Effect of carbon addition and predation on acetate-assimilating bacterial cells in groundwater

Krista Longnecker[1*], Andreia Da Costa[2,3], Maya Bhatia[4,5], and Elizabeth B. Kujawinski[1]

[1]Woods Hole Oceanographic Institution, Marine Chemistry and Geochemistry, Woods Hole, MA USA

[2]Rutgers University – Newark, NJ, USA, [3]Present address: Department of Psychiatry, St. James's Hospital, Dublin, Ireland

[4]Woods Hole Oceanographic Institution, Marine Geology and Geophysics, Woods Hole, MA 02543, USA

[5]MIT/WHOI Joint Program, Woods Hole MA 02543, USA

For submission to: *FEMS Microbiology Ecology*

Running title: Microbial community composition in groundwater

Keywords: groundwater microbiology, stable-isotope probing, microbial community structure

*Correspondence: Krista Longnecker, WHOI MS#4, Woods Hole, MA 02543, USA. Tel.: +1 508 289 2824; fax: +1 508 457 2164; e-mail: klongnecker@whoi.edu

Abstract

Groundwater microbial community dynamics are poorly understood due to the challenges associated with accessing subsurface environments. In particular, microbial interactions and their impact on the subsurface carbon cycle remain unclear. In the present project, stable isotope probing with uniformly-labeled [^{13}C]-acetate was used to identify metabolically-active and inactive bacterial populations based on their ability to assimilate acetate and/or its metabolites. Furthermore, we assessed whether substrate availability (bottom-up control) or grazing mortality (top-down control) played a greater role in shaping bacterial community composition by separately manipulating the organic carbon supply and the protozoan grazer population. A community fingerprinting technique, Terminal Restriction Fragment Length Polymorphism (T-RFLP), revealed that the bacterial community was not affected by changes in acetate availability but was significantly altered by the removal of protozoan grazers. In silico identification of terminal restriction fragments and 16S rDNA sequences from clone libraries revealed a bacterial community dominated by Proteobacteria, Firmicutes, and Bacteroidetes. Elucidation of the factors that structure the bacterial community will improve our understanding of the bacterial role in the carbon cycle of this important subterranean environment.

Introduction

Groundwater is important both as drinking water and as a source of freshwater flowing into coastal marine environments (Zektser & Everett, 2004; Mulligan & Charette, 2006). The groundwater microbial community is known to consist of prokaryotes, picoeukaryotes, and viruses (Ghiorse & Wilson, 1988; Fry *et al.*, 1997; Goldscheider *et*

al., 2006). The eukaryotic community within aquifers is dominated by small (< 5 μm) flagellates although amoebae, ciliates, and fungi can also be present (Sinclair & Ghiorse, 1987; Novarino *et al.*, 1997; Zarda *et al.*, 1998; Andrushchyshyn *et al.*, 2007; Brad *et al.*, 2008). Research on groundwater bacterial diversity has primarily focused on contaminated aquifers (Zarda *et al.*, 1998; Ludvigsen *et al.*, 1999; Haack *et al.*, 2004; Lehman *et al.*, 2004; Bowman *et al.*, 2006) or on comparisons between contaminated and pristine aquifers (Shi *et al.*, 1999; Griebler *et al.*, 2002). An endemic groundwater bacterial community has not been identified, although the groundwater microbial community is distinct from the microbial diversity observed in soil and other aquatic habitats (Griebler & Lueders, 2009). A recent study of the microbial community in a pristine aquifer revealed a spatially variable community that was dominated by Proteobacteria and that included both Euryarchaeota and Crenarchaeota (López-Archilla *et al.*, 2007). While no study has examined temporal variability of bacterial diversity in pristine aquifers, the abundance and size of bacterial cells in aquifers can be temporally variable (Velasco Ayuso *et al.*, 2009).

An important factor regarding the groundwater bacterial community is the extent to which the microorganisms are involved in biogeochemical processes. However, the presence of a microorganism is not always indicative of its role in biogeochemical cycles and the proportion of metabolically-active bacterial cells within groundwater is unknown. In marine ecosystems, metabolically-active bacterial cells can comprise between 1 and 80% of DAPI-stained cells (Karner & Fuhrman, 1997). Aquifers contain the same range of eutrophic and oligotrophic conditions as marine ecosystems and so it is likely that a similar range of metabolically-active cells will be observed in aquifers. Incubations with

labeled carbon substrates have revealed that some members of the groundwater bacterial community can incorporate ^{13}C-labeled organic compounds (toluene, benzene, or acetate) into polar lipid fatty acids with concomitant appearance of ^{13}C-labeled inorganic carbon (Pombo *et al.*, 2002; Kästner *et al.*, 2006) or methane (Pombo *et al.*, 2005). As a result, the bacterial community appears to consume and alter organic matter during transport through aquifers (DeFlaun *et al.*, 1997; Fuller *et al.*, 2000), suggesting a critical role for bacteria in the subsurface carbon cycle.

Bacterial community composition is generally determined by a balance between substrate availability (bottom-up control) and mortality due to grazing or viral lysis (top-down control). There are conflicting views as to whether top-down or bottom-up control prevails under different ecological conditions. Modeling results indicate top-down control is more important in both marine and freshwater eutrophic environments (Sanders *et al.*, 1992; Thelaus *et al.*, 2008). Conversely, model and experimental results indicate the opposite in oligotrophic marine ecosystems (Gasol *et al.*, 2002). Due to substantial variations in organic carbon and inorganic nutrient concentrations within aquifers, groundwater ecosystems can range from oligotrophic to eutrophic conditions. This range further complicates broad assessments of whether top-down or bottom-up control plays the larger role in structuring bacterial community composition within groundwater (Corno & Jürgens, 2008). While there is spatial variability in nutrient concentrations within our study area (Charette & Sholkovitz, 2006), the groundwater used for the present project is comparable to a pristine oligotrophic aquifer due to its low nutrient and organic carbon concentrations.

The interaction between grazers and the bacterial community in groundwater is not clear. Protozoan grazers are found in both pristine and contaminated aquifers, although their abundances are lower in pristine aquifers (Sinclair *et al.*, 1993; Novarino *et al.*, 1997). Grazing is known to occur in the subsurface and in flow-through columns designed to mimic in situ conditions (DeLeo & Baveye, 1997; Kinner *et al.*, 1997; Eisenmann *et al.*, 1998). In contaminated aquifers, interactions between grazers and their bacterial prey are linked to changes in the degradation of organic contaminants (Madsen *et al.*, 1991; Tso & Taghon, 2006). To our knowledge, only one study has examined the effect of grazers on bacterial community composition and its results were inconclusive because removing grazers either decreased or increased bacterial diversity depending on sampling depth within the aquifer (Nagaosa *et al.*, 2008).

Here we present the results of a study designed to examine patterns in the diversity of bacterial cells involved in acetate metabolism, and to further test factors which may be controlling microbial community structure in groundwater. There are differences in diversity between the free-living and sediment-associated bacterial community within groundwater (Lehman *et al.*, 2001; Lehman *et al.*, 2004). However, the present project focused on the free-living microbial community because of their ability to be transported to the coastal ocean at our study site (Cambareri & Eichner, 1998; Mulligan & Charette, 2006). In contrast to many coastal aquifers, temporal and spatial changes in groundwater chemistry at this site have been fairly well-characterized by other investigators (Valiela *et al.*, 1990; Charette *et al.*, 2001; Charette *et al.*, 2005; Bone *et al.*, 2006) and thus our results can be placed within the appropriate geochemical context.

Materials and Methods

Experimental setup

Groundwater was sampled from the freshwater zone of the aquifer at the Waquoit Bay National Estuarine Research Reserve (Figure 1). Polyethylene tubing lined with fluorinated ethylene propylene was inserted 2.4 m into the ground, and groundwater was pumped to the surface using a peristaltic pump at 100 ml min^{-1}. A YSI 556 MPS handheld sensor (YSI Incorporated, Yellow Springs, OH) was used to characterize the groundwater used for the experiment. The initial conditions were: temperature = 10.6 - 13.2ºC; salinity < 0.1, pH = 6.4 - 6.6; and dissolved oxygen concentration = 8.5 - 10.2 mg L^{-1}. The oligotrophic nature of our groundwater was confirmed by the low dissolved organic carbon concentrations (75 μM, see Table 2), and low nitrate (6 μM) and phosphate (< 0.1 μM) concentrations obtained from discrete, aqueous samples analyzed with a Lachat Instruments QuickChem 8000 Nutrient Analyzer.

Polycarbonate bottles (2.5 L bottles) were filled with 1.8 L of water, leaving about three cm of headspace within the bottles. At the bottom of each bottle was 500 g of autoclaved playground sand similar in texture to the *in situ* aquifer sand; the sediment covered the bottom three cm of each 25 cm high bottle. The sediment was homogenized after autoclaving and prior to adding to the bottles. Half of the bottles received whole, unfiltered, groundwater and half of the bottles received groundwater filtered through a 1 μm Polycap™ 36 AS filter. The filter unit (Whatman Inc. Florham Park, NJ) had been soaked in 10% hydrochloric acid followed by a rinse with 2 L of Milli-Q and 1 L of unfiltered groundwater prior to use. Examination of microbial cells stained with DAPI from a parallel experiment conducted with the same groundwater indicates that the protist

community is primarily comprised of flagellates between two and five μm long which would be captured by a 1 μm filter (data not shown).

The filled bottles were allowed to equilibrate in the dark for three days at room temperature (20ºC). Experimental bottles were then separated into three additional treatments: no carbon addition, a single addition of carbon, and multiple carbon additions (Table 1). Carbon was added as unlabeled acetate (Fisher, enzyme grade) or [^{13}C]-acetate (99% $^{13}CH_3$-$^{13}COOH$ from Cambridge Isotope Laboratories, Andover MA). Final acetate concentration in the bottles was 200 μM, which was added either as 200 μM on day four, or as 40 μM every other day for a total of 200 μM of acetate by day 12 of the incubation period.

