**Expression of acetate permease-like (*apl*) genes in subsurface communities of *Geobacter* species under fluctuating acetate concentrations**

Hila Elifantz[1]*, Lucie A. N'Guessan[1], Paula J. Mouser[1], Kenneth H. Williams[2], Michael J. Wilkins[3], Carla Risso[1], Dawn E. Holmes[1], Philip E. Long[4], and Derek R. Lovley[1]

[1] Department of Microbiology
University of Massachusetts
Amherst, MA 01003-9298

[2] Lawrence Berkeley National Laboratory
Earth Sciences Division
Berkeley, CA 94706

[3] Department of Earth and Planetary Science
University of California Berkeley
Berkeley, CA, 94720

[4] Pacific Northwest National Laboratory
Richland, WA 99354

* Corresponding author and current address: Hila Elifantz**,** Department of Microbial Ecology and Plant Science, The Alexander Silberman Institute of Life Sciences, The Hebrew University of Jerusalem, Givat Ram, Jerusalem, 91904, Israel. Phone:972-3-9683318 Fax: 972-3-9604017, e-mail:, helifantz@gmail.com

**Abstract**

The addition of acetate to uranium-contaminated aquifers in order to stimulate the growth and activity of *Geobacter* species that reduce uranium is a promising *in situ* bioremediation option. Optimizing this bioremediation strategy requires that sufficient acetate be added to promote *Geobacter* species growth. We hypothesized that under acetate-limiting conditions, subsurface *Geobacter* species would increase the expression of either putative acetate symporters genes (*apl*I and *apl*II). Acetate was added to a uranium-contaminated aquifer (Rifle, CO) in two continuous amendments separated by five days of groundwater flush to create changing acetate concentrations. While the expression of *apl*I in monitoring well D04 (high acetate) weakly correlated with acetate concentration over time, transcript levels for this gene were relatively constant in well D08 (low acetate). At the lowest acetate concentrations during the groundwater flush the transcript levels of *apl*II were the highest. Expression of *apl*II decreased 2-10 folds upon acetate reintroduction. However, the overall instability of acetate concentrations throughout the experiment could not support a robust conclusion regarding the role of *apl* genes in response to acetate limitation under field conditions, in contrast to previous chemostat studies, suggesting that the function of a microbial community cannot be inferred based on lab experiments alone.

**Introduction**

The microbial reduction of soluble U(VI) in contaminated aquifers to poorly soluble U(IV) has been successfully stimulated with the addition of organic electron donors in several uranium-impacted sites and the microbial community that may be involved in this process has been characterized (Anderson et al., 2003; North et al., 2004). Uranium contaminated sites are considered to be carbon/electron donor limited

(North et al., 2004) with a microbial community composition typical of low-nutrient environments (Holmes et al., 2002; North et al., 2004; Vrionis et al., 2005; Akob et al., 2007). When carbon sources and electron donors such as acetate, glucose, or ethanol are added, there is often a significant enrichment of *Geobacter* species which are considered to be the primary agent for U(VI) reduction (Anderson et al., 2003; North et al., 2004; Holmes et al., 2007; N'Guessan et al., 2008).

Promoting optimal U(VI) reduction by *Geobacter* species in the subsurface is a challenge. In uranium-contaminated subsurface environments *Geobacter* species grow primarily with Fe(III) oxides as an electron acceptor (Finneran et al., 2002). When Fe(III) oxides are depleted, the acetate-oxidizing sulfate-reducing microorganisms which are less effective in U(VI) reduction (Lovley et al., 1993) become more abundant (Vironis, et al., 2005**)**. Therefore, *Geobacter* species must move further down-gradient from the point of electron donor addition for available Fe(III) oxides (Anderson et al., 2003). This wastes acetate and results in a constantly moving front of U(IV) precipitation within the contaminated aquifer. Ideally, acetate or other electron donors, would be added at a level just high enough to promote sufficient activity of *Geobacter* species to remove U(VI) entering the acetate-amendment zone, but not deplete Fe(III) oxides any faster than is necessary.

Evaluating the physiological status of subsurface microorganisms may aid in optimizing groundwater bioremediation strategies (Lovley, 2003; Lovley et al., 2008). Previous studies have demonstrated that measurements of the abundance of gene transcripts of various genes may be useful for diagnosing the *in situ* metabolic state of *Geobacter* species in the subsurface. *In situ* transcript levels of citrate synthase genes and two genes encoding multi-copper containing proteins in *Geobacter* were related to differences in growth rates of *Geobacter* species in the subsurface (Holmes et al.,

2005; Holmes et al., 2008). The transcript levels of the nitrogen fixation gene, *nif*D, in petroleum-contaminated subsurface sediments and in uranium-contaminated groundwater were indicative of whether the *Geobacter* species were limited for fixed nitrogen (Holmes et al., 2004; Mouser at al., 2009b). High transcript levels of *feoB*, which encodes a protein involved in iron uptake in *Geobacter* species, suggested that the concentrations of dissolved Fe(II) in the subsurface may be insufficient to meet the high needs of *Geobacter* species for assimilation into iron-sulfur proteins (O'Neil et al., 2008).

The purpose of the current study was to determine whether quantitative analysis of the transcripts levels of acetate permease-like (*apl*) genes in the subsurface could serve as a guide for whether the metabolism of *Geobacter* species was limited by acetate availability during a biostimulation experiment. We hypothesized that when acetate concentrations are sufficient; there is a low transcription of *apl* genes, and vice versa. Acetate permease (ActP) was first identified in *Escherichia coli* and was found to be cotranscribed with the gene *acs*, coding for acetyl coenzyme A synthetase (Gimenez et al., 2003). A dependence on transmembrane electrochemical potential was suggested as the transport mechanism (Gimenez et al., 2003). A study with *Geobacter sulfurreducens* revealed four genes with homology to ActP (Risso et al., 2008). Continuous cultures of *G. sulfurreducens* grown under acetate-limiting conditions had higher transcript levels of *apl*I compared to cultures in which acetate was not limiting (Risso et al., 2008). Closely related genes were found in the genomes of other pure cultures of *Geobacter* species as well as in genomic DNA extracted from a subsurface environment undergoing *in situ* uranium bioremediation (Risso et al., 2008).

## Materials and Methods

### Environmental setup and sample collection

Groundwater samples from the uranium-contaminated aquifer were collected during the 2007 summer biostimulation experiment at the Old Rifle site, CO, as previously described (Mouser et al., 2009b and Fig.1). Average velocity of groundwater was 0.5m/d and the distance between injection gallery and first row of monitoring wells was 2.5m (Mouser et al., 2009b). In order to track the response of the microbial community to acetate limitation, the acetate injection was stopped after 10 days and the injection gallery was flushed with unamended groundwater. The injection of acetate was resumed when acetate concentrations were close to 0 mM in the first row of monitoring wells (D01-D04, Fig 1). Prior to sample collection, 12L of water was purged from each well in order to stabilize groundwater parameters as described previously (Anderson et al., 2003) and water samples were collected for geochemical analyses (Mouser et al., 2009b). For genomic DNA, about 1.5L of groundwater was filtered through a 0.22μm Sterivex filter (Millipore). For RNA, about 10L of groundwater was pumped and filtered every 2-3 days onto 293 mm 0.2 μm Supor membrane filter (Pall Life Sciences). Filtration time did not exceed 10 min and the filters were flash frozen in ethanol-dry ice bath and shipped on dry ice to the laboratory for further analyses. A total of 10 (D04), 9 (D08), and 10 (U-1) samples were analyzed for mRNA abundance throughout the experiment.

