Applied and Environmental Microbiology, July 1997, p. 2802–2813 Vol. 63, No. 7
0099-2240/97/$04.00+0
Copyright © 1997, American Society for Microbiology

Molecular Microbial Diversity of an Anaerobic Digestor as Determined by Small-Subunit rDNA Sequence Analysis

JEAN-JACQUES GODON,* EMMANUELLE ZUMSTEIN, PATRICK DABERT,
FRÉDÉRIC HABOUZIT, and RENÉ MOLETTA

Laboratoire de Biotechnologie de l'Environnement, Institut National de la Recherche Agronomique, 11100 Narbonne, France

Received 27 November 1996/Accepted 16 April 1997

The bacterial community structure of a fluidized-bed reactor fed by vinasses (wine distillation waste) was analyzed. After PCR amplification, four small-subunit (SSU) rDNA clone libraries of *Bacteria*, *Archaea*, *Procarya*, and *Eucarya* populations were established. The community structure was determined by operational taxonomic unit (OTU) phylogenetic analyses of 579 partial rDNA sequences (about 500 bp long). A total of 146 OTUs were found, comprising 133, 6, and 7 from the *Bacteria*, *Archaea*, and *Eucarya* domains, respectively. A total of 117 bacterial OTUs were affiliated with major phyla: low-G+C gram-positive bacteria, *Cytophaga-Flexibacter-Bacteroides*, *Proteobacteria*, high-G+C gram-positive bacteria, and *Spirochaetes*, where the clone distribution was 34, 26, 17, 6, and 4%, respectively. The other 16 bacterial OTUs represent 13% of the clones. They were either affiliated with narrow phyla such as *Planctomyces-Chlamydia*, green nonsulfur bacteria, or *Synergistes*, or deeply branched on the phylogenetic tree. A large number of bacterial OTUs are not closely related to any other hitherto determined sequences. The most frequent bacterial OTU represents less than 5% of the total bacterial SSU rDNA sequences. However, the 20 more frequent bacterial OTUs describe at least 50% of these sequences. Three of the six *Archaea* OTUs correspond to 95% of the *Archaea* population and are very similar to already known methanogenic species: *Methanosarcina barkeri*, *Methanosarcina frisius*, and *Methanobacterium formicicum*. In contrast, the three other *Archaea* OTUs are unusual and are related to thermophilic microorganisms such as *Crenarchaea* or *Thermoplasma* spp. Five percent of the sequences analyzed were chimeras and were removed from the analysis.

The anaerobic microbial degradation of organic matter to methane and carbon dioxide occurs naturally in a variety of anaerobic habitats such as intestinal tracts and sediments. Anaerobic digestion is exploited on a large scale as a simple and effective biotechnological process to reduce the pollution caused by organic wastes (2). The failure of many anaerobic digestors to operate reliably and with constant performance has underlined the need for more basic information on the biological aspects of the anaerobic digestion ecosystem. However, to date there has been no comprehensive analysis of the microbial species composition in terms of species present in any anaerobic digestor.

The use of molecular biological techniques, especially those that take advantage of the small-subunit (SSU) rRNA molecule, has eliminated the dependence on isolation of pure cultures as a means of studying the diversity and structure of microbial communities (1). The microbial consortia in soil (5, 28, 30), blanket bog peat (13), mine galleries (3), marine picoplankton (8, 11, 12), marine sediments (9), hydrothermal vents (18), human colonic biota (33), termite guts (20, 22), biodegraded wall paintings (23), oil fields (31), activated sludge (4), and anaerobic digestors (19) have been analyzed by this approach. However, most of these ecosystems are open-field systems, and only one sample was analyzed. Even with a representative sample, it is difficult to distinguish between endogenous and transient microorganisms. Others are closed ecosystems, but the organisms analyzed span either a small number of clones or only one taxon (9, 22).

An anaerobic digestor has the advantage of being a closed and comparatively homogeneous and stable ecosystem (27). Since little is known about this ecosystem, a molecular inventory is the first step to describe this dynamic microbial consortium without cultivation. In this paper, we provide an analysis of the microbial species diversity present in the biofilm of a fluidized-bed anaerobic digestor.

MATERIALS AND METHODS

Operation of the fluidized-bed digestor. The experimental apparatus consisted of a tubular reactor (0.6 m long, 0.1 m in diameter) on polyvinyl chloride with a working volume of 0.625 liter. The support material used for microbial fixation was granular pozzolana (powder of porous volcanic stone), with a diameter ranging from 315 to 500 μm. Fluidization was ensured by a circulation pump to produce an expansion rate of 25%. The reactor was fed by a peristaltic pump. The substrate was vinasses (wastes of wine distillation). The vinasses were collected at 80°C from the distillery and frozen at −20°C. The same stock of vinasses was used for all the experiments. The average composition was 10 g of total organic carbon (TOC)/liter and 25 g of chemical oxygen demand/liter. The organic load was 5.85 g of chemical oxygen demand/liter/day. The pH and temperature in the liquid phase were measured on line. The pH was maintained around 7.0 by neutralizing the influent with sodium hydroxide. The temperature was maintained at 35°C with a water jacket. The TOC in liquid phase and volatile fatty acids were analyzed twice a week by using a Dohrmann Carbon Analyzer and gas chromatography, respectively. The average measures were 800 mg/liter for TOC and 250 mg/liter for volatile fatty acids.

The inoculum represented 100% of the volume and came directly from a similar 15-liter fluidized-bed reactor which had previously been inoculated with anaerobic sludge from the lagoon of Narbonne's distillery and which had run continuously for 1 year (7). Under these conditions, the digestor became almost immediately stable, and the sample used for analysis was collected 40 days after inoculation.

Extraction and purification of total genomic DNA. A 4-ml sample was collected from the middle of the fluidized-bed reactor. This sample was concentrated by centrifugation at 4,000 × *g* for 10 min and resuspended in 4 ml of 4 M guanidine thiocyanate–0.1 M Tris (pH 7.5) and 600 μl of 10% *N*-lauroyl sarcosine. The pozzolana was ground with a mortar on ice, 250 μl of the ground material was transferred to a 2-ml screw-cap polypropylene microcentrifuge tube, and the remaining material was frozen. After the addition of 500 μl of 5%

* Corresponding author. Mailing address: Laboratoire de Biotechnologie de l'Environnement, INRA, Avenue des Etangs, 11100, Narbonne, France. Phone: 33 (0)4 68 42 51 54. Fax: 33 (0)4 68 42 51 60. E-mail: godon@ensam.inra.fr.

TABLE 1. Libraries of clones analyzed and distribution of clones among phylogenetic domains

SSU rDNA targeted	Library names[a]	Primers[b]	No. of clones analyzed	No. of clones sequenced	Domain distribution			
					Archaea	*Bacteria*	*Eucarya*	Chimera
Archaea total	CA, CB DA, DC	w003 + w002 w017 + w002	102	69	39	29	1	0
Bacteria total	BA, BB, BC, BD HA, HB	w001 + w002 w018 + w002	392	392	2	367	0	23
Procarya partial	IA	w024 + w002	88	88	23	62	0	3
Eucarya partial	EA	w016 + w002	96	30	2	1	27	0
Total clones			678	579	66	459	28	26
Total OTUs				146	6	133	7	

[a] The first letter corresponds to the set of primers used, and the second letter corresponds to an independent PCR amplification event.
[b] For primer sequences, see Table 2.

N-lauroyl sarcosine–0.1 M phosphate buffer (pH 8.0), the 2-ml tube was incubated at 70°C for 1 h. One volume (750 μl) of 0.1-mm-diameter silica beads (Sigma) previously sterilized by autoclaving was added, and the tube was shaken at maximum speed for 10 min in a Vibro shaker (Retsch). Polyvinylpolypyrrolidone (15 mg) was added to the tube, which was vortexed and centrifuged for 3 min at 12,000 × *g*. After recovery of the supernatant, the pellet was washed with 500 μl of TENP (50 mM Tris [pH 8], 20 mM EDTA [pH 8], 100 mM NaCl, 1% polyvinylpolypyrrolidone) and centrifuged for 3 min at 12,000 × *g*, and the new supernatant was added to the first supernatant. The washing step was repeated three times. The pooled supernatants (about 2 ml) were briefly centrifuged to remove particles and then split into two 2-ml tubes.

