Award: DE- DE-SC0006837

Institution: J Craig Venter Institute

Project Title: " The Study of Microbial Environmental Processes Related to the Natural Attenuation of Uranium at the Rifle Site using Systems-level Biology "

PI: Barbara Methe

Additional Authors

Dr. Mary Lipton- Pacific Northwest National Laboratory (PNNL)
Dr. Krishna Mahadevan- University of Toronto (UT)

Project Goals

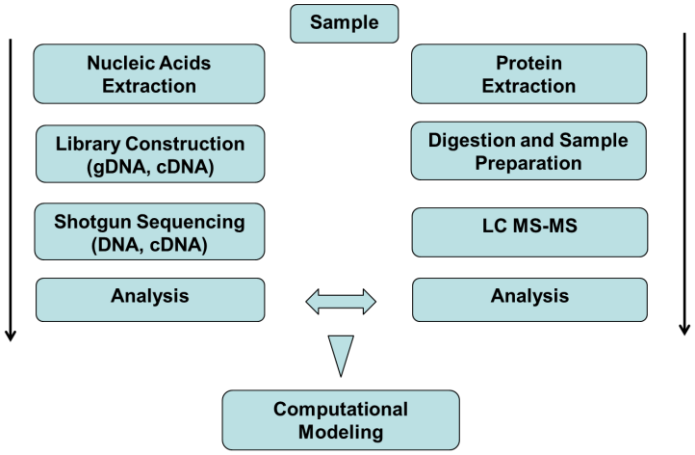
Diverse microbial communities exist in subsurface environments that possess significant metabolic potential to effect global carbon, nitrogen and metal cycles including the transformation of radionuclides. Objectives of this project are: 1) to apply systems-level biology through application of 'metaomics' approaches (collective analyses of whole microbial community DNA, RNA and protein) to the study of microbial environmental processes and their relationship to C, N and metals including the influence of microbial communities on uranium contaminant mobility in subsurface settings undergoing natural attenuation, 2) improve methodologies for data generation using metaomics (collectively metagenomics, metatranscriptomics and proteomics) technologies and analysis and interpretation of that data and 3) use the data generated from these studies towards microbial community-scale metabolic modeling.

Introduction

The objective of our study was to apply systems-level biology to the study of microbial environmental processes and their relationship to uranium contaminant mobility in subsurface environments undergoing natural attenuation at the Department of Energy (DOE) Rifle Integrated Field Research Challenge (RIFRC) site over spatial and temporal scales. Our hypothesis for the proposed study is that sediment zones undergoing natural attenuation of uranium possess microbiota and a metabolic capacity in the reservoir of genes and metabolic pathways that is critical to driving uranium reduction and is distinct from zones in which natural attenuation of uranium is not occurring. Alternatively, the genetic complement from sediment zones undergoing natural uranium attenuation is similar to sediment zones where this process is not taking place however; different environmental factors (e.g. bioavailability of electron donors and acceptors) influence the community genetic response. The nature and extent of this metabolic capacity and response can be identified, characterized and computationally modeled by a systems-level approach.

To meet the goals of this ongoing project, two subsurface sites from the RIFRC were interrogated using a suite of metaomic approaches. The first site consists of sediments from the Winchester 2007 gallery, 'JB" well locations and was chosen due to the occurrence of natural attenuation of uranium (uranium reduction in the absence of biostimulation or other remedial interventions). The second, more recent sites of study within this project were located in the Colorado River Floodplain (CORFP) and were chosen as analogs of naturally buried reduced sites.

Methods



This study utilized two field sties for analyses. The first site was the Winchester 2007 gallery, 'JB01-05' well location (consisting of large cobble and silt matrix) at a 4m depth were examined. Although biostimulation experiments via the addition of acetate to groundwater have not taken place here, uranium reduction has been verified by absorption spectroscopy. However, complete immobilization of uranium has not been noted. The second site was the Colorado River Floodplain

Figure 1. An overview of 'metaomics' approaches applied to samples from RIFRC sites.

(CORFP) sediments representing recent sediment depositions. Overbank deposits in the floodplain have become enriched in C, Fe and S minerals. Aggradation processes have led subsequent burial of these enriched sediments creating "hotspots" of biogeochemical activity which serve as analogs to the buried naturally reduced sediments at the JB sites.

The strategy for the examining samples is outlined in **Figure 1** and show the generation of sequence reads from microbial community DNA (metagenomics or whole genome shotgun sequencing (WGS)) and RNA (metatranscriptomcs or RNAseq) and protein information using protomics. Results were analyzed independently and through computational modeling.