The incubations lasted a total of 13 to 15 days. At the conclusion of the experiment one ml of water from each bottle was set aside for flow cytometry, and then up to one liter from each bottle was filtered through either combusted 0.2 μm Anodisc filters (Whatman International Ltd. Maidstone, England) or 0.22 μm Sterivex filters (Millipore Corp. Billerica, MA). Filters were stored at -80ºC until further processing. Water passed through the Anodisc filters (40 mL) was acidified using hydrochloric acid to pH~2 and stored at 4ºC for dissolved organic carbon (DOC) analysis.

Flow Cytometry

Water samples for flow cytometry were fixed with 0.2% w/v paraformaldehyde (final concentration), placed in the dark for at least 10 min at room temperature to harden cells, and stored at -80ºC until sample processing. A Becton-Dickinson FACSCalibur flow cytometer was used for cell enumeration. Heterotrophic cells were enumerated after staining with a 1x working stock of SYBR Green I (Invitrogen, Carlsbad, CA) for 15 min

following a protocol modified from Marie et al. (1997). Flow rates for cell abundance calculations were determined by the addition of known concentrations of 1 μm Fluoresbrite YG microspheres (Polysciences, Warrington, PA) to each sample prior to loading on the flow cytometer. The coefficient of variability between triplicate runs on the flow cytometer to determine bacterial abundances was <7%.

Dissolved organic carbon (DOC)

DOC concentrations were measured with a Shimadzu TOC-V_{CSH} total organic carbon analyzer using sucrose as a standard solution. DOC concentration was determined by subtracting the instrument blank area from the average peak area and dividing by the slope of the standard curve. Comparisons to low carbon water and deep-sea reference water provided by Prof. D. Hansell (University of Miami) were made daily. The coefficient of variability between triplicate injections was <1%.

DNA extractions

DNA was extracted from all filters using the UltraClean Soil DNA Kit (MoBio Laboratories, Inc. Carlsbad, CA), following the manufacturer's alternative protocol with the following additional modifications. The outer plastic shell of the Sterivex unit was broken. The filter was then removed and cut with flame-sterilized scissors prior to the DNA extraction. After addition of 200 μl Solution IRS, samples were vortexed in a Mini-Beadbeater™ (BioSpec Products, Inc. Bartlesville, OK) at 4800 rpm for 5 minutes. The volume of solution S3 was reduced to 1.25 ml, and the addition of Solution S4 and the subsequent centrifugation step was repeated twice.

Ultracentrifugation

Extracted DNA was mixed with cesium chloride and spun in a Beckman Coulter Optima L-80 XP Ultracentrifuge (Fullerton, CA) following protocols modified from Freitag et al. (2006). DNA was mixed with 1x TE buffer (10 mM Tris and 1 mM EDTA) to obtain a final concentration of 500 ng of DNA in 800 μL, and 4.25 mL of a cesium chloride solution (measured refractive index = 1.4143, corresponding to a calculated density of 1.859 g ml^{-1}) was added to 4.9 mL OptiSeal tubes. DNA was spun at 140,000 x *g* for 66 hours at 20ºC using a VTi 65.2 vertical rotor. Each batch of tubes included a standard to identify the proper separation of ^{12}C- and ^{13}C-labeled DNA. The standard was comprised of equal amounts of DNA extracted from *Halomonas halodurans* grown on unlabeled glucose or [U-^{13}C] glucose (99%; Cambridge Isotope Laboratories).

At the conclusion of the ultracentrifugation run, a series of 250 μL aliquots were removed from each tube using a pipette starting with the fraction at the top of the tube. Each fraction's refractive index was measured with an AR200 Digital Refractometer (Reichert, Inc. Depew, NY). The refractive index was converted to buoyant density using a linear regression calculated from tables translating refractive index to buoyant density. DNA was then precipitated following a protocol adapted from Griffiths et al. (2000) and Freitag et al. (2006). Two volumes of a 30% w/v polyethylene glycol 6000/1.6 M sodium chloride solution were added to each fraction. Fractions were incubated at 4ºC for 24 hours and then centrifuged at 20,000 x *g* for 15 minutes at 4ºC. The supernatant was discarded and 1 ml of cold 70% ethanol was added to each fraction followed by centrifugation at 20,000 x *g* for 10 minutes at 4ºC. This step was repeated twice more, discarding the supernatant each time for a total of three washings. The DNA pellet was

198 then dried and resuspended in 10 mM Tris. DNA from two adjacent fractions was then
199 combined resulting in eight fractions from each sample.
200 Two factors were used to determine which fractions contained DNA labeled with
201 ^{13}C and which contained ^{12}C-DNA. First, we examined the separation of ^{12}C-DNA and
202 ^{13}C-DNA in our standard tube with quantitative PCR (see below for details). The position
203 of the ^{12}C and ^{13}C peaks between the different ultracentrifugation runs varied <0.01 g ml^{-1}
204 within the cesium chloride gradient (n = 4, data not shown), implying that separation of
205 DNA was reproducible between individual ultracentrifugation runs. Second, we
206 examined variability in ^{12}C-DNA buoyant density. DNA was extracted from a no carbon
207 addition bottle, separated by ultracentrifugation, and processed in the same manner as the
208 ^{13}C-substrate DNA. As expected, PCR product was only obtained from the upper, ^{12}C-
209 DNA, region of the cesium chloride gradient. The density of fractions which did not
210 amplify with PCR were noted and used to constrain the ^{13}C-DNA region of the cesium
211 chloride gradient. Based on these two independent assessments, densities $\geq$ 1.7258 g ml^{-1}
212 contained ^{13}C-labeled DNA.
213 **Quantitative PCR**
214 Quantitative PCR was used to amplify a section of the 16S rRNA gene with the
215 ABsolute™ QPCR SYBR Green Mix (Thermo Fisher Scientific, Inc. Waltham, MA).
216 Reaction conditions were 1x master mix and 200 nM of each primer: 27F and 519R
217 (Operon Biotechnologies, Inc. Huntsville, AL). An enzyme activation cycle (95^{o}C, 15
218 minutes) was followed by 40 cycles of denaturation (95^{o}C, 15 seconds), annealing (55^{o}C,
219 30 seconds), and extension (72^{o}C, 30 seconds). This was followed by a melting curve
220 program (95^{o}C, 15 seconds, 55^{o}C, 15 seconds) concluding at 95^{o}C for 15 seconds.

Community fingerprinting

Terminal Restriction Fragment Length Polymorphism (T-RFLP) analysis was used to assess bacterial community composition in DNA extracted from the groundwater filling the sample bottles and DNA obtained from the ultracentrifugation fractions. DNA was amplified using the GoTaq Master Mix (Promega Corp. Madison WI). Reaction conditions were 2-10 ng of DNA, 700 nM FAM-labeled 27F and 700 nM 519R, and 1x GoTaq colorless master mix which contains 200 μM dNTPs (final concentration) and 1.5 mM $MgCl_2$ (final concentration). PCR conditions were an initial denaturation (95ºC for 5 minutes) followed by 35 cycles of denaturation (95ºC, 30 seconds), annealing (46ºC, 30 seconds), extension (72ºC, 90 seconds), and a final extension cycle of 72ºC for 5 minutes. Due to the low recovery of DNA precipitated from the cesium chloride, nested PCR was used to amplify DNA from the ultracentrifugation fractions. This involved an initial amplification with 15 cycles of the PCR program described above followed by transfer into a new PCR reaction and amplification for an additional 35 cycles of the same PCR program. At least two additional negative controls from the first PCR reaction were run in the second reaction to detect possible contamination due to the increased number of cycles.

PCR products were digested using 1 U of the restriction endonuclease *Hin*6I (Fermantas International, Inc. Burlington, Ontario) in Tango Buffer at 37ºC for 2 hours. The choice of the enzyme was based on the high number of possible restriction fragments obtained with a virtual digest at the Microbial Community Analysis web site (Shyu *et al.*, 2007). After the restriction digest, DNA was precipitated with 0.3 vol 2 M lithium chloride and 2 vol 100% ethanol, centrifuged 15 min at 3220 x *g*, washed twice with 2

vol cold 70% ethanol, and dried. After mixing with Hi-Di formamide (Applied Biosystems) and MegaBACE™ ET900-R size standard (GE Healthcare, Buckinghamshire, UK), the terminal restriction fragments (TRFs) were analyzed on an Applied Biosystems 3730XL capillary sequencer.

Chromatograms were analyzed using DAx Data Acquisition and Analysis software Version 8.0 (Van Mierlo Software Consultancy Eindhoven, the Netherlands). The ladder added to each well included 37 fragments of known size ranging from 60 to 900 base pairs, and the position of these peaks was used to determine the length of TRFs within each well. The position of TRFs between samples was aligned using Matlab code written by Liam Finlay (Finlay, L., J. Kitner, S.J. Giovannoni and E.B. Kujawinski, unpublished). The program gathers fragment lengths from all TRFs and applies a user-defined error (0.75 for the present project) to align TRFs from different wells on the same T-RFLP plate or between multiple plates. Peaks above an experimentally determined threshold value (100 fluorescent units for the present project) were then converted into a presence/absence matrix, thereby not considering differences in relative peak heights between samples.

In silico identification of TRFs

The web-based Microbial Community Analysis tool (MiCA, Department of Biological Sciences, University of Idaho, http://mica.ibest.uidaho.edu/) was used to obtain in silico identifications of the TRFs (Shyu *et al.*, 2007). The in silico analysis results in a list of 16S rRNA genes with restriction sites which would produce fragments the length of the user-provided TRFs. The web site allows the user to change the sensitivity of the analysis, and the following settings were used in the present project: one

mismatch within one base from the 5′ end of the primer and a window size of 2 base pairs. We wrote code in Matlab to place the taxonomic identity of each sequence within the Hugenholtz taxonomic outline of 16S rRNA genes available from greengenes.lbl.gov (DeSantis *et al.*, 2006). A single TRF may match different 16S rRNA genes. In order to eliminate TRFs matching 16S rRNA genes within different taxonomic groups, we considered a hit to be valid only if more than 50% of the sequences matching the TRF were from the same taxonomic group.