Nucleic acids were precipitated by the addition of 1 volume of isopropanol for 10 min at room temperature and centrifuged for 15 min at 20,000 × *g*. Pellets were resuspended and pooled in 450 μl of 100 mM phosphate buffer, pH 8, and 50 μl of 5 M potassium acetate. The tube was placed on ice for 90 min and centrifuged at 16,000 × *g* for 30 min. The supernatant was transferred to a new tube containing 20 μl of RNase (1 mg/ml) and incubated at 37°C for 30 min. Nucleic acids were precipitated by the addition of 50 μl of 3 M sodium acetate and 1 ml of absolute ethanol. The tube was incubated for 10 min at room temperature, and nucleic acids were recovered by centrifugation at 20,000 × *g* for 15 min. The DNA pellet was finally washed with 70% ethanol, dried, and resuspended in 400 μl of TE buffer. The DNA concentration (0.2 μg/μl) and size distribution (predominantly around 20 kb) were estimated by electrophoresis (data not shown).

Amplification, cloning, screening, and sequencing of SSU rDNA. Amplification of SSU rDNA genes from purified genomic DNA from a sample was carried out with primers for conserved domains. Six rDNA gene libraries were prepared and named B, C, D, E, H, and I according to the primers used (Tables 1 and 2). Group-specific primers were designed from alignment of the total SSU rRNA sequences available in the Ribosome Database Project (RDP, May 1995 update [16]) and from classical contingencies (secondary structures, GC content, etc.). *Archaea* primer w003 used in the first *Archaea* library presented a lack of specificity and was replaced by w017. In primer w003, the 3′-end bases do not discriminate between the *Archaea* and *Bacteria* domains, whereas in w017 the last two 3′-end bases do so (A/G-G instead of T-C in the *Bacteria* domain). Each reaction tube contained 0.2 μg of each primer (Table 2), 0.2 μg of purified template DNA, 1× Taq reaction buffer (Perkin-Elmer, Foster City, Calif.), 2.5 mM $MgCl_2$, 22 μM (each) deoxynucleoside triphosphate, and 1 U of *Taq* DNA polymerase (Perkin-Elmer), adjusted to a total volume of 100 μl. The reaction mixture was prepared on ice, covered with mineral oil, and placed in a thermocycler (Perkin-Elmer) at 94°C. After an initial denaturation at 94°C for 2 min, 25 temperature cycles were performed at either 50°C (libraries B, E, H, and I) or 55°C (libraries C and D) for 1 min, 72°C for 1 min, and 94°C for 1 min. The products were electrophoresed on a 0.7% agarose gel and viewed by ethidium bromide staining (data not shown). Bands of the proper size range (ca. 1,500 bases) were excised and eluted with a Qiaex II gel extraction kit (Qiagen, Hilden, Germany). The purified products were ligated into the pGEMt plasmid (Promega, Madison, Wis.). The ligation products were transformed into *Escherichia coli* TG1 competent cells with ampicillin selection and blue/white screening (25). Plasmid inserts were amplified by PCR with pGEMt primers T7 and SP6. The inserts were screened by comparing *Hae*III restriction endonuclease cleavage patterns. Only plasmids with different restriction patterns were further analyzed by DNA sequencing. Plasmid preparations for DNA sequencing were made with microcolumns as specified by the manufacturer (Qiagen). The nucleotide sequences of plasmid inserts were determined by automated DNA sequencing by using the dideoxy chain-termination method (26) and the ABI model 373A sequencer stretch (Applied Biosystems, Perkin-Elmer). Plasmid DNAs were sequenced with the dye-terminator cycle-sequencing ready-reaction kit, with AmpliTaq DNA polymerase FS kit buffer (Perkin-Elmer) and either the w015 or w024 SSU rDNA primer (Table 2). In a few cases, where these primers gave poor results, the backward primer w004 was used (Table 2). A partial sequence of at least 500 bp was performed for each clone.

Sequence analysis. An equal portion (about 500 bp) of SSU rDNA (*E. coli* positions 812 to 1307 [6]), was used for sequence analysis. Each sequence was compared with sequences available in databases (GenBank and RDP) (16). DNA sequence comparisons between samples were performed with Lasergene software (Dnastar, Madison, Wis.). Phylogenetic trees were calculated by Jukes-Cantor (14) and neighbor-joining (24) algorithms with the sequence alignment software VSM developed by R. Christen (Centre National de la Recherche Scientifique, Villefranche sur Mer, France). The presence of chimeric sequences was checked with the RDP Chimera Check program (16) and by comparing independently the alignments at the beginning of each sequence (*E. coli* positions 812 to 890 [6]) and at the end of each sequence (*E. coli* positions 1181 to 1307 [6]) and the alignments of the entire sequence (*E. coli* positions 812 to 1307 [6]).

Nomenclature. The first letter of each clone name (e.g., BA23) corresponds to the set of primers used in its amplification (Table 1); the second letter corresponds to an independent PCR amplification event, and the numeral corresponds to the clone number of the respective library. Clones with more than 96% sequence similarity were grouped into the same operational taxonomic unit (17). Each OTU was usually named according to the first clone identified that had the representative sequence.

Nucleotide sequence accession numbers. The nucleotide sequence data reported in this paper will appear in the EMBL, GenBank and DDBJ nucleotide sequence databases under accession no. U81640 to U81785 with the generic name of VADIN (vinasses anaerobic digestor of Narbonne).

TABLE 2. Sequences and target positions of the primers used in this study

Name	Sequence	Target	Position[a]
w001	AGAGTTTGATCMTGGCTC	SSU rRNA bacteria	F8
w002	GNTACCTTGTTACGACTT	SSU rRNA universal	R1509
w003	ATTCYGGTTGATCCYGSC	SSU rRNA archaea	F6
w004	TGACGGGCGGTGTGTRCAAG	SSU rRNA universal	R1408
w015	AGCRAACAGGATTAGATAC	SSU rRNA bacteria	F777
w016	CTTAATTTGACTCAACACGG	SSU rRNA eucarya	F955
w017	ATTCYGGTTGATCCYGSCRG	SSU rRNA archaea	F6
w018	GAGTTTGATCMTGGCTCAG	SSU rRNA bacteria	F9
w024	GCRAACVGGATTAGATAC	SSU rRNA procarya	F778

[a] The position corresponds to the primer 5′ end, using the *E. coli* SSU rRNA as a reference (6); F and R correspond to forward and reverse primer, respectively.

TABLE 3. Clones and OTU distribution

Group and subgroup	OTU[a]	Clone divergence (%)	No. (%) of clones: Libraries: *Bacteria*	*Archaea*	*Procarya*	Total	Frequency of OTU (%)[b]
Bacteria							
CFB	BA22	0.8	3	0	0	3	0.7
	BA24		1	0	0	1	0.2
	BA34	0.4	1	0	1	2	0.4
	BB12		1	0	0	1	0.2
	BB53	0.6	1	1	1	3	0.7
	BC07	0	2	0	3	5	1.1
	BC27	2.4	7	1	3	11	2.4
	BD07	0.8	2	0	0	2	0.4
	BE72		1	0	0	1	0.2
	DC01	0	0	2	0	2	0.4
	HA17	1.6	19	0	2	22[c]	4.8
	HA21	0.4	10	0	0	10	2.2
	HA28	0.8	11	0	2	13	2.8
	HA35	1	3	0	0	3	0.7
	HA45	0.8	4	0	2	6	1.3
	HA54	3	7	0	0	7	1.5
	HA61	1	15	1	1	17	3.7
	HB29		1	0	0	1	0.2
	HB56	0	3	0	3	6	1.3
	IA59		0	0	1	1	0.2
	IA66		0	0	1	1	0.2
CFB total	21		92 (25.1%)	5	20 (32.3%)	118[c] (25.7%)	
Deep	BA07*	0.4	6 (1.6%)	1	4 (6.5%)	11 (2.4%)	2.4
High GC	BA21	0.4	2	0	1	3	0.7
	BA25	3.1	3	0	0	3	0.7
	HA41	0.4	11	1	2	14	3.0
	HA55	0.4	5	1	0	6	1.3
	IA31		0	0	1	1	0.2
High GC total	5		21 (5.7%)	2	4 (6.5%)	27 (5.9%)	
Low GC	BA08	0.8	5	1	2	8	1.7
	BA19		1	0	0	1	0.2
	BA29	0.6	2	0	0	2	0.4
	BA36		1	0	0	1	0.2
	BA37		1	0	0	1	0.2
	BB14	0.2	4	0	0	4	0.9
	BB18		1	0	0	1	0.2
	BB22		1	0	0	1	0.2
	BB27	0.2	2	0	0	2	0.4
	BB42		1	0	0	1	0.2
	BB43		1	0	0	1	0.2
	BB60	0.8	3	0	6	9	2.0
	BB65	0.6	1	0	1	2	0.4
	BC06		1	0	0	1	0.2
	HA42	3.9	12	2	0	14	3.0
	BB35	6.1	7	1	0	8	1.7
	BC14	0.2	3	0	0	3	0.7
	BC24		1	0	0	1	0.2
	BC32		1	0	0	1	0.2
	BC50		1	0	0	1	0.2
	BC53		1	0	0	1	0.2
	BC56	4.6	2	0	0	2	0.4
	BD06		1	0	0	1	0.2
	BE02		1	0	0	1	0.2
	BE04	0.2	2	0	0	2	0.4
	BE16		1	0	0	1	0.2
	BE27	1.6	1	0	1	2	0.4
	BE55		1	0	0	1	0.2
	BE58		1	0	0	1	0.2
	BE65		1	0	0	1	0.2