Results

Metagenomics and Metatranscriptomics

Comparison of taxonomic profiles generated from metagenomic and metatranscriptomic reads from Rifle samples from the Winchester gallery revealed diverse profiles (**Figure 2**). Metagenomic reads (indicated as WGS **Figure 2'A'**) and RNA-seq (**Figure 2'B'**) reads. Taxonomic profiles were generated from high quality alignments (>60bp quality trimmed read length at >80% composite identity) to the NCBI NT database. Composite identity is the percent identity of the alignment times the percent of the quality trimmed read that was aligned. For both the WGS and RNA-seq data sets the best matches reveal the majority of reads are assigned to a few species with the most abundant taxonomic members belonging to the Gram negative beta- and gamma-subclasses of the Proteobacteria, and Gram positive Actinobacteria, respectively. For both data sets *Thiobacillus denitrificans* was the most abundant species. *T. denitrificans* is a facultatively anaerobic chemolithotroph, capable of coupling the oxidation of inorganic sulfur compounds to the reduction of oxidized nitrogen compounds (such as nitrate and nitrite) to dinitrogen. While there is interest in the use of *T. denitrificans* for use in the removal of nitrate and sulfide in environmental settings, this organism has also been shown to be capable of the oxidization of

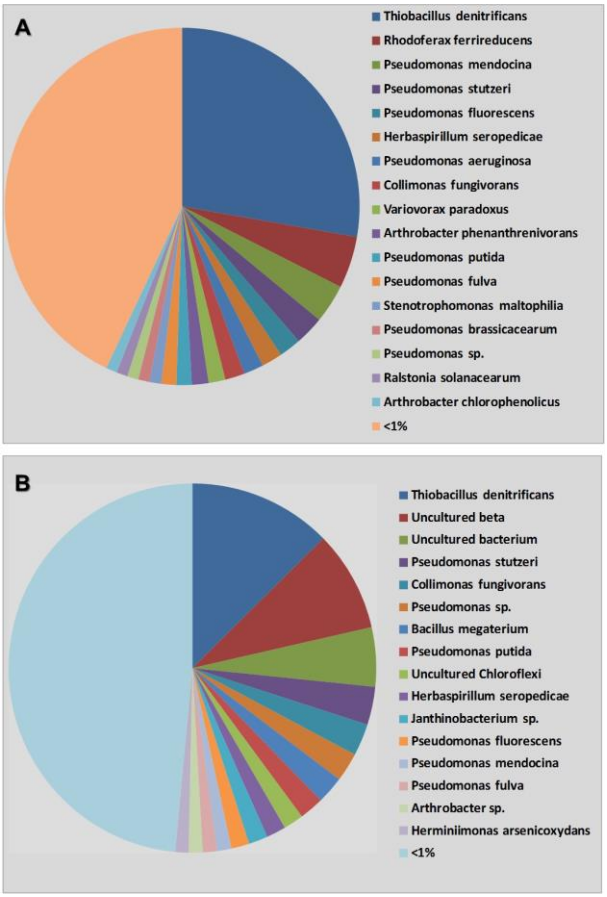


Figure 2. Taxonomic profiles from metagenomic reads (A) and metatranscripomic reads (B) from the JB site.

U(IV) to soluble U(VI) in the presence of nitrate. The relatively high abundance of *T. denitrificans* is an important finding in the lack of complete uranium immobilization at this site. However, evidence for the presence of metal reducing bacterial relatives (although not necessarily demonstrated to reduce uranium) such as *Rhodoferax ferrireducens*, and *Georgfuchsia toluolica*, an anaerobic beta-proteobacterium capable of degrading aromatic compounds with Fe(III), Mn(IV) or nitrate as an electron acceptor, were determined.

From assembled metagenomic and metatranscriptomic sequencing reads, metabolic pathways were also reconstructed (**Figure 3**). Using annotation of ORFs from the WGS (**Figure 3'A'**) and RNA-seq assemblies (**Figure 3'B'** KEGG Pathways were identified. Significant portions of central intermediary and energy metabolism were recovered including denitrification and partial pathways for carbon fixation and methane metabolism and a variety of pathways for the degradation of aromatic compounds including diooxygenase, monooxygenase and anaerobic pathways.

