APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Jan. 2004, p. 240–247 Vol. 70, No. 1
0099-2240/04/$08.00+0 DOI: 10.1128/AEM.70.1.240–247.2004

New Molecular Screening Tools for Analysis of Free-Living Diazotrophs in Soil

Helmut Bürgmann,[1]* Franco Widmer,[2] William Von Sigler,[1] and Josef Zeyer[1]

Soil Biology, Institute of Terrestrial Ecology, Swiss Federal Institute of Technology Zurich, 8952 Schlieren,[1] and Swiss Federal Research Station for Agroecology and Agriculture (FAL-Reckenholz), 8046 Zurich,[2] Switzerland

Received 14 March 2003/Accepted 29 September 2003

Free-living nitrogen-fixing prokaryotes (diazotrophs) are ubiquitous in soil and are phylogenetically and physiologically highly diverse. Molecular methods based on universal PCR detection of the *nifH* marker gene have been successfully applied to describe diazotroph populations in the environment. However, the use of highly degenerate primers and low-stringency amplification conditions render these methods prone to amplification bias, while less degenerate primer sets will not amplify all *nifH* genes. We have developed a fixed-primer-site approach with six PCR protocols using less degenerate to nondegenerate primer sets that all amplify the same *nifH* fragment as a previously published PCR protocol for universal amplification. These protocols target different groups of diazotrophs and allowed for direct comparison of the PCR products by use of restriction fragment length polymorphism fingerprinting. The new protocols were optimized on DNA from 14 reference strains and were subsequently tested with bulk DNA extracts from six soils. These analyses revealed that the new PCR primer sets amplified *nifH* sequences that were not detected by the universal primer set. Furthermore, they were better suited to distinguish between diazotroph populations in the different soils. Because the novel primer sets were not specific for monophyletic groups of diazotrophs, they do not serve as an identification tool; however, they proved powerful as fingerprinting tools for subsets of soil diazotroph communities.

Free-living prokaryotes with the ability to fix atmospheric dinitrogen (diazotrophs) are ubiquitous in soil, but our knowledge of their ecological importance and their diversity remains incomplete. The capacity for nitrogen fixation is widespread among *Bacteria* and *Archaea* (Fig. 1A) (41, 43). The great diversity of diazotrophs also extends to their physiological characteristics, as N fixation is performed by chemotrophs and phototrophs and by autotrophs as well as heterotrophs (19, 25, 30, 41).

In natural ecosystems, biological N fixation (by free-living, associated, and symbiotic diazotrophs) is the most important source of N (11, 12, 30). The estimated contribution of free-living N-fixing prokaryotes to the N input of soils ranges from 0 to 60 kg ha^{-1} $year^{-1}$ (11, 13, 21, 30). The contribution of asymbiotic relative to symbiotic N fixation varies greatly, but in some terrestrial ecosystems asymbiotic N fixation may be the dominant source (11). The ability of free-living diazotrophs to take advantage of their capacity to perform N fixation depends on a number of conditions that vary for each organism, such as the availabilities of C and N and oxygen partial pressures (19). Because of the direct link of diazotroph populations to the C/N balance of a soil and their high diversity associated with different physiological properties, they are of interest as potential bioindicators for the N status of soils. Reliable tools for the description of diazotroph communities would contribute greatly to our understanding of the role diazotrophs play in the soil N cycle.

Due to the physiological diversity of diazotrophs and the documented unculturability of many prokaryotes (20, 29), cultivation-based strategies have severe limitations for the description of the diversity of free-living soil diazotrophs. Therefore, molecular approaches have been developed and successfully applied to describe diazotroph communities in different soil systems, including forest soils (33, 36, 37, 40), pasture, agricultural soils (33), wetland soils (10, 31), and rhizospheres (17, 24). The molecular approach to study the diversity of diazotroph organisms is primarily based on PCR amplification of a marker gene (*nifH*) for N fixation. This gene codes for the enzyme nitrogenase reductase (the enzyme code for *Azotobacter vinelandii* is EC 1.18.6.1.1g1m) (4) and has been shown to contain phylogenetic information (18, 28). The phylogeny of *nifH* is in general agreement with the small-subunit (SSU) rRNA gene-based phylogeny (Fig. 1A) (43), although alternative nitrogenase systems (e.g., vanadium-dependent nitrogenase reductase) and multiple copies of the gene within one genome (5, 43) indicate that the identification of diazotrophs based on *nifH* must be treated with some caution (41). Many previous studies have been based on PCR amplification using universal primers for *nifH* (33, 36, 40, 45). In order to achieve universal amplification, these primers were designed to target *nifH* gene regions encoding highly conserved amino acid sequences. However, the degeneracy of the genetic code introduces considerable variability to the DNA sequence of these regions, and therefore these primers are either highly degenerate (40, 44) or target smaller subsets of the diazotroph community in order to avoid high degeneracy (31, 32, 36). Widmer et al. (40) developed a nested PCR scheme to achieve specific amplification of *nifH* from bulk soil DNA with a highly degenerate universal primer set (nifH-univ primer set, Fig. 1B). Their protocol was successfully applied to the amplification of

* Corresponding author. Mailing address: Soil Biology, Institute of Terrestrial Ecology, Swiss Federal Institute of Technology Zurich, 8952 Schlieren, Switzerland. Phone: 41-1-633 6046. Fax: 41-1-633 1122. E-mail: helmut.buergmann@cnv.ethz.ch.

