Spatial Heterogeneity of Crenarchaeal Assemblages Within Mesophilic Soil Ecosystems as Revealed by PCR-Single-Stranded Conformation Polymorphism Profiling

Marek K. Sliwinski

University of Northern Iowa

Copyright ©2004 American Society for Microbiology (ASM). The copyright holder has granted permission for posting.

Follow this and additional works at: https://scholarworks.uni.edu/bio_facpub

Part of the Biology Commons

Recommended Citation

This Article is brought to you for free and open access by the Faculty Work at UNI ScholarWorks. It has been accepted for inclusion in Faculty Publications by an authorized administrator of UNI ScholarWorks. For more information, please contact scholarworks@uni.edu.
Spatial Heterogeneity of Crenarchaeal Assemblages within Mesophilic Soil Ecosystems as Revealed by PCR–Single-Stranded Conformation Polymorphism Profiling

Marek K. Sliwinski1* and Robert M. Goodman1,2

Department of Plant Pathology1 and Gaylord Nelson Institute for Environmental Studies,2 University of Wisconsin—Madison, Madison, Wisconsin 53706

Received 10 July 2003/Accepted 2 December 2003

Microbial ecologists have discovered novel rRNA genes (rDNA) in mesophilic soil habitats worldwide, including sequences that affiliate phylogenetically within the division Crenarchaeota (domain Archaea). To characterize the spatial distribution of crenarchaeal assemblages in mesophilic soil habitats, we profiled amplified crenarchaeal 16S rDNA sequences from diverse soil ecosystems by using PCR–single-stranded-conformation polymorphism (PCR-SSCP) analysis. PCR-SSCP profiles provide a measure of relative microbial diversity in terms of richness (number of different phylotypes as estimated from the number of unique PCR-SSCP peaks) and evenness (abundance of each phylotype as estimated from the relative area under a peak). Crenarchaeal assemblages inhabiting prairie, forest, turf, and agricultural soils were characterized at six sampling locations in southern and central Wisconsin. Phytotype richness was found to be more stable than evenness among triplicate samples collected within 30 cm at each sampling location. Transformation of the PCR-SSCP data by principal-component analysis, followed by statistical testing (analysis of variance [P < 0.0001] and least-significant-difference analysis [α = 0.5]), supported the conclusion that each location exhibited a unique profile. To further characterize the spatial distribution of crenarchaeal assemblages at one location, additional soil samples (a total of 30) were collected from agricultural field plots at the Hancock Agricultural Research Station. PCR-SSCP revealed a patchy spatial distribution of crenarchaeal assemblages within and between these plots. This mosaic of crenarchaeal assemblages was characterized by differences in phylotype evenness that could not be correlated with horizontal distance (15 to 30 m) or with depth (0 to 20 cm below the surface). Crenarchaeal 16S rDNA clone libraries were produced and screened for unique SSCP peaks. Clones representing the dominant phylotypes at each location were identified, sequenced, and found to group phylogenetically with sequences in crenarchaeal clade C1b.

The combination of microbial ecology and molecular phylogeny has revealed that the majority of microbes in natural ecosystems have not been cultivated in the laboratory (19, 32). Of the possible millions of microbial species inhabiting Earth, fewer than 5,000 have been characterized (52), exemplifying the tremendous gap in our understanding of microbial life on this planet. Reports of novel 16S rRNA genes (rDNA) PCR amplified and cloned from the environment have become commonplace. But characterization of the corresponding organisms in their native habitats is still challenging, thus limiting our understanding of the niche most microbes occupy in the global ecosystem.

The growing collection of environmental 16S rDNA sequences includes members of the archaean division Crenarchaeota, which have been found in marine, aquatic, and terrestrial environments (for a review, see reference 10). Quantitative filter hybridization (5) and real-time PCR quantification (31) of 16S rRNA sequences from mesophilic soils have estimated that crenarchaeotes represent approximately 1% and 0.5 to 3.0%, respectively, of the total soil microbial community. Crenarchaeal 16S rDNA sequences from one monophyletic clade, C1b (9, 10), have been recovered from a variety of habitats around the globe, including earthworm casts (15), termite guts (14), a geothermal vent (57), a river estuary (8), freshwater sediments (17), an acid mine (50), and a variety of soils (3–6, 15, 25, 30, 31, 33, 36, 37, 39, 45, 53, 55). In Wisconsin, C1b rDNA sequences have been recovered from soil at the University of Wisconsin West Madison Agricultural Research Station (3) and crenarchaeotes have been found to colonize plant roots grown in soil from this same location (45).