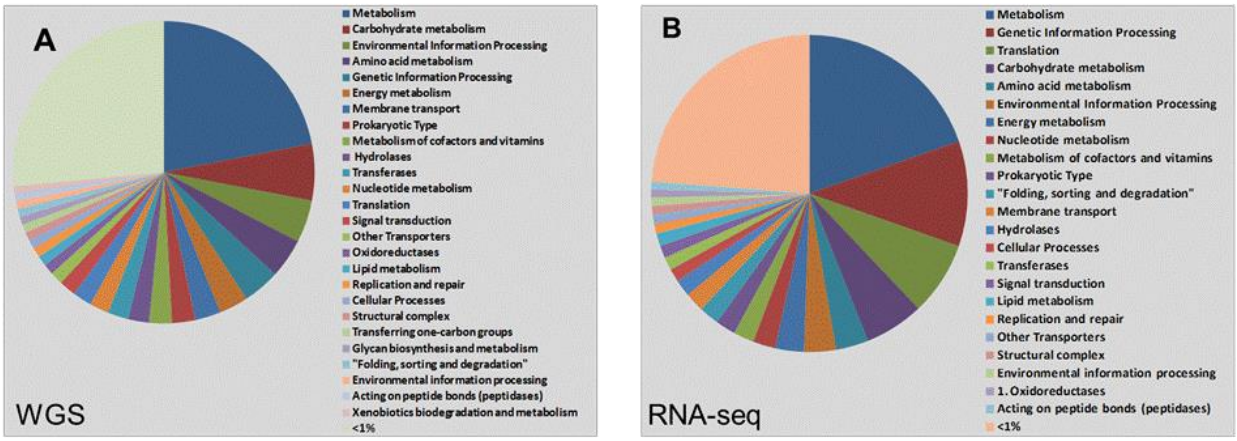


Figure 3. Functional profiles using ORFs derived from WGS (Panel 'A') and RNA-seq (Panel 'B') assemblies.

Comparison of RNAseq to WGS taxonomic profiles were compared and the results highlighted using an accumulation curve (**Figure 4**). As shown in accumulation curve the number of taxa shared between the top 100 taxa based on rank abundance were computed. Results indicates the strong concordance between the two profiles (~50% of the top 100 taxa are shared). These results further indicate the importance of producing both metagenomic and metranscriptomic data for studying in situ microbial communities.

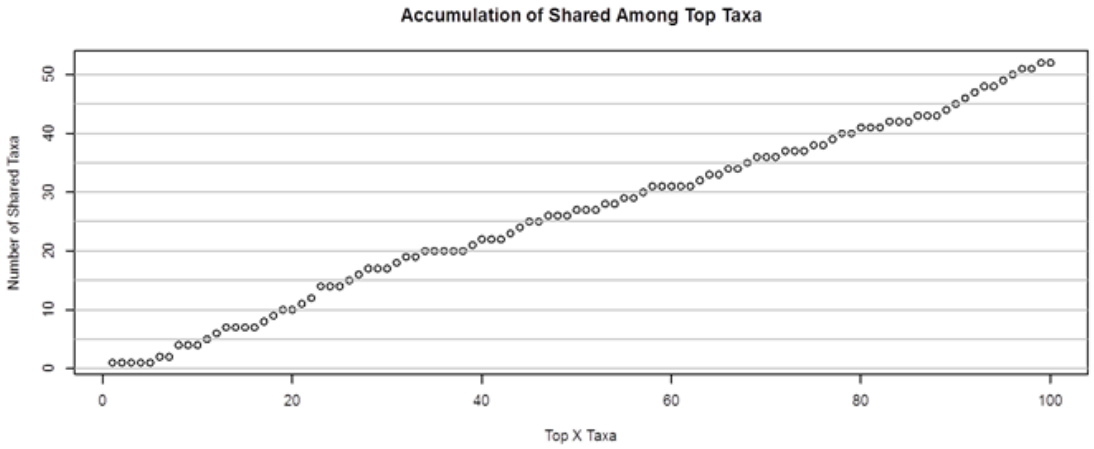
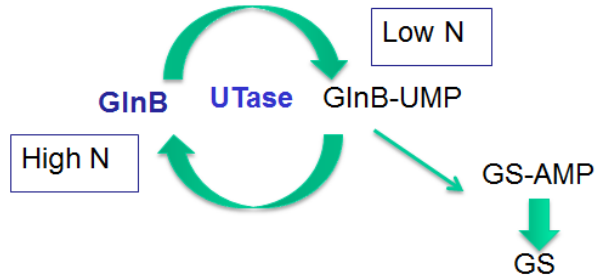


Figure 4. Comparison of RNAseq to WGS taxonomic profiles

Proteomics



Nitrogen cycling proteins GlnB/GlnK and uridylyltransferase/uridylyl-removing enzyme (UTase/UR) were found to be prominent in the protein profiles of *Rhodoferax ferrireducens* and *Thiobacillus denitrificans* indicating the importance of nitrogen regulation in the community.

An abundance of proteins were found that indicate the importance of nitrogen cycle regulation in *R. ferrireducens* and *T. denitrificans* as evidenced by the abundance of proteins derived from pathways of glutamine metabolism including GlnB, GlnK and other P-II proteins from these organisms (**Figure 5**).

Among the abundant proteins found, a partial TCA cycle has been recovered for *Anaeromyxobacter* (delta-protebacteria

Figure 5. Nitrogen cycling proteins determined from proteomic studies from CORFP sediments.

capable of coupling acetate and hydrogen oxidation to anaerobic respiration including U(VI)) as shown below. While not all the enzymes in the cycle are represented in the proteome, no evidence for the alternative glyoxylate cycle was detected.

