An integrated machine learning approach reveals geochemical controls on microbial electron-transfer protein abundance

JOY BUONGIORNO[1*], DONATO GIOVANNELLI[2], MAARTEN DEMOOR[3], PETER BARRY[4], MATTHEW SCHRENK[5], KAREN G. LLOYD[6], SHWETA NAKAR[7], SHAUNNA MORRISON[1], ROBERT HAZEN[1]

[1]Earth and Planets Division, Carnegie Institution of Washington, Washington, DC, USA
[2]Department of Biology, University of Naples "Federico II", Naples, Italy
[3]Volcanological and Seismological Observatory of Costa Rica Universidad Nacional, Costa Rica
[4]Marine Chemistry and Geochemistry Department, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts, USA
[5]Department of Earth and Environmental Sciences, Michigan State University, USA
[6]Department of Microbiology, University of Tennessee, Knoxville, USA
[7]Tetherless World Constellation, Rensselaer Polytechnic Institute, Troy, New York, USA
***email: jbuongiorno@carnegiescience.edu**

Planetary redox shifts on early Earth spawned microbial evolution of new electron-transfer processes mediated by oxidoreductase enzymes. Many oxidoreductase co-factors include transition metals, which are primarily sourced from volcanic emissions. Our understanding of how volcanically-mediated metals affected the trajectory of microbial evolution can be expanded by tracing the abundance of oxidoreductases in modern environments within the framework of volcanic regime and *in situ* transition metal chemistry.

Here, we conducted a comparative analysis of nearly 1,000 metagenomes collected from volcanically-influenced sites typified by differences in their magmatic water content, fluid evolution, and metal emission signatures. Co-correlation networks of several hundred oxidoreductases revealed that volcanic arcs are distinct from hydrothermal vents. Enzyme cliques were included in a machine learning approach alongside co-located transition metal chemistry to identify the key geochemical controls of oxidoreductase distribution. Results show that iron and manganese are crucial geochemical features.