Journal of Biotechnology 142 (2009) 38–49

Contents lists available at ScienceDirect

Journal of Biotechnology

journal homepage: www.elsevier.com/locate/jbiotec

Phylogenetic characterization of a biogas plant microbial community integrating clone library 16S-rDNA sequences and metagenome sequence data obtained by 454-pyrosequencing

Magdalena Kröber[a], Thomas Bekel[b], Naryttza N. Diaz[b], Alexander Goesmann[b], Sebastian Jaenicke[b], Lutz Krause[d], Dimitri Miller[a], Kai J. Runte[b], Prisca Viehöver[c], Alfred Pühler[a], Andreas Schlüter[a,*]

[a] Institut für Genomforschung und Systembiologie, Centrum für Biotechnology (CeBiTec), Universität Bielefeld, D-33615 Bielefeld, Germany
[b] Centrum für Biotechnologie (CeBiTec), Universität Bielefeld, D-33615 Bielefeld, Germany
[c] Lehrstuhl für Genomforschung, Universität Bielefeld, D-33615 Bielefeld, Germany
[d] Nestlé Research Center, BioAnalytical Science Department, CH-1000 Lausanne 26, Switzerland

ARTICLE INFO

Article history:
Received 17 December 2008
Received in revised form 10 February 2009
Accepted 17 February 2009

Keywords:
Biogas fermentation
Methanogenesis
Methane production
454-pyrosequencing
Metagenome

ABSTRACT

The phylogenetic structure of the microbial community residing in a fermentation sample from a production-scale biogas plant fed with maize silage, green rye and liquid manure was analysed by an integrated approach using clone library sequences and metagenome sequence data obtained by 454-pyrosequencing. Sequencing of 109 clones from a bacterial and an archaeal 16S-rDNA amplicon library revealed that the obtained nucleotide sequences are similar but not identical to 16S-rDNA database sequences derived from different anaerobic environments including digestors and bioreactors. Most of the bacterial 16S-rDNA sequences could be assigned to the phylum *Firmicutes* with the most abundant class *Clostridia* and to the class *Bacteroidetes*, whereas most archaeal 16S-rDNA sequences cluster close to the methanogen *Methanoculleus bourgensis*. Further sequences of the archaeal library most probably represent so far non-characterised species within the genus *Methanoculleus*. A similar result derived from phylogenetic analysis of *mcrA* clone sequences. The *mcrA* gene product encodes the α-subunit of methyl-coenzyme-M reductase involved in the final step of methanogenesis. BLASTn analysis applying stringent settings resulted in assignment of 16S-rDNA metagenome sequence reads to 62 16S-rDNA amplicon sequences thus enabling frequency of abundance estimations for 16S-rDNA clone library sequences. Ribosomal Database Project (RDP) Classifier processing of metagenome 16S-rDNA reads revealed abundance of the phyla *Firmicutes*, *Bacteroidetes* and *Euryarchaeota* and the orders *Clostridiales*, *Bacteroidales* and *Methanomicrobiales*. Moreover, a large fraction of 16S-rDNA metagenome reads could not be assigned to lower taxonomic ranks, demonstrating that numerous microorganisms in the analysed fermentation sample of the biogas plant are still unclassified or unknown.

1. Introduction

Energy production from fossil fuels becomes more and more problematic since these resources are running short and burning of coal, oil and natural gas is connected with emissions of the green-house gas carbon dioxide. For these reasons, use of renewable energies is promoted by national programs in many countries. For example, it is intended to enhance the proportion of biomass for energy production. Long-term objectives of this policy are to ensure future energy supply and to reduce green-house gas emissions. In this context biogas production from renewable resources or organic wastes is of importance (Weiland, 2003; Yadvika et al., 2004). Biogas consists of methane, carbon dioxide and low amounts of other gases. It is produced during anaerobic fermentation of organic substrates by specific microbial communities (Ohmiya et al., 2005). This process is biotechnologically exploited to produce biogas in commercial biogas plants (Angelidaki et al., 2003; Kleerebezem and van Loosdrecht, 2007). The energy plant maize in the form of maize silage mixed with liquid manure is mainly used as substrate for biogas production in Germany. Combined power and heat units convert the produced biogas to electric power and heat, respectively. Produced heat often is used as process heat and/or for community heating. Since the interaction of microorganisms for biogas production is not well understood, there is great interest in analysing the structure and functioning of endogenous microbial

* Corresponding author at: Lehrstuhl für Genetik, Universität Bielefeld, Postfach 100131, D-33501 Bielefeld, Germany. Tel.: +49 521 106 2036; fax: +49 521 106 5626.
E-mail address: Andreas.Schlueter@Genetik.Uni-Bielefeld.DE (A. Schlüter).