Continued on following page

TABLE 3—*Continued*

Group and subgroup	OTU[a]	Clone divergence (%)	No. (%) of clones: Libraries: *Bacteria*	*Archaea*	*Procarya*	Total	Frequency of OTU (%)[b]
	BE69		1	0	0	1	0.2
	BE78		1	0	0	1	0.2
	CA16	0.6	3	2	0	5	1.1
	CA17	0.8	2	1	1	4	0.9
	CA23		0	1	0	1	0.2
	DC09		0	1	0	1	0.2
	DC22		0	1	0	1	0.2
	DC38	0.2	1	1	0	2	0.4
	GA05	0.8	0	1	1	2	0.4
	HA01		1	0	0	1	0.2
	HA03		1	0	0	1	0.2
	HA05		1	0	0	1	0.2
	HA08	0.6	3	0	0	3	0.7
	HA10	0	2	0	0	2	0.4
	HA19	0	2	0	1	3	0.7
	HA31	0.4	7	0	0	7	1.5
	HA36		1	0	0	1	0.2
	HA67	1.8	6	0	1	7	1.5
	HA72	1	2	0	0	2	0.4
	HA75	0.2	1	0	2	3	0.7
	HB04	1	14	0	1	15	3.3
	HB09	3.2	3	0	0	3	0.7
	HB11	0	1	0	2	3	0.7
	HB33		1	0	0	1	0.2
	HB41		1	0	0	1	0.2
	HB44		1	0	0	1	0.2
	HB46	2	1	1	0	2	0.4
	HB54		1	0	0	1	0.2
	HB61		1	0	0	1	0.2
	HB77		1	0	0	1	0.2
	HB78		1	0	0	1	0.2
Low GC total	61		123 (33.5%)	13	19 (30.6%)	155 (33.7%)	
GNS	BA26	0.2	5	0	0	5	1.1
	BC11*	0.2	2	0	0	2	0.4
	BE97*		1	0	0	1	0.2
	HA48*	0.8	2	0	0	2	0.4
	HA64*	1	7	0	0	7	1.5
	HB28*		1	0	0	1	0.2
	HB65*		1	0	0	1	0.2
GNS total	7		19 (5.2%)	0	0 (0%)	19 (4.13%)	
Planctomyces-Chlamydia	BA30*	0.8	4	0	3	7	1.5
	BE38		1	0	0	1	0.2
	HA49*	0.2	2	0	1	3	0.7
	IA79*		1	0	0	1	0.2
Plan-Cla total	4		8 (2.2%)	0	4 (6.5%)	12 (2.6%)	
Proteobacteria	BA01	0.4	5	0	0	5	1.1
	BA15		1	0	0	1	0.2
	BA39		2	0	0	2	0.4
	BA44		1	0	0	1	0.2
	BB47		1	0	0	1	0.2
	BB56	1.8	4	0	1	5	1.1
	BB59	1.6	3	0	0	3	0.7
	BC38	1.4	12	1	3	16	3.5
	BD10		1	0	0	1	0.2
	BE06		1	0	0	1	0.2
	BE92		1	0	0	1	0.2
	DC43	0.8	2	2	0	4	0.9
	HA34		1	0	0	1	0.2
	HA40	1.4	12	0	0	12	2.6
	HA56		1	0	0	1	0.2

Continued on following page

TABLE 3—*Continued*

Group and subgroup	OTU[a]	Clone divergence (%)	No. (%) of clones: Libraries: *Bacteria*	*Archaea*	*Procarya*	Total	Frequency of OTU (%)[b]
	HA60	0.8	12	0	3	15	3.3
	HA62	0	2	0	0	2	0.4
	HA77		1	0	0	1	0.2
	HA84		1	0	0	1	0.2
	HB08		1	0	0	1	0.2
	HB18		1	0	0	1	0.2
	HB76		1	0	0	1	0.2
	HB86		1	0	0	1	0.2
Proteo total	23		68 (18.5%)	3	7 (11.3%)	78 (17.0%)	17.0
Spirochaetes	BA43	0.4	5	0	0	5	1.1
	BB62		1	0	0	1	0.2
	BC23	0	2	0	0	2	0.4
	BE82		1	0	0	1	0.2
	DC72	0	0	2	0	2	0.4
	HA24	0.4	4	0	0	4	0.9
	HA70	0	2	0	0	2	0.4
Spirochaetes total	7		15 (4.1%)	2	0 (0%)	17 (3.7%)	
Synergistes	BA23	0.2	2	0	0	2	0.4
	BB02		1	0	0	1	0.2
	CA02	2.6	8	2	3	13	2.8
	HA73	1	6	0	1	7	1.5
Synergistes total	4		17 (4.6%)	2	4 (6.5%)	23 (5.0%)	
Bacteria total	133		369	28	62	460[c]	100
Archaea							
	CA11	0	0	1	2	3	3.1
	CA25	0.6	0	18+(8)[d]	4	32[e]	32.7
	DA05	0.4	0	7+(7)	0	14	14.3
	DC06	0.6	2	12+(17)	12	43	43.9
	DC69	0.6	0	2	1	3	3.1
	DC79	0	0	2	1	3	3.1
Archaea total	6		2	42+(32)	20	98[e]	100

[a] *, OTU with uncertain affiliation.
[b] Frequency of the OTU compared with the total number of clones analyzed.
[c] One clone came from the *Eucarya* library (Table 1).
[d] Numbers in parentheses correspond to clones analyzed by restriction pattern.
[e] Two clones came from the *Eucarya* library (Table 1).

RESULTS

Clones analyzed. The 678 clones analyzed were obtained from four types of libraries: *Archaea* total (with primers whose 5′ end corresponds to *E. coli* SSU rRNA positions 6 to 1509 [6]), *Bacteria* total (*E. coli* positions 8 to 1509 [6]), *Procarya* partial (*E. coli* positions 778 to 1509 [6]), and *Eucarya* partial (*E. coli* positions 955 to 1509 [6]) (Table 1). These clones were grouped in 146 OTUs on the basis of having more than 96% sequence similarity within an OTU. The sequence divergence of clones belonging to the same OTU was generally low (between 2 and 0%) (Table 3). In one case, BB35 OTU (eight clones), the divergence between the most distant sequences was 6.1% whereas the divergence between the most related sequences was never more than to 3.5% (Table 3).

The bacterial clones were produced from six bacterial libraries (BA, BB, BC, BD, HA, and HB), one *Procarya* library (IA), and four *Archaea* libraries (CA, CB, DA, and DC) (Table 1). A wide variety of bacteria were identified with a total of 133 OTUs from 460 clones analyzed (Tables 1 and 3). All clones from the *Bacteria* and *Procarya* libraries were sequenced without prior restriction enzyme pattern screening (see Materials and Methods).

The *Archaea* clones were produced from four *Archaea* libraries (CA, CB, DA, and DC), one *Procarya* library (IA), two bacterial libraries (HA and HB), and one eukaryotic library (EA) (Table 1). The C and D libraries amplified with *Archaea*-specific primers (w003 and w017) (see Materials and Methods) nevertheless contain both archaeal (50 and 80%, respectively) and bacterial (50 and 20%, respectively) sequences whereas the I library with the *Procarya* primer (w024) contains 25% archaeal sequences. Six *Archaea* OTUs were found among 98 clones analyzed (Tables 1 and 3). The clones from the D libraries were screened by comparing restriction enzyme cleavage patterns prior to sequencing to avoid redundancy (see Materials and Methods).

