Taxonomic analysis of environmental DNA in coastal water samples from Denmark

Testing a PCR-free sequencing methodology

Adrián Gómez Repollés, 20180315

Master's Thesis

June 2020

Supervisors:

Phillip Francis Thomsen

Eva Egelyng Sigsgaard

Mikkel Heide Schierup

Aarhus University

Abstract

Environmental DNA (eDNA) became a relevant research method for the description of biological communities due to the development of sequencing technologies. In more detail, the metabarcoding methodology had been widely used in biodiversity surveys gaining great acceptance. However, multiple studies had highlighted flaws whereas other had reinforced its potential. The main biases in respect to its application are the sampling method, selected primers, PCR amplification and sequencing procedure. The purpose of this thesis is to generate a PCR-free methodology with the potential of reducing biases common in metabarcoding analyzes. Seawater surface samples were taken from different locations of the coast of Denmark and process with shotgun and nanopore sequencing. Moreover, a comparison of the effect of the filter pore size in the eDNA collection of the filtered water samples was done for the shotgun sequencing. In addition, a metabarcoding analysis associated to each sequencing procedure was generated to estimate the performance of the PCR-free methodology in the description of biological communities. The results suggested that the pore size affected the amount of DNA collected (p-value: 0.01) and, potentially, the taxa captured. The number of species obtained was $1,214$ and 662 for shotgun sequencing and its complementary metabarcoding analysis, respectively. Nonetheless, nanopore sequencing identified 367 whereas 735 (18S) and 679 (COI) for the metabarcoding analyzes. Our results exhibited a good performance of shotgun sequencing whereas nanopore sequencing did not show potential for species identification.

Acknowledgments

First of all, I would like to thank the people that took part in the development of the research. In this thesis, Philip Francis Thomsen performed the fieldwork. Katrine Bay Jensen performed the laboratory work in addition to Eva Egelyng Sigsgaard. Tobias Guldberg Frøslev (UCPH) developed the pipeline for the metabarcoding analysis. Eva performed the metabarcoding analysis (repeated by me) and, Mads Kristian Reinholdt Jensen helped me by discussing the research and providing interesting papers.

Second, I would like to thank my supervisors, Philip and Eva, for letting me become part of their research group and letting me learn about environmental DNA. A special mention for Eva for her dedication and constant effort, in addition to her undeniable help in the development of this thesis. Moreover, thank Mikkel Heide Schierup for teaching and helping me through the master's courses and second PiB.

Third, a huge thank you to my very special friends, Alejandro Corregidor Castro and Mateo Sokač, and to my roommates (Aram, Elena and Mathilde) without them I would not have been able to go all the way to finish the master. Sincere thanks to my master's colleagues, in special to Iker Rivas Gonzalez, for his constant help. I wish you the best in your scientific career.

Last but not least, thanks to my friends, you know who you are, to my brother and father, Daniel and Antonio, and my beloved family. In special to my grandmother Esperanza, your efforts have been fruitful.

Contents