PCR-based 16S rDNA profiling provides a rapid characterization of the most abundant phylotypes present within an environmental sample without the prerequisite of first culturing the microbe in the laboratory. While the comparison of two amplified 16S rDNA profiles provides an estimate of how alike or dissimilar two microbial assemblages may be, the results must be interpreted cautiously. Any PCR-based technique, including the comparison of 16S rDNA clone libraries, is subject to PCR biases (35, 48, 49) and may not accurately reflect cell counts of microbes in the environment owing to 16S rDNA copy number variability (12). In addition, 16S rDNA comprising less than approximately 0.1% of the available PCR template will be outcompeted during the reaction, thus eliminating detection of these rare phylotypes by PCR profiling (11, 26, 29, 56).

A variety of molecular profiling methods suitable for microbial ecology have been developed, including denaturant or

* Corresponding author. Present address: Department of Botany, University of Wisconsin—Madison, 430 Lincoln Dr., Madison, WI 53706. Phone: (608) 265-7929. Fax: (608) 262-7509. E-mail: mksliwin@wisc.edu.
temperature gradient gel electrophoresis (28, 29), ribosomal intergenic spacer analysis (4), terminal restriction fragment length polymorphism (24), and single-stranded-conformational polymorphism (SSCP) (23). The sensitivity of PCR-SSCP for detecting novel sequences in a population of small DNA fragments (<200 nucleotides [nt]) approaches that of sequencing (16, 44), and the technique has been successfully adapted for molecular ecology (47) to profile microbial 16S rDNAs (23, 51). PCR-SSCP is based on the separation of PCR fragments of the same length that differ in sequence composition. Each novel sequence has the potential to form a distinct single-stranded secondary structure that will migrate a unique distance through a polyacrylamide gel. The resulting peaks (each representing a phylotype) may thus represent one or possibly a few comigrating sequences. PCR-SSCP has been a useful tool for monitoring microbes within a variety of environments (22, 34, 40–43), including monitoring of community dynamics in bioreactors (2, 58).

The objective of this study was to analyze the spatial distribution of crenarchaeal assemblages in soil. We used PCR-SSCP to identify crenarchaeal phylotypes and to quantify their relative abundance in replicated soil samples collected at six locations in southern and central Wisconsin. We also sampled replicated plots in an agricultural field (within 350 m²) to determine the underlying spatial variability of crenarchaeal assemblages in relation to horizontal distance and vertical soil depth. Finally, clone libraries of 16S rDNA (~1,320 nt) were screened for sequences representing the dominant PCR-SSCP phylotypes amplified from the six sampling locations.

MATERIALS AND METHODS

Sample collection. Soil samples were collected from various locations in Wisconsin (Fig. 1), including a previously sampled site at the University of Wisconsin West Madison Agricultural Research Station (3, 45). Soil collection site descriptions and coordinates in meters (Universal Transverse Mercator, zone 16) are listed in Table 1. Three 2- to 5-g (fresh weight) soil samples were collected at each location with autoclave-sterilized implements. These three samples were collected along a transect, 15 cm apart at a 0- to 3-cm depth.

Additional samples from the University of Wisconsin Hancock Agricultural Research Station were collected in the same manner on 10 July 2001 from three plots, designated 12-4, 12-3, and 12-2 (see Fig. 4), that were part of a randomized, replicated field trial. Each plot was 3.7 by 7.6 m. Plot layout and crop history are described elsewhere (46). At the time of soil collection for this study, the plots had been fallow for 8 months. For depth comparisons, a 20-cm-deep soil core was collected from each plot and a 2-g (fresh weight) soil sample taken at 2, 6, 10, 15, and 20 cm below the ground was collected from the center of each core with autoclave-sterilized implements. All soil samples were immediately frozen in liquid nitrogen and stored in the laboratory at −80°C until processing.