Cloning and sequencing of 16S rRNA genes

Two samples were chosen for further analysis using clone libraries followed by DNA sequencing: one from a bottle with whole water (p02) and a second sample from a 1 μm-filtered bottle (p01). Both of the bottles had received a single addition of carbon. Almost full-length 16S rRNA genes were amplified and analyzed using conditions previously described (Longnecker & Reysenbach, 2001). Briefly, DNA was amplified using 27F and 1492R. The resulting PCR products were cloned using the pCR4-TOPO vector (TOPO-TA, Invitrogen) according to the manufacturer's directions. DNA was extracted from the resulting colonies using a modified alkaline lysis protocol (Sambrook *et al.*, 1989). Plasmid DNA was screened for appropriate-sized inserts with vector-specific primers. The PCR-amplified inserts were then digested with 1U of the restriction endonucleases *Msp*I and *Hin*PI following the manufacturer's instructions (New England Biolabs). The resulting products were separated by gel electrophoresis on a 3.5% NuSieve GTG agarose gel (Lonza) run in TBE buffer at 4°C. The clones were separated into different phylotypes based on the RFLP banding patterns.

289 At least one of each RFLP pattern was sequenced to 2x coverage by cycle
290 sequencing using Applied Biosystems Big Dye Terminator v3.1. The coverage of each
291 clone library was assessed using C = [1 – (n/N)]*100 where n = number of RFLP groups
292 with a single clone and N = the total number of clones examined (Good, 1953).
293 Sequences were assembled using Sequencher (Gene Codes Corporation), and the primer
294 and vector were trimmed from each sequence. Chimeras identified by Bellerophon
295 (Huber *et al.*, 2004) were removed from further analysis. GenBank sequence accession
296 numbers are FJ602393-FJ602433; clones are preceded by 'p02' (whole water bottle) or
297 'p01' (1 μm-filtered bottle).
298 The phylogenetic association of each clone was determined using the small subunit
299 rRNA taxonomy and alignment pipeline (STAP, Wu *et al.*, 2008). STAP gathers related
300 16S rDNA sequences, aligns them to the target sequence, and then automatically
301 generates phylogenetic trees to determine the taxonomic assignment for each DNA
302 sequence.
303 **Indicator Species Analysis**
304 Indicator Species Analysis (ISA) was used to identify TRFs unique to different
305 treatments (Dufrene & Legendre, 1997; McCune & Grace, 2002). ISA takes into account
306 the relative abundance and the relative frequency of each TRF within a pre-defined
307 group. The output from ISA is an indicator value for each TRF, with higher indicator
308 values assigned to TRFs that are most representative of a pre-defined group. The highest
309 indicator value possible is 100, obtained when a TRF is found within all samples of only
310 one pre-defined group. The significance of the indicator value was evaluated by Monte
311 Carlo simulations during which groups were randomized 1000 times and the indicator

312 value of the randomized data was compared to the real data. A cutoff of $p<0.05$ was used
313 to determine if peaks were statistically significant indicators of the pre-defined groups.
314 The use of relative abundance information in T-RFLP is problematic due to the potential
315 for PCR bias (Suzuki & Giovannoni, 1996), therefore only the presence or absence of a
316 TRF was used in the ISA calculations which reduced the number of TRFs with
317 significant indicator values. While the use of ISA is not common in microbial ecology, it
318 proved to be useful in identifying TRFs within different redox zones of tropical soil (Pett-
319 Ridge & Firestone, 2005).

320 **Statistical analysis**

321 Nonmetric multidimensional scaling (NMS) (Kruskal, 1964; Mather, 1976) was
322 used to analyze variability in bacterial community composition. NMS is a multivariate
323 statistical technique that can be used to examine similarities, or differences, between
324 samples by reducing the comparisons between samples from a multidimensional space to
325 fewer dimensions, preferably two or three. Differences between samples were calculated
326 based on the presence or absence of TRFs. The differences were then presented
327 graphically in a multidimensional space; samples which are close together in the
328 ordination are more similar than samples located further apart. Distances between
329 samples were calculated with a distance measure using the Fathom toolbox (David Jones,
330 University of Miami – Rosenstiel,
331 http://www.rsmas.miami.edu/personal/djones/matlab/matlab.html). The distance measure
332 used was either the Bray-Curtis distance measure or a relative Bray-Curtis distance
333 measure as noted in the results section. The relative Bray-Curtis distance measure
334 normalizes the distances by the total number of TRFs in each sample; this removes TRF

number as a factor in inter-sample variability. The statistics toolbox in Matlab was used to run the NMS analyses, and additional code was written to assess the dimensionality of the data set by comparing 40 runs with real data to 50 runs with randomized data. Additional axes were considered if the addition of the axis resulted in a significant improvement over the randomized data (at $p \leq 0.05$) and the reduction in stress was greater than 0.05. The p-values were calculated as the proportion of randomized runs with stress less than or equal to the observed stress which was calculated using Kruskal's stress formula 1; stress is a measure of goodness of fit used in NMS. The proportion of variation represented by each axis was assessed with a Mantel test to calculate the coefficient of determination (r^2) between distances in the ordination space and distance in the original space.

One-way analysis of similarity (ANOSIM) was used to assess if groups visualized by NMS were statistically significant. Matlab code for ANOSIM was also from the Fathom toolbox. The Bray-Curtis distance matrix calculated for the NMS was used for ANOSIM, and the distances were converted to ranked distances prior to ANOSIM calculations. The significance of each group was tested by 1000 randomizations of the dataset, and p-values were calculated to determine the probability of no difference between groups.

Results

Bacterial abundance and DOC concentrations

In the groundwater used to establish the experiment, the abundance of grazers was less than 100 cells ml^{-1}. By the conclusion of the experiment, both bacterial abundances and dissolved organic carbon concentrations were elevated relative to the initial

groundwater (Table 2). Measurements of the abundance of grazers present at the conclusion of the experiment were not conducted. Bacterial abundances in the water were highest in 1 μm-filtered bottles even when no carbon was added to the bottles. Dissolved organic carbon concentrations were higher in bottles with whole water compared to bottles with 1 μm-filtered water, and were higher in bottles which received multiple additions of carbon as opposed to a single addition of carbon. The increase in dissolved organic carbon in the bottles which did not receive added carbon was likely due to carbon leaching off the sediment added to each bottle, although we cannot discount the contribution of DOC exuded by bacterial cells.

Bacterial community analysis based on DNA extracted from filters

Variability in bacterial community composition within the groundwater filling the bottles was assessed using NMS analysis of the T-RFLP data; we did not assess the bacterial community composition in the sand remaining at the bottom of each sampling bottle. The first step in this process was to compare similarities between samples based on the presence or absence of TRFs using the Bray-Curtis distance measure. NMS was then used to visualize differences between samples as shown in Figure 2. The NMS calculation resulted in an ordination with a final stress of 0.1734, and $r^2 = 0.75$ with the variability split between the two axes (r^2 on axis 1 = 0.38, r^2 on axis 2 = 0.26). The bacterial community in bottles which received no carbon was not distinct from the bacterial community in bottles which received one or multiple carbon additions (Figure 2A). All of the bottles received the same aliquot of autoclaved sediment at the beginning of the experiment. There were also no significant differences in bacterial community composition between bottles which received unlabeled acetate compared to bottles with

labeled acetate (Figure 2B). The presence of grazers significantly altered the composition of the bacterial community (ANOSIM, $R = 0.48$, $p = 0.0010$). The separation of the whole water and grazer-free bottles is visually evident by the clustering of whole water samples on the left side of axis one, and grazer-free samples on the right side of axis one (Figure 2C).

The in silico analysis at the MiCA web site produced over 36,000 matches to the TRFs we obtained from our samples. The results from the in silico analysis for the bottles which received a single addition of acetate compared to bottles which received multiple acetate additions were combined because of the lack of significant differences in community composition between those two treatments. Taxonomic assignments restricted to the phylum level identified between 30 and 48% of the TRFs from the present study. At the lower taxonomic levels (class, order, etc.), the number of TRFs identified dropped further because of an increase in the number of ambiguous identifications (data not shown). Based on results from the in silico analysis, the bacterial community in the sample bottles at the conclusion of the experiment was dominated by Firmicutes and Proteobacteria, with smaller contributions from Actinobacteria and Bacteroidetes (Table 3). The initial groundwater community was also dominated by Firmicutes and Proteobacteria (Table 3 – column 1), but this conclusion is based on lower numbers of TRFs due to the smaller number of samples processed from the initial groundwater.

The two clone libraries provided more detailed phylogenetic information about the bacterial diversity in a whole water bottle (clone library p02) compared to a 1 μm-filtered bottle (clone library p01). A total of 192 clones were analyzed with 96 clones from each

clone library. After screening of the clone libraries using RFLP, 22 clones were fully sequenced from p01 while 28 clones were sequenced from p02. The clone library coverage of p01 was slightly higher than the coverage of p02 (91% and 87%, respectively). After removing chimeras, more than half of the sequences were from the Proteobacteria and the Bacteroidetes (Table 4) and most of the sequences were present in both clone libraries.

There were differences in phylogenetic diversity between the clone libraries and the in silico diversity assessment from the T-RFLP data. No Firmicutes were present in the clone libraries and only a single Actinobacteria sequence was observed. We used Probe Match at the Ribosomal Database Project (RDP, Cole *et al.,* 2009) to examine if the differences could be due to the presence of mismatches between our primers and the 16S rDNA sequences archived at RDP. There were 176,281 Firmicute sequences at RDP (as of December 2008), and the 519R primer we used for the T-RFLP matches 77%, 85%, or 87% of those sequences if zero, one, or two mismatches are allowed. For the 1492R primer used for the clone libraries, only 2%, 10%, or 12% of the Firmicute sequences match if zero, one, or two mismatches are allowed. A similar situation was observed for the Actinobacteria where 65-75% of the 54,945 Actinobacteria sequences at RDP match the 519R primer, but only 3-17% of Actinobacteria sequences match the 1492R primer. Thus, we conclude that the clone libraries were likely biased against Firmicutes and Actinobacteria. This comparison between DNA sequences and primers presupposes that the sequences available from RDP are broadly representative of the DNA sequences which are added to online databases at an exponential rate (Benson *et al.*, 2009).