Modeling

A framework for microbial community modeling was generated that couples collections of genome-scale models of bacterial metabolism from the subsurface ecosystem for the purpose of predicting how a complex microbial community may respond to changes in the subsurface environment. This effort is critical as it provides an important analytical strategy for integrating and interpreting metaomics data as well as the capacity to simulate various interactions between environmental conditions and microbial communities. Guided by results from previous metagenomic and metatranscriptomic analyses from a site identified as undergoing natural attenuation of uranium at the RIFRC (JB-05) and the CORFP site, collections of genome-scale metabolic models representing seven bacterial classes were chosen for metabolic modeling including:

(a) Beta-proteobacteria (representative species: *Thiobacillus denitrificans*, *Rhodoferax ferrireducens* and *Variovorax paradoxus*); (b) Delta-proteobacteria (representative species: *Geobacter metallireducens* and *Desulfovibrio vulgaris*); (c) Bacilli (representative species: *Bacillus subtilis*); (d) Gamma-proteobacteria (representative species: *Pseudomonas stutzeri*); (e) Alpha-proteobacteria (representative species: *Rhodopseudomonas palustris*); (f) Actinobacteria (*Arthrobacter* sp.) and (g) Clostridia (*Clostridium cellulolyticum*).

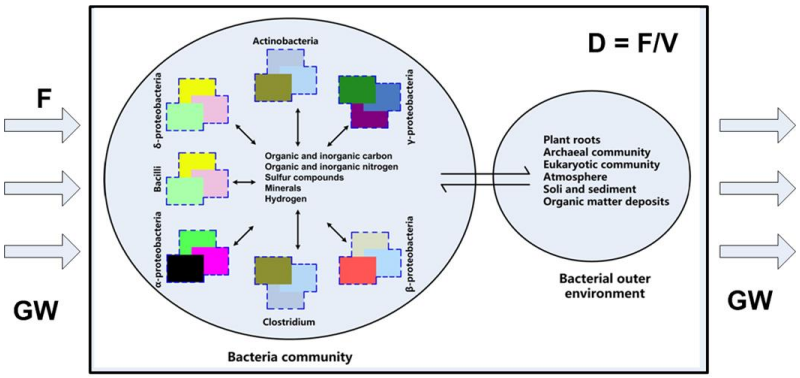


Figure 6. Conceptual overview of organisms in DMMM approach

This framework was then divided into two main components (**Figure 6**): 1) bacterial community 2) extracellular subsurface ecosystem.

The mixed microbial community model was placed in the subsurface ecosystem (outer environment), and interactions among community members as well as the exchange between the community and the outer environment were captured by a dynamic multi-species metabolic modeling (DMMM) approach.

The ecosystem model indicates that *T. denitrificans* may dominate the community at the JB site due to its ability to use inorganic electron donors for energy and fix CO_2 as its carbon source bypassing any limitations of bioavailable organic carbon (**Figure 7** Left Panel). Through electron transport with cytochrome bc1 complexes and NADH-Q oxidoreductase, a tight coupling between Fe(II) oxidation and nitrate reduction can be established to support CO_2 fixation as the main carbon source. This is in accordance with the bioinformatic analyses of taxonomic and functional profiles from the JB site indicating that *T. denitrificans* possesses a modified complete Calvin cycle for CO_2 fixation, and is the most abundant microorganism available among the seven bacterial classes. In contrast, the ecosystem model shows that in acetate amended sites, the delta-proteobacteria, *G. metallireducens*, would be numerically dominant due to its ability to fix N_2 to overcome limitations of organic nitrogen. (**Figure 7** Right Panel).

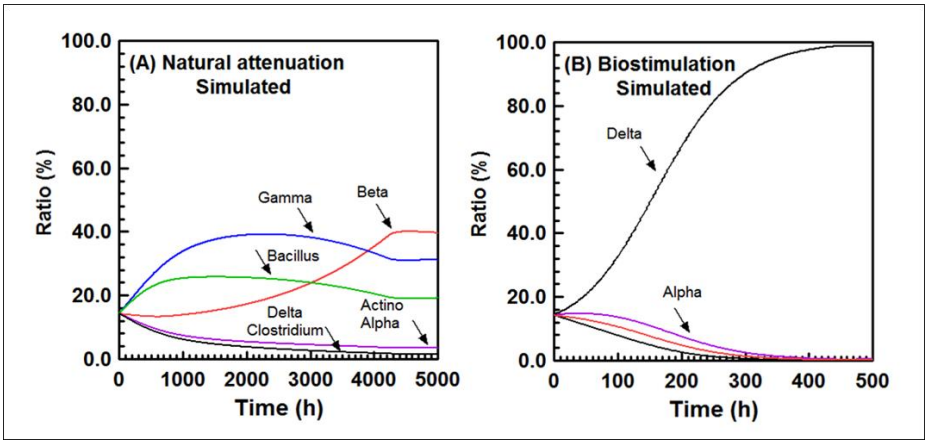


Figure 7. Simulation-based distribution of bacterial classes over time with geochemical conditions similar to sites undergoing natural attenuation (A) and with addition of acetate (B).

In addition to capture of the community structure in terms of the relative abundance approximating results measured in the RIFRC, the ecosystem model is also able to predict how the components of the microbial community may respond to changes in the environment they inhabit by altering conditions in the extracellular subsurface ecosystem component of the model.