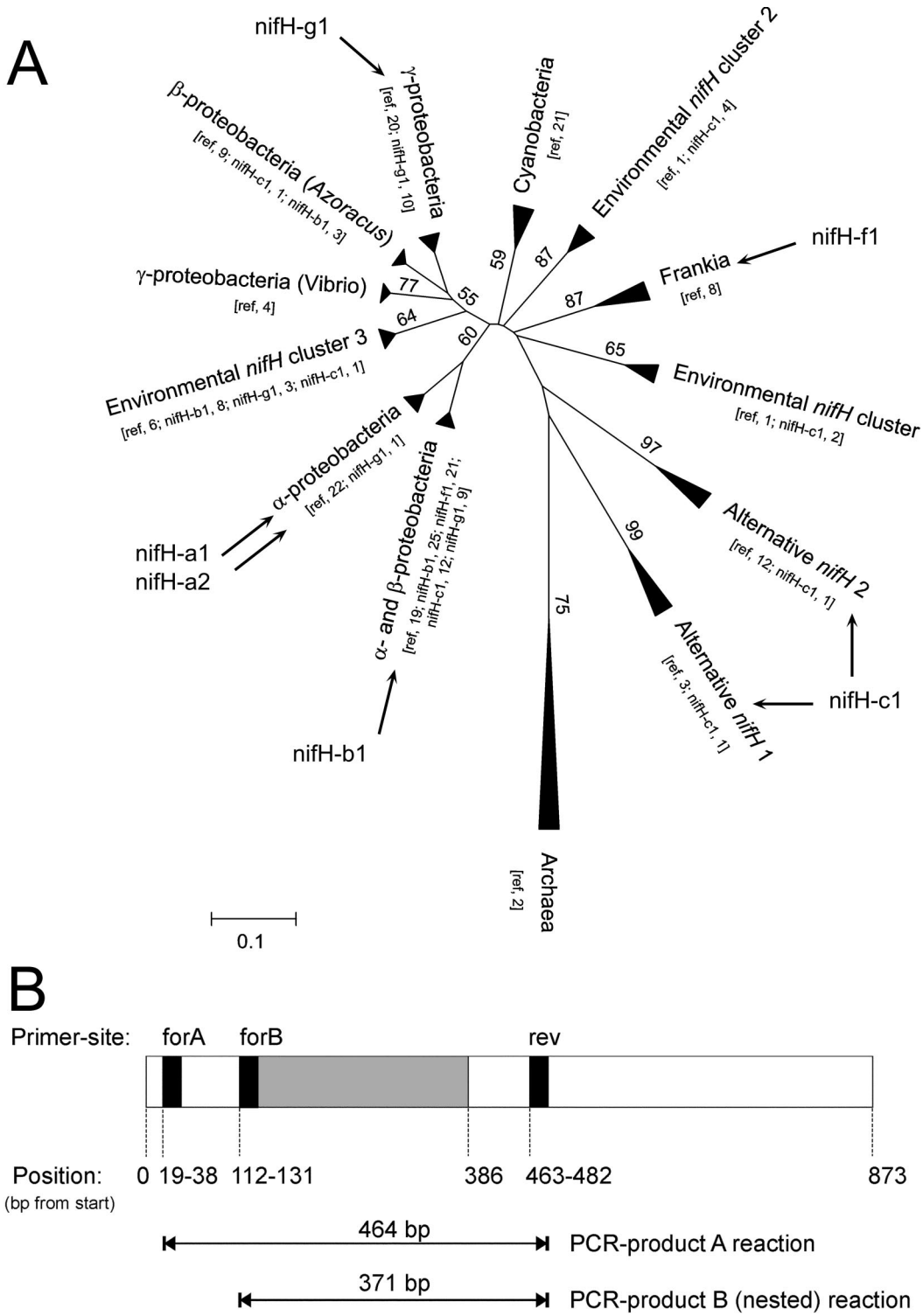

FIG. 1. (A) Overview of the NifH phylogeny inferred from an 85-amino-acid-residue fragment of protein sequences derived from 128 published *nifH* DNA sequences of cultivated organisms, selected environmental clones, and 103 *nifH* sequences cloned from bulk soil DNA in this study. The percentage of 300 bootstrap samplings which supported a cluster are indicated for values of more than 50%. Thirteen major clusters were collapsed for clarity. For each cluster the number of published reference sequences (ref) and sequences cloned with each primer set in this study are given in parentheses. The labeled arrows indicate the position of the target sequences used for the design of each primer set. (B) Schematic representation of the *nifH* gene, indicating the three primer sites (forA, forB, and rev) used in this study and resulting PCR products. The region that was used for phylogenetic analyses is marked in gray (amino acids 132 to 386). The positions refer to the *nifH* sequence of *A. vinelandii* (M20568).

TABLE 1. Description of soils used in this study

Soil[a]	Texture[c]	Land use	Total C (mg g [dry wt]$^{-1}$)	Total N (mg g [dry wt]$^{-1}$)	pH
HW	Silty loam[b]	Deciduous forest	410	31	5.5
HA	Silty clay[b]	Deciduous forest	49	5.5	6.8
AB	Loam[b]	Mixed forest	29	3.2	5.9
WI	Loamy sand[b]	Deciduous forest	26	2.7	4.6
PA	Sandy loam[b]	Agriculture (fallow)	16[b]	1.3[b]	7.5
GA	Loam[b]	Agriculture (fallow)	36	3.4	7.2

[a] Soils GA and PA are from the Rhône valley in southern Switzerland; all other soils are from northern Switzerland.
[b] From published data in references 14, 34, and 39.
[c] According to the U.S. Department of Agriculture soil taxonomy classifications.

nifH sequences from various soil samples and was shown to amplify *nifH* genes from diverse groups of organisms (9, 17, 37, 40). However, the use of highly degenerate primers combined with low-stringency amplification conditions may lead to biased results and may render the protocol less suited for studies with a more detailed and quantitative focus.

We studied the potential to use the phylogenetic information within the primer sites used by Widmer et al. (40) to construct complementary primers that allowed for a more focused and less biased amplification of subgroups of the diazotroph population in soil. We developed and evaluated new primer sets and optimized PCR protocols that all target the same fragment of *nifH* as the nifH-univ primer set. This fixed-primer-site approach allowed for direct comparative analysis of the PCR products with molecular fingerprinting techniques. The PCR protocols were tested and optimized on pure-culture DNA and subsequently were used to analyze soil DNA extracts.

MATERIALS AND METHODS

Soils. Soils were collected from four sites near Zurich and two sites in the upper Rhône valley of Switzerland (Table 1). At each site a block of soil (ca. 1 to 2 kg) was removed with a spade and the A horizons were separated and transported to the laboratory in sterile plastic bags. All soils were sieved (2.5-mm mesh size) and stored for up to 2 weeks at 10°C. The soils represented different ecological conditions: two seasonally wet forest soils (Hallwil [HW] and Abist [AB]), two well-drained forest soils (Hau [HA] and Winzlerboden [WI]), and two agricultural soils (Gartenacker [GA] and Pappelacker [PA]).

Total C and N contents of the soils were determined with a Leco CHNS-932 autoanalyzer (Leco, Düsseldorf, Germany). Soil pH was determined in 0.01 mM $MgCl_2$. Additional data were taken from the literature (14, 34, 39).

Development of group-specific *nifH* primers and PCR protocols. A database of 137 *nifH* DNA sequences of cultivated diazotrophs published in GenBank was assembled and manually aligned by using Bio-Edit version 5.0.9 (16). The DNA sequences of the primer sites (see Fig. 1B) defined by the nifH-univ primer set (40) were phylogentically analyzed by using the Kimura distance calculation (22) and the unweighted pair group method with arithmetic mean (UPGMA) clustering (27). New primer sets (Table 2) were designed by constructing consensus sequences from *nifH* genes of phylogenetically related organisms based on these cluster analyses. Primer set nifH-g1 targets *Azotobacter* spp. *nifH* sequences (GenBank accession numbers M11579, M20568, X13519, M73020, X03916). Set nifH-c1 targets *Clostridium* spp. *nifH* sequences (X07472, X07473, X07474, X07475, C07476, X07477, U59414). Set nifH-b1 targets *Herbaspirillum seropedicae nifH* sequences (U97121, Z54207). Set nifH-a2 targets *Azospirillum brasilense nifH* sequences (AF216882, M64344, X51500, X51609). nifH-f1 targets *Frankia* spp. *nifH* sequences (U78306, M21132, X57006, L41344, U53362, U53363 X73983, X17522, X76398). Set nifH-a1 targets a subset of *Rhizobiales* (J01781, L16503, M10587, M15941, M15942, M26961, V01215, Z95225, Z95226, Z95227, M55226, M55227, K01620, Z95230, M55229, M55231, M55232, Z95218, Z95219, Z95220, Z95229).