### Nucleic acids extraction

RNA was extracted from groundwater samples as previously described (Holmes et al., 2004) with a few modifications. tRNA was not added to the procedure of mRNA extraction. Linear acrylamide was excluded from the precipitation step.

Finally, additional DNAse treatment was added at the end of the extraction using a DNA-free™ kit (Ambion Inc, Austin, TX) for all samples. Genomic DNA was extracted using the BIO101 Kit (Bio 101 Inc., Vista, CA) according to the manufactures' instructions.

**Microbial community and *apl* genes diversity**

The microbial diversity was determined for the beginning of the experiment and for the maximum and minimum of acetate in each well. The microbial diversity was evaluated by amplifying the 16S rRNA gene from genomic DNA using the general primers 8F and 519R (Lane et al., 1985; Eden et al., 1991) with annealing temperature of 50˚C. The amplified fragments were resolved on an agarose gels and the band in the expected size was excised and cleaned with QIAquick Gel Extraction Kit (QIAGEN Sciences, MD, USA) according to the manufacture's instructions. The cleaned fragments were then cloned using the TOPO TA Cloning Kit for sequencing (Invitrogen) into TOP10 chemically competent *E. coli* (Invitrogen) according to the manufacturer's instructions. Ninety-six clones were sequenced from each library using ABI BigDye® Terminator v3.1 (Applied Biosystems) and an ABI DNA Analyzer 3730xl (Applied Biosystems). Base calling of the resulted electopherograms was done using the SEQMAN program (DNASTAR) and the good sequences were aligned to the 16S rRNA database using Arb (Ludwig et al., 2004). The affiliation of the various clones was determined by inspecting their location in the Arb phylogenetic tree and by BLAST analysis (Altschul et al., 1997). A detailed analysis of the deltaproteobacteria was achieved by creating a distance matrix in Arb and determining OTUs in Dotur (Schloss and Handelsman, 2005). Representatives of each OTU were blast against the NCBI database (Altschul et al., 1997). The 16S rRNA

gene sequences were deposited into Genebank under accession numbers GU559474 - GU559701.

Environmental *apl*I genes were amplified using primer set designed elsewhere (Risso et al., 2008). A primer set for *apl*II was designed based on the available genomes for *Geobacteraceae*. The degenerate primers were as follow: *apl*II-157F (5'-GCCAGYAACTGGATGAG) and apl*II*-420R (5'-KGSCGAKBCCCYTGAACTGT). Numbers in primers names represent locations in the equivalent sequences in *G. uraniireducens*. The diversity of *apl* genes was assessed in the genomic DNA from the beginning (4 days after acetate injection), middle (high acetate), and end (low acetate) of the experiment. The expressed fraction diversity (mRNA) was studied for the middle of the experiment at two monitoring wells (D04 and D08). The amplified *apl*I and II from either genomic environmental DNA or cDNA (constructed as described below) was resolved on 3% agarose gel and the band in the expected size for each *apl* gene was excised, cleaned, and cloned as described above. The sequences were inspected and the diversity indexes were calculated as described in detail elsewhere (Elifantz et al., 2008) and the distance matrix and a functional genes trees (trees are not shown) were used to determine clusters of environmental *apl*I and II in MEGA (Tamura et al., 2007). A cluster was defined by a similarity identity of above 90% between the sequences included in the cluster. The percentage similarities to genes closely related from *Geobacter* sp. M21 were determined as well using a distance matrix in MegAlign (DNASTAR). Some of the sequences were used for environmental QPCR primer design as described in the Results and Discussion sections. The *apl* sequences were deposited into Genebank as follows: *apl*I from genomic DNA - GU559128 - GU559248; *apl*I from mRNA - GU559249 -

GU559319; *apl*II from genomic DNA - GU559320 - GU559391; *apl*II from mRNA - GU559392 - GU559473.

**Reverse transcriptase - Quantitative PCR (QPCR)**

The reverse transcriptase enhanced avian kit (Sigma) was used to reverse complement RNA from environmental samples according to manufacturer's instructions. RNA was diluted 1:10 prior to the cDNA preparation in order to eliminate inhibition. Subsequently, the cDNA was diluted 1:10 and the dilution was used for the qPCR reaction with Power SYBR green PCR Master Mix (Applied Biosystems, UK) and 150nM of each of the primers. Primer sets were designed based on the dominant gene cluster for each of the *apl* genes as described in the results section. For *apl*I, a 120bp fragment was amplified with the primer set *apl*I-F29 (GGCCTGATCTCGCTCTA) and *apl*I-R149 (CCAAGGTGTATTTGCCGGCGTTT). For *apl*II, a 196bp fragment was amplified using the primer set *apl*II-F31 (GGGCTCGCCGGGATCATCTA) and *apl*I-R227 (ATGGAGATCATGGCGGCGAT). QPCR reactions were performed on 7500 ABI real time (Applied Biosystems) according to the ABI guide. Both *apl* genes transcripts were normalized to the housekeeping gene *recA* (Mouser et al., 2009b). The propagation of error for this normalization was calculated according to the following equation: Dz/z = ((Dx/x)^2+(Dy/y)^2)^0.5, where Dz is the standard error of the normalized value, z is the normalized value, Dx and Dy are the standard error of *apl* and *rec*A, respectively, and x and y are the values of transcribed *apl* and *rec*A respectively (http://www.rit.edu/cos/uphysics/uncertainties/Uncertaintiespart2.html).

## Results

### Acetate Delivery and Concentrations

Two monitoring wells, D04 and D08, were chosen for study due to the different concentrations of acetate each of the wells received during the experiment. These differences were the result of the direction of the groundwater flow throughout the field experiment (Mouser et al., 2009b). More acetate was delivered to D04 than D08, resulting in overall higher acetate concentrations at D04, but the general trend in acetate fluctuations was the same at both sites. To test the expression pattern of acetate permease-like (*apl*) genes, acetate was allowed to reach a maximum of 3.5mM and 0.6mM in wells D04 and D08, respectively (Figs. 2 and 3). On day 10, the acetate delivery was stopped and the plot was flushed with groundwater for seven days. In D04 the lowest acetate concentration detected during that time was 0.4 mM, while in D08 acetate was undetectable at day 23 (Figs. 2 and 3). After day 17 the acetate was re-introduced to the experimental plot and concentrations increased until the end of the experiment from day 23.

### Diversity analysis of the genomic and expressed *apl* genes over the course of the experiment

As previously reported, the acetate addition stimulated dissimilatory metal reduction with *Geobacter* species accounting for the majority of the subsurface microbial community (Mouser et al., 2009b). When acetate peaked in wells D04 and D08 most of *Geobacter*-like 16S rRNA genes were most similar to *Geobacter bemidjiensis* Bem (Table S1 of Supporting Information), an isolate from the Fe(III)-reducing zone of a petroleum-contaminated aquifer (Nevin et al., 2005) . As acetate decreased due to the groundwater flush, the community shifted to strains that were

more similar to *Geobacter uraniireducens* Rf4 (Table S1 of the Supporting Information), a bacterium that was isolated from the Rifle site (Shelobolina et al., 2008).

