the best pipeline to assess the genetic diversity and, in consequence, make one of the pipelines the main for the analysis of other marine metagenomes, providing, this way, a standard method. The total number of genes, functions, and taxonomic markers (rRNA-gene) will be analysed to determine whether we have reached the limits of microbial diversity.

Since this project is a basic science project, the immediate contributions are not clear, however, this characteristic also implies that there may be many contributions in the near future, for example, new genes with utility in biotechnology(bioprospection) may be found. To sum up, this is a basic science project that can lead to many contributions in several fields and does not require many funds, since the data has already been gathered and uses only computational methods.