

## **A gene-centric approach for integrating environmental genomics and biogeochemical models**

C. K. ALGAR<sup>1\*</sup>, D. C. REED<sup>2</sup>, J. A. HUBER<sup>1</sup>  
AND G. J. DICK<sup>2</sup>

<sup>1</sup>Josephine Bay Paul Center, Marine Biological Laboratory,  
Woods Hole, MA 02543

(\*correspondence: calgar@mbl.edu)

<sup>2</sup>Department of Earth and Environmental Sciences, University  
of Michigan, Ann Arbor, MI 48109

Rapid advances in molecular microbial ecology have yielded an unprecedented amount of data about the evolutionary relationships and functional traits of the microbial communities that regulate global geochemical cycles. Biogeochemical models, however, are trailing in the wake of the environmental genomics revolution and such models rarely incorporate explicit representations of bacteria and archaea, nor are they compatible with nucleic acid or protein sequence data.

Here, we present a functional gene-based framework for describing microbial communities in biogeochemical models [1]. The model can predict both traditionally modeled geochemical variables, such as nutrient concentrations, and also "-omics data", such as gene abundances that can be compared directly with (meta)genomic data. To demonstrate the approach in practice, nitrogen cycling in an oxygen minimum zone (OMZ) was modeled to examine key questions about "cryptic" sulfur cycling and dinitrogen production pathways in OMZs.

By directly linking geochemical dynamics to the genetic composition of microbial communities, the method provides mechanistic insights into the patterns and biogeochemical consequences of marine microbes, including the resolution of "cryptic" metabolic pathways, which cannot be resolved with geochemical data alone. Such an approach is critical for understanding the role microbes play in modulating Earth's biogeochemistry.

[1] Reed *et al* (2014) *PNAS* **111**(5) 1879-1884