Bacterial community composition following ultracentrifugation

DNA extracts from seven bottles which received ^{13}C-acetate were subjected to ultracentrifugation: two bottles from each of the carbon addition treatments listed in Table 1, except for the whole water bottles with multiple carbon additions where only one replicate was subjected to ultracentrifugation. A higher number of unique TRFs was obtained within the fractions from 1 μm-filtered bottles compared to fractions from the whole water bottles (Table 3). Identifications from the in silico analysis of the ^{13}C-labeled DNA fractions were predominantly Proteobacteria and Firmicutes (Table 3). Due to our conservative choices for the in silico analysis, only a portion of the TRFs were identified and therefore a component of the bacterial diversity remains unknown.

Patterns in the community composition of active (^{13}C-labeled) bacterial cells in whole and 1-μm filtered treatments were assessed using NMS. A relative Bray-Curtis distance measure was used for this NMS analysis. This removes the bias from the observed differences in the number of TRFs between the whole and 1 μm-filtered samples. The resulting NMS ordination had a final stress of 0.1289 and $r^2 = 0.8395$ with r^2 on axis 1 = 0.6224 and r^2 on axis 2 = 0.1582 (Figure 3). Similar to the observations for the DNA extracted from the filters, there was no significant difference in the composition of ^{13}C-labeled DNA from the bottles which received a single addition of carbon compared to bottles which received multiple carbon additions (ANOSIM, p-value > 0.05). However, there was a significant difference in the composition of active cells in the bottles with whole water compared to bottles with 1 μm-filtered groundwater (ANOSIM, R = 0.2573, p = 0.0234).

The output from the Indicator Species Analysis provided three TRFs with significant indicator values: a Bradyrhizobiales, a Fibrobacterales, and one TRF which could not be identified. Significant indicator values were only obtained for TRFs primarily found within the ^{13}C-labeled DNA. The Fibrobacterales (indicator value = 80, p-value = 0.004) and the unidentified TRF (indicator value = 80, p-value = 0.016) were primarily found within the ^{13}C-DNA from 1 μm-filtered bottles. Bradyrhizobiales had high indicator values for ^{13}C-DNA in both whole (indicator value = 100, p-value = 0.0090) and 1 μm-filtered bottles (indicator value = 72, p-value = 0.036), but not in the ^{12}C-DNA fractions. The indicator values for the Bradyrhizobiales are different between the whole water and 1 μm-filtered water because the analysis was run twice: once for the whole water bottles to compare ^{12}C- and ^{13}C-DNA, and a second time for the 1 μm-filtered bottles. Based on the higher indicator value, the Bradyrhizobiales is a better indicator of the ^{13}C-labeled community in the whole water bottles compared to the 1 μm-filtered bottles.

Discussion

Microorganisms in the subsurface represent a large component of bacterial biomass (Whitman *et al.*, 1998) and understanding the factors which control bacterial community composition and metabolic activity are critical to link the microbial community with subsurface biogeochemical cycles. We used a community fingerprinting technique, T-RFLP, to assess how variability in carbon supply and grazing affected the community structure of metabolically active bacterial cells in groundwater. Our use of T-RFLP was conservative in that we sought to explain patterns in bacterial community composition. Previous modeling studies and careful analysis of in situ samples has shown this to be a

471 valid use of T-RFLP data (Hartmann & Widmer, 2008; Zhang *et al.*, 2008). We also
472 recognize that manipulating the groundwater can alter both bacterial diversity (Massana
473 *et al.*, 2001; Agis *et al.*, 2007) and organic matter (Fuhrman & Bell, 1985). Thus, as with
474 any experimental manipulation, there are compromises with respect to experimental
475 design. In the present project we recognize that separating bacterial cells from grazers via
476 filtration artificially altered the in situ community. An alternative design would be to add
477 cultured flagellates to groundwater which has its own methodological concerns.
478 The active and inactive components of the bacterial community in our samples
479 were separately identified using stable isotope probing. While stable isotope probing
480 provides a window into the diversity and community composition of the metabolically
481 active community (Neufeld *et al.*, 2007a), it is not without methodological problems
482 (Neufeld *et al.*, 2007b). We attempted to minimize these issues with a number of
483 constraints. For example, we relied on incubations without added carbon to confirm the
484 location of ^{13}C-labeled DNA. Although this experiment was shorter than other studies
485 that rely on stable isotope probing, transfer of labeled carbon within the microbial
486 community was still likely during our incubations. As a result we do not claim that all the
487 organisms identified within the "active" community are directly involved in acetate
488 metabolism. Instead, we suggest that these organisms are participants in the broader
489 carbon cycle within groundwater. Finally, the experiment relied on a single carbon source
490 and our conclusion about the diversity of active cells is therefore only based on bacterial
491 cells able to assimilate acetate and/or its metabolites. Acetate was chosen as a substrate in
492 the present project to allow us to build on previous research in which acetate assimilation
493 was used to characterize bacterial activity in a hydrocarbon-contaminated aquifer (Pombo

494 *et al.*, 2002; Pombo *et al.*, 2005). Inferences about other carbon sources will require
495 additional research; however, our conclusions about the diversity of cells able to
496 assimilate acetate are an important step towards identifying the bacterial community
497 involved in the carbon cycle within groundwater.

498 Previous studies have shown that variability in the rate of substrate addition
499 affected bacterial diversity (Carrero-Colón *et al.*, 2006), and that microorganisms able to
500 respond rapidly to changes in substrate availability can out-compete microorganisms
501 adapted to steady nutrient concentrations (Pernthaler *et al.*, 2001). In contrast, in the
502 present study, variability in substrate supply (i.e. bottom-up control) did not affect
503 bacterial community composition. Although bacterial abundance increased during the
504 incubations suggesting that the bacterial community in the bottles was able to grow under
505 our experimental conditions, there were no shifts in bacterial community composition
506 which could be linked to substrate supply.

507 Experimental manipulation of the grazer population indicated that top-down control
508 affected bacterial community composition in the present study. In aquatic ecosystems, the
509 presence of grazers alters bacterial community composition (Suzuki, 1999; Jürgens &
510 Matz, 2002; Beardsley *et al.*, 2003; Vázquez-Domínguez *et al.*, 2005), but the mechanism
511 by which grazers select their prey and shape the bacterial community remains poorly
512 understood (Boenigk & Arndt, 2002; Weisse, 2002). In a sedimentary aquifer, Nagaosa et
513 al. (2008) found removal of grazers resulted in an increase in the number of operational
514 taxonomic units identified within DNA collected 2 m below the surface, but the opposite
515 situation was observed at 10 m depth. In the present study, we observed more restriction

516 fragments in ^{13}C-DNA from grazer-free bottles compared to the number of restriction
517 fragments found in ^{13}C-DNA from the whole water bottles.
518 We consider two possible explanations for grazer control of bacterial diversity.
519 First, grazers selectively consume more active bacterial cells (González *et al.*, 1990;
520 Sherr *et al.*, 1992; González *et al.*, 1993; del Giorgio *et al.*, 1996). In our whole water
521 incubations, the grazers could have removed the active bacterial cells resulting in the
522 observed decrease in bacterial diversity. This removal of active bacterial cells was
523 observed in our data as a lower number of TRFs when grazers were present.
524 Alternatively, the increased diversity in the grazer-free incubations could reflect bacterial
525 cells able to grow only when grazing pressure was relieved. Although the present project
526 only considered the presence or absence of different bacterial phylogenetic groups, higher
527 bacterial abundances were observed in the grazer-free incubations. Therefore data on the
528 abundance of metabolically active bacterial cells will be necessary in order to assess
529 which mechanism better explains the lower diversity in whole water incubations: growth
530 in the absence of grazers or grazing of active cells.
531 The combination of the in silico analysis of the T-RFLP data and the clone libraries
532 provided valuable information on the diversity of bacterial cells within the bottles.
533 However, in silico analyses and clone libraries have different limitations when used to
534 assess patterns in bacterial diversity. The in silico analyses ignore potential differences
535 between the observed and predicted length of the restriction fragments and whether
536 multiple organisms can have restriction fragments of the same length (Kitts, 2001; Thies,
537 2007). To reduce these issues, we chose to be conservative and we focus here on the
538 phylum level rather than present more detailed taxonomic information. Meanwhile, 16S

539 rDNA clone libraries offer less coverage of microbial diversity than is possible with
540 pyrosequencing (Sogin *et al.*, 2006) or metagenomic approaches (Rusch *et al.*, 2007). We
541 observed differences between the clone libraries and the in silico analysis which may be
542 due to the use of two different reverse primers: one for the T-RFLP analysis and one for
543 generating the PCR product used for the clone libraries. While both primers were
544 designed to target Bacteria, the presence of mismatches between the primer and a target
545 sequence decreases the likelihood a sequence will be amplified (Sipos *et al.*, 2007; Bru *et*
546 *al.*, 2008). Furthermore, even when in silico analyses reveal a high level of specificity for
547 a primer set, use of the primer set can result in amplification of a significant proportion of
548 non-target sequences (Morales & Holben, 2009). The mismatches between target
549 sequences and primers we observed using RDP's Probe Match likely biased our clone
550 library against the Firmicutes and Actinobacteria identified from the in silico analysis of
551 the T-RFLP data. Amplicon size, primer mismatch, cloning issues, and PCR-induced
552 errors can all bias bacterial diversity assessments (Huber *et al.*, 2009). Therefore, due to
553 the observed differences between the two methods used in the present study, we base our
554 conclusions on the diversity of bacterial cells in our incubations on the combination of
555 both datasets rather than relying solely on the T-RFLP data or the clone libraries.
556 The dominance of Proteobacteria in both the in silico analyses and the clone
557 libraries is consistent with previous observations of Proteobacteria within aquifers
558 (López-Archilla *et al.*, 2007; Blöthe & Roden, 2009). Proteobacteria have been observed
559 in regions of sulfate reduction, denitrification, and iron-based metabolisms (López-
560 Archilla *et al.*, 2007) suggesting this phylum may play a role in these biogeochemical
561 transformations. Proteobacteria are capable of many different metabolisms, and the

