

## Use of Length Heterogeneity PCR and Fatty Acid Methyl Ester Profiles To Characterize Microbial Communities in Soil†

NANCY J. RITCHIE, MARY E. SCHUTTER, RICHARD P. DICK, AND DAVID D. MYROLD\*

*Department of Crop and Soil Science, Oregon State University, Corvallis, Oregon 97331-7306*

Received 17 September 1999/Accepted 17 January 2000

**In length heterogeneity PCR (LH-PCR) a fluorescently labeled primer is used to determine the relative amounts of amplified sequences originating from different microorganisms. Labeled fragments are separated by gel electrophoresis and detected by laser-induced fluorescence with an automated gene sequencer. We used LH-PCR to evaluate the composition of the soil microbial community. Four soils, which differed in terms of soil type and/or crop management practice, were studied. Previous data for microbial biomass, nitrogen and carbon contents, and nitrogen mineralization rates suggested that the microbial characteristics of these soils were different. One site received two different treatments: no-till and conventional till perennial ryegrass. The other sites were no-till continuous grass plots at separate locations with different soil types. Community composition was characterized by assessing the natural length heterogeneity in eubacterial sequences amplified from the 5' domain of the 16S rRNA gene and by determining fatty acid methyl ester (FAME) profiles. We found that LH-PCR results were reproducible. Both methods distinguished the three sites. The most abundant bacterial community members, based on cloned LH-PCR products, were members of the  $\beta$  subclass of the class *Proteobacteria*, the *Cytophaga-Flexibacter-Bacteriodes* group, and the high-G+C-content gram-positive bacterial group. Strong correlations were found between LH-PCR results and FAME results. We found that the LH-PCR method is an efficient, reliable, and highly reproducible method that should be a useful tool in future assessments of microbial community composition.**

The study of microbial diversity in soil is hampered by limitations in isolation and culture techniques. It is estimated that we are able to culture less than 1% of the microbes present in soil (28). In order to overcome these problems, various methods to assess microbial diversity have been developed in order to circumvent the need for isolation. Many nonmolecular and molecular methods are available; these methods include fatty acid methyl ester (FAME) analysis (7, 8, 15, 36), phospholipid fatty acid analysis (5, 15, 19, 34, 35), Biolog substrate utilization profile analysis (7, 11, 19), guanine-plus-cytosine composition analysis (15, 16, 18, 26), amplified ribosomal DNA restriction analysis (10, 14), cloning and sequencing (4), and denaturing gradient gel electrophoresis (DGGE) (16, 23), as well as two more recent techniques, terminal restriction fragment length polymorphism (T-RFLP) analysis (3, 6, 21) and length heterogeneity PCR (LH-PCR) analysis (31).

LH-PCR analysis is similar to the more commonly used T-RFLP method. The difference between these two methods is that the T-RFLP method identifies PCR fragment length variations based on restriction site variability, whereas LH-PCR analysis distinguishes different organisms based on natural variations in the length of 16S ribosomal DNA sequences. T-RFLP analysis has been used successfully for a variety of environments, such as activated sludge, bioreactor sludge, aquifer sand, termite guts (21), aquatic environments (25), and mercury-contaminated soil (6). Use of the LH-PCR method has been more limited in studies of microbial diversity. To date, it has been used only for an aquatic environment (31).

One of the concerns with all PCR-based methods is biased representation of the community (27, 31). It is important that

the primers and DNA extraction procedures used minimize discrimination against community members. The first assurance is an effective DNA extraction method. We used a soil DNA extraction procedure that has been used successfully to characterize soil communities (4). Borneman et al. (4) performed a culture-independent analysis of microbial diversity in which cloning and sequencing techniques were used. They found that the soil DNA extraction method which they used was fast and efficient and effectively captured the diversity present in the microbial community. The effectiveness was based on the enormous levels of diversity found in the soils studied. In order to ensure that minimum bias occurred during amplification, we used universal primers and PCR protocols that have been used successfully to characterize complex communities (31).

The extractable FAME and phospholipid fatty acid contents of soils have been used extensively to characterize soil microbial communities (5, 7, 8, 15, 19, 35, 36). FAME analysis is relatively simple and fast, and its effectiveness for assessing community structure has been demonstrated (8; M. E. Schutter and R. P. Dick, submitted for publication).

To date, there has not been a study in which the researchers compared the results of a FAME profile analysis with the results of a DNA-based method for soil samples, although Fries et al. (14) performed a FAME analysis and an amplified ribosomal DNA restriction analysis with microbial communities from a toluene-, phenol-, and chlorinated hydrocarbon-contaminated aquifer. These workers found that most of the originally dominant microbial groups were still dominant after the last remediation treatment. Both methods successfully assessed community composition and succession. In this study, we used both FAME and LH-PCR analyses to study soil microbial community composition and evaluated the effectiveness of the two methods side by side for assessing composition.

The objectives of this study were (i) to determine the suitability and reproducibility of the LH-PCR method for measur-

\* Corresponding author. Mailing address: Department of Crop and Soil Science, Agric. Life Sci. Bldg., Rm 3017, Oregon State University, Corvallis, OR 97331-7306. Phone: (541) 737-5737. Fax: (541) 737-5725. E-mail: David.Myrold@orst.edu.

† Oregon Agricultural Experiment Station technical paper 11,614.

TABLE 1. Soil collection sites

Site	Management system	Location	Soil type
A	Fine fescue, no tillage	Silverton, Oreg.	Nekia silty clay loam (Xeric Haplohumult)
B	Tall fescue, no tillage	Lewisberg, Oreg.	Amity silty clay loam (Argiaquic Xeric Argialboll)
C1	Perennial ryegrass, conventional tillage	Tangent, Oreg.	Woodburn silt loam (Aqualtic Argixeroll)
C2	Perennial ryegrass, no tillage	Tangent, Oreg.	Woodburn silt loam (Aqualtic Argixeroll)

ing the microbial community composition of soil and (ii) to compare the LH-PCR method with the FAME method.

