

Metagenomics of hydrothermal fumaroles to study the evolution of early life at high temperatures

Beatrice Cobucci-Ponzano¹, Andrea Strazzulli¹, Rosa Giglio¹, Roberta Iacono¹, Yizhuang Zhou², Jin Xu², Mosè Rossi and Marco Moracci¹

¹Institute of Biosciences and BioResources – CNR, Via P. Castellino 111, 80131 Naples, Italy

²BGI Hong Kong, Limited, 16Dai Fu Street, Tai Po Industrial estate, Tai Po, New Territories, Hong Kong

The study of extremophilic Archaea, considered the most ancient forms of extant life on Earth, is of interest for the study of the origin and evolution of life in the Universe. The study of the communities populating extreme environments on Earth has relevant implications for astrobiology, since similar sites identified during the exploration of the solar system may host now, or may have hosted, organisms living in extreme conditions. Metagenomic analysis applied to extreme environments is a powerful tool to understand the composition of the communities populating these sites and to study their evolution and adaptation to geological events.

The Pisciarelli solfatara, located in the Phelgrea Fields (Naples, Italy) is characterized by a remarkable diversity in terms of temperature and pHs of the different mud pools and solfataric vents. Here several thermoacidophilic species, such as *Acidianus*, *Archaeoglobus* spp, *Pyrobaculum*, *Sulfolobus*, and many viruses have been isolated [1]. In order to map the composition of this extreme environment in terms of number and abundance of the different species, we decided to carry out a metagenomic deep sequencing project on soil and water/mud pools of Pisciarelli solfatara. Total DNA was extracted and analysed by using Illumina paired-end sequencing technology. In a single run, we produced ~3.0 Gb total nucleotides. These reads were assembled into 1,543 contigs with an N50 of 15863 bp, and a total of 17,652 Open Reading Frames (ORFs) were identified. These ORFs were subjected to BLAST similarity searches and annotated with KEGG, eggNOG and NR database respectively. The majority (77%) of the species found belonged to the Crenarchaeota phylum, which branches deeply in the phylogenetic tree of life, followed by Proteobacteria (22%), and, in much lesser amounts, Euryarchaeota, Actinobacteria, and others. In addition, among the different metabolic pathways identified, the ORFs involved in the carbohydrate metabolism were the majority. Details on the species composition and on the relative abundance of each ORF will be described and the importance of metagenomic studies in mapping extreme environments will be discussed.

This work was supported by the project ASI N.2014-026-R.0 "Esobiologia e ambienti estremi: dalla Chimica delle Molecole alla Biologia degli Estremofili - ECMB and by the project PON01_01966 "ENERBIOCHEM - Filiera agro-industriali integrate ad elevata efficienza energetica per la messa a punto di processi di Produzione Eco-compatibili di Energia e Bio-chemicals da fonte rinnovabile e per la valorizzazione del territorio" of the Programma Operativo Nazionale Ricerca e Competitività 2007-2013 - MIUR-Italy.

[1] R. Huber et al., 2000, FEMS Microbiol Rev., 24, 615