Universal nested PCR amplification of *nifH* was modified from Widmer et al. (40). The final PCR cocktails contained 1× reaction buffer (RPN0303X; Amersham Switzerland, Zürich, Switzerland), each deoxynucleoside triphosphate at a concentration of 200 μM, and each oligonucleotide primer at a concentration of 1 μM. For the first PCR (20 μl) we used 1 μl of DNA sample and bovine serum albumin at 5 mg ml^{-1} (SIGMA, Buchs, Switzerland). For the nested PCR (50 μl), 1 μl of the product from the first PCR was used along with bovine serum albumin at 0.3 mg ml^{-1}. After initial denaturation (5 min at 95°C), *Taq* polymerase (Life Technologies AG, Basel, Switzerland) in 5 μl of 1× reaction buffer was added to each reaction under hot-start conditions (80°C), resulting in a final *Taq* concentration of 0.02 U μl^{-1}. The cycling conditions used for the reactions were as described by Widmer et al. (40), except that annealing was performed for 8 s at 54°C and 30 s at 56°C for the first reaction and 8 s at 51°C and 30 s at 53°C for the nested reaction to reduce amplification of nonspecific by-products.

The specificity of the PCR protocols for the new primer sets was optimized on 14 diazotroph reference strains from several different phylogenetic groups. Tem-

TABLE 2. Sequences of primers used in this study

Name of primer set	Mismatched sequence[a] (forA/forB/rev/set)	Sequence (degeneracy)[b] of: forA site	forB site	rev site
nifH-univ[c]	NA[e]	GCIWTITAYGGNAARGGNGG[c] (128)	GGITGYGAYCCNAAVGCNGA[c] (96)	GCRTAIABNGCCATCATYTC[c] (48)
nifH-g1	-/0/2/0		GGTTGTGACCCGAAAGCTGA (0)	GCGTACATGGCCATCATCTC (0)
nifH-c1	-/0/4/0		GGWTGTGATCCWAARGCVGA (24)	GCATAYASKSCCATCATYTC (32)
nifH-b1	-/0/7/0		GGCTGCGATCCCAAGGCTGA (0)	GCGTACATGGCCATCATCTC (0)
nifH-a2	-/0/0/0		GGCTGCGATCCGAAGGCCGA (0)	GCGTAGAGCGCCATCATCTC (0)
nifH-f1	0/20/6/0	GCSTTCTACGGMAAGGGTGG (4)	GGBTGYGACCCSAASGCYGA (48)	GCGTACATSGCCATCATCTC (2)
nifH-a1	5/20/2/1[d]	GCRTTYTACGGYAARGGSGG (32)	GGMTGCGAYCCSAARGCSGA (32)	GCATAGAGCGCCATCATCTC (0)

[a] The number of nontarget sequences among the 137 investigated sequences that matched each primer (forA, forB, rev) or the primer set as a whole (set). Dashes indicate that no forA primer was used in this primer set.
[b] Degeneracy is indicated by standard conventions: K, G/T; M, A/C; R, A/G; S, C/G; W, A/T; Y, C/T; V, A/C/G; N, A/C/G/T; B, C/G/T.
[c] From Widmer et al. (40); I (inosine) is used to replace N at the 5′ portion of the degenerate primers.
[d] *Rhodospirillum rubrum* (M33774) matched all three primers of the set.
[e] NA, not applicable.

TABLE 3. Optimized PCR conditions for the different *nifH* primer sets

Name of primer set	Primer concn (μM)	Annealing temp. (°C)	$MgCl_2$ concn (mM)	No. of cycles
nifH-univ	2.0[a]/1.0[b]	56[a]/53[b]	2.0[a]/2.0[b]	30[a]/35[b]
nifH-g1	0.2[b]	60[b]	1.0[b]	50[b]
nifH-c1	0.8[b]	60[b]	1.4[b]	50[b]
nifH-b1	0.2[b]	60[b]	0.85[b]	50[b]
nifH-a2	0.2[b]	62[b]	0.8[b]	50[b]
nifH-f1	0.2[a]/0.2[b]	63[a]/64[b]	0.9[a]/0.9[b]	30[a]/35[b]
nifH-a1	0.2[a]/0.2[b]	60[a]/60[b]	1.2[a]/1.0[b]	30[a]/35[b]

[a] Values are for the first reaction of nested PCR protocols (using primer sites forA as forward primer and rev as reverse primer).
[b] Values are for direct or nested reaction (primer sites forB and rev).

plate concentrations were adjusted to yield a uniform amplification with bacterial SSU rRNA gene-specific PCR, using primers EUB338 (3) and uni-b-rev (6) with a previously described protocol (39). We adjusted $MgCl_2$ concentration, annealing temperature, cycle number, and primer concentration to achieve maximum specificity and sensitivity for each protocol (Table 3). Subsequently, optimized PCR protocols were applied to DNA extracted in duplicate from soil samples according to Bürgmann et al. (7) by processing 0.5 g of soil, 0.75 g of glass beads, and 1.25 ml of cetyltrimethylammonium bromide extraction buffer for 45 s at 5 m s^{-1} in a Fastprep bead beater (Bio101/Savant). Quality and quantity of DNA extracts and PCR products (6 μl) were analyzed by agarose gel electrophoresis.

Restriction fragment length polymorphism (RFLP) of *nifH* amplicons from soil. PCR products (45 μl) were precipitated and digested by using *Hae*III restriction endonuclease as described previously (40) in a total volume of 20 μl. In the case of weak initial amplification several reactions were pooled and concentrated by isopropanol precipitation. Digested PCR products were reprecipitated as described above and were resuspended in 10 μl of TE buffer (10 mM TRIS-HCl, 1 mM EDTA [pH 8]). Samples (5 μl) were electrophoresed in 12% acrylamide gels for 15 min at 30 V and for 3 h at 200 V at 35°C in a Dcode electrophoresis unit (Bio-Rad, Hercules, Calif.). Gels were stained with SybrGreen (Molecular Probes, Eugene, Oreg.) for 20 min and were analyzed with the GelDoc 2000 system (Bio-Rad). Bands and their respective fragment sizes were detected and quantified with the QuantityOne software (Bio-Rad). Band intensity values were standardized to the sum of band intensities in each lane. Cluster analysis of band intensity data was performed by using SysStat (Systat Software inc., Richmond, Calif.) by calculating euclidian distances and average linkage clustering.