Analysis of *apl*-like genes in the subsurface community, amplified with PCR primers designed to recover these genes from *Geobacter* species revealed a diversity of *apl* genes (Fig. 4). When these sequences were clustered by a similarity identity of above 90% it was apparent that there were more clusters of *apl*I than *apl*II. Only half of the *apl*I clusters contained representatives at all three sampling dates, but six of the eight *apl*II clusters contained genes in all sampling times. A dominant *apl*I cluster could be assigned for each of the sampling points; clusters L, A, and F for 4, 10, and 20 days after the initial acetate injection, respectively (Fig. 4A). In contrast, cluster S of the *apl*II gene family, was the major gene cluster at the beginning and end of the experiment, whereas cluster M was dominant in the middle of the experiment (Fig. 4B).

The genes and the gene transcripts retrieved on day 10, before the decrease in acetate concentration occurred were analyzed in more detail. Overall, *apl*I was more diverse than *apl*II at the genomic DNA level as suggested by the Shannon diversity index and the number of gene types (Table 1). In general, the environmental *apl* genes clustered separately from the known homologues in pure cultures (trees are not shown). Environmental *apl*I and *apl*II were 66-86% and 70-90% similar to the corresponding sequences from the isolate *Geobacter* sp. M-21 (genes 1542 and 1545 respectively) (Table 2). *Geobacter* sp. M-21 was isolated from the Rifle site from a sample recovered during the acetate-addition phase of a previous field experiment (Holmes et al., In Prep). Cluster analysis revealed that three *apl* clusters (1-A, 1-J, and 2-L Table 2) had more sequences in the cDNA library than in the genomic DNA

library. The rest of the clusters had either equal amount or fewer sequences in the transcribed fraction compared to the genomic DNA. Clusters 2-N and 2-O, which could not be detected in the genomic DNA, comprised about 22% of the transcripts (Table 2). Overall, the diversity of *apl*I and *apl*II genes was 2-2.2 fold higher than that of the transcripts for each *apl* super-family (Table 1).

**Expression of *apl* genes over time**

Monitoring the transcript abundance of the full diversity of *apl* genes over time in the field experiment was not feasible. Therefore, the genes in each *apl* family with the highest abundance of clones in the cDNA libraries (cluster 1-A for *apl*I and cluster 2-L for *apl*I, Table 2) were chosen for further investigation over time. Expression levels were normalized to the abundance of *Geobacter* transcripts for *recA*, a housekeeping gene whose expression is relatively constant under a variety of physiological conditions (Holmes et al., 2008; Mouser et al., 2009b). This normalization corrects for potential differences in mRNA extraction or other sampling discrepancies.

The two *apl* genes had slightly different response to the fluctuations in acetate concentrations over the course of the experiment. The relative abundance of *apl*I transcripts increased as acetate initially increased at D04 (Fig. 2). During the period of consistently high acetate concentrations between day 5 and 15 there was substantial variability (ca. 3-fold) in the *apl*I transcript abundance. A decrease in the acetate availability between day 15 and 17 was accompanied by a decrease in the relative expression of *apl*I. However, as acetate continued to decrease to day 20, an increase in *apl*I transcription was observed, and this increase continued until the end of the experiment. Overall, there was a weak positive correlation ($r = 0.36$, $p > 0.05$))

between *apl*I transcript abundance and acetate concentrations (Table 3 and Fig. S1). At D08 *apl*I transcript abundance remained relatively constant throughout the field experiment (Fig. 3). Most notably, when acetate concentrations became undetectable on day 23, the *apl*I transcript abundance was comparable to the abundance at previous time points when acetate was available, although in low concentration. Correlation between acetate and *apl*I transcript in D08 was low but significant ($r = 0.206$, $p < 0.05$).

The transcript abundance of *apl*II did not correspond with acetate concentrations in either sampling location. At D04, transcript abundance at some of the highest and lowest acetate concentrations was comparable (days 14 and 20) and there was a wide range of *apl*II transcript abundance during the period when acetate concentrations were the highest (Fig. 2). Similar to *apl*I, the decrease in acetate between day 15 and 17 was accompanied by a decrease in *apl*II transcripts levels. However, in contrast to *apl*I, the further decrease in acetate concentrations until day 20 was followed by an increase in *apl*II transcript levels (Fig. 2). There was an additional response to the acetate concentrations at the end of the experiment, as acetate was increasing at day 23, *apl*II levels decreased. At D08 *apl*II transcripts remained in a narrow range between days 5 and 23, despite changes in the acetate concentration (Fig. 3). The expression of *apl*II was negatively correlated with acetate ($r = - 0.63$, $p < 0.05$)) in that well (Table 3 and Fig. 1S). The expression of both *apl* genes could not be detected in a monitoring well located up-gradient of the injection wells (data is not shown).

**Discussion**

*In situ* bioremediation of uranium-contaminated groundwater has been studied for a while now (Finneran et al., 2002; Holmes et al., 2002; Anderson et al., 2003;

Istok et al., 2004) and at this stage of the research the challenge is to fine tune and find the optimal conditions for the remediation to occur efficiently and cost affectively (Holmes et al., 2008; N'Guessan et al., 2008; O'Neil et al., 2008; Mouser et al., 2009b). In order to do so, the physiology and ecology of the *Geobacter* community in-situ need to be investigated thoroughly. Although proteomics has become a useful tool for studying the physiology of microbial communities in the environment (Goltsman et al., 2009; Wilkins et al. 2009; Wilmes and Bond, 2009), methods to accurately quantify proteins in the environment have yet to be developed. Therefore, the physiology of a community can best be assessed by monitoring the transcription of key genes for various processes including nitrogen assimilation (Holmes et al., 2004; Mouser et al., 2009b), phosphate assimilation (N'Guessan et al, 2010), and other important elements (Holmes et al., 2008; Mouser et al., 2009a; O'Neil et al., 2008). The aim of this study was to evaluate if the transcription of *apl* genes might indicate the physiological state of the *Geobacter* community in regard to acetate availability.

As found in previous studies at the Rifle site, the microbial community in the groundwater changed over time according to the changing conditions during the field experiment (Anderson et al., 2003; Vrionis et al., 2005; Mouser et al., 2009b) and reflects the community on the sediments (Anderson et al, 2003). A detailed inspection of the bacteria in the deltaproteobacteria suggested that the majority belonged to the previously described *Geobacter* sub-surface clade 1 (Holmes et al., 2007). Within this clade, most of the OTUs were similar to either *Geobacter bemidjiensis* Bem or *Geobacter uraniireducens* Rf4. The observed variations in gene dominance could be expected as the community composition changes over time as environmental conditions change (Mouser et al., 2009b).