0168-1656/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.jbiotec.2009.02.010

communities residing in biogas-producing environments. Classically, environmental microbial communities are analysed by construction of 16S-rDNA clone libraries and subsequent sequencing of individual 16-rDNA clones. This approach has been applied for many biogas-producing microbial communities (Huang et al., 2005, 2002; Klocke et al., 2007, 2008; McHugh et al., 2003; Mladenovska et al., 2003; Shigematsu et al., 2006; Tang et al., 2007). In addition to the 16S-rDNA target, other marker genes such as *mcrA* encoding the α subunit of the methanogenesis enzyme methyl coenzyme-M reductase have been used to elucidate the composition of methanogenic consortia (Friedrich, 2005; Juottonen et al., 2006; Lueders et al., 2001; Luton et al., 2002; Rastogi et al., 2007). Amplification of marker genes from total community DNA preparations is usually biased by the selection of PCR primers. In addition, cloning biases cannot be excluded. In most studies based on clone library analyses, only limited amounts of clones were sequenced so that the whole complexity of the intrinsic community could not be disclosed. Development of high-throughput sequencing technologies such as 454-pyrosequencing nowadays offers the opportunity to effectively sequence metagenomes of microbial communities. Since library construction by classical cloning of fragmented DNA is not necessary for 454-pyrosequencing, biases should be relatively negligible when this technique is used for whole genome shotgun sequencing of microbial community metagenomes. Recently, the metagenome of a biogas-producing microbial community from a production-scale biogas plant fed with renewable primary products was analysed by applying the ultrafast 454-pyrosequencing technology (Schlüter et al., 2008). The isolated total community DNA was sequenced on a Genome Sequencer FLX system resulting in 616,072 reads with an average read length of 230 bases accounting for approximately 142 million bases sequence information. Assignment of obtained single reads to COG (Clusters of Orthologous Groups of proteins) categories revealed a genetic profile characteristic for an anaerobic microbial consortium conducting fermentative metabolic pathways. Community structure analysis of the fermentation sample revealed that *Clostridia* from the phylum *Firmicutes* is the most prevalent taxonomic class, whereas species of the order *Methanomicrobiales* are dominant among methanogenic *Archaea* (Krause et al., 2008b). Many sequence reads could be allocated to the genome sequence of the archaeal methanogen *Methanoculleus marisnigri* JR1. This result already indicated that species related to those of the genus *Methanoculleus* play a dominant role in hydrogenotrophic methanogenesis in the analysed fermentation sample (Schlüter et al., 2008).

Phylogenetic assignments based on 16S-rDNA sequences from clone libraries usually are precise due to relatively larger sequence lengths compared to the short read lengths resulting from 454-pyrosequencing. Recently, Wommack et al. (2008) provokingly stated that the use of short-read-length libraries is not appropriate for metagenomic characterization of microbial communities. On the other hand, other authors describe the phylogenetic classification of short environmental DNA fragments that were obtained by high-throughput sequencing technologies (Krause et al., 2008a; Manichanh et al., 2008).

In this study, a biogas plant microbial community is phylogenetically characterised by an analysis integrating clone library 16S-rDNA- and *mcrA*-sequences and metagenome sequence data obtained by 454-pyrosequencing.