#### MATERIALS AND METHODS

**Sampling procedures.** Soil samples were collected in November and December 1998 from four different grass fields in the Willamette Valley of Oregon (Table 1) that were part of the Sustainable Grass Seed Cropping System Research project, as described by J. J. Steiner (<http://pwa.ars.usda.gov/nfspr/steiner/steinersustain.htm>). The experimental design at each site consisted of a randomized block with four replicate plots per treatment (a total of 16 plots). A soil sample from each plot consisted of a composite of 25 to 30 2.5-cm-diameter cores taken from the top 5 cm of soil. All soil samples were sieved through a 2-mm screen and stored at  $-20^{\circ}\text{C}$  until they were used.

**DNA extraction.** Microbial DNA was extracted in triplicate from soil (0.5 g) from each plot by using a Fast DNA SPIN kit for soil (Bio 101, La Jolla, Calif.). This kit was designed to extract PCR-ready genomic DNA from bacteria, fungi, plants, and animals in a soil community. Briefly, cells were lysed with detergents, silica, and ceramic beads in a powerful homogenization unit (Fast DNA Prep System; Bio 101). Additional purification steps were performed in order to remove proteins and other contaminants. For the final purification step we used a GENCLEAN procedure that provided PCR-ready genomic DNA. The whole process took 30 to 45 min for six samples (the Fast DNA Prep System can hold up to 12 samples at a time). Purified DNA was stored at  $-20^{\circ}\text{C}$  until it was used.

**LH-PCR.** Purified DNA (10 ng) was amplified with a PCR by using fluorescently labeled forward primer 27F (5'-[6FAM] AGAGTTTGTATCTGGCTCA G-3' [17]) and unlabeled reverse primer 338R (5'-GCTGCCTCCCGTAGGAG T-3' [2]). Both primers are considered specific for eubacteria. The reactions were performed by using 50- $\mu\text{l}$  (final volume) mixtures containing  $1 \times$  PCR buffer, 0.06% bovine serum albumin, 1.5 mM  $\text{MgCl}_2$ , each deoxynucleoside triphosphate at a concentration of 0.2 mM, each primer at a concentration of 0.2  $\mu\text{M}$ , and 2 U of *Taq* DNA polymerase. Initial denaturation at  $94^{\circ}\text{C}$  for 3 min was followed by 25 cycles consisting of denaturation at  $94^{\circ}\text{C}$  for 45 s, annealing at  $55^{\circ}\text{C}$  for 45 s, and extension at  $72^{\circ}\text{C}$  for 2 min. There was a final extension step that consisted of  $72^{\circ}\text{C}$  for 7 min. LH-PCR samples were stored at  $-20^{\circ}\text{C}$  in the dark until they were used (usually less than 1 week).

**LH-PCR analysis.** Our initial results showed that the most consistent results were obtained when the same quantity of PCR products was used for analysis. The LH-PCR products (exactly 1 ng; approximately 10 fmol based on a product length of 335 bp) were separated by length by using Long Ranger (FMC, Rockland, Maine) polyacrylamide gel electrophoresis and an ABI automated DNA sequencer with GeneScan v2.1 software (Applied Biosystems, Inc., Foster City, Calif.) operated by workers at the Central Services Laboratory (Center for Gene Research and Biotechnology, Oregon State University, Corvallis). The software converted fluorescence data into electropherograms; the peaks represented fragments of different sizes, and the areas under the peaks were the relative proportions of the fragments (Fig. 1). Lengths (in base pairs) were calculated by using a size standard (GeneScan 400Rox or GeneScan 2500Rox; Applied Biosystems, Inc.) in conjunction with the no-smoothing option and the Local Southern method in GeneScan. The size standard was added to every lane of the gel for increased precision. We began by using GeneScan 2500Rox as our standard, but after initial tests we used GeneScan 400Rox because it gave more reproducible size calibration results. When we compared LH-PCR data with FAME analysis data, we used GeneScan Rox400. Relative peak areas were calculated by dividing an individual peak area by the total peak area (sum of the areas of all the peaks).

**LH-PCR reproducibility.** Based on the initial LH-PCR data, we chose the site with the greatest amount of variability for replicate plots to study reproducibility. We examined the following four different levels of reproducibility: plot, soil DNA extraction, PCR amplification, and GeneScan level. At the plot or within-site level, we extracted DNA from composite soil samples obtained from each of three replicate plots located at the site studied and compared the LH-PCR results. At the soil DNA extraction or within-plot level, we compared the LH-PCR results for three separate DNA extracts obtained from the composite soil sample from one plot. At the PCR amplification level, we compared the LH-PCR results obtained from three separate PCR amplifications of one soil DNA extract. At the GeneScan level, we compared the LH-PCR results obtained for three replicates during a single PCR amplification. This allowed us to determine which step in the LH-PCR procedure resulted in the greatest variability.

**Sequencing.** Unlabeled PCR fragments from site B were sequenced to verify that we were analyzing 16S rRNA genes in our LH-PCR analysis. First, the PCR product generated with the primers used for LH-PCR was cloned by using pGem-T Easy Vector System II (Promega Corp., Madison, Wis.). Clones were screened to determine whether inserts were present by using  $\alpha$ -complementation with X-Gal (5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactopyranoside) and IPTG (isopropyl- $\beta$ -D-thiogalactopyranoside). Positive clones were also screened by performing LH-PCR. Inserts from 33 random clones were subjected to an LH-PCR analysis, which allowed us to choose clones of specific lengths to sequence. Clones corresponding to LH-PCR peaks at 343 bp (clone 1), 345 bp (clone 2), 346 bp (clones 10 and 22), 353 bp (clone 18), 323 bp (clone 17), and 359 bp (clone 26) were sequenced by using *Taq* dye terminator chemistry and an ABI cycle sequencer (Central Services Laboratory, Center for Gene Research and Biotechnology, Oregon State University, Corvallis). The seven sequences were compared to sequences in the GenBank database by using BLAST 2.0 (1) and were interrogated for chimeric structures by using Check Chimera from the Ribosomal Database Project web page ([www.cme.msu.edu/RDP/html/index.html](http://www.cme.msu.edu/RDP/html/index.html)) (22). A phylogenetic analysis program, the Genetics Computer Group FileUp program, version 10, was used to align sequences (9). Indel-containing and ambiguous regions were excluded from the analyses. A phylogenetic tree was constructed by using neighbor joining with the Kimura two-parameter model method (20). Bootstrap analysis was performed by using 100 replicates to determine the confidence values for tree branches (12).