The *Eucarya* clones were produced from the *Eucarya* library (EA) and the *Archaea* library (CA). Seven *Eucarya* OTUs were isolated among 94 clones analyzed (Tables 1 and 3). Due to the

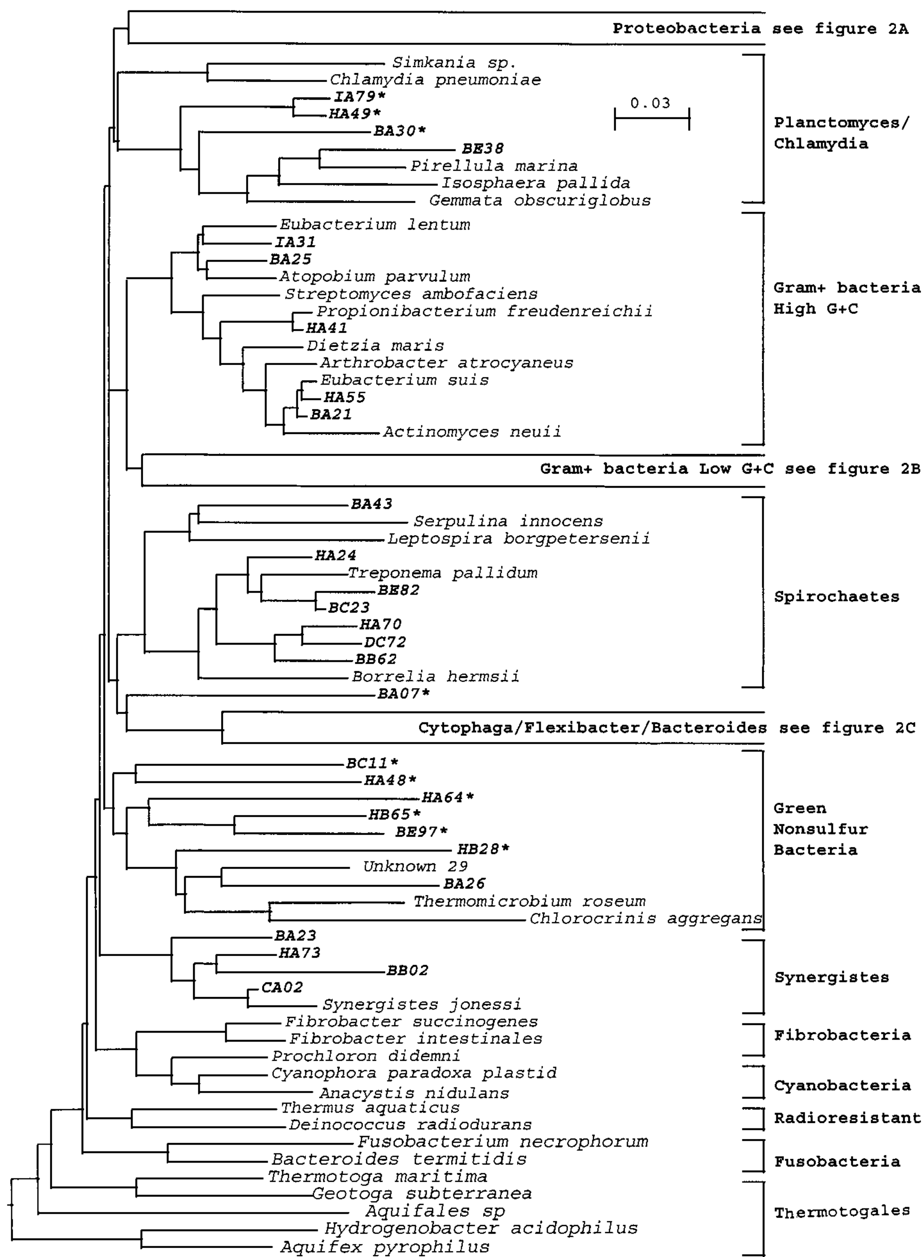

FIG. 1. Molecular phylogeny of partial SSU rDNA sequences from OTUs studied and sequences from identified bacteria in the databases. The distance matrices and phylogenetic trees were calculated by the Jukes-Cantor (14) and neighbor-joining (24) algorithms, respectively. Major taxonomic divisions are indicated. Differences between sequences are indicated by the sums of the horizontal lengths; the scale bar is in fixed nucleotide substitutions per sequence position. The archaeal *M. sarcina* sequence served as the outgroup for rooting the tree. Asterisks indicate OTUs with uncertain affiliations.

low diversity, the *Eucarya* clones were previously analyzed with restriction enzymes (see Materials and Methods).

Chimeras. Chimeric rDNA clones, composed of rDNAs from different organisms, can arise during PCR amplification of mixed DNA populations (15). A total of 26 chimeras were found among 486 bacterial clones (5.4%) and were removed from further analysis. Twenty-one of them resulted from recombination between two identified OTUs. Sequences found frequently in the OTUs were also common in chimeras (i.e., sequences belonging to the BC27 and HA61 OTUs occurred 12 and 16 times, respectively, in OTUs and 4 and 7 times, respectively, in chimeras). Five other chimeras resulted from the recombination between an identified OTU and a nonidentified OTU. Several OTUs, which appeared to be chimeras when analyzed by the CHECK_CHIMERA program (16), were considered to be OTUs because identical sequences were obtained from two or more independent PCR amplifications.

Phylogenetic analysis. (i) *Bacteria* domain. The distribution of the 133 OTUs into the bacterial phyla is as follows: 46% to low-G+C gram-positive bacteria, 16% to the *Cytophaga-Flexibacter-Bacteroides* group (CFB), 17% to the *Proteobacteria*, and 21% to the minor phyla (high-G+C gram-positive bacteria, green nonsulfur [GNS], *Spirochaetes*, *Synergistes*, *Planctomyces-Chlamydia*, and nonaffiliated) (Table 3). The bacterial phylo-