2. Materials and methods

2.1. Amplification and cloning of 16S-rDNA

To obtain PCR-compatible total community DNA from the fermentation sample of the analysed biogas plant reactor (Schlüter et al., 2008), the DNA-preparation obtained in August 2007 was further purified on silica membrane spin columns of the NucleoSpin Plant II Kit (Macherey-Nagel, Düren, Germany) according to the manual provided by the manufacturer and by gel-filtration using MicroSpin Columns (GE Healthcare, Buckinghamshire, UK). Amplification of 16S-rDNA fragments was carried out on the purified total community DNA by PCR using oligonucleotide primers that bind to conserved regions of bacterial or archaeal 16S-rDNA, respectively. Bacterial 16S-rDNA was amplified by means of the bacterial universal primers 27f and 1385r (Weidner et al., 1996). The 16S-rDNA of *Archaea* was amplified by the archaeal-specific primer w002 (5′-GNTACCTTGTTACGACTT) and the universal primer w017 (5′-ATTCYGGTTGATCCYGSCRG) (Huang et al., 2002). The conditions for PCR amplification for each approach were as described previously (Huang et al., 2002; Schlüter et al., 2007; Weidner et al., 1996). 16S-rDNA amplicons were cloned and transformed into competent *E. coli* DH5α cells using the pGEM-T Easy T-cloning kit according to the manufacturer's instructions (Promega). Recombinant pGEM-T Easy vectors were checked for their 16S-rDNA amplicon inserts by PCR using primers, the binding sites of which flank the pGEM-T Easy cloning site: FWD_PGEM (5′-ATGGCGGCCGCGGGAATTCG) and RV_PGEM (5′-CGGCCGCGAATTCSCTAGTG).

2.2. ARDRA analysis of 16S-rDNA amplicons

Amplified ribosomal DNA restriction analysis (ARDRA) was performed on amplified 16S-rDNA fragments using recombinant pGEM-T Easy vectors as template DNA. PCR products were digested with the restriction enzymes HpaII (*Bacteria*) and HhaI (*Archaea*). Resulting restriction patterns were analysed and clustered by means of the computer program GelCompar II (Applied Maths).

2.3. Amplification and cloning of mcrA gene fragments

Fragments of the *mcrA* gene encoding the α subunit of methyl coenzyme-M reductase were PCR amplified on total community DNA by using the *mcrA*-specific primer pairs ME1/ME2 and MLf/MLr as described previously (Juottonen et al., 2006). Obtained amplicons were cloned into the vector pGEM-T Easy as already outlined for the 16S-rDNA amplicons. The computer program GelCompar II (Applied Math) was used to analyse and cluster *mcrA* RFLP patterns that were obtained by HpaII restriction.

2.4. Sequencing and phylogenetic analysis

Based on ARDRA and RFLP clustering results, representative 16S-rDNA and *mcrA* clones were selected for sequencing using the primers FWD_PGEM and RV_PGEM. Sequencing of 16S-rDNA and *mcrA* amplicons was performed on an Applied Biosystems 3730*xl* DNA Analyzer (Applied Biosytems, Darmstadt, Germany) using a 96 capillary setup. For cycle-sequencing, BigDye Terminator (BDT) chemistry version 3.1 (Applied Biosystems) was used according to the protocol of the manufacturer. Cycling-sequencing was done in a GeneAmp 9700 PCR system (Applied Biosystems). Running parameters of the 3730*xl* DNA Analyzer were as follows: injection time for the samples was set to 15 s at a voltage of 1.5 kV. Separation of the fragments was done by capillary electrophoresis at 8.5 kV on a 50 cm array (running time: 2 h).

Consensus sequences for the analysed clones were generated by matching forward and reverse nucleotide sequences obtained from the same clone using the computer program DNA Baser (Heracle-Software). Nucleotide sequences of the 16S-rDNA amplicon clone library were deposited in the GenBank database under the Accession Nos. FJ205753–FJ205864, whereas *mcrA* clone sequences are available under the GenBank Accession Nos. FJ438965–FJ438994.

Sequences were compared to entries of the NCBI nucleotide sequence database using BLASTn. The 16S-rDNA amplicon sequences, *mcrA* sequences and selected reference sequences were aligned by means of ClustalW (http://align.genome.jp/clustalw/) using default settings. Aligned sequences were then cut to aligned blocks. Subsequently, gaps in aligned sequences were removed by applying the tool Gblocks as previously described (Manichanh et al., 2008). Phylogenetic trees were constructed by means of the software tool MEGA (Version 4, http://www.megasoftware.net/mega.html) using the Neighbor-Joining method combined with the Jukes–Cantor model and Bootstrap resampling analysis for 1000 repeats as described previously (Klocke et al., 2008).

