

## GeoMosaic: A flexible metagenomic pipeline combining biological and geochemical data to outline biosphere and geosphere interactions

Dr. Davide Corso, University of Naples Federico II



The interaction between the biosphere and geosphere is essential for their co-evolution and has contributed to keeping our planet habitable for more than 4 billion years. Metagenomic studies of the environmental microbial communities are essential to describe them and their adaptations necessary to colonize such environmental conditions. Integrating this information, along with the geochemical and geological context is fundamental to outlining a complete overview of the bio-geo interaction. Here we present GeoMosaic, a flexible and customizable metagenomics pipeline (work in progress) combining biological and environmental, and geochemical data.

It integrates different stream outcomes as reads-, assembly- and MAGs-based analysis to extract and optimize the data knowledge. Each stream contains different modules (e.g. pre-processing, assembly, binning, etc), in which are implemented one or multiple methods to perform the respective task. The designed modularity allows bioinformatic users to complete or update the workflow with new analysis modules, to improve the comprehensive range of analysis topics.

GeoMosaic is a publicly available pipeline, developed to be easy to use by choosing/ignoring both modules and methods for each task. The major input for the users is an Excel tabular file describing the name of the sequencing that reads R1 and R2 and the corresponding sample name.

Metagenomics is a widely used technique to study the taxonomic composition and metabolic function of microbial communities. Using such sequencing data, the GeoMosaic workflow provides assimilated results of metagenomics and geochemistry steps in ready-to-use data and plots for downstream interpretations.

Per ulteriori informazioni: estermaria.eckert@cnr.it Fruibile in streaming attraverso la piattaforma TEAMS