DNA extraction. For nucleic acid extraction, 0.1 to 0.5 g of each soil sample was transferred to a 2-ml screw-cap centrifuge tube containing 0.6 mg of 0.1-mm-diameter zirconium-silica beads, 0.6 mg of 1.0-mm-diameter zirconium-silica beads, and one 2.5-mm diameter silica bead (BioSpec Products Inc., Bartlesville, Okla.). To each tube, 1 ml of TEND (50 mM Tris, 50 mM EDTA, 100 mM NaCl, 1 M guanidine thiocyanate, 20 mM EDTA, 0.00002% phenol red sodium salt), in a 96-well assay block (2 ml per well). Samples were transferred, with multichannel pipettes, to a Whatman UniFilter GF/B microplate 800 (Whatman Inc., Ann Arbor, Mich.). DNA was bound to the GF/B filter at the bottom of each well by pulling the liquid sample through the filter with a microplate vacuum filtration apparatus. The GF/B filters were rinsed once with 200 µl of MEGP and twice with 200 µl of MENE rinse solution (10 mM MOPS [pH 7.4], 1 mM EDTA, 100 mM NaCl, 55% ethanol). GF/B filters were dried under a vacuum for 5 min. DNA was eluted twice with 30 µl of 10 mM Tris, pH 8.0, that had been heated to 65°C. The two DNA eluates were pooled.

(ii) Microspin column chromatography. Sepharose CL-2B resin was rinsed five times with 1 volume of sterile, MilliQ (SMQ)-purified water to remove the 20% (vol/vol) ethanol resin storage solution. After rinsing, the resin was resuspended in SMQ-purified water to produce a 60% (vol/vol) slurry. Spin columns were packed as follows. Eight hundred microliters of slurry was transferred to each well of a UniFilter GF/B-800 microplate and centrifuged at 30 × g for 1 min. Two hundred microliters of additional slurry per column was packed twice by subsequent centrifugations at 30 × g for 1 min. Column eluates were discarded after each packing centrifugation. After packing, 50 µl of DNA solution was transferred to the top of each column and centrifuged at 30 × g for 1 min to collect the first eluate. Additional eluates were collected after addition of 50 µl of SMQ-purified water to the top of each column, followed by centrifugation at 30 × g for 1 min. Most of the DNA was present in the second, third, and fourth eluates, which were pooled for a total volume of 150 µl. Silica binding, followed by CL-2B spin column purification, was performed twice for each soil sample.
PCR-SSCP. Each 20-μl PCR mixture consisted of 1× Promega buffer A, 2 mM MgCl₂, 50 μM deoxyribonucleoside triphosphate (Promega, Madison, Wis.), 10 pmol of each primer, 10 μg of nonacetylated bovine serum albumin (Ambion Inc., Austin, Tex.), 1.0 U of Taq polymerase (Promega), and 1 μl of purified soil DNA. Reactions were performed in a PTC-Hotplate 8 Robocycler Gradient 96 Thermocycler (Stratagene) with a 1 cycle of 94°C for 1 min and 35 cycles of 94°C for 30s, 55°C for 1.5 min, and 72°C for 1.5 min, followed by 1 cycle of 72°C for 5 min. The crenarchaeal biased primers used for PCR-SSCP were 133FN6F (5’-GCGGAT CGCGGGCGCTCAAGACGATCTCAAC-3’; modified from reference 45) and 248R5P (5’-GCGTTACCTACCAAGC-3’; this study). The 133FN6F primer includes a NotI restriction site and a fluorescent label (6-FAM) attached to the 5’ end. 248R5P is phosphorylated at the 5’ end. PCR products were checked by electrophoresis of 1/10 of the reaction mixture on a 2% agarose gel (38), followed by ethidium bromide staining. PCR negative controls containing SMQ-purified water in place of soil DNA routinely produced no detectable band.