2.5. BLASTn analysis of 16S-rDNA metagenome reads versus 16S-rDNA clone library sequences

To identify metagenome reads carrying 16-rRNA sequence information, a homology search to the RDP database (release 10.5) (Cole et al., 2007, 2008) by means of BLASTn was conducted. This approach led to the identification of 2781 sequence reads out of 616,072 reads producing significant alignments with an *e*-value of $1e^{-6}$ or better and a minimum alignment length of 50 bp. Aligned read sequences were extracted from the dataset and further analysed by means of BLASTn with the low complexity filter turned off. Significant hits to the set of the 109 16S-rDNA amplicon sequences obtained during this study with an alignment length of at least 60 bases were classified according to the observed identity values (95–98% or 98–100%).

All reads that could not be assigned to any of the amplicon sequences (in total 2114 reads) were grouped into OTUs according to their nearest neighbor (best matching BLAST-hit) in the RDP database version 10.7. Only BLAST hits with an *e*-value below $1e^{-5}$, hit-length above 50 bp and pairwise sequence-identity of at least 80% were included.

3. Results and discussion

3.1. Community structure analysis based on sequencing of 16S-rDNA clone libraries

Total community DNA was recently prepared from a fermentation sample obtained from the first fermenter of an agricultural biogas plant by means of a CTAB-based DNA isolation protocol (Schlüter et al., 2008). To analyse the community structure of the sample, 16S-rDNA clone libraries were constructed covering bacterial as well as archaeal 16S-rDNA fragments. Purified total community DNA was used as template in polymerase chain reactions (PCRs) to generate 16S-rDNA amplicons by means of a universal primer pair (1385r and 27f) and the *Archaea*-specific primers w002 and w017 (Huang et al., 2002). Bacterial and archaeal clone libraries were constructed using the vector pGEM-T Easy. The bacterial clone library comprises 100 clones as verified by reamplification of cloned 16S-rDNA amplicons (approximately 1.3 kb in length) by means of the vector primers FWD_pGEM and RV_pGEM, whereas the archaeal library consists of 94 clones which also have insert lengths of approximately 1.3 kb. To characterise the two clone libraries, *Amplified Ribosomal DNA Restriction Analysis* (ARDRA) was applied using the restriction enzymes HpaII and HhaI, respectively. Obtained ARDRA profiles were clustered by applying the computer program GelCompar II (results not shown). Representative clones were selected from each ARDRA cluster and subsequently sequenced yielding 72 bacterial 16S-rDNA sequences and 37 archaeal sequences.

3.2. Bacterial 16S-rDNA sequences are related to clones previously obtained from different anaerobic environments

BLASTn analysis of the obtained bacterial sequences revealed that they are similar but not identical to corresponding 16S-rDNA sequences that were previously established for microbial communities residing in solid waste digestors, anaerobic sediments, different bioreactors and chemostats for degradation of organic material, landfills, leachates of closed municipal solid waste landfills, rumina and fecal samples (see Table 1). All environments and locations listed above represent habitats where organic material is digested under anaerobic conditions and biogas is formed to a greater or lesser extent. Most hits, namely twenty-one, were found for 16S-rDNA sequences originating from a thermophilic, anaerobic solid waste digester. Unfortunately, the corresponding study has not been published yet. Eight clones correspond to non-*Planctomycetes* operational taxonomic units (OTUs) encountered in a consortium of the anaerobic, sulfide- and sulfur-rich Zodletone spring source sediment (Elshahed et al., 2007). Species terms could not be assigned to these clones but five of them are also related (87–93% identity) to the 16S-rDNA sequence of Candidatus *Cloacamonas acidaminovorans*, which was recently identified in an anaerobic digester of a municipal wastewater treatment plant (Pelletier et al., 2008). The genome of a representative of this species was reconstructed from metagenomic data and based on the obtained genome sequence it was predicted that *C. acidaminovorans* probably is a hydrogen-producing syntrophic bacterium able to derive its carbon and energy from the fermentation of amino acids (Pelletier et al., 2008). Five 16S-rDNA sequences from the analysed biogas plant are similar to sequences from a packed-bed reactor that was continuously operated with an artificial garbage slurry (Sasaki et al., 2007). These clones could be classified as belonging to the *Clostridia* and to the phyla *Bacteroidetes* and *Chloroflexi*, respectively. Interestingly, three clones are 94–97% identical to a 16S-rDNA sequence from a thermophilic methanogenic bioreactor performing cellulosic waste digestion (Shiratori et al., 2006). The reference sequence in this case clusters with *Clostridium algidicarnis* which is known to cause spoilage of meat (Lawson et al., 1994). Relatedness of three sequences to a clone in MSW (Municipal Solid Waste) Cluster 1 defined in a study analysing the phylogenetic diversity of microorganisms in a thermophilic anaerobic municipal solid waste digester (Tang et al., 2004) again demonstrates that many bacteria from biogas-producing communities are only distantly related to cultured strains. MSW Cluster 1 is closely related to a *Clostridium* cluster and the genus *Desulfotomaculum* but no rRNA sequences from cultured strains could be assigned to this cluster (Tang et al., 2004). The three sequences that correspond to clones from a leachate of a closed municipal solid waste landfill were grouped, respectively, within the low-G + C Gram-positive bacteria and the *Cytophaga/Flexibacter/Bacteroides* (CFB) cluster (Huang et al., 2005). Further sequences are related to, respectively, an unknown cluster from a mesophilic propionate-degrading methanogenic consortium, a *Clostridium* sp. from the community in granular sludge treating brewery wastewater (Liu et al., 2002) and unspecified isolates from human and primate fecal samples (Ley et al., 2008; Mai et al., 2006). Only one sequence could be assigned to a defined species, namely *Ruminofilibacter xylanolyticum* which is a rumen bacterium involved in digestion of xylan. However, most of the 16S-rDNA amplicon sequences from the analysed biogas plant are related to clones from anaerobic environments which do not correspond to known, well-characterised species suggesting that the majority of microorganisms in the analysed biogas-producing community is still unexplored.