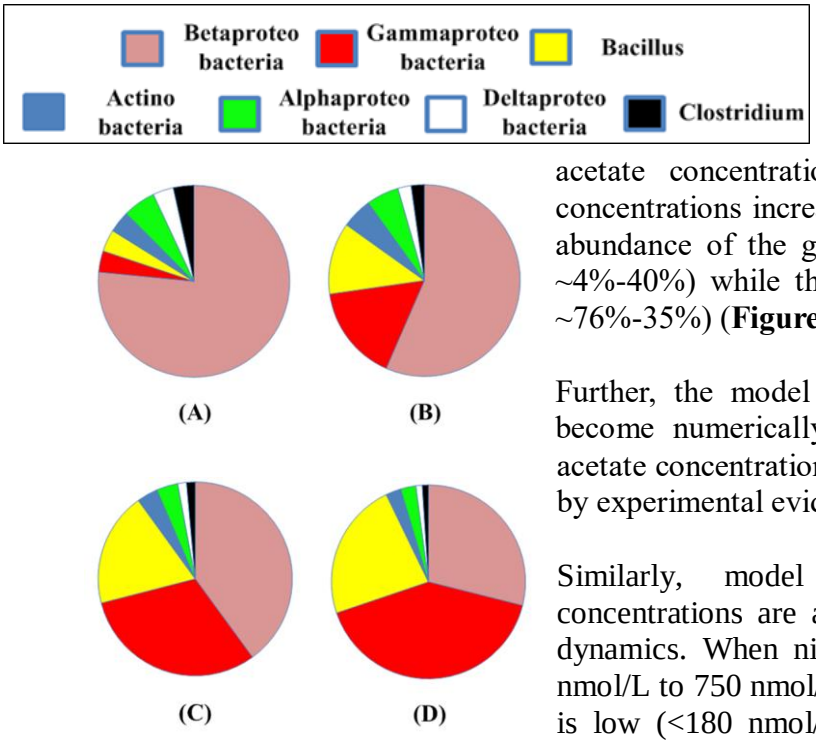


Figure 8. The predicted effect of acetate availability on the distribution of bacterial classes based on model simulations.

Model simulations show for example, that the community structure is highly responsive to acetate concentrations in the subsurface. As acetate concentrations increase (0 nmol/L to 300 nmol/L) relative abundance of the gamma-proteobacteria increases (from ~4%-40%) while the beta-proteobacteria decrease (from ~76%-35%) (**Figure 8**).

Further, the model suggests that *R. ferrireducens* may become numerically dominant over *T. denitrificans* as acetate concentrations increase (which has been supported by experimental evidence).

Similarly, model simulations show that nitrate concentrations are also important in driving community dynamics. When nitrate concentrations are increased (0 nmol/L to 750 nmol/L) and organic carbon such as acetate is low (<180 nmol/L) the model predicts a community structure similar to that measured using metaomic approaches at the JB site where beta-proteobacteria, in particular *T. denitrificans*, are numerically dominant (increasing to ~43%) (**Figure 9**). Distribution of CO_2 fixation genes at JB-05 (**Figure 10**) and at the CORFP site (**Figure 11**) were also examined for their presence and activity in the data sets based on modeling results

.