562 possibility that they are involved in iron-based metabolisms is intriguing due to the
563 presence of iron in both groundwater (Charette & Sholkovitz, 2006) and sediments
564 (Charette *et al.*, 2005) sampled within 100 meters of our sampling location. Furthermore,
565 most probable number (MPN) assays have enumerated bacterial cells participating in
566 both iron oxidation and iron reduction in groundwater (Blöthe & Roden, 2009). In that
567 same study, clone libraries generated from the groundwater and the MPN cultures were
568 entirely Proteobacteria, with Betaproteobacteria representing more than 50% of the 16S
569 rDNA sequences (Blöthe & Roden, 2009). While we cannot definitely state that the
570 Proteobacteria identified within our bottle incubations were involved in iron-based
571 metabolisms, evaluating microbial involvement in the subsurface iron cycle at our study
572 site will be an interesting area of future research.
573 Two other phylogenetic groups, Saprospiraceae and Firmicutes, were dominant in
574 our assessment of bacterial diversity. However, their importance is harder to assess
575 because the Saprospiraceae were only found in the clone libraries while the Firmicutes
576 were only identified with the in silico analysis of the T-RFLP data. To our knowledge,
577 Saprospiraceae have not previously been observed in aquifers though they are found in
578 freshwater lakes and ponds (Schauer & Hahn, 2005), in activated sludge (Kong *et al.*,
579 2007; Xia *et al.*, 2008), and in an experiment examining carbon cycling in activated
580 sludge (Ginige *et al.*, 2004). Firmicutes have also been identified within aquifers
581 although they are generally more abundant in contaminated aquifers. Two studies have
582 found a small proportion of Firmicutes within pristine aquifers (Miyoshi *et al.*, 2005;
583 López-Archilla *et al.*, 2007). Within hydrocarbon-contaminated aquifers, Firmicutes can
584 account for 30-60% of DNA sequences obtained from the bacterial community

585 (Richardson *et al.*, 2002; Gu *et al.*, 2004; Macbeth *et al.*, 2004; Bowman *et al.*, 2006).
586 Due to the small number of studies on oligotrophic aquifers, we cannot make any
587 inferences as to what causes such variability in the prevalence of the Firmicutes.
588 However, Firmicutes were present in groundwater prior to the onset of the present
589 project, and their presence within the Waquoit Bay aquifer indicates that much remains to
590 be understood about bacterial diversity within non-hydrocarbon contaminated aquifers.
591 To our knowledge, the presence of Firmicutes in groundwater has not been correlated
592 with any single microbial metabolism, and therefore we cannot speculate on their role in
593 the ecosystem.

594 Indicator Species Analysis provided two bacterial groups on which to focus future
595 research efforts. A TRF identified as Fibrobacterales was diagnostic of the grazer-free
596 bacterial community able to assimilate acetate. Fibrobacterales is a small phylum within
597 the Bacteria and the few which have been described are responsible for degradation of
598 cellulose within ruminants (Krause & Russell, 1996; Montgomery *et al.*, 1998). Since
599 Fibrobacterales were found within the ^{13}C-DNA community, they do not appear to have
600 been passively transported through the subsurface but were instead actively involved in
601 the carbon cycle within the grazer-free incubations. The second TRF with a significant
602 indicator value was a Bradyrhizobiales within the Alphaproteobacteria. This group is
603 often associated with root nodules, and includes diverse metabolisms such as nitrogen-
604 fixation, and both aerobic and anaerobic respiration (Kuykendall, 2005). Future work will
605 be needed to assess how these two groups are involved in biogeochemical cycles within
606 groundwater.

In conclusion, the present project revealed that the availability of carbon did not structure community composition in a groundwater bacterial community. Instead, the community composition of acetate-assimilating bacterial cells shifted in response to the removal of grazers in our experimental manipulations. Phylogenetic identifications from our incubation experiments included microbial groups known to be present in groundwater as well as the first identification of the Saprospiraceae which indicates that we lack a complete picture of bacterial diversity in groundwater. The present study further provides an important step in determining the factors which control the bacterial community and carbon cycling in groundwater. Further studies will be needed to assess whether these patterns are unique to the Waquoit Bay aquifer or extend to other aquifers.

ACKNOWLEDGMENTS

We thank Meagan Eagle Gonneea for help installing the groundwater well, Ann Mulligan for discussions regarding groundwater fluxes in the area, Mark Dennett for access to the flow cytometer, Ann Tarrant for access to Sequencher, Katie Barott for assistance with qPCR and T-RFLP analysis, Paul Henderson for the nutrient analysis, and Mar Nieto-Cid for help with dissolved organic carbon measurements. We also thank the staff at the W. M. Keck Ecological and Evolutionary Genetics Facility at the Josephine Bay Paul Center at the Marine Biological Laboratory. We extend a special note of gratitude to the Waquoit Bay National Estuarine Research Reserve who provided space and permission to install the well. Funding was provided by NSF grant EAR-0525166 to EBK and the WHOI Summer Student Fellowship Program to ADC.

628 **References**

629 Agis M, Granda A & Dolan JR (2007) A cautionary note: Examples of possible microbial
630 community dynamics in dilution grazing experiments. *J Exp Mar Biol Ecol* **341**: 176-183.

631 Andrushchyshyn OP, Wilson KP & Williams DD (2007) Ciliate communities in shallow
632 groundwater: seasonal and spatial characteristics. *Freshw Biol* **52**: 1745-1761.

633 Beardsley C, Pernthaler J, Wosniok W & Amann R (2003) Are readily culturable bacteria
634 in coastal North Sea waters suppressed by selective grazing mortality? *Appl Environ*
635 *Microbiol* **69**: 2624-2630.

636 Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J & Sayers EW (2009) GenBank.
637 *Nucleic Acids Res* **37**: D26-31.

638 Blöthe M & Roden EE (2009) Microbial iron redox cycling in a circumneutral-pH
639 groundwater seep. *Appl Environ Microbiol* **75**: 468-473.

640 Boenigk J & Arndt H (2002) Bacterivory by heterotrophic flagellates: community
641 structure and feeding strategies. *Antonie Leeuwenhoek* **81**: 465-480.

642 Bone SE, Gonneea ME & Charette MA (2006) Geochemical cycling of arsenic in a
643 coastal aquifer. *Environ Sci Technol* **40**: 3273-3278.

644 Bowman KS, Moe WM, Rash BA, Bae H-S & Rainey FA (2006) Bacterial diversity of
645 an acidic Louisiana groundwater contaminated by dense nonaqueous-phase liquid
646 containing chloroethanes and other solvents. *FEMS Microbiol Ecol* **58**: 120-133.

Brad T, Braster M, van Breukelen BM, van Straalen NM & Roling WFM (2008) Eukaryotic diversity in an anaerobic aquifer polluted with landfill leachate. *Appl Environ Microbiol* **74**: 3959-3968.

Bru D, Martin-Laurent F & Philippot L (2008) Quantification of the detrimental effect of a single primer-template mismatch by real-time PCR using the 16S rRNA gene as an example. *Appl Environ Microbiol* **74**: 1660-1663.

Cambareri TC & Eichner EM (1998) Watershed delineation and ground water discharge to a coastal embayment. *Ground Water* **36**: 626-634.

Carrero-Colón M, Nakatsu CH & Konopka A (2006) Effect of nutrient periodicity on microbial community dynamics. *Appl Environ Microbiol* **72**: 3175-3183.

Charette MA & Sholkovitz ER (2006) Trace element cycling in a subterranean estuary: Part 2. Geochemistry of the pore water. *Geochim Cosmochim Acta* **70**: 811-826.

Charette MA, Buesseler KO & Andrews JE (2001) Utility of radium isotopes for evaluating the input and transport of groundwater-derived nitrogen to a Cape Cod estuary. *Limnol Oceanogr* **46**: 465-470.

Charette MA, Sholkovitz ER & Hansel CM (2005) Trace element cycling in a subterranean estuary: Part 1. Geochemistry of the permeable sediments. *Geochim Cosmochim Acta* **69**: 2095-2109.

Cole JR, Wang Q, Cardenas E, *et al.* (2009) The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Res* **37**: D141-D145.

667 Corno G & Jürgens K (2008) Structural and functional patterns of bacterial communities
668 in response to protist predation along an experimental productivity gradient. *Environ*
669 *Microbiol* **10**: 2857-2871.

670 DeFlaun MF, Murray CJ, Holben W, *et al.* (1997) Preliminary observations on bacterial
671 transport in a coastal plain aquifer. *FEMS Microbiol Rev* **20**: 473-487.

672 del Giorgio PA, Gasol JM, Vaqué D, Mura P, Agustí S & Duarte CM (1996)
673 Bacterioplankton community structure: Protists control net production and the proportion
674 of active bacteria in a coastal marine community. *Limnol Oceanogr* **41**: 1169-1179.

675 DeLeo PC & Baveye P (1997) Factors affecting protozoa predation of bacteria clogging
676 laboratory aquifer microcosms. *Geomicrobiol J* **14**: 127-149.

677 DeSantis TZ, Hugenholtz P, Larsen N, *et al.* (2006) Greengenes, a chimera-checked 16S
678 rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* **72**:
679 5069-5072.

680 Dufrene M & Legendre P (1997) Species assemblages and indicator species: The need
681 for a flexible asymmetrical approach. *Ecol Monogr* **67**: 345-366.

682 Eisenmann H, Harms H, Meckenstock R, Meyer EI & Zehnder AJB (1998) Grazing of a
683 *Tetrahymena* sp. on adhered bacteria in percolated columns monitored by in situ
684 hybridization with fluorescent oligonucleotide probes. *Appl Environ Microbiol* **64**: 1264-
685 1269.

686 Freitag TE, Chang L & Prosser JI (2006) Changes in the community structure and
687 activity of betaproteobacterial ammonia-oxidizing sediment bacteria along a freshwater-
688 marine gradient. *Environ Microbiol* **8**: 684-696.