SSCP nondenaturing polyacrylamide gels (1× MDE; BMA, Rockland, Maine) were cast in accordance with the manufacturer’s instructions. Glycerol was not added to the MDE gels as it decreased sensitivity by producing broader peaks (data not shown). Sample preparation for SSCP included a fivefold dilution of the PCR with SMQ-purified water, followed by desalting of a 10-μl aliquot of each reaction mixture with Sephadex G50 microspin columns (1) packed in a UniFilter GF/B 800 microplate. After the removal of excess salts, 1 μl of each PCR product was mixed with 1 μl of SSCP loading buffer (60% deionized formamide, 25% ROX, GeneFlo 625 standard [ChimerX, Milwaukee, Wis.], 7.5% 100 mM NaOH, 7.5% blue dextran EDTA [blue dextran at 50 mg/ml, EDTA at 50 mM]). Samples were denatured at 94°C for 2 min and then immediately transferred onto ice. The reproducibility of gel profiles (replicates of the same amplification products) was high. Initial experiments used lambda exonuclease (New England Biolabs, Beverly, Mass.) to produce single-stranded products by selective digestion of the 5’ phosphorylated strand produced with primer 248R5P. This digestion step was found to be unnecessary and was discontinued for later samples. Immediate cooling of the denatured PCR products produced enough single-stranded molecules for sufficient detection. Also, the single-stranded molecules detected in this study could be identified without exonuclease digestion since they all migrated more slowly than any double-stranded or heteroduplex molecules. A 0.5-μl aliquot of each sample was spotted onto a membrane comb. The MDE gel was loaded onto an ABI PRISM 377 DNA Sequencer connected to an external cooling water bath. The SSCP gel temperature was maintained at 23°C during the entire 5-h run. The GS RUN 60W CHILLER module was selected in the ABI PRISM 377 collection software to maintain a constant wattage. The peak size (relative migration distance) and peak area (area under the curve) were determined with GeneScan software version 3.1 (Applied Biosystems, Foster City, Calif.). Peak height cutoffs ranged from 5 to 50 relative fluorescence units, depending on the signal intensity. Samples were aligned on the basis of the internal size standard present in each lane.

Statistical analyses. Data files containing peak size and peak area were exported from GeneScan software. Files were adjusted manually when GeneScan misaligned or incorrectly identified peaks common to multiple electropherograms. Statistical analysis was conducted in SAS (release 8.2; SAS Institute, Cary, N.C.). Principal-component analysis (PCA) and analysis of variance were performed by the PRINCOMP and GLM procedures, respectively.

Clone library construction and screening. PCR for clone libraries was conducted as described above, except that primers 133F (45) and 1492R (5’-GGY TACCTTGTTGACCACTT-3’) (27) were used. PCRs for each sampling location (a total of six) were purified with a QIAquick PCR purification kit (QIAGEN, Valencia, Calif.) and T/A cloned with a pGEM-T Vector System I kit (Promega) in accordance with the manufacturers’ instructions. Individual clones were screened via colony PCR with 133FN6F and 248R5P. Reaction conditions were the same as those stated above for environmental templates. PCR products were prepared and separated on MDE gels as described above.

Sequencing and phylogenetic analysis. Sequencing was accomplished with the ABI Prism BigDye Terminator system. Each clone was sequenced with forward and reverse primers (T7 and SP8) spanning the multiple cloning site to achieve at least twofold coverage over the entire cloned fragment. Areas of ambiguity were resequenced with internal primers 530F (21) and 755R (45) as necessary until a consensus sequence was obtained. Sequences were checked for chimeras with the CHECK_CHEMERA program of the Ribosomal Database Project (7) and by partial treecing as previously described (20). Two putative chimeric sequences were detected and removed from further analysis.

Phylogenetic analysis. For phylogenetic tree construction, only sequences of >1,300 nt were aligned. The multiple-sequence alignment was refined manually to remove regions of ambiguity, and any columns containing unresolved gaps or
missing sequence were removed from the data set used for tree construction. The tree was inferred from 1,228 positions of the aligned sequences by the maximum-likelihood method, FASTDNAML, within the ARB software environment (http://www.arb-home.de/). Phylogenetic trees were also inferred by using evolutionary distance, ARB neighbor joining (with Olsen correction), and maximum parsimony (DNAPARS from the PHYLIP software package, v. 3.55) (13). Results from these methods were in close agreement. The outgroups used in generating the tree have been omitted from Fig. 8 for clarity. These outgroups derived from these methods were in close agreement. The outgroups used in parsimony (DNAPARS from the PHYLIP software package, v.3.55) (13).