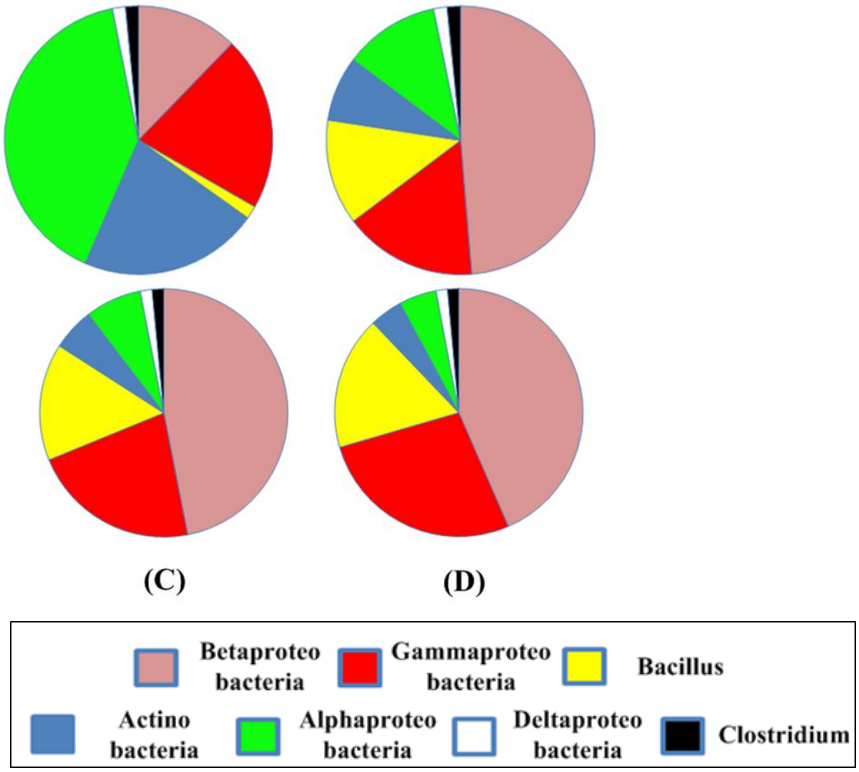
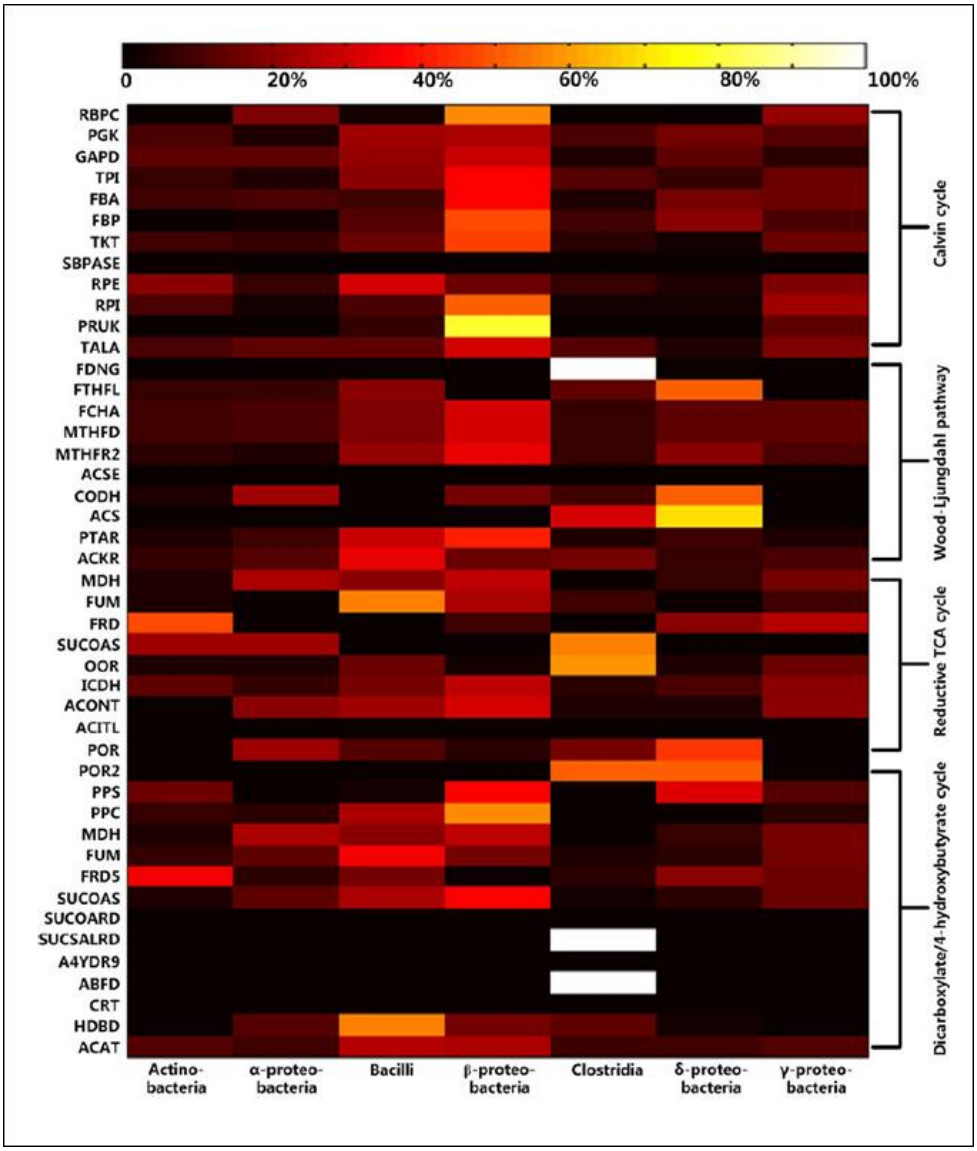
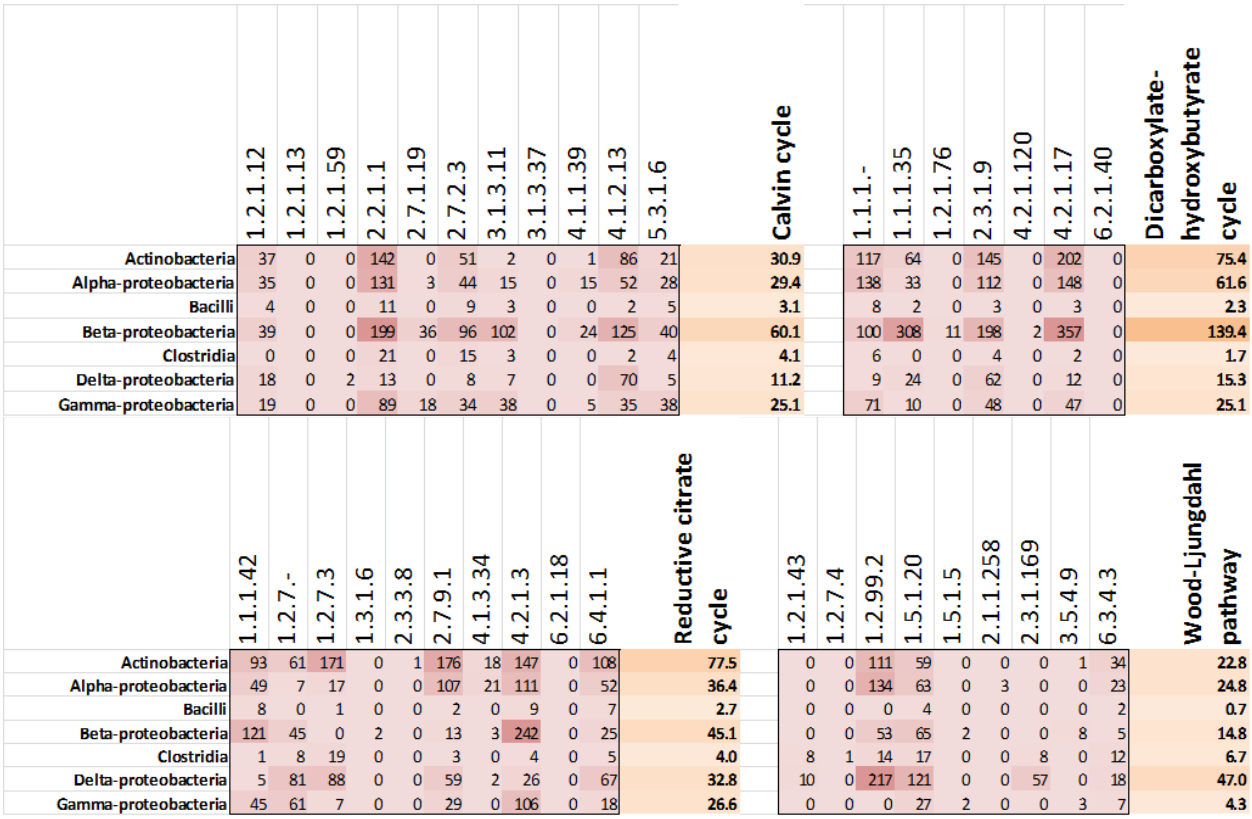