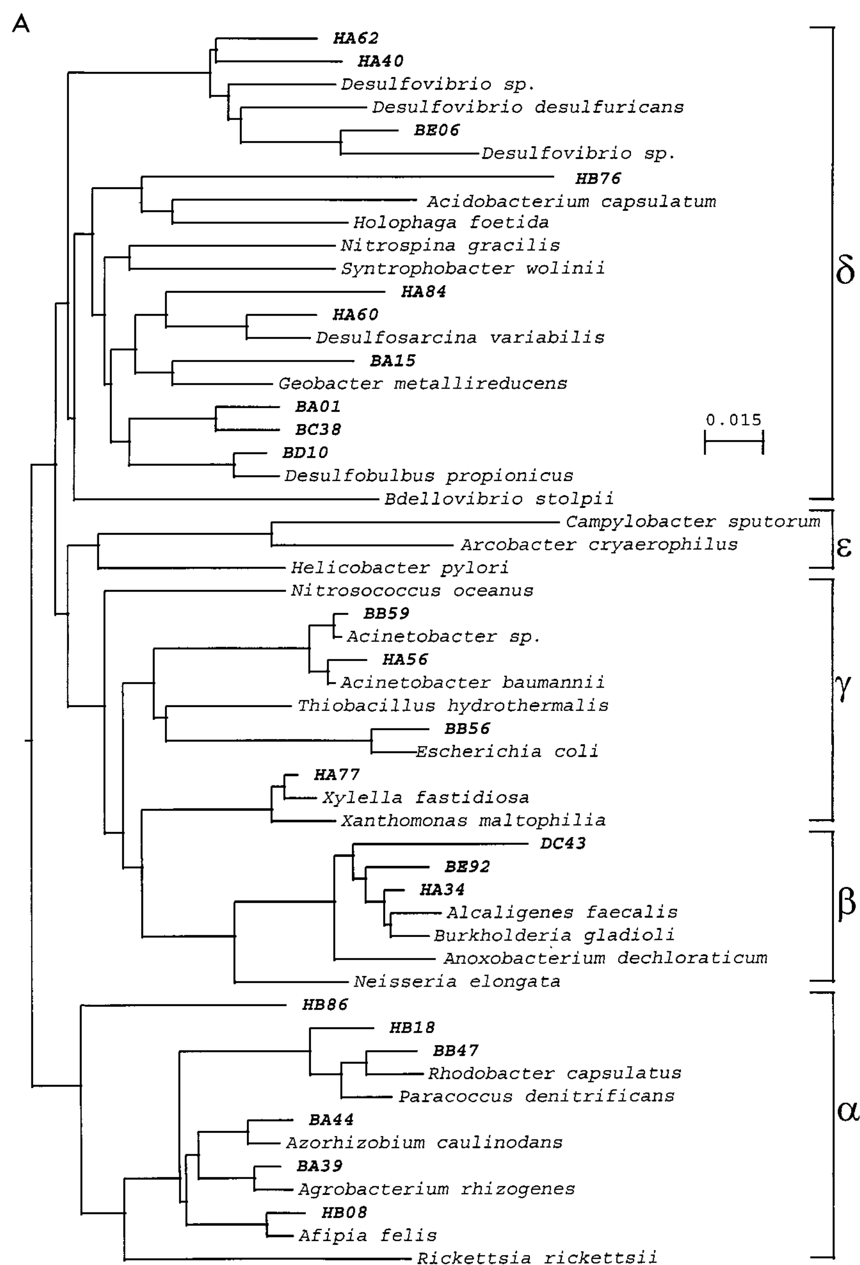

FIG. 2. Molecular phylogeny of partial SSU rDNA sequences from OTUs studied and sequences from identified bacteria in the databases. The distance matrices and phylogenetic trees were calculated by Jukes-Cantor (14) and neighbor-joining (24) algorithms, respectively. (A) Phylogenetic tree showing the relationships between the low-G+C gram-positive bacterial sequences characterized. The *Streptomyces ambofaciens* sequence served as the outgroup for rooting the tree. (B) Phylogenetic tree showing the relationships between the CFB sequences characterized. The *Chlorobium vibrioforme* sequence served as the outgroup for rooting the tree. (C) Phylogenetic tree showing the relationships between the proteobacterial sequences characterized. The *Bacillus subtilis* sequence served as the outgroup for rooting the tree. The scale bar is in fixed nucleotide substitutions per sequence position.

genetic tree is shown in Fig. 1 and 2. The five high-G+C gram-positive bacteria OTUs form three subgroups. Four OTUs appear to belong to the *Planctomyces-Chlamydia* phylum (BA30, BE38, HA49, and IA79). Four *Planctomyces* SSU rRNA sequence signatures are present in the 500 bp sequenced (A at positions 933 and 1109, C at position 955, and an additional U at position 983 [34]). The OTU BE38 presents all of them, whereas the other three OTUs present only two (A at positions 933 and 1109). Seven OTUs belong to the *Spirochaetes* phylum. Four OTUs are clustered around the sequence of a rumen bacterium, *Synergistes jonesii*. This unaffiliated sequence is deeply branched in the phylogenetic tree. Seven OTUs are weakly clustered with sequences belonging to the green nonsulfur phylum. The OTU BA07, which represents 2.4% of the population, is individually deeply branched in the phylogenetic tree and cannot be affiliated.

The 23 OTUs belonging to the *Proteobacteria* phylum fall into four of the five classical subclasses: α, β, γ, and δ (Fig. 2A).

Figure 2B shows the phylogenetic tree of the low-G+C gram-positive bacteria phylum, which contains 61 OTUs distributed among several groups of the phylum. Eight OTUs are clustered around the *Eubacterium plautii* sequence, represent-