Revolutionary distance, ARB neighbor joining (with Olsen correction), and maximum likelihood method, FASTDNAML, within the ARB software environment (http://www.arb-home.de/). Phylogenetic trees were also inferred by using evolutionary distance, ARB neighbor joining (with Olsen correction), and maximum parsimony (DNAPARS from the PHYLIP software package, v. 3.55) (13). Results from these methods were in close agreement. The outgroups used in generating the tree have been omitted from Fig. 8 for clarity. These outgroups included the C1a sequences SAGMA-10 (accession no. AB050240), SAGMA-11 (accession no. AB050241), SAGMA-D (accession no. AB050208), SAGMA-X, (accession no. AB050229), APA3-0 cm (accession no. AF119136), CRAB-11 cm (accession no. AF119127), and Crenarchaeum symbiosum (accession no. U51469); C2 sequence pGrfC26 (accession no. U59986); cultured Crenarchaeota sequences Sulfolobus shibatae (accession no. M32504) and Sulfolobus solfataricus (accession no. D26490); and Euryarchaeota sequences Archaeoglobus fulgidus (accession no. X05567), Halobacterium halobium (accession no. M11583), Methanococcos jannaschii (accession no. U67473), and Methanopyrus kandleri (accession no. M59932).

Nucleotide sequence accession numbers. Sequences have been deposited in GenBank under accession numbers AY522861 to AY522902.

RESULTS

PCR-SSCP profiles of six disparate sampling locations. Following purification by silica binding and Sepharose CL-2B spin column chromatography, greater than 90% of the extracted DNA samples yielded PCR products detectable by agarose gel electrophoresis. A second round of purification was necessary to obtain a sufficiently pure template for the remaining samples. To maintain uniformity, all samples were purified by two rounds of silica binding, followed by Sepharose CL-2B spin column chromatography. After two rounds of purification, every soil sample evaluated in this study produced a 133FN6F-248R5P PCR product of the expected size.

To assess the distribution of crenarchaeal phylotypes in disparate soil ecosystems, four sampling locations (Table 1) in Madison and two others in south central Wisconsin were chosen (Fig. 1). The variability of phylotype richness (number of unique PCR-SSCP peaks) and evenness (relative peak area within an electropherogram) for each soil were estimated from three independent soil samples collected 15 cm apart along a transect (all within 0.3 m). The PCR-SSCP primers used in this study, 133FN6F and 248R5P, span a polymorphic region (positions 133 to 248 of the Escherichia coli numbering) near the 5′ end of the crenarchaeal 16S rDNA. This pair of primers is not specific for all known mesophilic soil crenarchaeal sequences but does match the crenarchaeal sequences previously recovered at the West Madison Agricultural Research Station (3, 45) without mismatches. Each of the six sampling locations produced a distinct PCR-SSCP profile (Fig. 2).

The area under each unique SSCP peak in the data set was calculated from data tables generated by GeneScan software. This information was transformed to a relative peak area for each electropherogram, making possible the direct comparison of independent samples amplified in different PCRs. The transformed data for each electropherogram were organized into a matrix for PCA. The matrix columns represented each unique phylotype present in the data set, and matrix rows designated each soil sample. The numerical variables within the matrix were the calculated relative abundance values for each PCR-SSCP peak. The first principal component (prin1) explained 25.6% of the variability present in the data set, and prin2 explained a further 20.2%. Analysis of variance indicated

FIG. 2. PCR-SSCP electropherograms of independent soil samples collected at six locations. Designations: turf, West Madison Agricultural Research Station (W); forest, Picnic Point (B); prairie, Curtis prairie (C); agricultural field, Arlington Agricultural Research Station (A); agricultural field, Hancock Agricultural Research Station (H). Three independent soil samples collected 15 cm apart at each location are numbered 1, 2, and 3. For comparisons between samples, individual electropherograms were aligned on the basis of a common internal lane standard. The y axis was adjusted for each electropherogram and represents relative fluorescence intensity, the x axis represents the relative migration distance of the PCR-SSCP DNA fragment. Representative clones for the dominant phylotypes, labeled 1 through 14 across all samples, are MWS38, MCS34, MCS7, MBS11, MCS48, MAS3, MHS15, MAS18, MWS12, MGS4, MGS313, MCS4, MHS2, and MHS12, respectively.
significant differences ($P < 0.0001$) for both prin1 and prin2. To determine which locations were distinct, least-significant-difference (LSD) values were calculated. LSD t-test groupings ($\alpha = 0.5$) supported the hypothesis that each sampling location was significantly different from every other location (Fig. 3).

