Study of modular denitrification in microbial communities through sequence analysis

Alburuj R. Rahman, James Shapleigh, Armanda Roco[1], Joseph Yavitt[2]

Department of Microbiology[1], Department of Natural Resources[2], Cornell University

Introduction

Denitrification is one way bacteria carry out nitrate respiration (5). It can be described as the reduction of nitrate to nitrogen gas, in four reduction steps (Fig. 1).

Earlier studies show that bacteria possess all four reductases needed to carry out each step of denitrification. New research suggests the denitrifier community is made up of incomplete denitrifiers, organisms that carry out only certain steps of denitrification (5).

Denitrification is common in soil ecosystems. One of which is Hubbard Brook (1,2), an experimental forest that contains six watersheds used for carrying out ecosystem studies.

PURPOSE:
To examine the denitrifying community and assess the presence of both complete and partial denitrifiers.

The taxonomic classifications and functional characteristics of the sequenced DNA will be analyzed in order to test for presence of modularity.

Methods

Soil was collected from locations with different pH and elevation. All samples were taken from one of three elevations, three depths per elevation, and two watersheds, which did or did not contain calcium to raise pH.

DNA sequenced from samples through:
- use of Mo-Bio PowerSoil DNA Isolation kits (4)
- processing of DNA into 180 million reads, each 100 nucleotides long, at Columbia

Using the CLC, a high-throughput sequence analysis tool, data was:
- sequenced and analyzed with Illumina HiSeq Next Generation (4)
- trimmed into reads of 18-30 x 10^6 sequences and assessed based on quality of reads

These trimmed paired reads were then:
- sent to a metagenomics automated analysis platform: MG-RAST (3)
- separately analyzed with blastx using custom databases for NAP and NIR

Developed programs to:
- parse hits from NCBI blastx for organismal classification
- match sequences hits from queries to original trimmed reads
- paired reads matched (eventually)
- re-blast sequences to find associated organisms and assoc. biological func.

Results

As an initial approach to establish differences between communities, a principle coordinate analysis, which examines variance in data using eigenvalues, was used to find correlation among samples based on elevation. These analyses showed that the samples clustered by depth and not by location (Fig. 3).

For denitrification analysis, more organisms have had hits – that is, their sequences were linked to possessing a certain reductase – to Nor than they did to Nir. Also, the most common organisms with hits to Nor, were not found to have hits to Nir. This is shown through analysis of NirK and Nor hits for samples CJ001-03 (~pH 4) and CJ010-012 (~pH 3) (Fig. 4).

Through blastx searches on CJ001 using custom Nap and Nir databases, different organismal classification was found than that from MG-RAST.

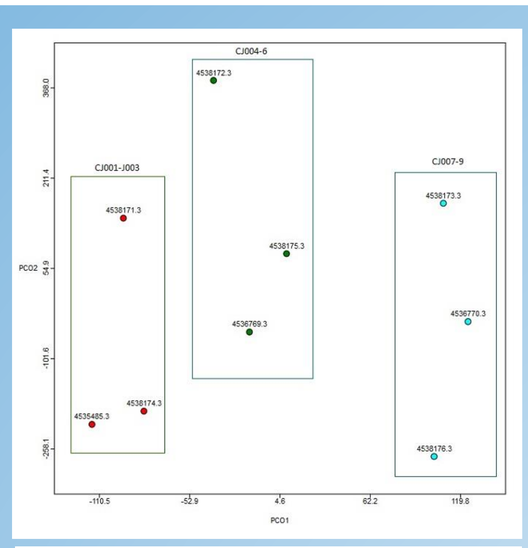

Fig. 3: A principal coordinate analysis that shows grouping of samples based on elevation.

Discussion

A different number of hits or abundance of organisms associated with particular reductases and different organismal classifications were found to be associated with those hits (5). This all goes to reinforces the notion of modularity.

MG-RAST was found to misassign many of the proteins labeled as nitrate reductase, likely because it uses non-specialized databases (5). And so we will work on building more appropriate databases and validation tools to allow reasonable determination of nitrate reductase distribution at Hubbard Brook.

To increase confidence in taxonomic assignments, we plan to analyze both ends of the paired end sequence generated by Illumina sequencing. This will require development of programs to facilitate extraction of missing sequences from trimmed reads and then do additional sequence comparisons. This will be used to support initial taxonomic assignments.

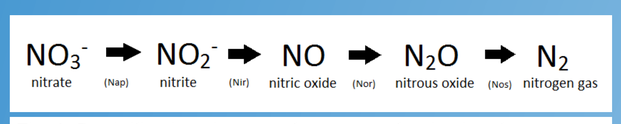

Fig. 1: Denitrification shown as four steps of reduction, from nitrate to nitrogen gas.

Fig. 2: A representation of Hubbard Brooke Experimental Forest's watersheds.

Fig. 4: The abundance of hits plotted against organism classification for NirK and Nor for CJ001-3 (higher pH) and CJ010-12 (lower pH) using MG-RAST databases.

References

1. Dittman, J.A., C.T. Driscoll, P. M. Groffman and T. J. Fahey. 2007. Dynamics of nitrogen and dissolved organic carbon at the Hubbard Brook Experimental Forest. Ecology 88:1153-1166.

2. Judd, K.E., G.E. Likens and P.M. Groffman. 2007. High nitrate retention during winter in soils of the Hubbard Brook Experimental Forest. Ecosystems 10:217-225.

3. Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E. M., Kubal, M., Paczian, T., ... Edwards, R. A. (January 01, 2008). The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *Bmc Bioinformatics, 9.*

4. Roco, Armanda. Personal Communication. Nov. 2013.

5. Shapleigh, JP. 2012. The denitrifying prokaryotesIn E. Rosenberg, E. Stackebrandt, F. Thompson, E. DeLong, and S. Lory (eds.), The Prokaryotes. Springer, New York, NY.