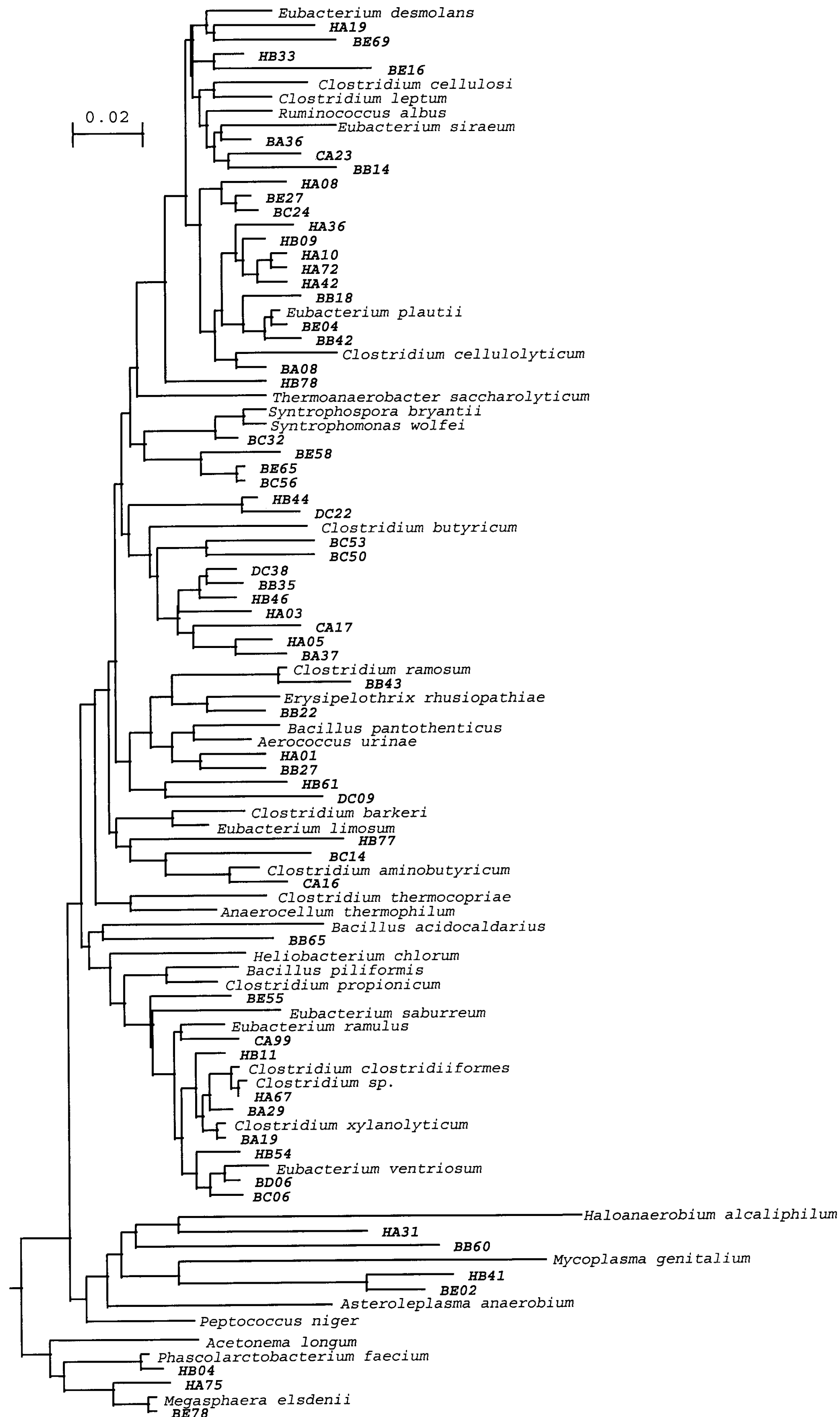

FIG. 2—*Continued.*

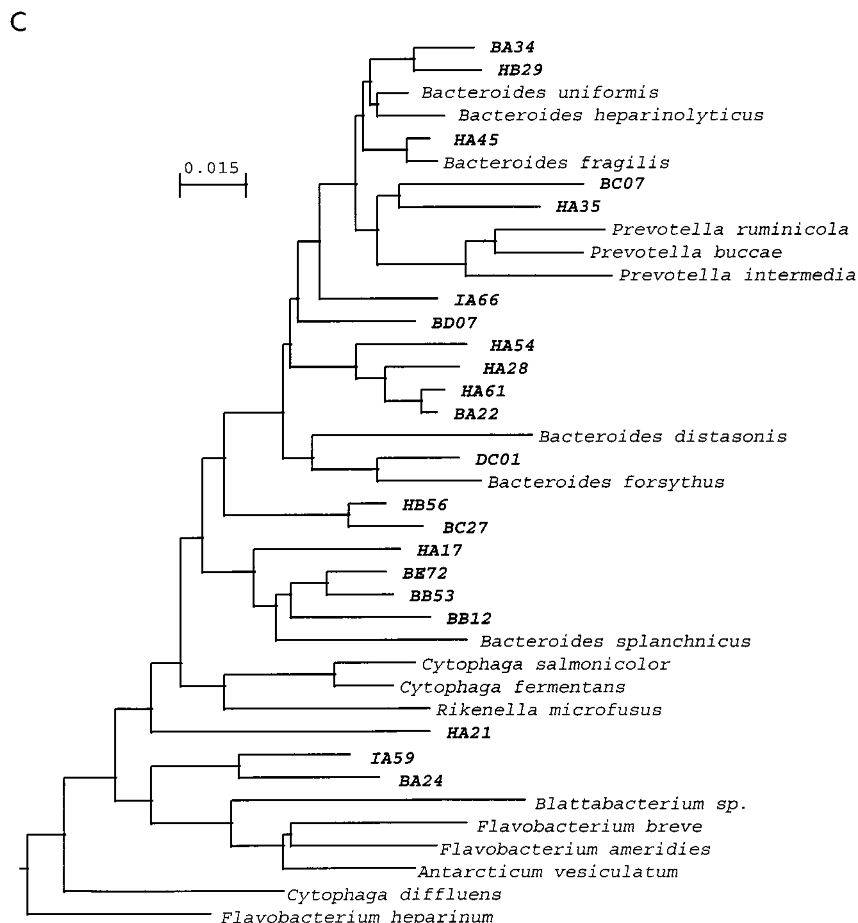

FIG. 2—*Continued.*

ing 17% of the low-G+C gram-positive bacterial clones. Four OTUs (HA31, HB41, BE02, and BB60) seem to be affiliated to the *Haloanaerobium-Mycoplasma* cluster. These OTUs show some features of the subgroup: a high level of sequence divergence and a U residue at position 888 (except for HA31) (32).

Figure 2C shows the phylogenetic relationships of the CFB phylum. The 21 OTUs clearly belong to the phylum, although several are deeply branched.

None of the bacterial SSU rRNA sequences has less than 3% divergence from sequences found in the databases. It is generally admitted that 3% divergence is the threshold for cutoff at the species level based on rRNA sequence.

(ii) *Archaea* and *Eucarya* domains. The phylogenetic positions of the six *Archaea* OTUs placed in the *Archaea* domain are shown in Fig. 3. Three of them can be considered similar to already known methanogenic SSU rRNA sequences: CA25 presents 1.2% divergence with *Methanosarcina barkeri*, DA05 presents 0.2% divergence with *Methanosarcina frisius*, and DC06 presents 1.2% divergence with *Methanobacterium formicicum*. In contrast, the other three *Archaea* OTUs are very atypical. Their sequences are remotely related to the thermophilic microorganisms *Crenarchaea* (DC69) or *Thermoplasma* sp. (CA11 and DC79).

A comparison of the seven eukaryotic OTUs with their closest known relatives is presented in Table 4. Two OTUs seem to be composed of yeast (EA02 and EA32), three are amoebae (EA13, EA83, and EA88) and one is affiliated with *Tritrichomonas foetus* (EA12). However, except for three of them (EA02, EA12, and EA32), the S_{ab} value (similarity coefficient for query and matching sequences) is low (less than 0.55) (16). The affiliation is therefore doubtful. The sequence divergence of EA02, EA12, and EA32 is respectively 2.0% with *Pichia anomala*, 9.7% with *Tritrichomonas foetus*, and 14.9% with *Candida lusitaniae.*

Distribution of clones. The phylogenetic analysis presented above describes the diversity of the population on the basis of OTUs. The following analysis is presented at the clone level to estimate the percentage of each OTU in the population.

Of the 678 clones analyzed 460, 98, and 94 belonged to the *Bacteria*, *Archaea*, and *Eucarya* domains, respectively, and the other 26 were chimeras. The distribution of bacterial clones presented in Table 3 combined results mainly from the *Archaea* and *Procarya* libraries. The bacterial clone distribution within their phyla is 34% low-G+C gram-positive bacteria, 26% CFB, 17% *Proteobacteria*, and 24% other phyla (Table 3). The most frequent bacterial OTU contains 22 clones, which correspond to 5% of the bacterial clones (Table 3). Thirteen OTUs contain either nine or more clones, and their frequencies are higher than 2%. Seventeen OTUs contain five to eight clones, and their frequencies are between 1 and 2%. Thirty-nine OTUs contain two to four clones, and their frequencies are between 0.2 and 0.9%. Sixty-three OTUs contain only one clone, and their frequencies are lower than 0.2%.

The distribution of *Archaea* clones presented in Table 3

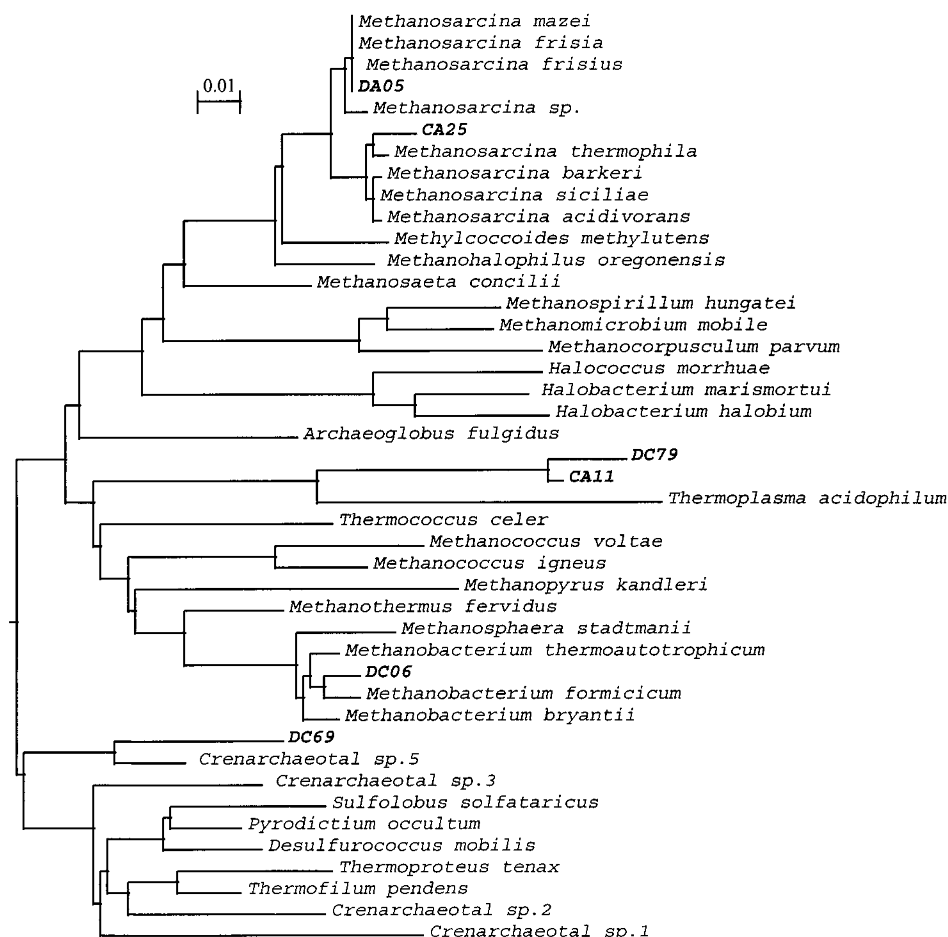

FIG. 3. Molecular phylogeny of the partial SSU rDNA sequence from OTUs studied and sequences from identified members of the *Archaea* domain in the databases. The distance matrices and phylogenetic trees were calculated by the Jukes-Cantor (14) and neighbor-joining (24) algorithms, respectively. The phylogenetic tree shows the relationships between the *Archaea* sequences characterized. The scale bar is in fixed nucleotide substitutions per sequence position. The *E. coli* sequence served as the outgroup for rooting the tree.

combined results from the *Bacteria*, *Archaea*, and *Procarya* libraries. Of the *Archaea* clones, 91% belong to three major *Archaea* OTUs.