Figure 9. The predicted effect of nitrate availability on the distribution of bacterial classes based on model simulations.

Figure 10. Distribution of CO_2 fixation genes at JB-05. Key genes that participate in CO_2 fixation pathways were identified in the metagenomic datasets generated from the JB-05 site. Their abundances in the heat map are given as the relative abundances of each gene found across the 7 bacterial classes, normalized by row. Results show that CO_2 fixation pathways predominate in the beta-proteobacteria.





	1.2.1.12	1.2.1.13	1.2.1.59	2.2.1.1	2.7.1.19	2.7.2.3	3.1.3.11	3.1.3.37	4.1.1.39	4.1.2.13	5.3.1.6	Calvin cycle	1.1.1.-	1.1.1.35	1.2.1.76	2.3.1.9	4.2.1.120	4.2.1.17	6.2.1.40	Dicarboxylate-hydroxybutyrate cycle
Actinobacteria	37	0	0	142	0	51	2	0	1	86	21	30.9	117	64	0	145	0	202	0	75.4
Alpha-proteobacteria	35	0	0	131	3	44	15	0	15	52	28	29.4	138	33	0	112	0	148	0	61.6
Bacilli	4	0	0	11	0	9	3	0	0	2	5	3.1	8	2	0	3	0	3	0	2.3
Beta-proteobacteria	39	0	0	199	36	96	102	0	24	125	40	60.1	100	308	11	198	2	357	0	139.4
Clostridia	0	0	0	21	0	15	3	0	0	2	4	4.1	6	0	0	4	0	2	0	1.7
Delta-proteobacteria	18	0	2	13	0	8	7	0	0	70	5	11.2	9	24	0	62	0	12	0	15.3
Gamma-proteobacteria	19	0	0	89	18	34	38	0	5	35	38	25.1	71	10	0	48	0	47	0	25.1