689 Fry NK, Fredrickson JK, Fishbain S, Wagner M & Stahl DA (1997) Population structure
690 of microbial communities associated with two deep, anaerobic, alkaline aquifers. *Appl*
691 *Environ Microbiol* **63**: 1498-1504.

692 Fuhrman JA & Bell TM (1985) Biological considerations in the measurement of
693 dissolved free amino acids in seawater and implications for chemical and microbiological
694 studies. *Mar Ecol Prog Ser* **25**: 13-21.

695 Fuller ME, Dong H, Mailloux BJ, Onstott TC & DeFlaun MF (2000) Examining bacterial
696 transport in intact cores from Oyster, Virginia: Effect of sedimentary facies type on
697 bacterial breakthrough and retention. *Water Resour Res* **36**: 2417-2431.

698 Gasol JM, Pedrós-Alió C & Vaqué D (2002) Regulation of bacterial assemblages in
699 oligotrophic plankton systems: results from experimental and empirical approaches.
700 *Antonie Leeuwenhoek* **81**: 435-452.

701 Ghiorse WC & Wilson JT (1988) Microbial ecology of the terrestrial subsurface.
702 *Advances in Applied Microbiology* **33**: 107-172.

703 Ginige MP, Hugenholtz P, Daims H, Wagner M, Keller J & Blackall LL (2004) Use of
704 stable-isotope probing, full-cycle rRNA analysis, and fluorescence in situ hybridization-
705 microautoradiography to study a methanol-fed denitrifying microbial community. *Appl*
706 *Environ Microbiol* **70**: 588-596.

707 Goldscheider N, Hunkeler D & Rossi P (2006) Review: Microbial biocenoses in pristine
708 aquifers and an assessment of investigative methods. *Hydrogeol J* **14**: 926-941.

709 González JM, Sherr EB & Sherr BF (1990) Size-selective grazing on bacteria by natural
710 assemblages of estuarine flagellates and ciliates. *Appl Environ Microbiol* **56**: 583-589.

711 González JM, Sherr EB & Sherr BF (1993) Differential feeding by marine flagellates on
712 growing vs starving, and on motile vs non-motile, bacterial prey. *Mar Ecol Prog Ser* **102**:
713 257-267.

714 Good IJ (1953) The population frequencies of species and the estimation of population
715 parameters. *Biometrika* **40**: 237-264.

716 Griebler C & Lueders T (2009) Microbial biodiversity in groundwater ecosystems.
717 *Freshw Biol* **54** 649-677.

718 Griebler C, Mindl B, Slezak D & Geiger-Kaiser M (2002) Distribution patterns of
719 attached and suspended bacteria in pristine and contaminated shallow aquifers studied
720 with an in situ sediment exposure microcosm. *Aquat Microb Ecol* **28**: 117-129.

721 Griffiths RI, Whiteley AS, O'Donnell AG & Bailey MJ (2000) Rapid method for
722 coextraction of DNA and RNA from natural environments for analysis of ribosomal
723 DNA- and rRNA-based microbial community composition. *Appl Environ Microbiol* **66**:
724 5488-5491.

725 Gu AZ, Hedlund BP, Staley JT, Strand SE & Stensel HD (2004) Analysis and
726 comparison of the microbial community structures of two enrichment cultures capable of
727 reductively dechlorinating TCE and cis-DCE. *Environ Microbiol* **6**: 45-54.

728 Haack SK, Fogarty LR, West TG, *et al.* (2004) Spatial and temporal changes in microbial
729 community structure associated with recharge-influenced chemical gradients in a
730 contaminated aquifer. *Environ Microbiol* **6**: 438-448.

731 Hartmann M & Widmer F (2008) Reliability for detecting composition and changes of
732 microbial communities by T-RFLP genetic profiling. *FEMS Microbiol Ecol* **63**: 249-260.

733 Huber JA, Morrison HG, Huse SM, Neal PR, Sogin ML & Welch DBM (2009) Effect of
734 PCR amplicon size on assessments of clone library microbial diversity and community
735 structure. *Environ Microbiol* **11**: 1292-1302.

736 Huber T, Faulkner G & Hugenholtz P (2004) Bellerophon: a program to detect chimeric
737 sequences in multiple sequence alignments. *Bioinformatics* **20**: 2317-2319.

738 Jürgens K & Matz C (2002) Predation as a shaping force for the phenotypic and
739 genotypic composition of planktonic bacteria. *Antonie Leeuwenhoek* **81**: 413-434.

740 Karner M & Fuhrman JA (1997) Determination of active marine bacterioplankton: a
741 comparison of universal 16S rRNA probes, autoradiography, and nucleoid staining. *Appl*
742 *Environ Microbiol* **63**: 1208-1213.

743 Kästner M, Fischer A, Nijenhuis I, *et al.* (2006) Assessment of microbial in situ activity
744 in contaminated aquifers. *Eng Life Sci* **6**: 234-251.

745 Kinner NE, Harvey RW & Kazmierkiewicz-Tabaka M (1997) Effect of flagellates on
746 free-living bacterial abundance in an organically contaminated aquifer. *FEMS Microbiol*
747 *Rev* **20**: 249-259.

748 Kitts CL (2001) Terminal restriction fragment patterns: a tool for comparing microbial
749 communities and assessing community dynamics. *Curr Issues Intest Microbiol* **2**: 17-25.

750 Kong Y, Xia Y, Nielsen JL & Nielsen PH (2007) Structure and function of the microbial
751 community in a full-scale enhanced biological phosphorus removal plant. *Microbiology*
752 **153**: 4061-4073.

753 Krause DO & Russell JB (1996) How many ruminal bacteria are there? *J Dairy Sci* **79**:
754 1467-1475.

755 Kruskal JB (1964) Multidimensional scaling by optimizing goodness of fit to a nonmetric
756 hypothesis. *Psychometrika* **29**: 1-27.

757 Kuykendall LD (2005) Genus Bradyrhizobium. *The Alpha-, Beta-, Delta- and*
758 *Epsilonproteobacteria, The Proteobacteria, Part C, Bergey's Manual of Systematic*
759 *Bacteriology*, Vol. 2 (Brenner, Krieg, Staley & Garrity, eds.), pp. 438-443. Springer, New
760 York.

761 Lehman RM, O'Connell SP, Banta A, Fredrickson JK, Reysenbach A-L, Kieft TL &
762 Colwell FS (2004) Microbiological comparison of core and groundwater samples
763 collected from a fractured basalt aquifer with that of dialysis chambers incubated in situ.
764 *Geomicrobiol J* **21**: 169-182.

765 Lehman RM, Roberto FF, Earley D, *et al.* (2001) Attached and unattached bacterial
766 communities in a 120-meter corehole in an acidic, crystalline rock aquifer. *Appl Environ*
767 *Microbiol* **67**: 2095-2106.

768 Longnecker K & Reysenbach A-L (2001) Expansion of the geographic distribution of a
769 novel lineage of ε-Proteobacteria to a hydrothermal vent site on the Southern East Pacific
770 Rise. *FEMS Microbiol Ecol* **35**: 287-293.

771 López-Archilla AI, Moreira D, Velasco S & López-García P (2007) Archaeal and
772 bacterial community composition of a pristine coastal aquifer in Doñana National Park,
773 Spain. *Aquat Microb Ecol* **47**: 123-139.

774 Ludvigsen L, Albrechtsen HJ, Ringelberg DB, Ekelund F & Christensen TH (1999)
775 Distribution and composition of microbial populations in a landfill leachate contaminated
776 aquifer (Grindsted, Denmark). *Microb Ecol* **37**: 197-207.

777 Macbeth TW, Cummings DE, Spring S, Petzke LM & Sorenson KS, Jr. (2004) Molecular
778 characterization of a dechlorinating community resulting from in situ biostimulation in a
779 trichloroethene-contaminated deep, fractured basalt aquifer and comparison to a
780 derivative laboratory culture. *Appl Environ Microbiol* **70**: 7329-7341.

781 Madsen EL, Sinclair JL & Ghiorse WC (1991) In situ biodegradation: microbiological
782 patterns in a contaminated aquifer. *Science* **252**: 830-833.

783 Marie D, Partensky F, Jacquet S & Vaulot D (1997) Enumeration and cell cycle analysis
784 of natural populations of marine picoplankton by flow cytometry using the nucleic acid
785 stain SYBR Green I. *Appl Environ Microbiol* **63**: 186-193.

786 Massana R, Pedrós-Alió C, Casamayor EO & Gasol JM (2001) Changes in marine
787 bacterioplankton phylogenetic composition during incubations designed to measure
788 biogeochemically significant parameters. *Limnol Oceanogr* **46**: 1181-1188.

789 Mather PM (1976) *Computational methods of multivariate analysis in physical*
790 *geography*. J. Wiley & Sons, London.

791 McCune BM & Grace JB (2002) *Analysis of Ecological Communities*. MjM Software
792 Design, Gleneden Beach, Oregon.

793 Miyoshi T, Iwatsuki T & Naganuma T (2005) Phylogenetic characterization of 16S
794 rRNA gene clones from deep-groundwater microorganisms that pass through 0.2-
795 micrometer-pore-size filters. *Appl Environ Microbiol* **71**: 1084-1088.

796 Montgomery L, Flesher B & Stahl D (1998) Transfer of *Bacteroides succinogenes*
797 (Hungate) to *Fibrobacter* gen. nov. as *Fibrobacter succinogenes* comb. nov. and
798 description of *Fibrobacter intestinalis* sp. nov. *Int J Syst Bacteriol* **38**: 430-435.

799 Morales SE & Holben WE (2009) Empirical testing of 16S rRNA gene PCR primer pairs
800 reveals variance in target specificity and efficacy not suggested by in silico analysis. *Appl*
801 *Environ Microbiol* **75**: 2677-2683.

802 Mulligan AE & Charette MA (2006) Intercomparison of submarine groundwater
803 discharge estimates from a sandy unconfined aquifer. *J Hydrol* **327**: 411-425.