The ratio of *Archaea* to *Bacteria* can be estimated to be 1:4 from the prokaryotic library data. The distribution of eukaryotic clones presented in Table 4 comes almost exclusively from the *Eucarya* library. A total of 46% of the eukaryotic clones are yeasts, 16% are amoebae, and 37% are affiliated to *Tritrichomonas foetus.*

DISCUSSION

This work gives an overall description of the population of an anaerobic digestion ecosystem by using SSU rDNA identification. The fluidized-bed reactor fed with vinasses has several advantages for this analysis. A fluidized bed generates a biofilm-based ecosystem, which is very stable over time. The almost sterile substrate guarantees that the microorganisms described are not in transit. This allows us to investigate a stable and closed microbial community.

This ecosystem is composed of more than 146 different organisms belonging to the three domains *Bacteria*, *Eucarya*, and *Archaea* (35). The representation of the *Bacteria* domain was extremely diverse and contained more than 133 OTUs distributed among at least eight of the major groups. Phylogenetic analysis showed that 122 bacterial OTUs (>90%) fell obviously into well-defined phyla that compose the bacterial lineage (Table 3; Fig. 2 and 3) (21). Ten OTUs were on the border of known phyla, and one was too deeply branched to be easily affiliated. Of these 10 OTUs, 5 came from separate PCR events and were considered to be nonchimeric sequences. This work has permitted us, within the limitation of the method, to quantify the representation of 70 bacterial OTUs that were present at a frequency up to 0.3% within the overall population of the sample studied. A total of 63 OTUs include only one clone, and their frequencies could not be deduced. The number of OTUs present at a frequency less than 0.3% in the sample remains unknown. However, after 460 clones were analyzed, the discovery of new OTUs was less frequent (Fig. 4). At this point, the identification of a new OTU occurred only once in every 10 clones sequenced. In spite of the large bacterial diversity, few OTUs represent a large percentage of clones. For example, the 20 most frequent bacterial OTUs represent 50% of the total clones.

The *Archaea* domain contained six OTUs distributed within the two archaeal kingdoms *Euryarchaeota* and *Crenarchaeota* (35). The three major *Archaea* OTUs represented 20% of the prokaryotic clones, and the DC06 OTU represented approximately 14% of the prokaryotic clones. The *Archaea* libraries

TABLE 4. Eucarya clones and OTU distribution

OTU	% Divergence	S_{ab} value[a]	Organism with the best matching sequence[b]	Taxon of the best matching sequence[b]	No. of clones[c]	Frequency of OTU (%)[e]
EA02	0.2	0.792	*Pichia anomala*	Hemiascomycetes	5 + (10)	16.0
EA12	0.2	0.647	*Tritrichomonas foetus*	*Parabasalidae*	7 + (28)[d]	37.2
EA13[f]		0.47	*Phreatamoeba balamuthi*	Phreatamoebids	1	1.1
EA32	0.2	0.595	*Candida lusitaniae*	Hemiascomycetes	6 + (22)	29.8
EA34[f]	0.2	0.537	*Scypha ciliata*	*Porifera*	2	2.1
EA83[f]	3	0.451	*Acanthamoeba palestinensis*	*Acanthamoebidae*	4 + (5)	9.6
EA88[f]	0.8	0.397	*Phreatamoeba balamuthi*	Phreatamoebids	3 + (1)	4.3
					28 + (66)	100

[a] Similarity coefficients for query and matching sequences (16).
[b] The best matching sequence was found with SIMILARITY_RANK (16).
[c] Numbers in parentheses correspond to the number of clones analyzed by restriction pattern.
[d] One clone came from the *Archaea* library (Table 1).
[e] Frequency of the OTU compared with the total number of the eukaryotic clones analyzed.
[f] OTU with uncertain affiliation.

contain a high percentage of bacterial clones (50 and 20% in the C and D libraries, respectively), whereas the converse is not true: two *Archaea* clones only in the bacterial libraries. There is no obvious explanation for this cross-amplification. The 3′ ends of the primers used distinguished clearly between the *Bacteria* and *Archaea* domains, and the stringency during the PCR amplification was higher for *Archaea* libraries than for the others. It cannot be explained by a low frequency of the *Archaea* rDNA matrices in the sample, since the *Archaea* clones represented 25% of the *Procarya* library when prokaryotic primers were used (Table 1). The primers w003 and w017, which, in combination with w002, allowed us to amplify the total *Archaea* SSU rRNA sequences, are not optimal.

The domain *Eucarya* contains seven OTUs distributed among several kingdoms. The frequencies of eukaryotic clones cannot be correlated to those of the clones of other domains because a separate primer set was used.

The anaerobic digestion ecosystem was described by classical methods (for a review, see reference 2), and several genera involved in the different steps of the trophic pathway were identified. Phylogenetic characterization described in this work has identified members of all the genera or taxons expected with known functions, e.g., *Bacteroides*, *Eubacterium*, *Clostridium*, *Proteobacteria* delta (sulfate-reducing bacteria), *Syntrophomonas*, and methanogenic members of the domain *Archaea*. We also identified several genera whose function was unknown, including *Spirochaetes*, GNS, *Planctomyces*, *Synergistes*, nonmethanogenic members of the *Archaea*, and most of the eukaryotic OTUs.

A contrast can also be seen between the bacterial diversity (133 OTUs) and the diversity of *Archaea* and *Eucarya* (6 and 7 OTUs, respectively). This may be related to the fact that only two types of methanogen are known within the trophic chain: the acetophilic and hydrogenophilic *Archaea* (2). Three of the *Archaea* OTUs are nearly identical to acetophilic methanogenic members of the *Archaea* (*Methanosarcina barkeri* and *M. frisius*) and hydrogenophilic methanogenic members of the *Archaea* (*Methanobacterium formicicum*, which is the most frequent prokaryotic organism of this community). Members of the *Bacteria* take over all the other functions in the trophic web, and the bacterial diversity is just a picture of these functions diversity due to the substrate complexity.

The significance of these results depends on the bias introduced by this approach (10, 28, 29, 33). It is probably not correct to assume that the OTU distribution in the sample is the same as the species distribution into the reactor. Several parameters could affect our ability to detect the actual species distribution (G+C contents of SSU rDNA, differential cell lysis, *Taq* polymerase specificity, primer specificity, chimera formation, and copy number of the SSU rRNA genes). Nevertheless, the following results obtained suggest that these disruptive factors were minimized. (i) The OTUs identified belong to several bacterial taxons with various G+C contents and cell wall structures. (ii) The low-stringency PCR conditions allowed the pairing of primer and template sequences with mismatches, as demonstrated by the cross-amplification between *Bacteria* and *Archaea* within specific libraries. (iii) Different sets of primers were used to amplify bacterial clones (Table 1). The phylogenetic distribution did not reveal a significant bias in the phyla where the number of clones is not too small (Table 3). (iv) Chimeras due to recombination during the PCR amplification remain the most tricky problem. The presence of similar clones from independent PCR amplifications leads to easy chimera detection. Twenty-six chimeras were found (5.4% of the clones) on the 500 bp of the SSU rDNA sequenced. Undetected chimera sequences remain possible in the 70 OTUs containing a unique SSU rDNA sequence. However, since only 5% of the chimeras detected result from the recombination between an identified OTU and an unknown OTU, the frequency of recombination between two unknown sequences is expected to be rare.

Although the frequency of each microorganism within the

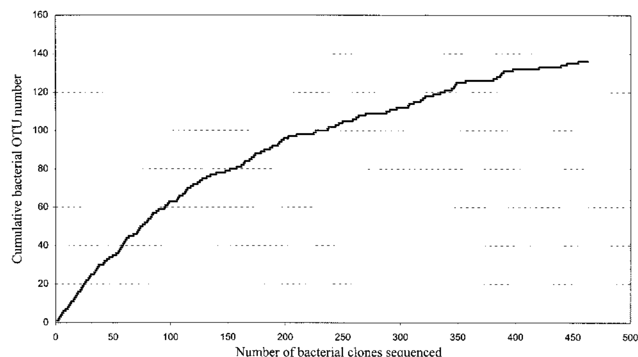

FIG. 4. Estimation of diversity in the bacterial anaerobic digestor community. The sequential detection of cumulative OTUs following sequence analysis of 460 bacterial SSU rDNA clones is represented. The clone order is alphabetical.

ecosystem given by this molecular culture-free approach is probably not completely exact, it can be used as starting point for further investigations. Group-specific or OTU-specific fluorescence-labeled probes could be used to determine the OTU distribution in the population by in situ hybridization.