**FAME analysis.** Fatty acids were extracted from soil samples by using the ester-linked method developed by Rhae Drijber, University of Nebraska (R. Drijber, personal communication). This method is described by Schutter and Dick elsewhere (submitted). Briefly, 3 g of soil was mixed with 15 ml of 0.2 M KOH in methanol, and the preparation was incubated for 1 h at  $37^{\circ}\text{C}$ , during which ester-linked fatty acids were released from soil microorganisms and methylated. FAMES were extracted into a hexane organic phase, and the sample was centrifuged at  $480 \times g$  for 10 min to separate the aqueous and hexane phases. The hexane layer was transferred to a clean tube, and the hexane was evaporated off, after which FAMES were resuspended in 0.5 ml of hexane-methyl *tert*-butyl ether (1:1) for analysis. Individual FAMES were separated and quantified by gas chromatography by using a model 5890 Series II instrument (Hewlett-Packard Co., Palo Alto, Calif.) equipped with a flame ionization detector. The temperature was programmed to increase from 170 to  $270^{\circ}\text{C}$  at a rate of  $5^{\circ}\text{C}$  per min. The temperature was increased to  $270^{\circ}\text{C}$  for 2 min between samples in order to clean the column. Fatty acids were identified and relative peak areas were determined by using standard MIDI procedures (30).

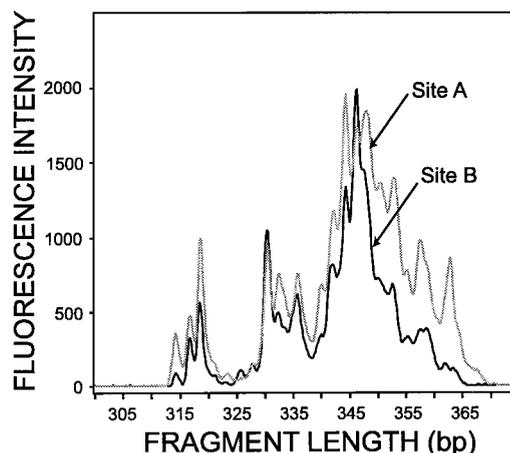


FIG. 1. Example of electrophoretic output from GeneScan software. The fragment lengths and relative fluorescence intensities of LH-PCR products amplified from DNA extracted from soil samples obtained at sites A and B are shown. Although neither site contained LH-PCR fragments that had unique lengths, the relative abundances of the fragments varied markedly.

**Statistical analysis.** The reproducibility of LH-PCR results was evaluated by comparing coefficients of variation (CVs) of the different DNA peaks at each level of analysis (e.g., plot, soil DNA extraction, etc.). The CVs were calculated from the relative area of each peak of three replicates, and the CVs of the individual peaks were averaged to produce an overall CV for each level.

After checking that the distributions of the relative peak areas were normally distributed, a necessary assumption for using principal component analysis (PCA), we analyzed the relative peak areas for LH-PCR and FAME results by using the PCA option in the statistical analysis package PC-ORD (24).

A Mantel test was performed in order to examine similarities between LH-PCR and FAME results (PC-ORD) (24). The Mantel test evaluates the null hypothesis of no correlation between two distance matrices that contain the same set of sample units. In our case, it was used to test the significance of the correlation between FAME-based community structure and LH-PCR-based community structure. The Mantel test was performed by using Sorensen's distance measure and the randomization (Monte Carlo) procedure with 1,000 randomized runs.

Relationships between LH-PCR fragments and FAME peaks were assessed by performing a linear correlation analysis (24).

Three different measures of diversity were used. We calculated richness, evenness, and diversity for FAME and LH-PCR results by using PC-ORD (24) and compared these indices by using a one-way analysis of variance (29).

## RESULTS

**LH-PCR.** Because kinetic bias is a concern with any PCR-based method, we tested for it by analyzing LH-PCR samples after 15, 20, and 25 cycles. There were no observable differences in the number or relative abundances of peaks based on cycle number when GeneScan analysis was used (data not shown); consequently, we used 25 cycles for all samples.

LH-PCR results were highly reproducible, and 19 to 23 different fragments per sample were obtained (Fig. 1). Within a site, the number of LH-PCR peaks obtained was the same for each plot for all of the sites except site B; one unique peak was obtained for one of the site B replicate plots.

Site B was the site with the greatest amount of variability; therefore, reproducibility was studied by using this site. Variability was measured at the plot, soil DNA extraction, PCR amplification, and GeneScan analysis levels as described above. For most peaks, the variation decreased from the plot level through the GeneScan analysis level (Fig. 2); e.g., the CV of the 346-bp fragment, which was the most abundant PCR product and accounted for about one-fourth of the total peak area, steadily decreased from 8.5 to 2.5%. Less abundant peaks exhibited greater variability, but generally the results followed the same trend. When average values for all LH-PCR products were calculated, the CVs decreased from the plot level (14%) to the soil DNA extraction level (13%) to the PCR amplification level (12%) to the GeneScan analysis level (8%). Thus, the greatest amount of variability was at the plot level, as expected.