Cloning and sequencing. PCR products were cloned by using the pGEM-T easy kit (Promega, Madison, Wis.) according to the manufacturer's instructions. Clone libraries were screened by transferring cells from white colonies to PCR tubes containing the appropriate *nifH* primers and PCR mix. PCR was performed as described above. PCR products from *nifH*-positive clones were subjected to *Hae*III RFLP profiling. At least one clone representing each restriction fragment pattern observed during screening was sequenced. Plasmids were prepared with the QIAprep Spin Miniprep kit (Qiagen, Basel, Switzerland). Sequencing was performed on both strands by using the BigDye Terminator Cycle Sequencing kit (Applied Biosystems, Rotkreuz, Switzerland). Sequences were determined on an ABI PRISM 310 Genetic Analyzer (Applied Biosystems).

Phylogenetic analyses. The cloned sequences were aligned with 128 *nifH* sequences of cultured organisms and environmental clones obtained from GenBank. Phylogenetic analyses were performed on a fragment that excluded both primer sites and additional bases at the 3′ end of the PCR product (Fig. 1B) to allow phylogenetic analysis with a maximum number of published reference sequences. The programs Protdist and Neighbor from the PHYLIP package (version 3.5c) (15) were used to perform Kimura-based distance matrix calculations (22) and UPGMA clustering (27). Seqboot and Consense from the same package were used to construct consensus cladograms based on 300 bootstrap samplings. The derived phylogenetic inference trees were edited in the program Tree-Explorer (23). Phylogenetically unusual *nifH* sequences were checked for chimeric characteristics by performing BLAST searches (2) with full-length and partial sequences.

Selected *nifH* sequences, including cloned and reference sequences, were subjected to in silico RFLP typing with the Webcutter tool (version 2.0; M. Heimann [http://www.firstmarket.com/cutter/cut2.htm]) to predict *Hae*III restriction fragment sizes of the amplified fragment.

Nucleotide sequence accession numbers. The 103 *nifH* sequences isolated from soils were submitted to GenBank and are available under accession numbers AY196363 to AY196465.

RESULTS

Primer design. New primer sets were designed by using the primer sites forA, forB, and rev defined by the nifH-univ primer set (40) (Fig. 1B). Phylogenetic analyses of the primer regions of 137 reference sequences indicated that new primer sets that target defined subgroups of *nifH* sequences can be designed for these sites (data not shown). Based on this analysis we identified six groups of sequences from phylogenetically related diazotrophs and designed a new primer set for each group (Table 2). Although the individual primers of a set were allowed to match nontarget sequences, the combination was specific for the selected target sequences. The only exception was for primer set nifH-a1 that also matched the *nifH* sequence of *Rhodospirillum rubrum* (M33774) in addition to the targeted *Rhizobium*-related sequences (Table 2). For the selected *Frankia* and *Rhizobium nifH* sequences direct PCR was not specific. Therefore, additional primers for the forA site (Fig. 1B) and nested PCR protocols were designed for the nifH-f1 and nifH-a1 primer sets (Table 2 and 3).

Development and optimization of PCR protocols. For each of the new *nifH* primer sets (Table 2) optimized amplification conditions (Table 3) were determined by use of DNA from selected diazotroph reference strains. All optimized PCR protocols revealed strong selectivity for *nifH* sequences from respective target reference strains, although some weak amplification from certain nontarget strains could not be entirely eliminated (Fig. 2A). More stringent protocols severely reduced amplification efficiency and were therefore not used (data not shown). The nifH-univ primer set did not amplify all reference strains.

Validation of new *nifH* PCR protocols on soil DNA extracts. Amplification of DNA extracts from six different soils (Table 1) with the nifH-univ PCR protocol and bacteria-specific SSU rRNA gene PCR resulted in strong PCR products of the expected size (*nifH*, approximately 370 bp; SSU rRNA, approximately 1,070 bp) for all soils (Fig. 2B). The amount of PCR product obtained with the new *nifH* primer sets varied among primer sets and soils. Primer sets nifH-g1, nifH-c1, nifH-b1, and nifH-f1 amplified fragments of the correct size from several different soils (Fig. 2B). However, primer sets nifH-a1 and nifH-a2 yielded only weak or no PCR product when applied to the soil DNA extracts (Fig. 2B). Due to the weak amplification products obtained from soils AB and WI and with primer sets nifH-a1 and nifH-a2, these samples were excluded from further analysis.

RFLP analysis. PCR products amplified from soils HW, HA, PA, and GA with primers nifH-g1, nifH-c1, nifH-b1, and nifH-f1 were subjected to *Hae*III RFLP analysis, which yielded highly reproducible fingerprints. Each primer set resulted in a different RFLP pattern for the same bulk soil DNA extract, as shown for soil PA in Fig. 3 (other data not shown).

The RFLP patterns obtained from the nifH-univ primer set were dominated by a number of poorly resolved bands between 170 and 185 bp, and some soils yielded a separate strong band at around 135 bp. Compared to the patterns from the new primer sets, the nifH-univ pattern contained a more intense