The variations in *Geobacter* species, as well as the shifts in the community composition throughout the experiment were also reflected by the diversity observed for the *apl* genes. Overall, there was a higher diversity of *apl*I genes compared to *apl*II genes in the environmental genomic DNA. This result is in accordance with *Geobacteraceae* genomes wherein there are several copies of *apl*I and only one copy of *apl*II per genome (Risso et al., 2008). However, the similarity to the *apl*I and *apl*II genes of isolates from the Rifle site was not more than 90%. It is not unusual that genes with similar functions are different between cultured organisms and environmental samples (Zehr et al., 2003; LeCleir et al., 2004; Elifantz et al., 2008). But, we expected higher similarity between genes of isolates and genes in the community of the environment from which they were isolated. As many microorganisms have not yet been cultivated (Rappe and Giovannoni, 2003), it is possible that the diversity of *Geobacter* species reflected by the 16S rRNA genes is also the reason for the high diversity in *apl* genes in the environment.

Although the diversity of functional genes can be informative regarding the potential of a community to perform a certain activity (Gray and Head, 2001; Zehr et al., 2003), it does not mean that all the gene types present in the environment are transcribed. Some of the *apl* genes were transcribed as expected from their abundance in the DNA library, while others where over-expressed. In a couple of cases (2-N and 2-O, table 2) the gene type was detectable only at the transcription level. Similarly, only a portion of the 16S rRNA genes was expressed at another uranium-contaminated site (Akob at al., 2007). The authors suggested that while the lack of 16S rRNA genes in the genomic DNA-based library was a result of low abundance that could not be detected, occurrence of the transcripts in the mRNA-based library implies that these bacteria were highly active. Therefore, the frequency of *apl* clones

in the genomic and mRNA based libraries may indicate which clusters of the *apl* genes were important in the environment under fluctuating acetate conditions. The gene clusters that had higher amounts of mRNA than DNA based sequences may belong to bacteria that were more sensitive to acetate availability and therefore more active.

In general, studies with pure cultures of *Geobacter* species have provided good insights into how gene expression in subsurface communities of *Geobacter* may be used to help diagnose the physiological status of the subsurface *Geobacter* species (Holmes et al., 2004; Holmes et al, 2008; O'Neil et al., 2008; Mouser et al., 2009a; Mouser et al., 2009b). However, the study presented here demonstrates that results from chemostat studies with *Geobacter sulfurreducens* (Risso et al., 2008) may not always translate to subsurface communities. The expression of *apl*I genes were substantially higher when *Geobacter sulfurreducens* was grown under acetate-limiting conditions than under conditions in which acetate was not limiting, leading to the suggestion that quantifying *apl*I transcripts during bioremediation of uranium-contaminated groundwater would make it possible to detect whether cells had sufficient acetate or were acetate-limited (Risso et al., 2008). However, there was no correspondence between the abundance of the predominant and most actively transcribed cluster of *Geobacter apl*I genes in the subsurface and acetate availability. However, the negative, significant correlation between *apl*II and acetate concentrations in D08 may suggest that this gene may respond to low acetate concentrations in the environment.

There are many potential explanations for the differences in the response between the pure cultures and the subsurface community. The pure cultures were grown under steady-state conditions whereas in the subsurface conditions were far

from steady state. The subsurface community was perturbed with the initial inputs of high concentrations of acetate, followed by turning off and then turning back on the acetate inputs. The results of previous studies suggest that the availability of Fe(III), the primary electron acceptor supporting growth, is likely to have been changing throughout the field experiment (Finneran et al., 2002; Michalsen et al., 2006). This and other geochemical changes associated with high rates of acetate oxidation coupled to Fe(III) reduction result in a constantly changing environment for the subsurface *Geobacter* species. Furthermore, factors other than external acetate concentrations may control the expression of acetate transporter genes in *Geobacter* species. For example, the initial addition of acetate to the subsurface promotes a rapid growth of *Geobacter* species. High demand for acetate to meet the requirements for rapid growth may increase the expression of acetate transporters even when external acetate concentrations are high. In addition, other bacteria in the subsurface may compete for the acetate, adding to this constantly changing environment.

While the main hypothesis put to the test in this study could not be confirmed, some valuable information was collected nevertheless. The high diversity in *apl* genes combined with the community composition by 16S rRNA clone libraries may suggest the presence of a much more complex and versatile community in the subsurface than previously considered. The results also suggest that the investigation of expression patterns of highly diverse genes such as the *apl* genes may not be as feasible as it is for other genes such as citrate synthase and *nif*D (Holmes et al., 2004; Holmes et al., 2005). Future metagenomics work should make it possible to relate the different species with functional genes, which may help in finding better gene targets to be used for physiology studies.

**Acknowledgments**

We thank Joy Ward and Betsy Blunt for the technical support. We thank our collaborators at the Integrated Field Challenge program in Old Rifle, CO. This research was supported by the Office of Science (BER), U.S. Department of Energy, under Contract DE-AC02-05CH11231 and Cooperative Agreement No. DE-FC02-02ER63446 and Grant No. DE-FG02-07ER64377.

**References**

Akob DM, Mills HJ, and Kostka JE (2007) Metabolically active microbial communities in uranium-contaminated subsurface sediments. *FEMS Microbiol Ecol* **59**: 95-107

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, and Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**:3389-3402.

Anderson, RT, Vrionis, HA, Ortiz-Bernard, I. Resch, CT, Long, PE, Dayvault, R. Karp K, Marutzky S, Metzler DR, Peacock AD, White DC, Lowe M, Lovley DR (2003) Stimulating the in situ activity of *Geobacter* species to remove uranium from the groundwater of a uranium-contaminated aquifer. *Appl Environ Microbiol* **69**: 5884–5891

Eden, PE, Schmidt, TM, Blakemore, RP, and Pace, NR (1991) Phylogenetic analysis of Aquaspirillum magnetotacticum using polymerase chain reaction-amplified 16S rRNA-specific DNA. *Int J Syst Bacteriol* **41**: 324-325.

Elifantz, H, Waidner, LA, Michelou, VK, Cottrell, MT, and Kirchman, DL (2008) Diveristy and abundance of glycosyl hydrolase family 5 in the North Atlantic Ocean. *FEMS Microbiol Ecol* **63**: 316-327.

Finneran, KT, Anderson, RT, Nevin, KP, and Lovley, DR (2002) Potential for bioremediation of uranium-contaminated aquifers with microbial U(VI) reduction. *Soil and Sediment Contam* **11**: 339-357.

Gimenez, R, Nuñez, MF, Badia, J, Aguilar, J, and Baldoma, L (2003) The gene *yjc*G, cotranscribed with the gene *acs*, encodes an acetate permease in *Escherichia coli*. *J Bacteriol* **185**: 6448-6455

Goltsman, DSA, Denef, VJ, Singer, SW, VerBerkmoes, NC, Lefsrud, M, Mueller, RS, Dick, GJ, Sun, CL, Wheeler, KE, Zemla, A, Baker, BJ, Hauser, L, Land, M, Shah, MB, Thelen, MP, Hettich, RL, Banfield, JF (2009) Community genomics and proteomics analyses of chemoautotrophic iron-oxidizing "*Leptospirillum rubarum*" (group II) and "*Leptospirillum ferrodiazotrophum*" (group III) bacteria in acid mine drainage biofilms. *Appl Environ Microbiol* **75**: 4599-4615

Gray, ND, and Head, IM (2001) Linking genetic identity and function in communities of uncultured bacteria. *Environ Microbiol* **3**: 481-492.