	1.1.1.42	1.2.7.-	1.2.7.3	1.3.1.6	2.3.3.8	2.7.9.1	4.1.3.34	4.2.1.3	6.2.1.18	6.4.1.1	Reductive citrate cycle	1.2.1.43	1.2.7.4	1.2.99.2	1.5.1.20	1.5.1.5	2.1.1.258	2.3.1.169	3.5.4.9	6.3.4.3	Wood-Ljungdahl pathway
Actinobacteria	93	61	171	0	1	176	18	147	0	108	77.5	0	0	111	59	0	0	0	1	34	22.8
Alpha-proteobacteria	49	7	17	0	0	107	21	111	0	52	36.4	0	0	134	63	0	3	0	0	23	24.8
Bacilli	8	0	1	0	0	2	0	9	0	7	2.7	0	0	0	4	0	0	0	0	2	0.7
Beta-proteobacteria	121	45	0	2	0	13	3	242	0	25	45.1	0	0	53	65	2	0	0	8	5	14.8
Clostridia	1	8	19	0	0	3	0	4	0	5	4.0	8	1	14	17	0	0	8	0	12	6.7
Delta-proteobacteria	5	81	88	0	0	59	2	26	0	67	32.8	10	0	217	121	0	0	57	0	18	47.0
Gamma-proteobacteria	45	61	7	0	0	29	0	106	0	18	26.6	0	0	0	27	2	0	0	3	7	4.3

Figure 11. Distribution of CO_2 fixation genes at the CORFP site. Key genes that participate in CO_2 fixation pathways were identified in the metatranscriptomic (RNA) datasets generated from the CORFP site. These results reveal the importance of CO_2 fixation as a driver of carbon cycling in subsurface communities. For example, analogous to the JB-05 site, the Calvin cycle predominates in the beta-proteobacteria, while the Wood-Ljungdahl pathway is predominated by members of the delta-proteobacteria.

Counts are the number of hits to UniProt90 clusters out of $4x10^6$ annotatable reads that were randomly subsampled across multiple biological and technical replicates. EC's not exclusive to a single CO_2 fixation pathway were excluded. Pathways are based on KEGG modules (MO). The tan columns contain the average number of reads across the EC's for each bacterial class.

The ecosystem model indicates that *T. denitrificans* may dominate the community at the JB site due to its ability to use inorganic electron donors for energy and fix CO_2 as its carbon source effectively bypassing any limitations of bioavailable organic carbon. Through electron transport with cytochrome bc1 complexes and NADH-Q oxidoreductase, a tight coupling between Fe(II) oxidation and NO_3^- reduction can be established to support CO_2 fixation as the main carbon source. This is in accordance with the bioinformatic analyses of taxonomic and functional profiles from the JB site indicating that *T. denitrificans* possesses a modified complete Calvin cycle for CO_2

fixation, and is the most abundant microorganism available among the seven bacterial classes. In contrast, the ecosystem model shows that in acetate amended sites, the delta-proteobacteria, *G. metallireducens*, would be numerically dominant over other bacterial classes, due to its ability to fix N_2 to complement or overcome limitations of organic nitrogen allowing for rapid growth during biostimulation.

In addition to capture of the community structure in terms of the relative abundance approximating the RIFRC, the ecosystem model is also able to predict how the components of the microbial community may respond to changes in the environment they inhabit by altering conditions in the extracellular soil ecosystem component of the model. Results from the model simulations show for example, that the community structure is highly responsive to acetate concentrations in the subsurface. As acetate concentrations increase (from 0 nmol/l to 300 nmol/l) the relative abundance of the gamma-proteobacteria also increases (from~4%-40%) while that of the beta-proteobacteria decreases (from ~76%-35%). Further, the model suggests that *R. ferrireducens* may become numerically dominant over *T. denitrificans* as acetate concentrations increase, a finding which has been supported by experimental evidence. Similarly, the model simulations show that nitrate concentrations are are also important in driving the community dynamics. When nitrate concentrations are increased (0 nmol/l to 750 nmol/l) and organic carbon such as acetate is low (180 nmol/l) the model predicts a community structure similar to that measured using metaomic approaches at the JB site in which the beta-proteobacteria class (in particular *T. denitrificans*) are numerically dominant over the other bacterial classes (increasing to ~43%).

Overall, the community model presented here captures, at least in part, the microbial community structure that was observed using metaomic approaches at RIFRC sites and provides an important framework for continued community modeling development. The model as created here is capable of predicting the response of the community structure in changing environments such as anoxic/oxic conditons or limitations by carbon or nutrients which is critical to understanding carbon and energy flows in an ecosystem leading to improved predictions that can be used to design more efficient remediation and management strategies and better understand the implications of environmental perturbations such as climate change.

Conclusions

- The ecosystem model presented here captures, at least in part, the microbial community structure that was observed using metaomic approaches at RIFRC sites.

- The approaches presented here provide an important framework for metaomic data integration and continued community modeling development.

- We have extended the methodology based on the DMMM approach so that it can be used to model microbial community metabolism.

- The ecosystem model indicates that *T. denitrificans* may dominate the community at the JB site due to its ability to use inorganic electron donors for energy and fix CO_2 as its carbon source effectively bypassing any limitations of bioavailable organic carbon.

- The ecosystem model is also able to predict how the components of the microbial community may respond to changes, such as anoxic/oxic conditions or limitations by carbon or nutrients in the environment they inhabit by altering conditions in the extracellular subsurface ecosystem component of the model.

- Metagenomic and metatranscriptomic datasets reveal that the CORFP sites share many metabolic processes with the JB sites, especially those related to carbon cycling.