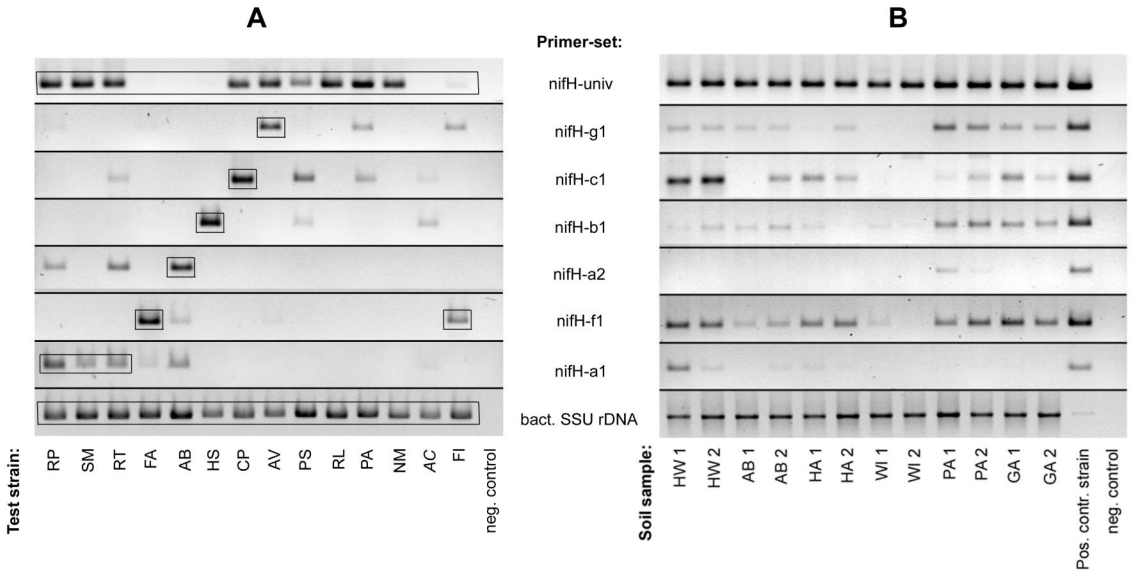

FIG. 2. Amplification of *nifH* genes using the six new primer sets developed in this study in comparison to universal *nifH* amplification and amplification of bacterial SSU rRNA genes. (A) PCR amplification from reference strain DNA. Agarose gel (2%) electrophoresis of PCR products (6 μl) amplified with different primer sets (rows) from DNA of 14 reference strains (columns). Only the regions of the gels that contained the approximately 370-bp *nifH* fragment amplified by the *nifH* PCR with the indicated primer set or the 1,070-bp fragment amplified by the bacterial (bact.) SSU rRNA gene amplification are shown. Boxed positions indicate expected targets according to DNA sequence analysis. RP, *Rhizobium phaseoli*; SM, *Sinorhizobium meliloti*; RT, *Rhizobium tropici*; FA, *Frankia alni (*strain Ag45/Mut15); AB, *Azospirillum brasilense* (DSM 1690); HS, *Herbaspirillum seropedicae* (DSM 6445); CP, *Clostridium pasteurianum* (DSM 525); AV, *Azotobacter vinelandii* (DSM 85); PS, *Pseudomonas stutzeri* (DSM 4166); RL, *Rhizobium leguminosarum* (DSM 30132); PA, *Paenibacillus azotofixans* (DSM 5976); NM, *Nostoc muscorum* (PCC 7120); AC, *Azoarcus communis*; FI, *F. alni* (strain I.3). (B) Amplification of soil DNA extracts. Agarose gel (2%) electrophoresis of PCR products (6 μl) amplified with the same protocols as in panel A from DNA extracted in duplicate from six soils (in the columns, replicate 1 and 2 are indicated).

smear of weaker, unresolved bands resulting in a higher baseline. The patterns obtained with the new primer sets differed from the nifH-univ patterns and from each other (Fig. 3). The patterns obtained with nifH-g1, for example, were dominated either by a single strong band of approximately 170 bp or by two distinct bands of 170 and 185 bp. The nifH-c1-derived patterns all contained a number of large (>300 bp) and small (<100 bp) fragments that were not present in any other pattern. nifH-f1 and nifH-b1 also yielded characteristic bands of smaller (50 to 90 bp) fragment sizes. It should be noted that bands that were dominant in some patterns, e.g., the strong band of approximately 135 bp in the nifH-univ pattern (Fig. 3), were absent from the patterns obtained with other primer sets, such as with nifH-c1 and nifH-f1.

Patterns obtained with the different primer sets furthermore indicated differences between the *nifH* gene pools of the studied soils, as revealed by cluster analysis of the RFLP data. Cluster analyses of the RFLP patterns derived from the nifH-univ amplification did not correctly group all samples: the replicates of the agricultural soils GA and PA did not group together, and in general the calculated distances were small (data not shown). In contrast, cluster analysis of the RFLP patterns obtained with nifH-g1 primers clearly indicated an improved differentiation between soils by grouping the patterns of all replicate soils and correctly separating the similar soils GA and PA (data not shown).

Sequence analysis. In silico RFLP analysis of DNA sequences of soil-derived clones indicated that the major bands of the respective RFLP patterns were represented by the isolated clones (data not shown). However, comparison of the results of phylogenetic analysis and in silico *Hae*III RFLP analysis of the sequenced soil clones and published sequences in our database revealed that single RFLP fragments were not indicative of phylogeny.

Phylogenetic analysis of 128 representative published *nifH*

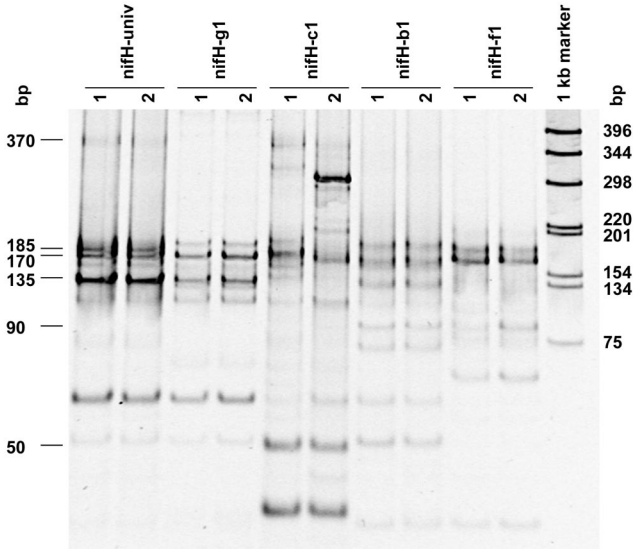

FIG. 3. *Hae*III RFLP of *nifH* PCR products obtained from duplicate DNA extracts (replicates 1 and 2) of PA soil, using nifH-univ primer set and newly developed primer sets nifH-g1, nifH-c1, nifH-b1, and nifH-f1. Numbers on the left indicate fragment sizes of important RFLP fragments discussed in the text, and numbers to the right indicate the fragment sizes of the 1-kb marker.

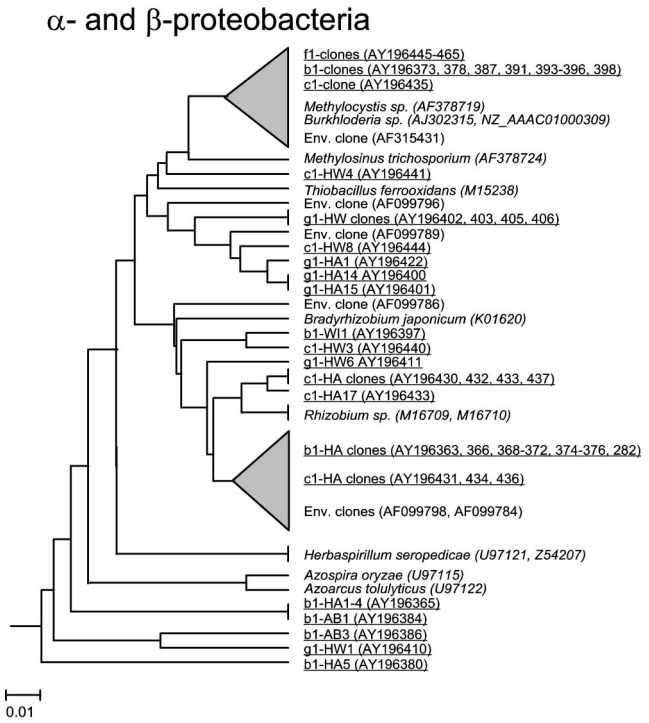

FIG. 4. Sub-tree of NifH sequences belonging to the α- and β-proteobacteria cluster (see Fig. 1). Details of the phylogenetic inference tree are based on Kimura distances and UPGMA clustering of an 85-amino-acid fragment of NifH (see Fig. 1). Sequences were derived from published *nifH* sequences of known organisms (in italics), environmental clones (plain text), and sequences cloned in this study (underlined). GenBank accession numbers of the *nifH* sequences are given in parentheses. Two subclusters of environmental clones were collapsed for clarity.