Holmes, DE, Nevin, KP, and Lovley, DR (2004) In situ expression of nifD in *Geobacteraceae* in subsurface sediments. *Appl Environ Microbiol* **70**: 7251-7259.

Holmes, DE, Finneran, KT, O'Neil, RA, and Lovley, DR (2002) Enrichment of members of the family *Geobacteraceae* associated with stimulation of dissimilatory metal reduction in uranium-contaminated aquifer sediments. *Appl Environ Microbiol* **68**: 2300-2306.

Holmes, DE, Nevin, KP, O'Neil, RA, Ward, JE, Adams, LA, Woodard, TL, Vrionis, HA, and Lovley, DR (2005) Potential for quantifying expression of the *Geobacteraceae* citrate synthase gene to assess the activity of *Geobacteraceae* in the subsurface and on current-harvesting electrodes. *Appl Environ Microbiol* **71**: 6870-6877.

Holmes, DE, Mester, T, O'Neil, RA, Perpetua, LA, Larrahondo, MJ, Glaven, R, Sharma, ML, Ward, JE, Nevin, KP, and Lovley, DR (2008) Genes for two multicopper proteins required for Fe(III) oxide reduction in *Geobacter sulfureducens* have different expression patterns both in the subsurface and on energy-harvesting electrodes. *Microbiology* **154**: 1422-1435.

Holmes, DE, O'Neil, RA, Vrionis, HA, N'Guessan, AL, Ortiz-Bernard, I, Larrahondo, MJ, Adams LA, Ward JA, Nicoll JS, Nevin KP, Chavan MA, Johnson JP, Long PE, and Lovley DR. (2007) Subsurface clade of *Geobacteraceae* that predominates in a diversity of Fe(III)-reducing subsurface environments. *ISME J* **1**: 663-677.

Istok, JD, Senko, JM, Krumholz, LR, Watson, D, Bogle, MA, Peacock, AD, Chang YJ, White DC (2004) In situ bioremediation of technetium and uranium in a nitrate-contaminated aquifer. *Environ Sci Technol* **38**: 468-475.

Lane, DJ, Pace, B, Olsen, GJ, Stahl, D, Sogin, ML, and Pace, NR (1985) Rapid determination of 16S ribosomal RNA sequences for phylogenetic analysis. *P Natl Acad Sci USA* **82**: 6955-6959.

LeCleir, GR, Buchan, A, and Hollibaugh, JT (2004) Chitinase gene sequences retrieved from diverse aquatic habitats reveal environment-specific distributions. *Appl Environ Microbiol* **70**: 6977-6983.

Lovley, DR (2003) Cleaning up with genomics: applying molecular biology to bioremediation. *Nature Rev Microbiol* **1**: 35-44.

Lovley, DR, Mahadevan, R, and Nevin, KP (2008) Systems biology approach to bioremediation with extracellular electron transfer. In *Microbial biodegradation, Genomics and molecular biology*. Diaz, E. (ed). Norfolk, UK: Caister Academic Press, pp. 71-96.

Lovley, DR, Roden, EE, Phillips, EJP, and Woodward, JC (1993) Enzymatic iron and uranium reduction by sulfate-reducing bacteria. *Mar Geol* **113**: 41-53.

Ludwig, W, Strunk, O, Westram, R, Richter, L, Meier, H, Yadhukumar. et al. (2004) ARB: a software environment for sequence data. *Nucleic Acids Res* **32**: 1363-1371.

Michalsen, MM, Goodman, BA, Kelly, SD, Kemner, KM, McKinley, JP, Stucki, JW, and Istok, JD (2006) Uranium and technetium bio-immobilization in intermediate-scale physical models of an in situ biobarrier. *Environ Sci Technol* **40**: 7048-7053.

Mouser, PJ, Holmes, DE, Perpetua, LA, DiDonato, R, Postier, B, Liu, A, and Lovley, DR (2009) Quantifying expression of *Geobacter* spp. oxidative stress genes in pure culture and during in situ uranium bioremediation. *ISME J* **3**:454-65.

Mouser, PJ, N'Guessan, AL, Elifantz, H, Holmes, DE, Williams, KH, Wilkins, MJ, Long PE, and Lovley DR (2009) Influence of heterogeneous ammonium availability on bacterial community structure and the expression of nitrogen fixation and ammonium transporter genes during in situ bioremediation of uranium-contaminated groundwater. *Environ Sci Technol* **43**: 4386-4392.

Nevin, KP, Holmes, DE, Woodard, TL, Hinlein, ES, Ostendorf, DW, and Lovley, DR (2005) *Geobacter bemidjiensis* sp. nov. and *Geobacter psychrophilus* sp. nov., two new Fe(III)-reducing subsurface isolates. *Int J Syst Evol Microbiol* **55**: 1667-1674

N'Guessan, AL, Elifantz, H, Nevin, KP, Mouser, PJ, Methe, BA, Woodward, TL, Manley, K, Williams, KH, Wilkins, MJ, Larsen, JT, Long, PE, and Lovley, DR (2010) Molecular analysis of phosphate limitation in *Geobacteraceae* during the bioremediation of a uranium-contaminated aquifer. ISME J. **4**: 253-266

N'Guessan, AL, Vrionis, HA, Resch, CT, Long, PE, and Lovley, DR (2008) Sustained removal of uranium from contaminated groundwater following stimulation of dissimilatory metal reduction. *Environ Sci Technol* **42**: 2999-3004.

North, NN, Dollhopf, SL, Petrie, L, Istok, JD, Balkwill, DL, and Kostka, JE (2004) Changes in bacterial community structure during in situ biostimulation of subsurface sediment co-contaminated with uranium and nitrate. *Appl Environ Microbiol* **70**: 4911-4920.

O'Neil, RA, Holmes, DE, Coppi, MV, Adams, LA, Larrahondo, MJ, Ward, JE Nevin KP, Woodard TL, Vrionis HA, N'Guessan AL, Lovley DR (2008) Gene transcript analysis of assimilatory iron limitation in *Geobacteraceae* during groundwater bioremediation. *Environ Microb* **10**: 1218-1230.

Rappe, MS, and Giovannoni, SJ (2003) The uncultured microbial majority. *Annu Rev Microbiol* **57**: 369-394.

Risso, C., Methe, BA, Elifantz, H, Holmes, DE, and Lovley, DR (2008) Highly conserved genes in *Geobacter* species with expression patterns indicative of acetate limitation. *Microbiology* **154**: 2589-2599.

Schloss, PD, and Handelsman, J (2005) Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. *Appl Environ Microbiol* **71**: 1501-1506.

Shelobolina, ES, Vrionis, HA, Findlay, RH, and Lovley, DR (2008) *Geobacter uraniireducens* sp. nov., isolated from subsurface sediment undergoing uranium bioremediation. *Int J Syst Evol Microbiol* **58**: 1075-1078.

Tamura, K, Dudley, J, Nei, M and Kumar, S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* **24**: 1596-1599

Vrionis, HA, Anderson, RT, Ortiz-Bernard, I, O'Neil, RA, Resch, CT, Peacock,

AD. Dayvault R, White DC, Long PE, Lovley DR. (2005) Microbiological and geochemical heterogeneity in an in situ uranium bioremediation field site. *Appl Environ Microbiol* **71**: 6308-6318.