**Spatial distribution of crenarchaeotes within one sampling location.** Replicated fallow control plots were sampled in a field experiment (46) at the Hancock Agricultural Research Station (Fig. 4). Soil samples A to I (Fig. 5) were collected at plot 12-4 and were within 30 m$^2$ of each other. Two other plots were sampled in this field, 12-3 (soil samples J to L) and 12-2 (soil samples M to O). All samples were collected within 350 m$^2$. In general, evenness was found to vary more than richness in the samples from this agricultural field, and this variation was not correlated with distance. Samples A to C and D to F produced similar electropherograms within short distances, while samples G to I, J to L, and M to O exhibited more variability within short distances.

To explore the influence of soil depth on crenarchaeal assemblages, we collected independent soil cores at plots 12-4, 12-3, and 12-2 and sampled 2, 6, 10, 15, and 20 cm below the surface (Fig. 6). Soil cores 1, 2, and 3 were collected from plots 12-4, 12-3, and 12-2, respectively (Fig. 5). There was marked variability between soil cores. Soil core 1 contained a greater percentage of peak 3 (Fig. 2 numbering) than did soil core 2 or 3. In general, there was no reproducible effect of depth on crenarchaeal assemblage. The set of electropherograms for each core appeared more similar to each other than to electropherograms from different cores taken at similar depths.

This data set suggests that variation in crenarchaeal assemblages by depth is less pronounced than the variation observed over horizontal distance in this agricultural soil.

**Construction and screening of clone libraries.** Primers 133F and 1492R were used to amplify nearly full-length 16S rDNA sequences. A sample of 222 clones from all libraries was screened for unique SSCP peaks by colony PCR with SSCP primers 133FN6F and 284R5P. PCR products were separated on MDE gels, and 44 clones including representatives from the 14 dominant phylotypes identified in Fig. 2 were sequenced. Two putative chimeric sequences were identified and removed from further analysis. Representative clones for the 14 phylotypes were identified by overlaying PCR-SSCP profile (soil profiles and clone profiles) with GeneScan software to identify peaks that migrated the same distance through the MDE gel. The clones are as follows (numbers correspond to those in Fig. 2): 1, MWS38; 2, MCS34; 3, MCS7; 4, MBS11; 5, MCS48; 6, MAS3; 7, MHS15; 8, MAS18; 9, MWS12; 10, MGS4; 11, MGS13; 12, MCS4; 13, MHS2; 14, MHS12.

We identified four clones that generated abnormal SSCP profiles (Fig. 7). MHS6 generated a broad peak that may represent a single-stranded secondary structure that is unstable under the conditions used in this study. MCS12, MHS1, and MCS34 generated a doublet consisting of a major SSCP peak followed by a slower-migrating minor peak. Each of these clones produced a characteristic doublet and migrated a unique distance through the gel. This result may indicate two stable single-stranded secondary structures for each clone. These four abnormal SSCP patterns were observed on multiple MDE gels and from multiple PCRs, including a PCR with purified plasmid as the template.

**Phylogenetic analysis.** The forty-two sequences identified in this study were all found to group phylogenetically in C1b (Fig. 7). The forty-two sequences identified in this study were all found to group phylogenetically in C1b (Fig. 7).
A number of clones grouped within previously identified clades C1b.A1, C1b.A2, C1b.B1, and C1b.B2 (45). Clones MHS15, MBS11, MWS36, MBS20, MWS31, MGS1, MWS12, MHS36, MAS5, MAS46, MAS20, and MAS18 grouped within C1b.A1. Clones MGS6, MBS2, MCS27, MHS20, MGS5, and MGS4 grouped within C1b.A2. Clones MWS22, MWS38, MCS46, and MCS7 grouped within C1b.B1. Clones MHS29 and MCS3 placed as outgroups to C1b.B1 and may represent an extended monophyletic clade C1b.B1. Clones MCS4, MCS41, and MHS2 grouped within C1b.B2. Overall tree topology was confirmed by comparing trees generated by the evolutionary-distance, maximum-parsimony, and maximum-likelihood methods. The three clones producing a PCR-SSCP doublet, MCS12, MHS1, and MCS34, were found to cluster together in the phylogenetic tree, although they did not form a monophyletic clade.