804 Nagaosa K, Maruyama T, Welikala N, *et al.* (2008) Active bacterial populations and
805 grazing impact revealed by an in situ experiment in a shallow aquifer. *Geomicrobiol J* **25**:
806 131 - 141.

807 Neufeld JD, Wagner M & Murrell JC (2007a) Who eats what, where and when? Isotope-
808 labelling experiments are coming of age. *ISME J* **1**: 103-110.

809 Neufeld JD, Dumont MG, Vohra J & Murrell JC (2007b) Methodological considerations
810 for the use of stable isotope probing in microbial ecology. *Microb Ecol* **53**: 435-442.

811 Novarino G, Warren A, Butler H, *et al.* (1997) Protistan communities in aquifers: a
812 review. *FEMS Microbiol Rev* **20**: 261-275.

813 Pernthaler A, Pernthaler J, Eilers H & Amann R (2001) Growth patterns of two marine
814 isolates: Adaptations to substrate patchiness. *Appl Environ Microbiol* **67**: 4077-4083.

815 Pett-Ridge J & Firestone MK (2005) Redox fluctuation structures microbial communities
816 in a wet tropical soil. *Appl Environ Microbiol* **71**: 6998-7007.

817 Pombo SA, Pelz O, Schroth MH & Zeyer J (2002) Field-scale ^{13}C-labeling of
818 phospholipid fatty acids (PLFA) and dissolved inorganic carbon: tracing acetate
819 assimilation and mineralization in a petroleum hydrocarbon-contaminated aquifer. *FEMS*
820 *Microbiol Ecol* **41**: 259-267.

821 Pombo SA, Kleikemper J, Schroth MH & Zeyer J (2005) Field-scale isotopic labeling of
822 phospholipid fatty acids from acetate-degrading sulfate-reducing bacteria. *FEMS*
823 *Microbiol Ecol* **51**: 197-207.

824 Richardson RE, Bhupathiraju VK, Song DL, Goulet TA & Alvarez-Cohen L (2002)
825 Phylogenetic characterization of microbial communities that reductively dechlorinate
826 TCE based upon a combination of molecular techniques. *Environ Sci Technol* **36**: 2652-
827 2662.

828 Rusch DB, Halpern AL, Sutton G, *et al.* (2007) The Sorcerer II Global Ocean Sampling
829 expedition: Northwest Atlantic through Eastern Tropical Pacific. *PLoS Biology* **5**: e77.

830 Sambrook J, Fritsch EF & Maniatus T (1989) *Molecular Cloning: A Laboratory Manual*.
831 Cold Spring Harbor Laboratory Press, Plainview, NY.

832 Sanders RW, Caron DA & Berninger U-G (1992) Relationships between bacteria and
833 heterotrophic nanoplankton in marine and fresh waters: an inter-ecosystem comparison.
834 *Mar Ecol Prog Ser* **86**: 1-14.

835 Schauer M & Hahn MW (2005) Diversity and phylogenetic affiliations of
836 morphologically conspicuous large filamentous bacteria occurring in the pelagic zones of
837 a broad spectrum of freshwater habitats. *Appl Environ Microbiol* **71**: 1931-1940.

838 Sherr BF, Sherr EB & McDaniel J (1992) Effect of protistan grazing on the frequency of
839 dividing cells in bacterioplankton assemblages. *Appl Environ Microbiol* **58**: 2381-2385.

840 Shi Y, Zwolinski MD, Schreiber ME, Bahr JM, Sewell GW & Hickey WJ (1999)
841 Molecular analysis of microbial community structures in pristine and contaminated
842 aquifers: field and laboratory microcosm experiments. *Appl Environ Microbiol* **65**: 2143-
843 2150.

844 Shyu C, Soule T, Bent S, Foster J & Forney L (2007) MiCA: A web-based tool for the
845 analysis of microbial communities based on terminal-restriction fragment length
846 polymorphisms of 16S and 18S rRNA genes. *Microb Ecol* **53**: 562-570.

847 Sinclair JL & Ghiorse WC (1987) Distribution of protozoa in subsurface sediments of a
848 pristine groundwater study site in Oklahoma. *Appl Environ Microbiol* **53**: 1157-1163.

849 Sinclair JL, Kampbell DH, Cook ML & Wilson JT (1993) Protozoa in subsurface
850 sediments from sites contaminated with aviation gasoline or jet fuel. *Appl Environ*
851 *Microbiol* **59**: 467-472.

852 Sipos R, Székely AJ, Palatinszky M, Révész S, Márialigeti K & Nikolausz M (2007)
853 Effect of primer mismatch, annealing temperature and PCR cycle number on 16S rRNA
854 gene-targetting bacterial community analysis. *FEMS Microbiol Ecol* **60**: 341-350.

855 Sogin ML, Morrison HG, Huber JA, *et al.* (2006) Microbial diversity in the deep sea and
856 the underexplored "rare biosphere". *Proc Natl Acad Sci USA* 0605127103.

857 Suzuki MT (1999) Effect of protistan bacterivory on coastal bacterioplankton diversity.
858 *Aquat Microb Ecol* **20**: 261-272.

859 Suzuki MT & Giovannoni SJ (1996) Bias caused by template annealing in the
860 amplification of mixtures of 16S rRNA genes by PCR. *Appl Environ Microbiol* **62**: 625-
861 630.

862 Thelaus J, Haecky P, Forsman M & Andersson A (2008) Predation pressure on bacteria
863 increases along aquatic productivity gradients. *Aquat Microb Ecol* **52**: 45-55.

Thies JE (2007) Soil microbial community analysis using terminal restriction fragment length polymorphisms. *Soil Sci Soc Am J* **71**: 579-591.

Tso S-F & Taghon G (2006) Protozoan grazing increases mineralization of naphthalene in marine sediment. *Microb Ecol* **51**: 460-469.

Valiela I, Costa J, Foreman K, Teal JM, Howes B & Aubrey D (1990) Transport of groundwater-borne nutrients from watersheds and their effects on coastal waters. *Biogeochemistry* **10**: 177-197.

Vázquez-Domínguez E, Casamayor EO, Català P & Lebaron P (2005) Different marine heterotrophic nanoflagellates affect differentially the composition of enriched bacterial communities. *Microb Ecol* **49**: 474-485.

Velasco Ayuso S, Guerrero MC, Montes C & López-Archilla AI (2009) Spatiotemporal distribution of microbial communities in a coastal, sandy aquifer system (Doñana, SW Spain). *Geobiology* **7**: 66-81.

Weisse T (2002) The significance of inter- and intraspecific variation in bacterivorous and herbivorous protists. *Antonie Leeuwenhoek* **81**: 327-341.

Whitman WB, Coleman DC & Wiebe WJ (1998) Prokaryotes: The unseen majority. *Proc Natl Acad Sci USA* **95**: 6578-6583.

Wu D, Hartman A, Ward N & Eisen JA (2008) An Automated Phylogenetic Tree-Based Small Subunit rRNA Taxonomy and Alignment Pipeline (STAP). *PLoS ONE* **3**: e2566.

Xia Y, Kong Y, Thomsen TR & Halkjaer Nielsen P (2008) Identification and ecophysiological characterization of epiphytic protein-hydrolyzing *Saprospiraceae* ("*Candidatus* Epiflobacter" spp.) in activated sludge. *Appl Environ Microbiol* **74**: 2229-2238.

Zarda B, Mattison G, Hess A, Hahn D, Hohener P & Zeyer J (1998) Analysis of bacterial and protozoan communities in an aquifer contaminated with monoaromatic hydrocarbons. *FEMS Microbiol Ecol* **27**: 141-152.

Zektser IS & Everett LG (2004) Groundwater resources of the world and their use. United Nations Educational, Scientific and Cultural Organization, IHP-VI, Series on Groundwater No. 6.

Zhang R, Thiyagarajan V & Qian P-Y (2008) Evaluation of terminal-restriction fragment length polymorphism analysis in contrasting marine environments. *FEMS Microbiol Ecol* **65**: 169-178.

899 Table 1. Experimental bottles were divided based on differences in carbon addition and
900 whether the groundwater was filtered to remove grazers. All of the carbon addition
901 bottles received the same total concentration of carbon either as a single addition of 200
902 μM carbon (single addition), or five additions of 40 μM carbon (multiple addition). For
903 each carbon addition treatment, two of the bottles received unlabeled acetate and the
904 remaining bottles received [^{13}C]acetate.

Carbon	**Filtered**	**# of bottles**
none	Whole	2
none	1 μm-filtered	2
Single addition	Whole	4
Single addition	1 μm-filtered	5
Multiple addition	Whole	4
Multiple addition	1 μm-filtered	5

905

906

907 Table 2. Summary of ancillary data collected from selected bottles at the conclusion of
908 the experiment. Values are means ± standard deviations. Only single measurements are
909 available for the no carbon added bottles.

	Initial	**Whole**	**1 µm-filtered**
Bacterial abundance (x 10^4 cells ml^{-1})	1.61 ± 0.02		
Multiple carbon addition		28.5 ± 13.4	73.5 ± 37.5
Single carbon addition		19.6 ± 18.9	27 ± 11.3
No carbon addition		18	57
Dissolved organic carbon (µM)	75 ± 4		
Multiple carbon addition		198 ± 61	139 ± 4
Single carbon addition		138 ± 3	107 ± 10
No carbon addition		145	94

910

911 **Table 3.** The number of terminal restriction fragments (TRFs) for the different phylogenetic groups identified using the in silico
912 analysis. Not all of the TRFs were identified, and the number of TRFs putatively identified was further reduced as described in the
913 methods. 'Unclassified sequences' are environmental sequences which have been deposited in GenBank, but have not received any
914 further taxonomic assignment. Data in the table are from DNA extracted from filters or DNA removed from the cesium chloride
915 gradient following ultracentrifugation. Identification based on DNA removed from the cesium chloride gradient is further split
916 between the whole water bottles and the 1 μm-filtered bottles, there is a subsequent division based on whether or not the TRF was
917 found within the ^{12}C- or the ^{13}C-region of the cesium chloride gradient.