Wilkins, MJ, VerBerkmoes, NC, Williams, KH, Callister, SJ, Mouser, PJ, Elifantz, H, N'Guessan, AL, Thomas, BC, Nicora, CD, Shah, MB, Abraham, P, Lipton, MS, Lovley, DR, Hettich, RL, Long, PE, and Banfield, JF (2009) Proteogenomic monitoring of *Geobacter* physiology during stimulated uranium bioremediation. *Appl Environ Microbiol* **75**: 6591-6599.

Wilmes, P, and Bond, PL (2009) Microbial community proteomics: elucidating the catalysts and metabolic mechanisms that drive the Earth's biogeochemical cycles. *Curr Opin Microbiol* **12**: 310-317

Zehr, JP, Jenkins, BD, Short, SM, and Steward, GF (2003) Nitrogenase gene diversity and microbial community structure: a cross-system comparison. *Environ Microbiol* **5**: 539-554

Figure legends:

Figure 1: Site layout at the Rifle, CO. experiment in 2007. The closed circles represent the studied wells (U-01, D04, and D08). The long arrow indicates the direction of groundwater flow.

Figure 2: Acetate concentrations and transcripts levels of *apl*I and *apl*II normalized to *rec*A in well D04 over the course of the biostimulation experiment at Old Rifle, CO. Error bars represent propagation of errors of triplicates QPCR reactions. Closed symbols – *apl* genes, open squares – acetate. Black star indicate beginning of groundwater flush, white star indicate re-introduction of acetate.

Figure 3: Acetate concentrations and transcripts levels of *apl*I and *apl*II normalized to *rec*A in well D08 over the course of the biostimulation experiment at Old Rifle, CO. Error bars represent propagation of errors of triplicates QPCR reactions. Closed symbols – *apl* genes, open squares – acetate. Black star indicate beginning of groundwater flush, white star indicate re-introduction of acetate.

Figure 4: *apl* I (A) and *apl* II (B) gene clusters from three sampling times: 4 days (08/12/07), 10 days (08/18/07), and 20 days (08/28/07) after initial acetate injection. Clusters were determined by 90% similarity at the nucleic acid level. UN = Sequences that were unaffiliated with either cluster.

Figure S1: Transcripts levels of *apl*I (A) and *apl*II (B) as a function of acetate concentration in the uranium-contaminated aquifer in Old Rifle, CO. Error bars

represent propagation of errors of triplicates QPCR reactions. Correlation coefficient for each gene and well is indicated within the panels.

Table 1: Diversity analysis of *apl* genes from environmental genomic DNA and mRNA at the 90% similarity. Shannon diversity index is given with 95% confidence levels in parenthesis.

| | *apl* I | | *apl* II | |
|---|---|---|---|---|
| | DNA | cDNA | DNA | cDNA |
| Shannon diversity index | 3.35 (±0.18) | 2.19 (±0.29) | 2.65 (±0.27) | 2.18 (±0.17) |
| Number of Gene types[a] | 45 | 20 | 27 | 13 |
| Evenness[b] | 0.88 | 0.73 | 0.80 | 0.84 |
| Coverage[c] | 0.64 | 0.75 | 0.64 | 0.84 |

[a] Number of Gene types were calculated using a rarefaction curves at the 90% similarity (equivalent to operational taxonomic units).

[b] Evenness was calculated according to H/lnS, H is Shannon index, S is number of gene types.

[c] Coverage was calculated according to C=1-(n/N). n is number of gene types, N is the number of clones.

Table 2: Clustering of *apl*I and *apl*II from gnomic DNA and cDNA libraries. Clusters were determined based on percent similarity of more than 90%. The numbers of sequences in each cluster are given in parenthesis. Note – clusters names do not correspond to the cluster names in Figure 2 as a separate cluster analysis was done for the genomic data in Figure 2 and for the DNA and mRNA data in this table.

| | *apl* super-family I | | | | *apl* super-family II | | |
|---|---|---|---|---|---|---|---|
| Cluster | % DNA sequences | % RNA sequences | % identity to M21 | Cluster | % DNA sequences | % RNA sequences | % identity to M21 |
| **1-A** | **16.19 (17)** | **55.68 (49)** | **79-80** | **2-L** | **14.49 (10)** | **48.15 (39)** | **86-89** |
| 1-B | 6.67 (7) | 1.14 (1) | 81 | 2-M | 11.59 (8) | 1.23 (1) | 86-90 |
| 1-C | 6.67 (7) | 3.41 (3) | 79-81 | **2-N** | **0 (0)** | **9.88 (8)** | **81-82** |
| 1-D | 4.76 (5) | 5.68 (5) | 76-78 | **2-O** | **0 (0)** | **12.35 (10)** | **70-71** |
| 1-E | 4.76 (5) | 3.41 (3) | 75-76 | 2-P | 26.09 (18) | 1.23 (1) | 85-86 |
| 1-F | 4.76 (5) | 0 (0) | 69-70 | 2-Q | 8.70 (6) | 3.70 (3) | 83 |
| 1-G | 3.81 (4) | 4.55 (4) | 76-79 | 2-R | 20.29 (14) | 19.75 (16) | 82-85 |
| 1-H | 2.86 (3) | 1.14 (1) | 64-66 | 2-U* | 18.84 (13) | 3.70 (3) | |
| 1-I | 7.62 (8) | 2.27 (2) | 66-73 | | | | |
| **1-J** | **0.95 (1)** | **11.36 (10)** | **85-86** | | | | |
| 1-K | 15.24 (16) | 3.41 (3) | 80-83 | | | | |
| 1-U* | 25.71 (27) | 7.95 (7) | | | | | |
| Total | 100 (105) | 100 (88) | | Total | 100 (69) | 100 (81) | |

* 1-U and 2-U- Unaffiliated sequences - those were not included in any of the clusters due to large differences in similarities.

Table 3: Correlation coefficients between acetate concentrations and *apl* genes transcript levels.

| Monitoring well | *apl* I | | *apl* II | |
|---|---|---|---|---|
| | r | significance | r | significance |
| **D04** | 0.36 | $p > 0.05$ | 0.037 | $p > 0.05$ |
| **D08** | 0.206 | $p < 0.05$ | -0.63 | $p < 0.05$ |

Supplementary tables

Table S1: Microbial community composition based on 16S rRNA genes. Numbers in table are percentage of total clones. Affiliation of 16S rRNA genes was determined using Arb (Ludwig et al., 2004).

| | D04 | | | D08* | | |
|---|---|---|---|---|---|---|
| **Days from acetate addition** | **0** | **8** | **18** | **0** | **14** | **23** |
| ***Alphaproteobacteria*** | 5.56 | 1.11 | 1.32 | 1.23 | | 7.78 |
| ***Betaproteobacteria*** | 16.67 | 6.67 | 7.89 | 35.80 | 7.14 | 5.56 |
| ***Dechloromonas* (*Betaproteobacteria*)** | 12.50 | 2.22 | 5.26 | 13.58 | | 11.11 |
| ***Gammaproteobacteria*** | 4.17 | 4.44 | 5.26 | 11.11 | 4.76 | 7.78 |
| ***Deltaproteobacteria*** | 9.72 | 81.11 | 61.84 | 13.58 | 73.81 | 55.56 |
| ***Cytophaga-Flavobacteria-Bacteriodetes*** | 13.89 | 1.11 | 1.32 | 4.94 | | |
| ***Firmicutes*** | 12.50 | | 1.32 | 1.23 | | 1.11 |
| **Candidate divisions OD1-OP11-WS6-TM7** | 9.72 | 3.33 | 6.58 | 8.64 | 3.57 | 3.33 |
| **Candidate division OP5** | 2.78 | | 6.58 | 1.23 | 1.19 | 4.44 |
| **All other bacteria** | 12.50 | | 2.63 | 8.64 | 9.52 | 3.33 |

* Data adapted from (Mouser et al., 2009b).