**DISCUSSION**

To explore crenarchaeal assemblages inhabiting disparate ecosystems, replicated independent soil samples were collected at six locations representing prairie, forest, turf, and agricultural ecosystems. The PCR-SSCP electropherograms from three of these locations, West Madison turf (W), Picnic Point forest (B), and Curtis prairie (C) generated very low within-site variability (Fig. 2), suggesting that crenarchaeal relative diversity is stable within the sampled area at these locations. In general, richness was consistent within each location while...
evenness varied at some locations, namely, the Hancock Agricultural Research Station (H), Arlington Agricultural Research Station (A), and Walnut St. turf (G) locations. Ordination of prin1 and prin2 showed that each sampling location produced a unique cluster that was significantly different (LSD $\alpha = 0.5$) from every other cluster (Fig. 3). For the analysis of crenarchaeal PCR-SSCP profiles, we avoided calculated diversity indices as they have been reported to produce inaccurate estimates for samples containing low species numbers. For a review of the suitability of ecological diversity indices applied to molecular soil microbiology, see reference 18.

The effects of a number of different variables on crenarchaeal assemblages were considered in this data set. First, comparisons between two turf sites and between two agricultural sites revealed different PCR-SSCP profiles, suggesting that macroscopic habitat descriptions are not predictive of the crenarchaeal assemblage inhabiting the soil. Furthermore, the two most similar sets of profiles, turf at West Madison (W) and forest at Picnic Point (B), represent different ecosystems yet group by PCA more closely than the two turf locations (W and G). Second, soil classification could not predict the similarity of crenarchaeal assemblages. The replicated soils, Typic Hapludoll (C and D) and Typic Argiudoll (W and A), did not form distinct clusters in PCA. Finally, the spatial distance between these sampling locations does not predict the similarity of crenarchaeal assemblages. The two closest sites, B and G (0.8 km apart), were less similar than B and W (5.7 km apart). Of the variables considered, the best predictor of crenarchaeal assemblage was found to be small-scale distance (within 0.3 m). That is, soil samples collected within 0.3 m provide the best prediction of the crenarchaeal assemblage profile for an unknown soil sample. To explore the spatial variation within a single uniform ecosystem, soil samples were collected from plots within an agricultural field at the Hancock Agricultural Research Station. The soil type, Plainfield sand (Typic Udipsamment), was consistent throughout all of the plots, and the plots were treated identically (46). These samples revealed a heterogeneous distribution of phylotypes that was not correlated with horizontal distance. Of the five sets of samples collected along a 0.3-m transect, only two exhibited nearly identical profiles, A to C and D to F (Fig. 5). The other three sets of 0.3-m samples, G to I, J to L, and M to O, showed greater changes in phylotype evenness. Two of the soil samples collected the farthest apart, J and M, were more similar to each other (in both richness and evenness) than to the samples closest to them, samples K and L and samples N and O, respectively. In addition, no reproducible effect of depth on crenarchaeal assemblage was detected (Fig. 6). Although some variability among these agricultural soil samples (horizontal and depth samples) was observed, in general, the Hancock electropherograms (Fig. 5 and 6) appeared more similar to each other than to any of the electropherograms from the other sampling locations (Fig. 2).

The Hancock soil data suggest that individual crenarchaeal phylotypes are able to dominate undefined patches of soil throughout this uniform agricultural field. The result is a mosaic distribution of patches inhabited by different dominant crenarchaeal phylotypes. No single variable considered in this study could sufficiently explain the observed differences between electropherograms. The diversity of crenarchaeotes within a single soil sample may thus be the result of a complex interaction of environmental and biological factors. For instance, soil type alone may not be the primary determinant of crenarchaeal diversity, but it may be an important factor when combined with the presence of a specific bacterial clade. In future experiments, PCR-SSCP profiling would be a suitable tool to estimate the influence of different environmental factors on crenarchaeal diversity by providing the means to rapidly screen permutations of many environmental conditions.