To taxonomically characterise the 16S-rDNA amplicon sequences from the biogas reactor, the Ribosomal Database Project Classifier (http://rdp.cme.msu.edu/classifier/classifier.jsp)

Table 1
Comparison of 16S-rDNA and *mcrA* clone library sequences from the biogas plant to corresponding NCBI nucleotide sequence database entries by means of BLASTn.

Habitat/Source/Microorganism[a]	Hits[b]	Identity[c]	Reference/Accession No.
Bacteria (16S-rRNA)			
Thermophilic anaerobic solid waste digestor	22	89–99%	EF558949, EF558951, EF558987, EF558991, EF559035, EF559060, EF559066
Anaerobic, sulfide- and sulfur-rich zodletone spring, Zodletone Spring source sediments	8	88–93%	Elshahed et al. (2007)
Biomass adhering to supporting material in a packed-bed reactor degrading organic solid waste (garbage slurry)	5	88–99%	Sasaki et al. (2007)
Thermophilic digestor of biowastes, anaerobically digested sludge	5	95–99%	AM947548, AM947555
Household biogas digester	5	84–95%	EU407216
Thermophilic anaerobic solid waste bioreactor	4	93–99%	DQ887942, DQ887948, DQ887962
Anaerobic solid waste digester fed with methanol	3	88–99%	EF585987, EF586001
Thermophilic anaerobic municipal solid waste digester, methane fermentation	3	92–99%	Tang et al. (2004)
Thermophilic methanogenic bioreactor performing cellulosic waste digestion	3	94–97%	Shiratori et al. (2006)
Leachate of a closed municipal solid waste landfill	3	92–97%	Huang et al. (2005)
Mesophilic propionate-degrading methanogenic consortium in chemostat cultivation	1	90%	Shigematsu et al. (2006)
Landfill environment, soil, 5.5 ft. depth	1	98%	EU219952
Leachate of a closed municipal solid waste landfill	1	99%	Huang et al. (2003)
Granular sludge treating brewery wastewater	1	90%	Liu et al. (2002)
Anaerobic, mesophilic sludge digester, municipal wastewater treatment plant	1	99%	Chouari et al. (2005)
Biogas-producing co-fermentation of maize silage and bovine manure, biogas-producing laboratory-scale completely stirred tank reactor	1	92%	EF686964
Ruminofilibacter xylanolyticum strain S1 from rumen	1	99%	DQ141183
Thermophilic anaerobic sludge enriched by starch	1	97%	AB239189
Mesophilic anaerobic solid waste digestor	1	97%	EF559105
Homo sapiens fecal sample	1	95%	Mai et al. (2006)
Sumatran orangutan feces	1	92%	Ley et al. (2008)
Archaea (16S-rRNA)			
Landfill environment	17	97–99%	EU369613, EU369621, EU369622
Leachate of a closed municipal solid waste landfill	11	98–99%	Huang et al. (2003)
Thermophilic anaerobic municipal solid waste digester, methane fermentation	4	99%	Tang et al. (2004)
Mesophilic anaerobic bovine serum albumin digester	2	99%	Tang et al. (2005)
Deep subsurface groundwater from sedimentary rock milieu, fault-bordered aquifer	1	99%	Shimizu et al. (2006)
Methanoculleus marisnigri JR1	1	84%	CP000562
Archaea (*mcrA*)			
Commercial biogas plant utilizing herbal biomass; mesophilic CSTR supplied with cattle liquid manure and maize silage	28	94–99%	EU636851, EU636862, EU636863, EU636865, EU636867, EU636868, EU636873, EU636874, EU636877, EU636878, EU636881, EU636883, EU636884, EU636885, EU636891, EU636892, EU636894
Archaeal community on rice roots; uncultured *Methanomicrobiaceae* archaeon	2	96%	AY125650, AY125652