Average relative peak areas were determined by using two separate DNA extracts per plot and were compared by using PCA (Fig. 3A). The first principal component (PC-1) explained 45% of the variability, and the second principal component (PC-2) explained 20% of the variability. Site C was separated from sites A and B by PC-1, and PC-2 separated site A from site B. Thus, the LH-PCR data allowed us to distinguish the microbial communities at the three different sites. The microbial communities present in soils that were subjected to different tillage practices at site C (site C1 versus site C2) appeared to differ slightly when this method was used. This suggests that the LH-PCR method may be useful for distinguishing soil microbial communities that are quite similar.

**Sequencing.** All seven sequences obtained were 16S rRNA sequences based on BLAST search comparisons with GenBank sequences. One of the seven sequences (peak length, 323 bp; clone 17) appeared to be chimeric, so no further investigations were performed with this sequence. The remaining six

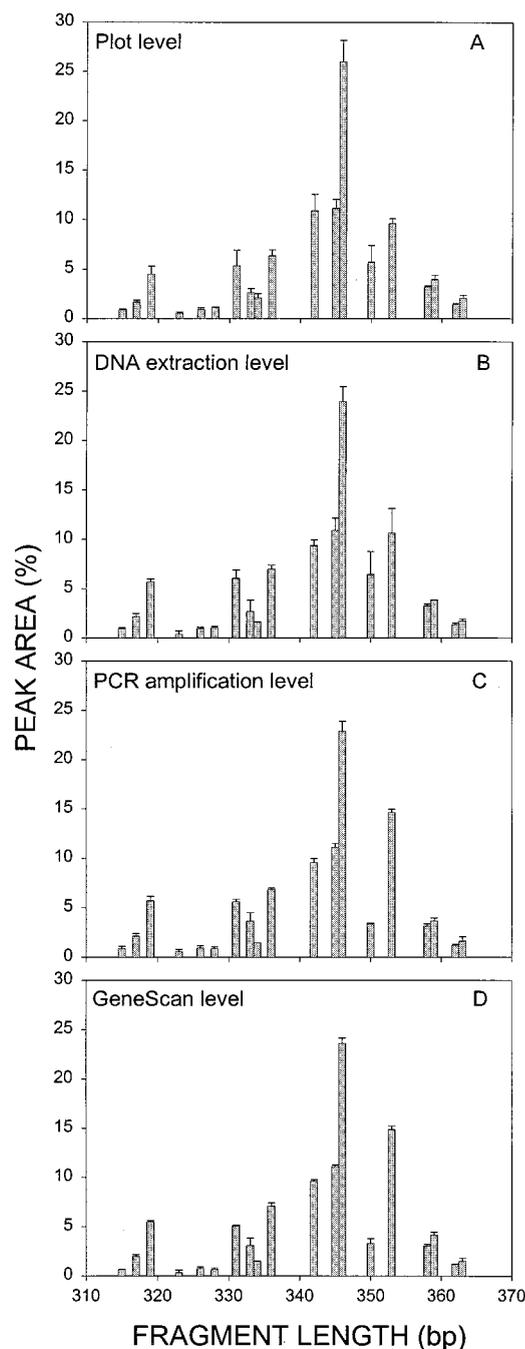


FIG. 2. Comparison of variability at the plot (A), DNA extraction (B), PCR amplification (C), and GeneScan analysis (D) levels. The amounts of variability for three replicates at each level (total number of samples, 12) are shown. The error bars indicate the standard deviations based on three replicates. The average CVs for all of the peaks were 14, 13, 12, and 8% for the plot, DNA extraction, PCR amplification, and GeneScan analysis levels, respectively.

sequences, clones 1, 2, 10, 18, 22, and 26, were aligned with similar sequences (based on BLAST hits) and were clustered in a neighbor-joining tree (Fig. 4). Both clone 1 and clone 10 clustered with members of the  $\beta$  subclass of the class *Proteobacteria* ( $\beta$ -*Proteobacteria*), and the clone 1 sequence was identical to the sequence of a *Nitrosospira* sp. Clone 2 clustered with the high-G+C-content gram-positive bacteria, and clone

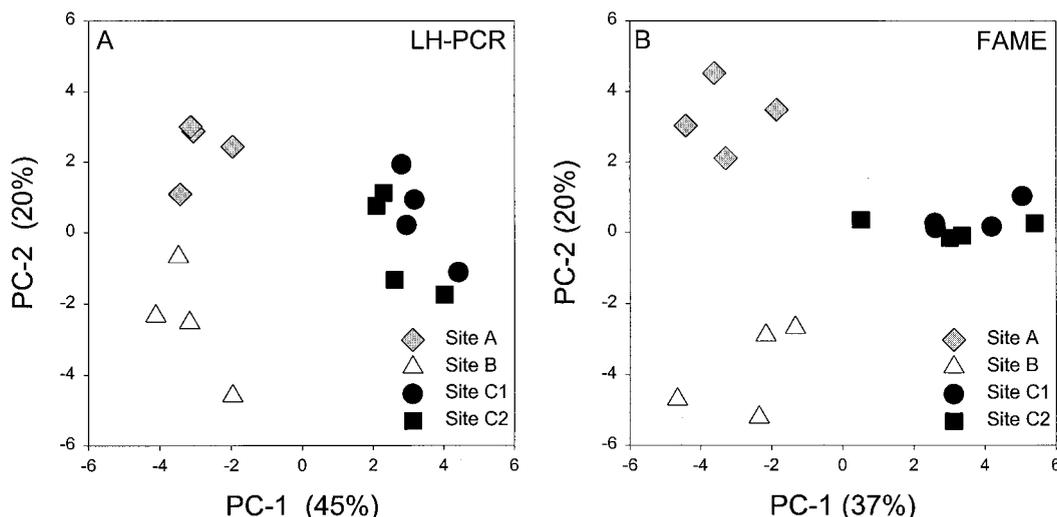


FIG. 3. Average relative peak areas obtained with two separate DNA extracts (A) and two separate FAME extracts (B) per plot, compared by using PCA. Sites A and B are long-term, no-till fescue plots that are located in geographically distinct areas. Sites C1 and C2 are long-term ryegrass plots that were either conventionally tilled (site C1) or not tilled (site C2). When LH-PCR data were used, PC-1 explained 45% of the variability and PC-2 explained 20% of the variability. For the FAME analysis, PC-1 explained 37% of the variability and PC-2 explained 20% of the variability. Within a site, the variation among the data for the four replicate plots was one indication of variability.