Table S2: *Deltaproteobacteria* species composition based on 16S rRNA genes. Clustering to OTUs at the 97% similarity level was done using Dotur (Schloss and Handelsman, 2005), and each OTU affiliation was determined by blast analysis. Numbers in table are percentage of total *Deltaproteobacteria* and numbers in parenthesis are the number of OTUs.

| | | **D04** | | | **D08** | | **% Similarity** |
|---|---|---|---|---|---|---|---|
| **Days from acetate injection** | **0** | **8** | **18** | **0** | **14** | **23** | |
| **Number of clones in the library** | **7** | **72** | **46** | **11** | **59** | **47** | |
| ***Geobacter bemidjiensis* Bem** | | 72.22 (6)* | 13.04 (4)* | | 76.27 (6)* | 31.91 (3)* | 96% |
| ***Geobacter uraniireducens* Rf4** | | 5.56 (1) | 58.70 (2)* | 9.09 (1) | 5.08 (1) | 38.30 (2)* | 96% |
| ***Geobacter psychrophilus* strain P35** | | 11.11 (4)* | 10.87 (5)* | | 10.17 (1) | 6.38 (3)* | 96% |
| ***Geobacter* sp. M18** | | 2.78 (2)* | 6.52 (2)* | | 5.08 (1) | 10.64 (1) | 96% |
| ***Geobacter* sp. M21** | | 1.39 (1) | 6.52 (1) | | | | 96% |
| ***Geobacter* sp. G02** | | | 2.17 (1) | | | | 96% |
| **Other *deltaproteobacteria*** | 100 (4) | 6.94 (4) | 2.17 (1) | 90.9 1 (5) | 3.39 (2) | 12.77 (5) | 83-95% |

* These OTUs were affiliated with the same *Geobacter* strain according to blast analysis.

Figure 1:

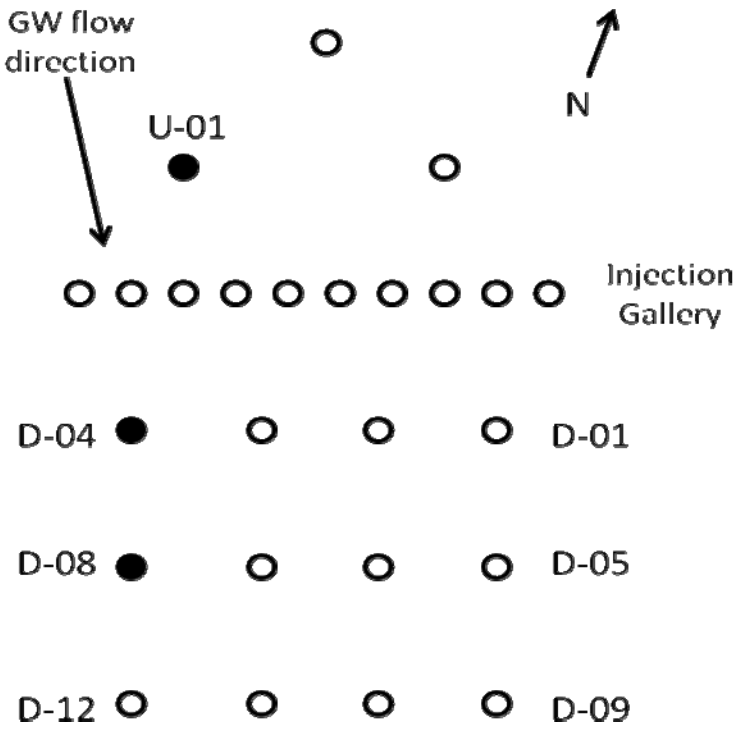

GW flow direction
U-01
N
Injection Gallery
D-04
D-01
D-08
D-05
D-12
D-09

Figure 2:

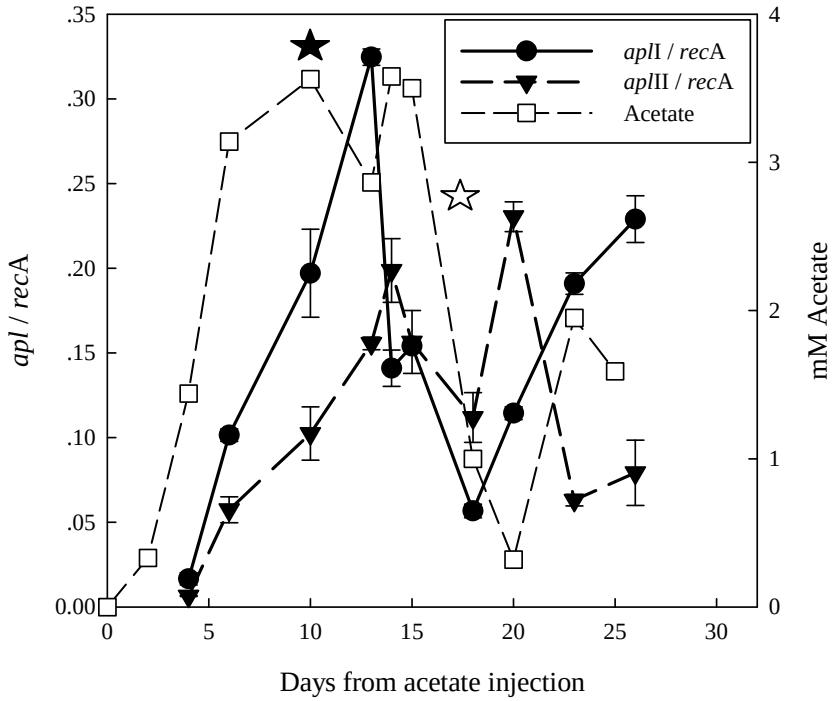

.35
.30
.25
.20
.15
.10
.05
0.00
0
5
10
15
20
25
30
4
3
2
1
0
apl / recA
mM Acetate
Days from acetate injection
aplI / recA
aplII / recA
Acetate

Figure 3:

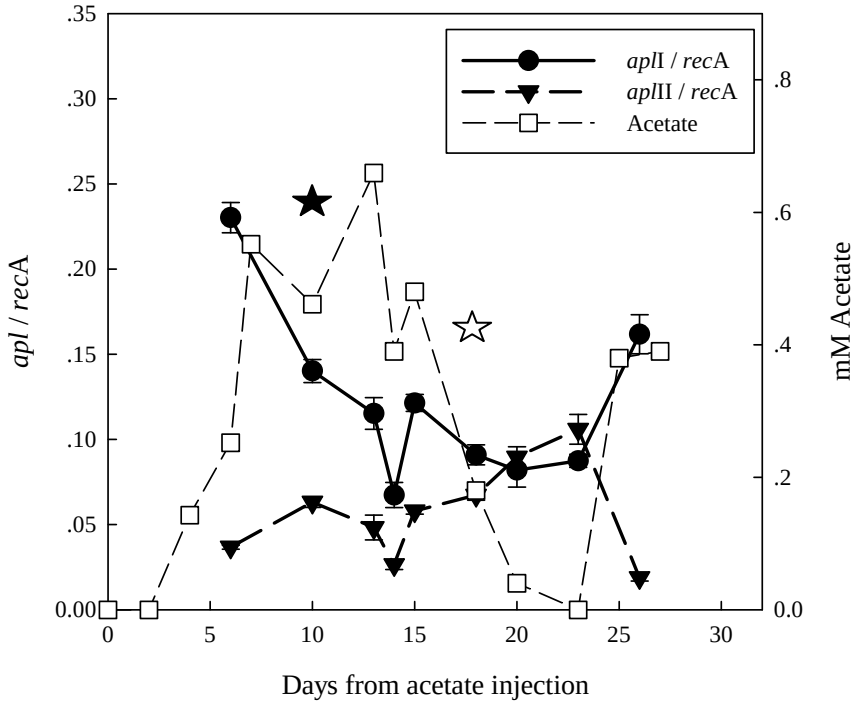

.35
.30
.25
.20
.15
.10
.05
0.00
.8
.6
.4
.2
0.0
0
5
10
15
20
25
30
apl / recA
mM Acetate
Days from acetate injection
aplI / recA
aplII / recA
Acetate

Figure 4:

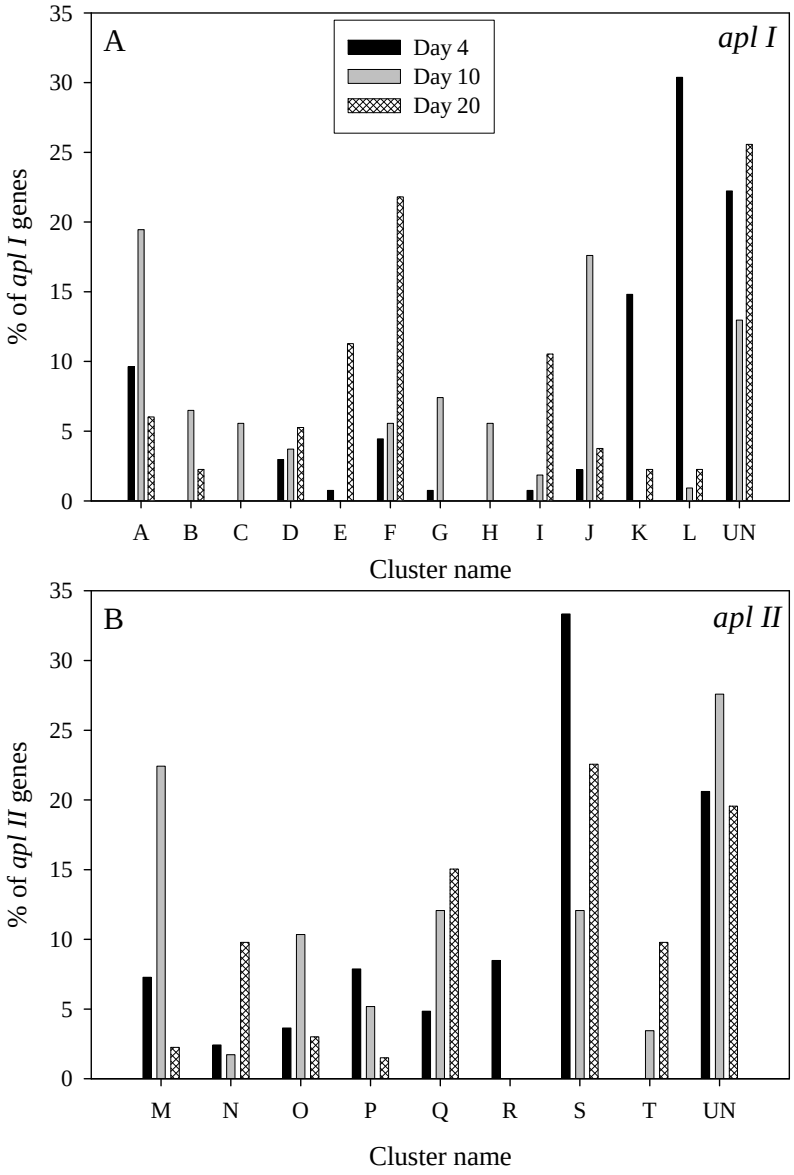
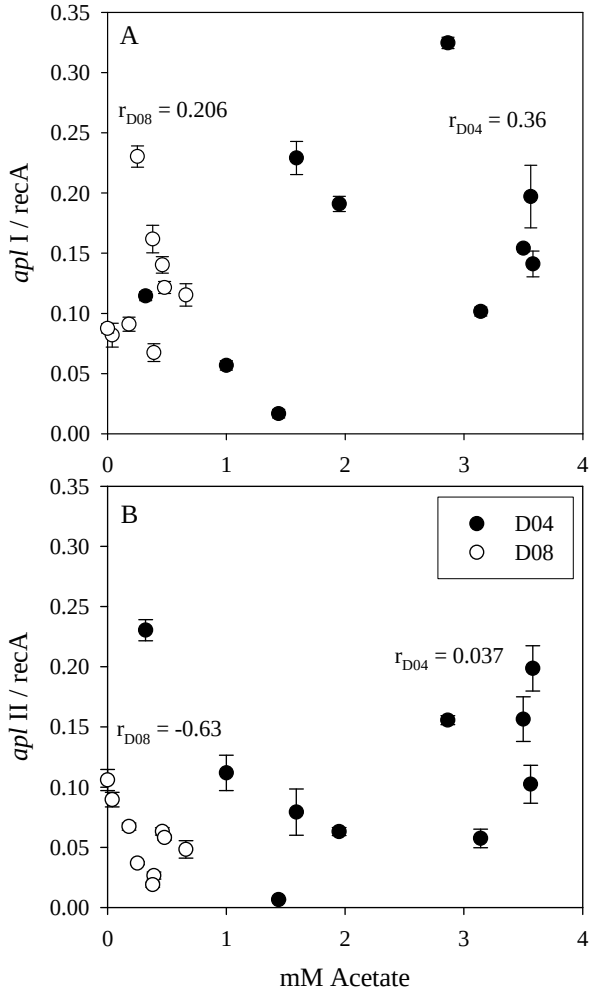

A
apl I
Day 4
Day 10
Day 20
% of apl I genes
0
5
10
15
20
25
30
35
A B C D E F G H I J K L UN
Cluster name
B
apl II
% of apl II genes
0
5
10
15
20
25
30
35
M N O P Q R S T UN
Cluster name

Figure S1:


A
B
D04
D08
$r_{D08}$ = 0.206
$r_{D04}$ = 0.36
$r_{D04}$ = 0.037
$r_{D08}$ = -0.63
*apl* I / recA
*apl* II / recA
mM Acetate
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0
1
2
3
4