[a] 16S-rDNA and *mcrA* clone library sequences were compared to entries of the NCBI nucleotide sequence database by means of BLASTn. The habitat, source or microorganism of the best-hit database clone sequence is listed.
[b] Number of hits to the respective habitat, source or microorganism.
[c] Range of nucleotide sequence identity over aligned regions of clone library 16S-rDNA and *mcrA* sequences to corresponding database clone sequences.

was applied (see Fig. 1A). This naïve Bayesian classifier was developed for rapid assignment of rRNA sequences to the current bacterial taxonomy (Wang et al., 2007). This analysis revealed that 71 of the 16S-rDNA amplicons were of eubacterial origin, whereas only one amplicon could be assigned to the superkingdom *Archaea*. Most of the sequences could be allocated to the phylum *Firmicutes* with the most abundant class *Clostridia* (Fig. 1A). In addition to this defined class, other sequences belong to so far unclassified *Firmicutes*. Further sequences were assigned to the class *Bacteroidetes* and the order *Bacteroidales*. The phylum *Chloroflexi* and the class *Anaerolineae* were identified as non-abundant taxa. Only few sequences could be allocated to lower taxonomic ranks such as 'family' or 'genus' suggesting that most sequences originate from unknown, so far unclassified species. The phylogeny of the bacterial clones from the biogas fermenter is visualised in Fig. 2. The phyla *Firmicutes*, *Bacteroidetes*, *Chloroflexi* and an unknown bacterial group are marked in Fig. 2. In most cases it was not possible to assign closely related reference species to the different clusters within the different phyla, which again demonstrates that many species residing in the biogas-producing community are unknown. Seven clones (E10, E43, E56, E64, E74, E78 and E87) are in one cluster with *Clostridium algidicarnis*, which is the causative agent of meat spoilage. Another *Firmicutes* cluster is characterised by the clostridial species *C. straminisolvens*, *C. thermocellum* and *C. sporosphaeroides*. *C. thermocellum* is known for its cellulolytic activity (Zverlov et al., 2005). This species produces so-called cellulosomes, which are supramolecular protein complexes facilitating concerted cellulose degradation (Liu et al., 2002; Bayer et al., 2007, 1998a,b, 2004; Madkour and Mayer, 2003). Eight clones are in the cluster of unknown *Bacteria* and could not be assigned to lower taxonomic ranks. All of these clones are related to 16S-rDNA clones obtained from the sediments of a mesophilic spring (Elshahed et al., 2007). Five of these clones are also related to the 16S-rDNA sequence of Candidatus *Cloacamonas acidaminovorans* (see above).

