



UNIVERSIDAD ADOLFO IBÁÑEZ
FACULTAD DE INGENIERÍA Y CIENCIAS
PhD IN COMPLEX SYSTEMS ENGINEERING

STUDY ON THE USE OF GENETIC FUNCTIONAL TRAITS TO CHARACTERIZE
MICROBIAL COMMUNITIES

Salvador Ramírez Flandes

Thesis Director: Dr. Bernardo González
Thesis Co-director: Dr. Osvaldo Ulloa

Submitted as a requirement for the degree of PhD in Complex Systems
Engineering at Universidad Adolfo Ibáñez, April 2019

Contents

| | |
|--|------------|
| Abstract | 2 |
| Chapter 1. Introduction | 3 |
| Chapter 2. Redox traits characterize the organization of global microbial communities | 6 |
| Manuscript text | 6 |
| Supplementary Information Appendix (SI Appendix) | 20 |
| Chapter 3. Protein coding potential of nucleotide sequences based on Kmers | 115 |
| Manuscript text | 115 |
| Supplementary Information Appendix (SI Appendix) | 128 |
| Chapter 4. Conclusion | 138 |
| References | 139 |
| Published papers with affiliation to this program | 142 |

ABSTRACT

Biological communities are conventionally described as assemblages of species, whose ecological roles are known or predictable from their observable morphology. In microbial ecology, however, such a taxonomic approach is hindered by our limited knowledge of the functions of most microorganisms, which often alter their genetic material through diverse mechanisms. To tackle these problems, microbial ecologists have used culture-independent genetic approaches to study the whole pool of functional genes at the community level. However, this approach requires dealing with gene categories not necessarily related to the ecology of the organisms, such as functions associated with DNA replication or cellular division. In this work it is demonstrated that genes encoding oxidoreductases characterize the microbial communities better than other categories of genes, including those associated with taxonomy. Additionally, with this approach, the role of microbial communities of the different ecosystems in biogeochemical cycles becomes readily apparent. The importance of this result is, however, limited by the coverage of known genetic functions over the total pool of metagenomic genes, which is currently around the fifty percent in sampled metagenomes. To help increasing this reduced coverage of known functions, a methodology for the recognition of the protein-coding potential of DNA sequences is proposed. Alternative applications of this methodology are discussed. The results of this study pave the way for a better assessment and evaluation of microbial ecosystem services from different environments of our planet. This improved diagnostic of microbial ecosystems can be possible, for example, by focusing directly on the diversity of redox functions encoded in the metagenomes of microbial communities, rather than on their taxonomic structures. Thus, this approach should help in developing better management and conservation policies that effectively include not only iconic species or colonies, such as polar bears or coral reefs, but also microorganisms.