Six 16S rDNA clone libraries were generated to survey the crenarchaeal sequences present at each sampling site and to identify individual clones corresponding to the dominant phylotypes present in the soil samples. To include longer sequences for phylogenetic tree construction, the libraries were generated with primers 133FN6F and 1492R (−1,320 nt amplified) and then screened by PCR-SSCP (133FN6F and 248R5P, −115 nt amplified) to identify unique phylotypes. Most of the 16S rDNA clones sequenced (38 of 42) generated an SSCP profile consisting of a single sharp peak (Fig. 7 and 9), suggesting that this region of the crenarchaeal 16S rDNA often forms stable single-stranded molecules and is well suited for PCR-SSCP. However, grouping of the PCR-SSCP profiles by phylogeny (Fig. 9) reveals that most peaks cannot be matched to distinct phylogenetic clades. Cloned sequences grouping together within the phylogenetic tree did not always form distinct clusters of PCR-SSCP peaks. Some phylogenetic clades can be distinguished from each other, i.e., C1b.B1 and C1b.B2, but most groupings cannot be distinguished, i.e., C1b.A1 and C1b.A2.

This study exemplifies the importance of collecting replicated samples in the study of mesophilic soil Crenarchaeota. Pooling multiple samples provides a means to limit misrepresentation of the diversity in environmental samples, but to be effective, the pooling strategy should be based on an understanding of the underlying variability. For instance, a single pooled soil sample composed of A and B (Fig. 5) would produce a much different estimate of crenarchaeal diversity within a Hancock Agricultural Research Station field than a pooled sample of N and O. The importance of replicated sampling is further exemplified by the data from the soil depth profiles (Fig. 6). Individual cores did produce differences in crenar-
FIG. 8. Inferred phylogenetic tree of *Crenarchaeota* C1b sequences. The clones identified in this study are in boldface. Clades A1, A2, B1, and B2 have been identified previously (45), and the groups numbered 1 to 10 correspond to clone groupings designated in Fig. 9. Bootstrap values (100 resamplings within DNAPARS) for branch positions are indicated as filled circles for strong support (>75% bootstrap value) and as unfilled circles for moderate support (50 to 74% bootstrap values). Branch points without circles are unresolved (<50% bootstrap values). GenBank accession numbers are included after the clone names. The scale bar represents 0.01 change per nucleotide.
The y axis was adjusted for each electropherogram and represents relative fluorescence intensity, and the x axis represents relative migration distance.

chaeal assemblages according to depth. For instance, soil core sample 3e contains a slightly higher relative proportion of phylotype 3 (Fig. 2 numbering) at 20 cm below the surface. Individually, this result could be interpreted as preferential colonization of this phylotype, but comparison to soil core 1, samples 1a to 1e, reveals that this phylotype can be a dominant member at all of the depths sampled. These results highlight the importance of replicated sampling to distinguish the aberrant chance event from a true representation of the microbial community in the environment.

In conclusion, characterizing many independent samples allows the preservation of environmental variability, which may provide insights into the ecological significance of crenarchaeotes. Without adequate sampling, meaningful biological variation may be obscured. The results described in this paper provide a basis for designing sampling strategies to test hypotheses about crenarchaeal diversity in mesophilic soils. For instance, to determine whether crenarchaeal assemblages in the rhizosphere are distinct from bulk soil in native environments, many samples at different locations would be necessary to account for variation between sampling sites.

ACKNOWLEDGMENTS

This work was supported by the Storkan-Hanes Foundation, Novartis AG, and the McKnight Foundation.

We thank Dick Wolkowski for soil classifications, the University of Wisconsin Biotechnology Center for sharing C. R. Kuske. Bret Larget and Teri Balser for statistical consultations, and Holly Simon and two anonymous reviewers for comments on an early draft. Holly Simon is also thanked for providing unpublished data.

REFERENCES