26 grouped with a *Syntrophomonas* sp., a low-G+C-content gram-positive bacterium. Both clone 18 and clone 22 clustered with the *Cytophaga-Flexibacter-Bacteriodes* group.

**FAMES.** Twenty-seven to 35 distinct fatty acids were observed per site. The average relative peak areas were determined by using two FAME extracts per plot and were compared by using PCA (Fig. 3B). The amounts of variability explained by PC-1 (37%) and PC-2 (20%) when FAME data

were used were similar to the amounts of variability observed when the LH-PCR method was used. PC-1 separated site C from sites A and B, and PC-2 separated site A from site B. Thus, by using FAMES we were able to distinguish the microbial communities at the three different sites; however, the FAME data did not differentiate between the soil microbial communities that were subjected to the different tillage practices at site C.

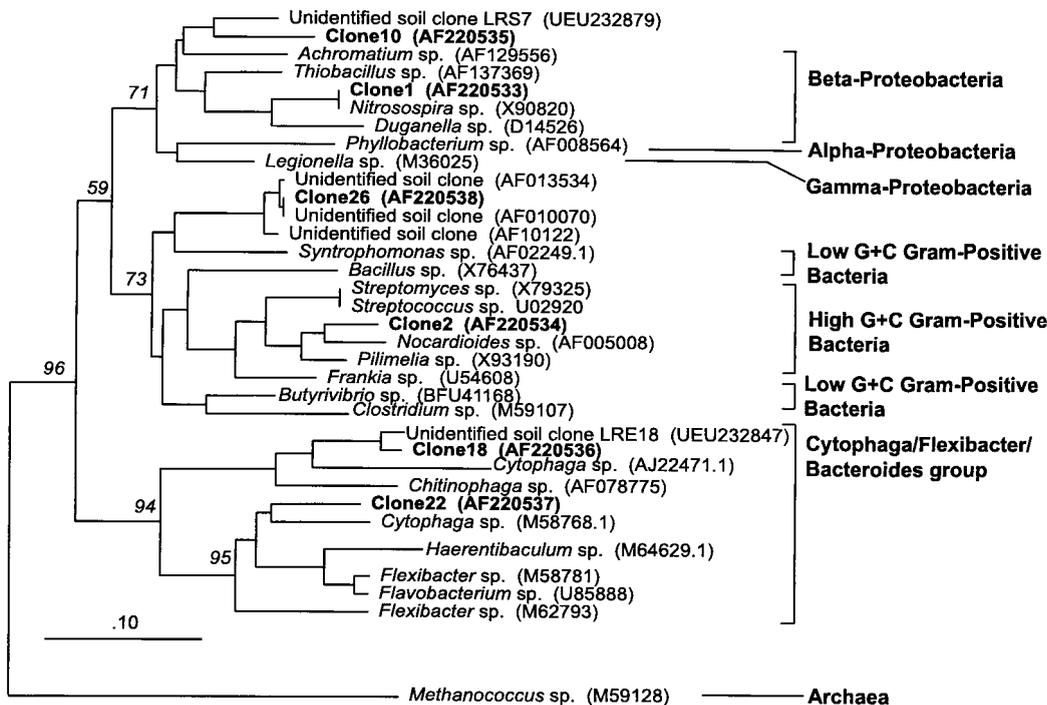


FIG. 4. Neighbor-joining tree for LH-PCR fragment clones obtained by using 16S ribosomal DNA sequence data. *Methanococcus* sp. was used to root the tree. The sequences generated in this study are in boldface type. The numbers above the branches are bootstrap values. The numbers in parentheses are GenBank accession numbers.

TABLE 2. Diversity indices calculated from LH-PCR and FAME data

Method	Site	Richness <sup>a</sup>	Evenness <sup>b</sup>	Diversity <sup>c</sup>
LH-PCR	A	20.0 ± 0.0c <sup>d</sup>	0.91 ± 0.01b	2.73 ± 0.02c
	B	19.3 ± 0.5d	0.90 ± 0.01c	2.65 ± 0.04d
	C1	22.0 ± 0.0b	0.91 ± 0.01b	2.80 ± 0.02b
	C2	23.0 ± 0.0a	0.92 ± 0.00a	2.90 ± 0.01a
FAME	A	34.0 ± 0.8a	0.85 ± 0.00a	3.00 ± 0.02a
	B	30.5 ± 2.9b	0.83 ± 0.02b	2.82 ± 0.09b
	C1	31.8 ± 1.0ab	0.86 ± 0.00a	2.96 ± 0.02a
	C2	32.4 ± 1.0ab	0.85 ± 0.01a	2.96 ± 0.02a

<sup>a</sup> Richness is equal to the number of LH-PCR or FAME peaks.

<sup>b</sup> Evenness is equal to diversity/ln (richness).

<sup>c</sup> Diversity (Shannon index) ( $H$ ) was calculated as follows:  $H = -\sum p_i \ln(p_i)$ , where  $p_i$  is the relative abundance of a given LH-PCR or FAME peak.

<sup>d</sup> Values are means ± standard deviations ( $n = 4$ ). For each index, different letters indicate that values for different sites determined by the same method were significantly different ( $P < 0.05$ ), as determined by one-way analysis of variance and least significant difference.