sequences and the 103 new *nifH* clones isolated in this study identified several branches in the dendrogram that could be phylogenetically classified as well as a number of branches that contained only environmental clones of yet-uncultivated organisms (Fig. 1A). The primer sets nifH-g1, nifH-c1, and nifH-b1 allowed the amplification of sequences closely related to their intended target sequences but also allowed the amplification of a number of clones belonging to other phylogenetic groups (Fig. 1) (target sequences for nifH-g1, γ-proteobacteria; for nifH-c1, alternative *nifH* 1 and 2; for nifH-b1, α- and β-proteobacteria). In the case of nifH-g1 all clones that clustered close to the target *nifH* sequences from *A. vinelandii* and *A. chroococcum* strains were obtained from soil PA, while other soils yielded clones that clustered with the α- and β-proteobacteria cluster or environmental *nifH* cluster 3. In addition to two sequences similar to that of clostridial *nifH*, the nifH-c1 set amplified a number of clones grouping with various proteobacterial branches and in addition was the only set that yielded clones that grouped with environmental clones in the branches environmental *nifH* clusters 1 and 2 (Fig. 1A). The nifH-b1 set amplified many clones that grouped loosely around the *Herbaspirillum* target sequence (Fig. 1A and 4). While these clones did not appear to be directly related to *Herbaspirillum*, most clustered with other β-proteobacteria in the α- and β-proteobacteria and β-proteobacteria (*Azoarcus*) clusters (Fig. 1A). Amplification with nifH-f1 primers yielded no sequences related to known *Frankia nifH* sequences. However, all 21 clones obtained from soils PA, HA, and HW were highly similar (>88% identity) and formed a tight cluster with two *nifH* sequences from the β-proteobacterial genus *Burkholderia* (Fig. 4). A number of clones obtained with various primer sets clustered with environmental clones in the environmental *nifH* cluster 3 (Fig. 1A).

DISCUSSION

Detection of *nifH* in environmental samples. PCR primer sets for *nifH* with a broad amplification range (26, 32, 36, 38, 40) or sets that target specific diazotroph groups or species (1, 35, 44) have been developed previously. However, these protocols amplify *nifH* fragments of variable length, which makes study-to-study comparisons difficult. Furthermore, because tests on pure cultures (38, 40) or environmental DNA (1, 26) are lacking, the range of *nifH* sequences amplified by many of these primer sets has not been thoroughly evaluated. In some cases tests on pure cultures and environmental DNA were performed, but environmental amplicons were not thoroughly characterized by DNA sequencing (32, 35).

Previously published studies of *nifH* sequences in soil indicate a great diversity and great variation of diazotroph community compositions in different soils. Poly et al. reported sequences closely related to the α- and β-proteobacteria cluster and one clone that clustered closest to our β-proteobacteria (*Azoarcus*) group (Fig. 1A, (33). Piceno et al., using a different primer set (31), found *nifH* sequences mostly related to γ-proteobacteria in the rhizosphere of smooth cordgrass (*Spartina alterniflora*) (24). Rösch et al., again using a different primer set, cloned 10 different *nifH* sequences related to the α- and β-proteobacteria and the α-proteobacteria clusters from acid forest soils (36). The nifH-univ primer set that was also used in the present study was applied to study soil diazotroph communities in several previous studies (17, 37, 40). Widmer et al. (40) showed that *nifH* sequences derived from forest soil and litter clustered with the α- and β-proteobacteria cluster, α-proteobacteria cluster, and environmental *nifH* cluster 1 (Fig. 1A). Hamelin et al. (17) isolated a large number of clones from the rhizosphere of the grass *Molina coerulea* that clustered with environmental *nifH* cluster 3, environmental *nifH* cluster 2, and the alternative *nifH* clusters 1 and 2 (Fig. 1A). These results indicated that the nifH-univ primer set amplified phylogenetically diverse *nifH* sequences from a broad range of diazotrophs. However, the results presented in this study showed that this primer set did not amplify all diazotroph reference strains, even though the primers for detecting those strains were present in the nifH-univ PCR (Fig. 2A). Others have also found that the amplification range of many primer sets is lower than theoretically expected (32). This may be due to reduced amplification efficiency of certain primer combinations of the degenerate primer sets.

Due to certain limitations, the specificity of PCR amplification from environmental samples cannot be deduced from tests on reference strains or database searches. First, the subset of *nifH* sequences available for primer design might not be sufficient to adequately assess primer specificity. Second, some weak nonspecific amplification (Fig. 2A) could result in coamplification of environmental nontarget templates. Third, in

contrast to the constant concentration of pure-strain DNA used for PCR optimization, the relative abundances of different *nifH* target sequences in bulk soil DNA extracts may vary over several orders of magnitude.

Improved *nifH* detection with new homologous primer sets. In this study we have used a fixed-primer-site approach to develop six new primer sets and optimized PCR protocols designed to amplify a homologous fragment of the *nifH* gene from different subgroups of diazotrophs (Fig. 1A and Table 2 and 3). The homologous nature of the fragments allowed us to directly compare PCR products obtained with the different primer sets by means of various downstream methods, such as molecular fingerprinting with RFLP (Fig. 3) and sequencing (Fig. 1A and 4). When used in combination the new primer sets allowed for a greatly improved coverage of the diazotroph diversity compared to that of the nifH-univ primer set (40), which was unable to amplify *nifH* from certain cultures that were amplified by the nifH-g1, nifH-b1, and nifH-f1 primer sets (Fig. 2A).

Four of the specific amplification protocols resulted in a strong amplification of *nifH* genes from the investigated soils (Fig. 2B). For two of the soils no sufficiently strong amplification was achieved with any of the specific primer sets, despite successful amplification with the nifH-univ primer set. This indicated that the nifH-univ protocol amplified a considerable portion of the overall diazotroph community that was not amplified by any of the specific protocols. This was expected because the new primer sets do not represent the full extent of the degeneracy contained in nifH-univ. However, combined use of the new primer sets and the nifH-univ primer set greatly increases the observable diazotroph diversity. Apart from their use as fingerprinting tools for diazotroph communities, the new protocols are also useful for studies on pure cultures. This was recently demonstrated in a study on *nifH* mRNA expression in soil using the nifH-g1 primer set (8).