This work reveals that even a mesophilic, hospitable, and common ecosystem has a high biodiversity. Several of the OTUs identified belong to deep branches of the SSU rRNA classification. The inventory of the diversity described here on this stable and closed ecosystem is wide but not out of reach, in contrast to similar approaches applied to open ecosystems.

ACKNOWLEDGMENTS

We thank Pierre Buffière for his help in establishing the anaerobic digestor. We are grateful to Richard Christen for the software facilities and to Suzan K. Amundsen for critical reading of the manuscript.

REFERENCES

1. **Amann, R. I., W. Ludwig, and K.-H. Schleifer.** 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiol. Rev. **59:**143–169.
2. **Archer, D. B., and B. H. Kirsop.** 1990. The microbiology and control of anaerobic digestion, p. 43–89. *In* A. Wheatley (ed.), Anaerobic digestion: a waste treatment technology. Elsevier Science Publishing Ltd., London, England.
3. **Boivin-Jahns, V., A. Bianchi, R. Ruimy, J. Garcin, S. Daumas, and R. Christen.** 1995. Comparison of phenotypical and molecular methods for the identification of bacterial strains isolated from a deep subsurface environment. Appl. Environ. Microbiol. **61:**3400–3406.
4. **Bond, P. L., P. Hugenholtz, J. Keller, and L. L. Blackall.** 1995. Bacterial community structures of phosphate-removing and non-phosphate-removing activated sludges from sequencing batch reactors. Appl. Environ. Microbiol. **61:**1910–1916.
5. **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.
6. **Brosius, J., T. J. Dull, D. D. Sleeter, and H. F. Noller.** 1981. Gene organization and primary structure of a ribosomal RNA operon from *Escherichia coli*. J. Mol. Biol. **148:**107–127.
7. **Buffière, P., R. Moletta, and C. Fonade.** 1995. Continuous operations of fluidized bed bioreactor for anaerobic digestion: residence time effect on degradation kinetics. Biotechnol. Lett. **17:**833–838.
8. **Delong, E. F., K. Y. Wu, B. B. Prézelin, and R. V. M. Jovine.** 1994. High abundance of *Archaea* in Antarctic marine picoplankton. Nature **371:**695–697.
9. **Devereux, R., and G. W. Mundfrom.** 1994. A phylogenetic tree of 16S rRNA sequences from sulfate-reducing bacteria in a sandy marine sediment. Appl. Environ. Microbiol. **60:**3437–3439.
10. **Farrelly, V., F. A. Rayney, and E. Stackebrandt.** 1995. Effect of genome size and *rrn* gene copy number on PCR amplification of 16S rRNA genes from a mixture of bacterial species. Appl. Environ. Microbiol. **61:**2798–2801.
11. **Fuhrman, J. A., K. McCallum, and A. A. Davis.** 1993. Phylogenetic diversity of submarine microbial communities from the Atlantic and Pacific oceans. Appl. Environ. Microbiol. **59:**1294–1302.
12. **Giovannoni, S. J., T. B. Britschgi, C. L. Moyer, and K. G. Field.** 1990. Genetic diversity in Sargasso Sea bacterioplankton. Nature **345:**60–62.
13. **Hales, B. A., C. Edwards, D. A. Ritchie, G. Hall, R. W. Pickup, and J. R. Saunders.** 1996. Isolation and identification of methanogen-specific DNA from blanket bog peat by PCR amplification and sequence analysis. Appl. Environ. Microbiol. **62:**668–675.
14. **Jukes, T. H., and C. R. Cantor.** 1969. Evolution of protein molecules, p. 21–132. *In* H. N. Munro (ed.), Mammalian protein metabolism. Academic Press, Inc., New York, N.Y.
15. **Kopckzynski, E. D., M. M. Bateson, and D. M. Ward.** 1994. Recognition of chimeric small-subunit ribosomal DNAs composed of genes from uncultivated microorganisms. Appl. Environ. Microbiol. **60:**746–748.
16. **Maidak, B. L., N. Larsen, M. J. McCaughey, R. Overbeek, G. J. Olsen, K. Fogel, J. Blandy, and C. R. Woese.** 1994. The Ribosomal Database Project. Nucleic Acids Res. **22:**3485–3487.
17. **Moyer, C. L., F. C. Dobbs, and D. M. Karl.** 1994. Estimation of diversity and community structure through restriction fragment length polymorphism distribution analysis of bacterial 16S rRNA genes from a microbial mat at an active hydrothermal vent system, Loihi Seamount, Hawaii. Appl. Environ. Microbiol. **60:**871–879.
18. **Moyer, C. L., F. C. Dobbs, and D. M. Karl.** 1995. Phylogenetic diversity of the bacterial community from a microbial mat at an active hydrothermal vent system, Loihi Seamount, Hawaii. Appl. Environ. Microbiol. **61:**1555–1562.
19. **Ng, A., W. T. Melvin, and P. N. Hobson.** 1994. Identification of anaerobic digester bacteria using a polymerase chain reaction method. Bioresour. Technol. **47:**73–80.
20. **Ohkuma, M., and T. Kudo.** 1996. Phylogenetic diversity of the intestinal bacterial community in the termite *Reticulitermes speratus*. Appl. Environ. Microbiol. **62:**461–468.
21. **Olsen, G. J., C. R. Woese, and R. Overbeek.** 1994. The winds of (evolutionary) change: breathing new life into microbiology. J. Bacteriol. **176:**1–6.
22. **Paster, B. J., F. E. Dewhirst, S. M. Cooke, V. Fussing, L. K. Poulsen, and J. A. Breznak.** 1996. Phylogeny of not-yet-cultured spirochetes from termite guts. Appl. Environ. Microbiol. **62:**347–352.
23. **Rölleke, S., G. Muyzer, C. Wawer, G. Wanner, and W. Lubitz.** 1996. Identification of bacteria in a biodegraded wall painting by denaturing gradient gel electrophoresis of PCR-amplified gene fragments coding for 16S rRNA. Appl. Environ. Microbiol. **62:**2059–2065.
24. **Saitou, N., and M. Nei.** 1987. The neighbor-joining method: a new method for constructing phylogenetic trees. Mol. Biol. Evol. **4:**406–425.
25. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
26. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-termination inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
27. **Shieh, W. K., and J. D. Keenan.** 1986. Fluidized bed biofilm reactor for wastewater treatment. Adv. Biochem. Eng. Biotechnol. **33:**132–186.
28. **Stackebrandt, E., W. Liesack, and B. M. Goebel.** 1993. Bacterial diversity in a soil sample from a subtropical Australian environment as determined by 16S rDNA analysis. FASEB J. **7:**232–236.
29. **Suzuki, M. T., and S. T. Giovannoni.** 1996. Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. Appl. Environ. Microbiol. **62:**625–630.
30. **Ueda, T., Y. Suga, and T. Matsuguchi.** 1995. Molecular phylogenetic analysis of a soil microbial community in a soybean field. Eur. J. Soil Sci. **46:**415–421.
31. **Voordouw, G., S. M. Armstrong, M. F. Reimer, B. Fouts, A. J. Telang, Y. Shen, and D. Gevertz.** 1996. Characterization of 16S rRNA genes from oil field microbial communities indicates the presence of a variety of sulfate-reducing, fermentative, and sulfide-oxidizing bacteria. Appl. Environ. Microbiol. **62:**1623–1629.
32. **Weisburg, W. G., J. G. Tully, D. L. Rose, J. P. Petzel, H. Oyaizu, D. Yang, L. Mandelco, J. Sechrest, T. G. Lawrence, J. Van Etten, J. Maniloff, and C. R. Woese.** 1989. A phylogenetic analysis of the mycoplasmas: basis for their classification. J. Bacteriol. **171:**6455–6467.
33. **Wilson, K. H., and R. B. Blitchington.** 1996. Human colonic biota studied by ribosomal DNA sequence analysis. Appl. Environ. Microbiol. **62:**2273–2278.
34. **Woese, C. R.** 1987. Bacterial evolution. Microbiol. Rev. **51:**221–271.
35. **Woese, C. R., O. Kandler, and M. L. Wheelis.** 1990. Towards a natural system of organisms: proposal for the domains *Archaea*, *Bacteria*, and *Eucarya*. Proc. Natl. Acad. Sci. USA **87:**4576–4579.