**Comparison of methods.** Table 2 shows several diversity indices for FAME and LH-PCR data. The richness (or number of peaks) was greater but more variable when the FAME analysis was used than when the LH-PCR analysis was used, which was also reflected in the greater evenness observed with LH-PCR data than with FAME data. There were significant differences among all sites when both richness and diversity indices were used with the LH-PCR data. When evenness was used, sites A and C1 were not significantly different, but all other sites were significantly different when LH-PCR data were used. When FAMES were used, there were fewer significant differences among sites. The richness indices for sites A and B were significantly different when FAMES were used. The fatty acid diversity and evenness for site B were significantly different than the fatty acid diversity and evenness for any of the other sites. Both methods showed that site B had a less diverse microbial community than the other sites.

Several significant correlations ( $P < 0.01$ ) were found between 19 of the LH-PCR fragments and 22 of the extracted fatty acids (Table 3), and most (66%) of these correlations were positive. In general, a greater proportion (74%) of positive correlations was found between LH-PCR fragments and fatty acids associated with bacteria, and a greater proportion (56%) of negative correlations was associated with fatty acids known as fungal markers. These trends are consistent with the fact that the LH-PCR fragments were generated by using eubacterium-specific primers.

## DISCUSSION

For both methods, PCA explained a large amount of the variability in the data (65% for LH-PCR data and 57% for FAME data when the first two principal components were used). Both FAME data and LH-PCR data could be used to separate the microbial communities from the three different locations. LH-PCR data appeared to do a better job of separating sites that were subjected to different tillage practices, however. The ordination patterns generated by the FAME and LH-PCR analyses were highly correlated ( $r = 0.69$ ;  $P = 0.001$ ; Mantel test) and supported the similar distribution of sites observed with PCA (Fig. 3).

Many, but not all, of the fatty acids extracted from the soils which we studied have been characterized. For example, fatty acid i15:0 is a fatty acid that is most commonly associated with gram-positive bacteria (30). This fatty acid exhibited a signifi-

cant positive correlation with five of the LH-PCR fragments, which suggests that these fragments may be associated with members of gram-positive genera. Fatty acid i15:0 also was negatively correlated with three of the LH-PCR fragments, which suggests that these fragments are probably not from members of gram-positive genera.

One potential use of comparing LH-PCR fragments with fatty acid peaks may be identification of fatty acids that originate from unknown sources. For example, the source of 16:0 N alcohol in community FAME profiles is not clear. A search of the MIDI aerobe library (30) revealed that relatively large amounts (3 to 4.5%) of 16:0 N alcohol are present in the FAME profiles of a few *Moraxella* species, which are gram-negative coccobacilli that belong to the  $\beta$ -*Proteobacteria*. It is not known whether the 16:0 N alcohol found in the soil community FAME profiles was from these organisms, but sequencing of the 354-bp LH-PCR fragment that was highly correlated ( $r = 0.74$ ) with the presence of 16:0 N alcohol may reveal its source. Unfortunately, we did not clone sequences that were this length.

When we examined the fragments that were sequenced, however, we did not find a correspondence with phylogenetic relationships based on LH-PCR fragment lengths and FAME correlations. For instance, the LH-PCR peak at 345 bp was positively correlated with fungal FAME markers and negatively correlated with gram-positive bacterial or actinomycete FAME markers, whereas the LH-PCR fragment that was sequenced (clone 2) was closely associated with the high-G+C-content gram-positive bacteria. One reason for this discrepancy may be that one microbe can contribute to more than one FAME peak, yet a single LH-PCR fragment or FAME peak can correspond to several different microbes. A good example of this was observed with the 346-bp LH-PCR fragment. The two clones that corresponded to this fragment, clones 10 and 22, were not the same. Clone 10 clustered with members of the  $\beta$ -*Proteobacteria*, and clone 22 clustered with members of the *Cytophaga-Flexibacter-Bacteroides* group. Clearly, more extensive comparisons of sequences of LH-PCR fragments and FAME peaks will have to be done to evaluate the utility of these correlations.

Our LH-PCR sequence results were consistent with the results which other researchers have obtained with soil microbial communities (4, 10). For instance, the most abundant bacterial community members at the three study sites, based on sequenced peak lengths of 345 and 346 bp, were members of the  $\gamma$ -*Proteobacteria* and the *Cytophaga-Flexibacter-Bacteroides* group. Members of the other bacterial groups that were sequenced, the high-G+C-content gram-positive bacteria and the low-G+C-content gram-positive bacteria, have also been commonly found in soils (4, 10). Furthermore, our phylogenetic groups for specific fragment lengths were similar in most instances to the groups described by Suzuki et al. (31). It should be noted, however, that Suzuki et al. (31) found members of several phyla that had fragments which were the same sequence length.

Overall, we found that the LH-PCR technique was effective for studies of soils. It was easy, quick, and reproducible. The whole process, from preparation of DNA extracts to electrophoresis of fluorescently labeled products on a polyacrylamide gel, could be done in an 8-h day, and 12 to 24 samples per day could be analyzed. This technique is much easier and less time-consuming than other commonly used molecular techniques, including rRNA intergenic spacer analysis, DGGE, and T-RFLP PCR analysis. Both rRNA intergenic spacer analysis and DGGE involve cumbersome polyacrylamide gel electrophoresis followed by silver staining, which requires large



quantities of DNA and must be done manually. Automated rRNA intergenic spacer analysis (ARISA) has been developed recently, however (13). Disadvantages of the T-RFLP PCR method include the time involved in an additional restriction digestion (usually overnight) and the potential complications associated with the process (e.g., incomplete digestion and the need for secondary purification steps).