***nifH* gene pools detected with new primer sets.** The results of RFLP pattern analyses indicated that each new primer set amplified a different subset of phylotypes from the total *nifH* gene pool and successfully discriminated against other, possibly more abundant phylotypes (Fig. 3). Cluster analyses of the RFLP data showed that the amplification of a subset of the total population resulted in improved differentiation between soils. The limited capacity for sample differentiation of PCR protocols targeting broad phylogenetic groups versus those focusing on narrower groups is frequently observed in comparisons of 16S ribosomal DNA fingerprints of the bacterial domain versus class- or genus-level fingerprints (6). Sequence analyses confirmed the results of RFLP as each of the new primer sets amplified different subsets of *nifH* sequences (Figs. 1A and 4), which makes them very useful for fingerprinting diazotroph communities. However, while the primers did preferentially amplify *nifH* sequences related to their intended target pure cultures (Fig. 1A and 2A), environmental conditions (e.g., for nifH-g1 in soils HW and HA or for nifH-c1 and nifH-f1 in all soils) resulted in the amplification of a number of nontarget sequences as well, probably due to a low abundance of target sequences. In the case of the nifH-f1 primer set designed to amplify *nifH* sequences of *Frankia* spp., our protocol did not yield *nifH* sequences from the target phylum (Fig. 1A) at all. However, its demonstrated capability to amplify a narrow subgroup of the α- and β-proteobacterial cluster (Fig. 4) warrants further investigation of the group specificity of this primer set as deduced from the published sequences and *Frankia* cultures.

A large number of apparently nontarget *nifH* sequences amplified with various primer sets grouped within the α- and β-proteobacteria cluster (Fig. 1A and 4) and environmental *nifH* cluster 3 (Fig. 1A). We assume that a high abundance of these sequences in soil is responsible for their frequent detection by the new PCR protocols. Sequences belonging to *nifH* cluster 3 were found to be abundant in grass rhizosphere (Cluster A in Hamelin et al. [17]) and were found in other environmental clone libraries (38, 42). We have identified an *nifH* homolog in a published genomic sequence of *Geobacter metallireducens* (AAAS01000020), which clusters with these sequences. To our knowledge, this is the first δ-proteobacterial *nifH* gene sequence that does not cluster with the alternative *nifH* cluster 1 or 2 (Fig. 1A). Whether the environmental *nifH* cluster 3 includes other δ-proteobacterial *nifH* sequences remains open until more sequence data from cultured diazotroph δ-proteobacteria become available.

Conclusions. The new primer sets developed in this study can be used to supplement the nifH-univ primer set to target groups of interest with higher resolution. As for the nifH-univ universal primer set, these primer sets clearly cannot be used as direct diagnostic tools for the identification of specific diazotrophs in soil or to deduce a quantitative measure of diazotroph community composition. However, reduction of the complexity and direct comparability of genetic fingerprints obtained with the new primer sets makes them more powerful for detection of differences or changes in diazotroph communities in environmental samples.

ACKNOWLEDGMENTS

This research was supported by COST Action 831 "Biotechnology of soil: monitoring, conservation and remediation" grant C99.0036 and the Indo-Swiss Collaboration in Biotechnology (project SA 5).

René Schaubhut is acknowledged for his help with primer design and technical assistance. Angela Sessitsch (ARC Seibersdorf Research GmbH) is acknowledged for providing *Rhizobium* reference strains.

REFERENCES

1. **Aguilar, O. M., D. H. Grasso, P. M. Riccillo, M. V. Lopez, and E. Szafer.** 1998. Rapid identification of bean *Rhizobium* isolates by a *nifH* gene PCR assay. Soil Biol. Biochem. **30:**1655–1661.
2. **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.
3. **Amann, R. I., W. Ludwig, and K. Schleifer.** 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiol. Rev. **59:**143–169.
4. **Bairoch, A.** 1993. The ENZYME data bank. Nucleic Acids Res. **21:**3155–3156.
5. **Bishop, P. E., and R. Premakumar.** 1992. Alternative nitrogen fixation systems, p. 736–762. *In* G. Stacey, R. H. Burris, and H. J. Evans (ed.), Biological nitrogen fixation. Chapman & Hall, New York, N.Y.
6. **Bundt, M., F. Widmer, M. Pesaro, J. Zeyer, and P. Blaser.** 2001. Preferential flow paths: biological 'hot spots' in soils. Soil Biol. Biochem. **33:**729–738.
7. **Bürgmann, H., M. Pesaro, F. Widmer, and J. Zeyer.** 2001. A strategy for optimizing quality and quantity of DNA extracted from soil. J. Microbiol. Methods **45:**7–20.
8. **Bürgmann, H., F. Widmer, W. V. Sigler, and J. Zeyer.** 2003. mRNA extraction and reverse transcription-PCR protocol for detection of *nifH* gene expression of *Azotobacter vinelandii* in soil. Appl. Environ. Microbiol. **69:** 1928–1935.
9. **Burke, D. J., E. P. Hamerlynck, and D. Hahn.** 2002. Interactions among plant species and microorganisms in salt marsh sediments. Appl. Environ. Microbiol. **68:**1157–1164.
10. **Chelius, M. K. L. J. E.** 1999. Restriction fragment length polymorphism