918

919

	Identification of TRFs in DNA extracted from filters		Identification of TRFs in fractions following ultracentrifugation			
Phylum	**Prior to experiment**	**Final diversity within bottles**	**Whole water**		**1 µm-filtered water**	
			^{12}C-DNA	**^{13}C-DNA**	**^{12}C-DNA**	**^{13}C-DNA**
Proteobacteria	11	65	9	12	22	14
Firmicutes	12	119	9	15	25	19
Actinobacteria	3	29	2	2	5	8
Bacteroidetes	5	18	2	5	6	7
Chlorobi		1				
Fibrobacteres		3				1
Acidobacteria		6		1	1	1
OP3		1				
Aminanaerobia		1				
Planctomycetes		1				
Gemmatimonadetes		1				1
Haloanaerobiales	1	1				
OP9_JS1		1				
Unclassified sequences	3	19	1	4	4	5
# of TRFs with good taxonomic assignments	35	266	23	39	63	56
Total # of TRFs	89	548	77	109	186	166

920

921

922 Table 4. Summary of phylogenetic information from the 16S rDNA sequences from the
923 clone libraries: one from a bottle with whole water and one from a bottle with 1 μm-
924 filtered water.

Phylum	Class Order (-ales) / Family (-aceae) / Genus	% of clones	
		Whole water	**1 μm-filtered water**
Actinobacteria	Microbacteriaceae | Aureo-Microbacterium		1
Proteobacteria	Alphaproteobacteria		
	Rhizobiales | Rhizobiaceae | Rhizobium/Agrobacterium	4	
	Sphingomonadales | Novosphingobium	4	
	Bradyrhizobiales		
	Methylobacteriaceae		4
	Unclassified		7
	Caulobacterales	6	
	Consistiales | Caedibacteraceae		5
	Betaproteobacteria		
	Burkholderiales		
	Comamonadaceae	29	16
	Oxalobacteraceae	3	14
	Ralstoniaceae | Cupriavidus/Wautersia	3	
	Ralstoniaceae | Unclassified		19
	Gammaproteobacteria		
	Moraxellaceae | Acinetobacter	6	
	Epsilonproteobacteria		
	Helicobacteraceae | Wolinella	4	
Bacteroidetes	Saprospiraceae | Unclassified	36	34
	Flavobacteriales | Unclassified	3	

925

926

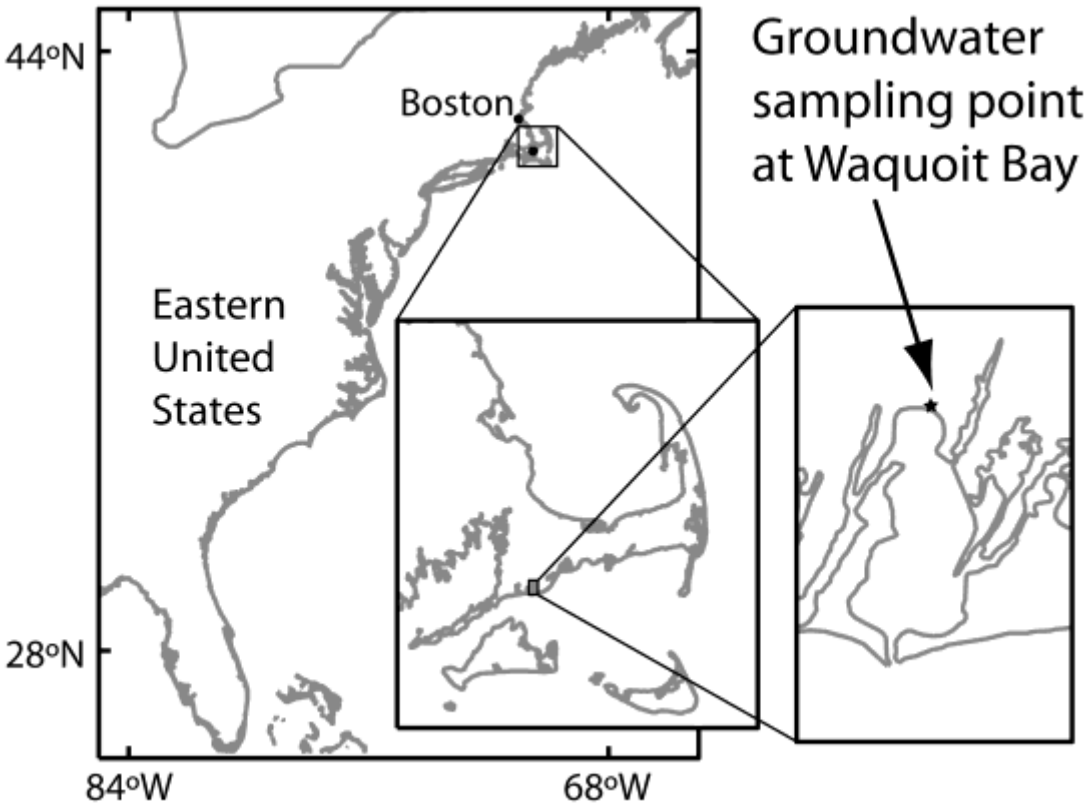
Figure legends

Figure 1. Map of groundwater sampling region. Groundwater was pumped from a well installed at the Waquoit Bay National Estuarine Research Reserve. The well was within the freshwater region of the aquifer above the zone where freshwater and saltwater mix within the subterranean estuary (Charette *et al.*, 2005).

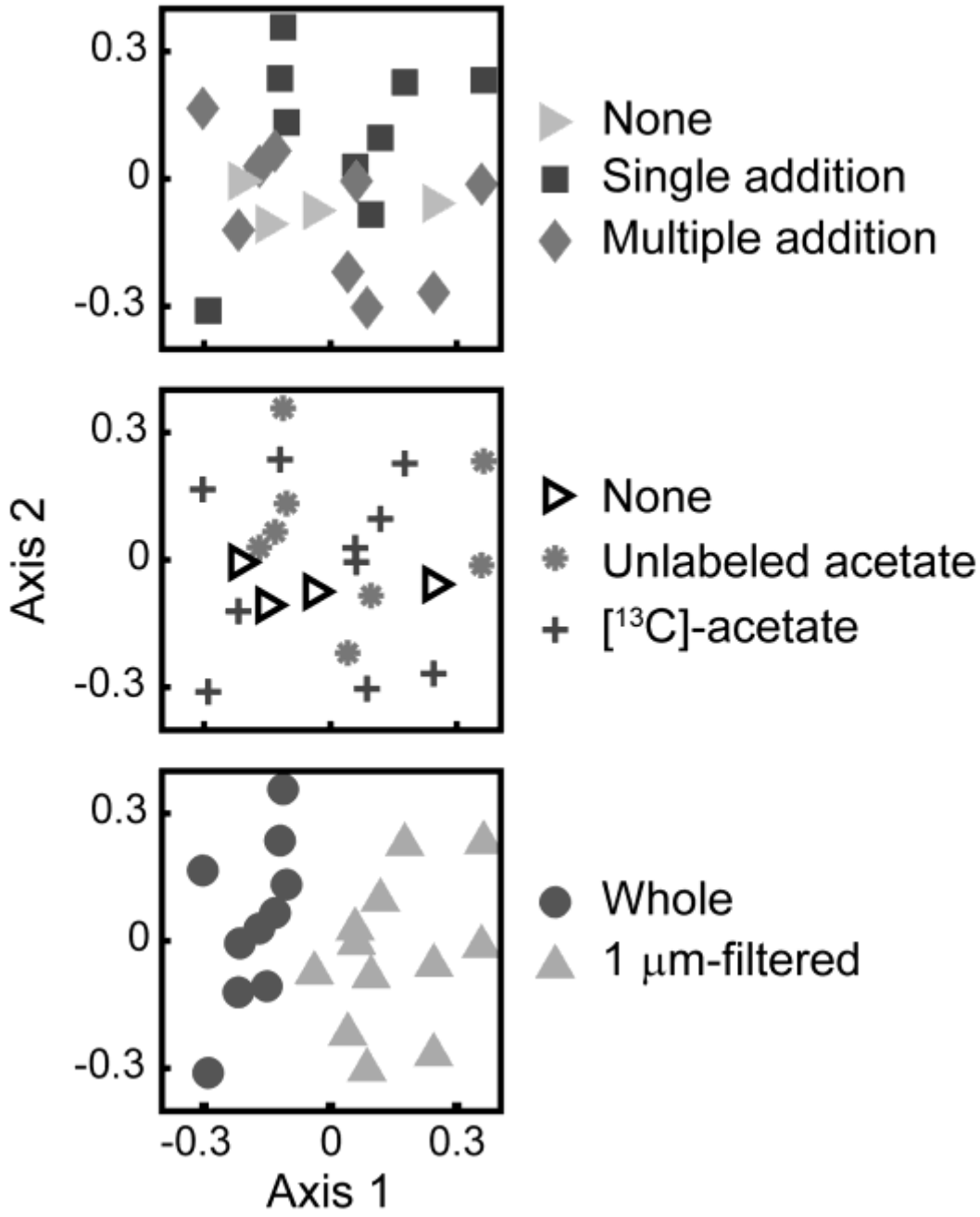
Figure 2. Non-metric multidimensional scaling (NMS) analysis showing the differences in bacterial community composition between samples based on T-RFLP analysis of DNA extracted from filters prior to ultracentrifugation. Each point within the figure is a single sample. Points that are closer together are more similar, while points that are located further apart display greater differences. All three panels of the figure contain the same NMS results with the samples coded differently to highlight (A) the rate of carbon addition, (B) the type of carbon added, and (C) filtration to remove grazers.

Figure 3. Patterns in the community composition of ^{13}C-labeled DNA ('active' bacterial cells based on their ability to assimilate acetate and/or its metabolites) in the incubations based on the NMS analysis. (A) and (B) contain the same NMS results with (A) coded to high the effect of carbon addition and (B) revealing the differences between the whole and 1 μm-filtered bottles.

Longnecker et al.
Figure 1



Longnecker et al.
Figure 2



Longnecker et al.
Figure 3