Because LH-PCR analysis is a new method, there are several considerations that require further examination. First, there is not a one-to-one correspondence between the LH-PCR fragment lengths and the lengths of sequenced fragments. In our study, the LH-PCR fragment sizes determined by GeneScan analysis were 1 to 2 bp smaller or larger than the sizes based on sequence analysis. This was due in part to the contribution of the fluorophore (6FAM) to the lengths of the LH-PCR fragments, which increased the size by a small but variable amount (unpublished data). Another factor may be related to the inherent level of precision that can be achieved with standard curves based on fluorescent size standards. For example, we obtained better standard curves when we used GeneScan 400Rox than when we used GeneScan 2500Rox (unpublished data). If these difficulties can be overcome, then it should be possible to directly compare and associate LH-PCR fragment lengths with a sequence fragment length database. At present, such information is limited to marine bacteria (31). It may also be possible to use 16S ribosomal DNA sequences in the GenBank and Ribosomal Database Project databases; however, for this to occur it would have to be assumed that matching LH-PCR primers by base complementarity accurately reflects what occurs in an empirical PCR with environmental DNA templates. Thus, a second consideration is the need for additional cloning and sequencing of LH-PCR fragments in order to determine which fragment length corresponds to which microorganism(s). We note that both the T-RFLP PCR method and the ARISA method have these same two problems.

A third consideration is the level of taxonomic or phylogenetic resolution of the LH-PCR method. Members of more than one taxonomic group can have LH-PCR products that are the same size, although some fragment lengths may be unique to a given taxon (31). Because T-RFLP PCR analysis and ARISA are likely to produce more fragments, one of these methods may provide better taxonomic discrimination than can be obtained with LH-PCR. Soil communities may, however, produce T-RFLP or ARISA patterns that are too complex to analyze. This was the case when we tried to perform a T-RFLP PCR analysis with our soil samples (data not shown).

In addition to the concerns described above, biases involved in PCR need to be addressed. We attempted to minimize bias by limiting the cycle number in order to ensure that kinetic bias as described by Suzuki et al. (32) did not occur. Other biases to consider are the biases inherent in the PCR process. Only dominant, active members of the community may be amplified, and hence many rare members of the community may not be detected. Estimates based on DNA reassociation data suggest that there may be more than 4,000 species in a typical soil (33). In this study, the LH-PCR analysis detected only 19 to 23 fragments having different sizes, and the FAME analysis detected only 27 to 35 different fatty acids. Thus, we did not capture the complete diversity of the microbial communities at our study sites. Further investigation of the relationship between the numbers of fragments in LH-PCR and FAME profiles and overall microbial diversity is needed.

The LH-PCR and FAME techniques were very useful in this study. The most appropriate use of these methods is for rapid assessment of microbial communities. This can be done prior to or in conjunction with more accurate and time-consuming

techniques (such as cloning and sequencing). The LH-PCR technique may be enhanced by using more specific primers or by using this method with a different gene. Such analyses may facilitate discrimination at lower taxonomic levels, as discussed above.

We have shown that the LH-PCR and FAME techniques are comparable in terms of distinguishing microbial communities in soil. They are quick, easy, and effective. The LH-PCR method may be more useful because it is easy to identify unknown peaks by subsequent cloning and sequencing techniques and because discrimination can be increased by using more specific primers or by including a restriction digestion step. Both methods should have great utility in future studies of this kind.

#### ACKNOWLEDGMENTS

We thank Anne Bernhard and Kevin Vergin for use of their equipment and their advice and criticism; Jeffrey Steiner and other members of the Sustainable Grass Seed Cropping Systems Research team for establishing the research plots used in this study; Robert Christ for providing soil samples; Ray Shimabuka of the U.S. Environmental Protection Agency for analysis of FAME samples; and Caprice Rosato for assistance with GeneScan analyses.

#### REFERENCES

1. Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**:403–410.
2. Amann, R. L., L. Krumholz, and D. A. Stahl. 1990. Fluorescent-oligonucleotide probing of whole cells for determinative, phylogenetic, and environmental studies in microbiology. *J. Bacteriol.* **172**:762–770.
3. Avannis-Aghajani, E., K. Jones, D. Chapman, and C. Brunk. 1994. A molecular technique for identification of bacteria using small subunit ribosomal RNA sequences. *Bio/Technology* **17**:144–149.
4. Borneman, J., P. W. Skroch, K. M. O'Sullivan, J. A. Palus, N. G. Rumjanek, J. L. Jansen, J. Nienhuis, and E. W. Triplett. 1996. Molecular microbial diversity of an agricultural soil in Wisconsin. *Appl. Environ. Microbiol.* **62**:1935–1943.
5. Bossio, D. A., K. M. Scow, N. Gunapala, and K. J. Graham. 1998. Determinants of soil microbial communities: effects of agricultural management, season, and soil type on phospholipid fatty acid profiles. *Microb. Ecol.* **36**:1–12.
6. Bruce, K. D. 1997. Analysis of *mer* gene subclasses within bacterial communities in soils and sediments resolved by fluorescent-PCR-restriction fragment length polymorphism profiling. *Appl. Environ. Microbiol.* **63**:4914–4919.
7. Buyer, J. S., and L. E. Drinkwater. 1997. Comparison of substrate utilization and fatty acid analysis of soil microbial communities. *J. Microbiol. Methods* **30**:3–11.
8. Cavigelli, M. A., G. P. Robertson, and M. J. Klug. 1995. Fatty acid methyl ester (FAME) profiles as measures of soil microbial community structure. *Plant Soil* **170**:99–13.
9. Devereux, J., P. Haeberli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12**:387–395.
10. Dunbar, J., S. Takala, S. M. Barns, J. A. Davis, and C. R. Kuske. 1999. Levels of bacterial community diversity in four soils compared by cultivation and 16S rRNA gene cloning. *Appl. Environ. Microbiol.* **65**:1662–1669.
11. El Fantroussi, S., L. Verschuere, W. Verstraete, and E. M. Top. 1999. Effect of phenylurea herbicides on soil microbial communities estimated by analysis of 16S rRNA gene fingerprints and community-level physiological profiles. *Appl. Environ. Microbiol.* **65**:982–988.
12. Felsenstein, J. PHYLIP: phylogeny inference package, version 3.5. University of Washington, Seattle.
13. Fisher, M., and E. W. Triplett. 1999. Automated approach for ribosomal intergenic spacer analysis of microbial diversity and its application to freshwater bacterial communities. *Appl. Environ. Microbiol.* **65**:4630–4636.
14. Fries, M. R., G. D. Hopkins, P. L. McCarty, L. J. Forney, and J. M. Tiedje. 1997. Microbial succession during a field evaluation of phenol and toluene as the primary substrates for trichloroethene cometabolism. *Appl. Environ. Microbiol.* **63**:1515–1522.
15. Griffiths, B. S., K. Ritz, N. Ebbelwhite, and G. Dobson. 1999. Soil microbial community structure: effects of substrate loading rates. *Soil Biol. Biochem.* **31**:145–153.
16. Harris, D. 1994. Analysis of DNA extracted from microbial communities, p. 111–118. *In* K. Ritz, J. Dighton, and K. E. Giller (ed.), *Beyond the biomass*. John Wiley and Sons, West Sussex, United Kingdom.
17. Hicks, R. E., R. I. Amann, and D. A. Stahl. 1992. Dual staining of natural