analysis of PCR-amplified *nifH* sequences from wetland plant rhizosphere communities. Environ. Technol. **20:**883–889.
11. **Cleveland, C. C., A. R. Townsend, D. S. Schimel, H. Fisher, R. W. Howarth, L. O. Hedin, S. S. Perakis, E. F. Latty, J. C. Von Fischer, A. Elseroad, and M. F. Wasson.** 1999. Global patterns of terrestrial biological nitrogen (N_2) fixation in natural ecosystems. Glob. Biogeochem. Cycle **13:**623–645.
12. **Dart, P. J., and S. P. Wani.** 1982. Non-symbiotic nitrogen fixation and soil fertility, p. 3–27. 12th International Congress of Soil Science: managing soil resources to meet the challenges of mankind, vol. I. Indian Society of Soil Science, New Delhi, India.
13. **Day, J. M., D. Harris, P. J. Dart, and P. Van Berkum.** 1975. The Broadbalk experiment. An investigation of nitrogen gains from non-symbiotic nitrogen fixation, p. 71–84. *In* W. D. P. Stewart (ed.), Nitrogen fixation by free-living micro-organisms, vol. 6. Cambridge University Press, London, United Kingdom.
14. **Favre, C.** 1982. Contribution à la connaissance des sols de la plaine du Rhône du valais central. Terre Valaisanne 30 Janvier**:**25–26.
15. **Felsenstein, J.** 1989. PHYLIP – phylogeny inference package (version 3.2). Cladistics **5:**164–166.
16. **Hall, T. A.** 1999. Bio-Edit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. **41:**95–98.
17. **Hamelin, J., N. Fromin, S. Tarnawski, S. Teyssier-Cuvelle, and M. Aragno.** 2002. *nifH* gene diversity in the bacterial community associated with the rhizosphere of *Molinia coerulea*, an oligonitrophilic perennial grass. Environ. Microbiol. **4:**477–481.
18. **Hennecke, H., K. Kaluza, B. Thony, M. Fuhrmann, W. Ludwig, and E. Stackebrandt.** 1985. Concurrent evolution of nitrogenase genes and 16S ribosomal-RNA in *Rhizobium* species and other nitrogen-fixing Bacteria. Arch. Microbiol. **142:**342–348.
19. **Hill, S.** 1992. Physiology of nitrogen fixation in free-living heterotrophs, p. 87–134. *In* G. Stacey, R. H. Burris, and H. J. Evans (ed.), Biological nitrogen fixation. Chapman & Hall, New York, N.Y.
20. **Hugenholtz, P., B. M. Goebel, and N. R. Pace.** 1998. Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. J. Bacteriol. **180:**4765–4774.
21. **Kennedy, I. R., and N. Islam.** 2001. The current and potential contribution of asymbiotic nitrogen fixation to nitrogen requirements on farms: a review. Aust. J. Exp. Agric. **41:**447–457.
22. **Kimura, M.** 1983. The neutral theory of molecular evolution. Cambridge University Press, Cambridge, United Kingdom.
23. **Kumar, S., K. Tamura, I. B. Jakobsen, and M. Nei.** 2001. MEGA2: molecular evolutionary genetics analysis software. Bioinformatics **17:**1244–1245.
24. **Lovell, C. R., Y. M. Piceno, J. M. Quattro, and C. E. Bagwell.** 2000. Molecular analysis of diazotroph diversity in the rhizosphere of the smooth cordgrass, *Spartina alterniflora*. Appl. Environ. Microbiol. **66:**3814–3822.
25. **Malik, K. A., and H. G. Schlegel.** 1981. Chemolithoautotrophic growth of bacteria able to grow under N_2-fixing conditions. FEMS Microbiol. Lett. **11:**63–67.
26. **Marusina, A. I., E. S. Boulygina, B. B. Kuznetsov, T. P. Tourova, I. K. Kravchenko, and V. F. Gal'chenko.** 2001. A system of oligonucleotide primers for the amplification of *nifH* genes of different taxonomic groups of prokaryotes. Microbiology **70:**73–78.
27. **Michener, C. D., and R. R. Sokal.** 1957. A quantitative approach to a problem in classification. Evolution **11:**130–162.
28. **Normand, P., and J. Bousquet.** 1989. Phylogeny of nitrogenase sequences in *Frankia* and other nitrogen-fixing microorganisms. J. Mol. Evol. **29:**436–447.
29. **Pace, N. R.** 1996. New perspective on the natural microbial world: molecular microbial ecology. ASM News **62:**463–470.
30. **Paul, E. A., and F. E. Clark.** 1996. Soil microbiology and biochemistry, 2nd ed. Academic Press, San Diego, Calif.
31. **Piceno, Y. M., P. A. Noble, and C. R. Lovell.** 1999. Spatial and temporal assessment of diazotroph assemblage composition in vegetated salt marsh sediments using denaturing gradient gel electrophoresis analysis. Microb. Ecol. **38:**157–167.
32. **Poly, F., L. J. Monrozier, and R. Bally.** 2001. Improvement in the RFLP procedure for studying the diversity of *nifH* genes in communities of nitrogen fixers in soil. Res. Microbiol. **152:**95–103.
33. **Poly, F., L. Ranjard, S. Nazaret, F. Gourbiere, and L. J. Monrozier.** 2001. Comparison of *nifH* gene pools in soils and soil microenvironments with contrasting properties. Appl. Environ. Microbiol. **67:**2255–2262.
34. **Richard, F., and P. Lüscher.** 1983. Physikalische Eigenschaften von Böden der Schweiz, Band 3, vol. 3. Professur für Bodenphysik-ETH Zürich; Eidgenössische Anstalt für das forstliche Versuchswesen, Zürich, Birmensdorf, Switzerland.
35. **Rosado, A. S., G. F. Duarte, L. Seldin, and J. D. Van Elsas.** 1998. Genetic diversity of *nifH* gene sequences in *Paenibacillus azotofixans* strains and soil samples analyzed by denaturing gradient gel electrophoresis of PCR-amplified gene fragments. Appl. Environ. Microbiol. **64:**2770–2779.
36. **Rösch, C., A. Mergel, and H. Bothe.** 2002. Biodiversity of denitrifying and dinitrogen-fixing bacteria in an acid forest soil. Appl. Environ. Microbiol. **68:**3818–3829.
37. **Shaffer, B. T., F. Widmer, L. A. Porteous, and R. J. Seidler.** 2000. Temporal and spatial distribution of the *nifH* gene of N_2 fixing bacteria in forests and clearcuts in western Oregon. Microb. Ecol. **39:**12–21.
38. **Ueda, T., Y. Suga, N. Yahiro, and T. Matsuguchi.** 1995. Remarkable N_2-fixing bacterial diversity detected in rice roots by molecular evolutionary analysis of *nifH* gene sequences. J. Bacteriol. **177:**1414–1417.
39. **Widmer, F., A. Fliessbach, E. Laczko, J. Schulze-Aurich, and J. Zeyer.** 2001. Assessing soil biological characteristics: a comparison of bulk soil community DNA-, PLFA-, and Biolog-analyses. Soil Biol. Biochem. **33:**1029–1036.
40. **Widmer, F., B. T. Shaffer, L. A. Porteous, and R. J. Seidler.** 1999. Analysis of *nifH* gene pool complexity in soil and litter at a Douglas fir forest site in the Oregon Cascade Mountain Range. Appl. Environ. Microbiol. **65:**374–380.
41. **Young, J. P. W.** 1992. Phylogenetic classification of nitrogen-fixing organisms, p. 43–86. *In* G. Stacey, R. H. Burris, and H. J. Evans (ed.), Biological nitrogen fixation. Chapman & Hall, New York, N.Y.
42. **Zani, S., M. T. Mellon, J. L. Collier, and J. P. Zehr.** 2000. Expression of *nifH* genes in natural microbial assemblages in Lake George, New York, detected by reverse transcriptase PCR. Appl. Environ. Microbiol. **66:**3119–3124.
43. **Zehr, J. P., B. D. Jenkins, S. M. Short, and G. F. Steward.** 2003. Nitrogenase gene diversity and microbial community structure: a cross-system comparison. Environ. Microbiol. **5:**539–554.
44. **Zehr, J. P., and L. A. McReynolds.** 1989. Use of degenerate oligonucleotides for amplification of the *nifH* gene from the marine cyanobacterium *Trichodesmium thiebautii*. Appl. Environ. Microbiol. **55:**2522–2526.
45. **Zehr, J. P., M. T. Mellon, and S. Zani.** 1998. New nitrogen-fixing microorganisms detected in oligotrophic oceans by amplification of nitrogenase (*nifH*) genes. Appl. Environ. Microbiol. **64:**3444–3450.