- bacterioplankton with 4',6-diamidino-2-phenylindole and fluorescent oligonucleotide probes targeting kingdom-level 16S rRNA sequences. *Appl. Environ. Microbiol.* **58**:2158–2163.
18. **Holben, W. E., V. G. Calbrese, D. Harris, J. O. Ka, and J. M. Tiedje.** 1993. Analysis of structure and selection in microbial communities by molecular methods, p. 367–370. *In* R. Guerrero and C. Pedrós-Alió (ed.), *Trends in microbial ecology*. Spanish Society for Microbiology, Barcelona, Spain.
  19. **Ibekwe, A. M., and A. C. Kennedy.** 1998. Phospholipid fatty acid profiles and carbon utilization patterns for analysis of microbial community structure under field and greenhouse conditions. *FEMS Microbiol. Ecol.* **26**:151–163.
  20. **Kimura, M.** 1990. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**:111–120.
  21. **Liu, W., T. L. Marsh, H. Cheng, and L. J. Forney.** 1997. Characterization of microbial diversity by determining terminal restriction fragment length polymorphism of genes encoding 16S rRNA. *Appl. Environ. Microbiol.* **63**:4516–4522.
  22. **Maidak, B. L., J. R. Cole, C. T. Parker, Jr., G. M. Farrity, N. Larsen, B. Li, T. G. Lilburn, M. J. McCaughey, G. J. Olsen, R. Overbeek, S. Pramanik, T. M. Schmidt, J. M. Tiedje, and C. R. Woese.** 1999. A new version of the RDP (Ribosomal Database Project). *Nucleic Acids Res.* **27**:171–173.
  23. **McCaig, A. E., C. J. Phillips, J. R. Stephen, G. A. Kowalchuk, S. M. Harvey, R. A. Herbert, T. M. Embley, and J. I. Prosser.** 1999. Nitrogen cycling and community structure of proteobacterial  $\beta$ -subgroup ammonia-oxidizing bacteria within polluted marine fish farm sediments. *Appl. Environ. Microbiol.* **65**:213–220.
  24. **McCune, B., and M. J. Mefford.** 1997. PC-ORD. Multivariate analysis of ecological data, version 3.0. MjM Software, Gleneden Beach, Oreg.
  25. **Moeseneder, M. M., J. M. Arrieta, G. Muyzer, C. Winter, and G. J. Herndl.** 1999. Optimization of terminal-restriction fragment length polymorphism analysis for complex marine bacterioplankton communities and comparison with denaturing gradient gel electrophoresis. *Appl. Environ. Microbiol.* **65**:3518–3525.
  26. **Nüsslein, K., and J. M. Tiedje.** 1999. Soil bacterial community shift correlated with change from forest to pasture vegetation in a tropical soil. *Appl. Environ. Microbiol.* **65**:3622–3626.
  27. **Polz, M. F., and C. M. Cavanaugh.** 1998. Bias in template-to-product ratios in multitemplate PCR. *Appl. Environ. Microbiol.* **64**:3724–3730.
  28. **Rozsak, D. B., and R. R. Colwell.** 1987. Survival strategies of bacteria in the natural environment. *Microbiol. Rev.* **51**:365–379.
  29. **SAS Institute.** 1996. SAT/STAT user's guide, version 6.12. SAS Institute, Cary, N.C.
  30. **Sasser, M.** 1990. Identification of bacteria by gas chromatography of cellular fatty acids. Technical note 101. Microbial ID, Newark, Del.
  31. **Suzuki, M., M. S. Rappé, and S. J. Giovannoni.** 1998. Kinetic bias in estimates of coastal picoplankton community structure obtained by measurements of small-subunit rRNA gene PCR amplicon length heterogeneity. *Appl. Environ. Microbiol.* **64**:4522–4529.
  32. **Suzuki, M., and S. J. Giovannoni.** 1996. Bias caused by template annealing in the amplification mixtures of 16S rRNA genes by PCR. *Appl. Environ. Microbiol.* **62**:625–630.
  33. **Torsvik, V., J. Goksøyr, and F. L. Daee.** 1990. High diversity in DNA of soil bacteria. *Appl. Environ. Microbiol.* **56**:782–787.
  34. **Vestal, J. R., and D. C. White.** 1989. Lipid analysis in microbial ecology. *BioScience* **8**:535–541.
  35. **Wander, M. M., D. S. Hendrick, D. Kaufman, S. J. Traina, B. R. Stinner, S. R. Kehrmeier, and D. C. White.** 1995. The functional significance of the microbial biomass in organic and conventionally managed soils. *Plant Soil* **170**:87–97.
  36. **Zelles, L., and Q. Y. Bai.** 1994. Fatty acid patterns of phospholipids and lipopolysaccharides in environmental samples. *Chemosphere* **28**:391–411.