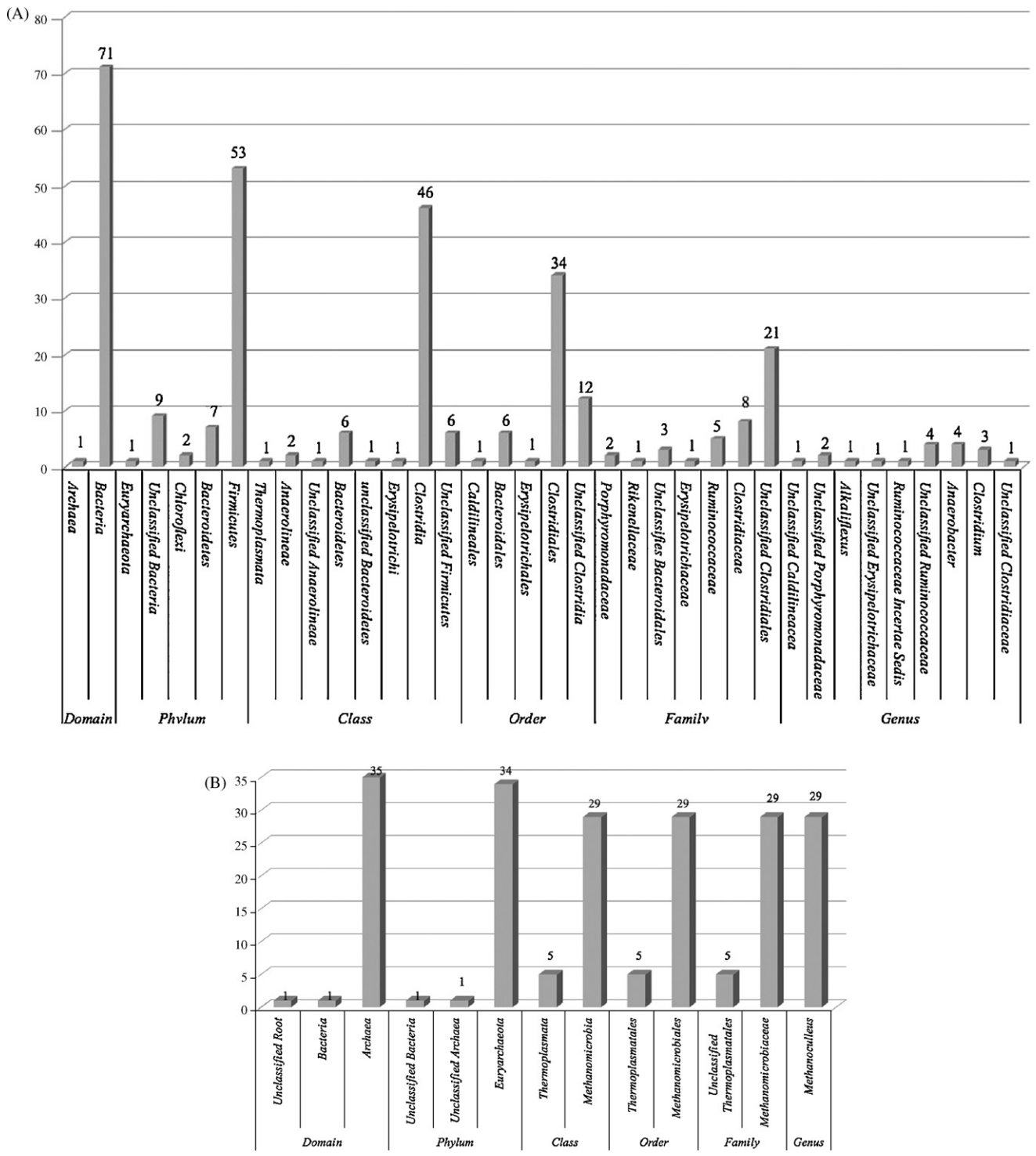

Fig. 1. A and B. Ribosomal Database Project (RDP) Classifier analysis of bacterial (A) and archaeal (B) 16S-rDNA clone library sequences. 16S-rDNA sequences were assigned to the following taxa: domain, phylum, class, order, family and genus. Only assignments with confidence values above 80% were considered for this chart.

This analysis suggests that species of the phylum *Firmicutes* dominate the community of the analysed fermentation sample.

3.3. Most archaeal 16S-rDNA sequences are related to sequences of the genus Methanoculleus

Sequencing of representative clones from the archaeal 16S-rDNA clone library resulted in thirty-seven valid nucleotide sequences (see Table 1). BLASTn analysis of these sequences revealed that seventeen of them are very similar to clones from a landfill environment (Accession Nos. EU369613, EU369621, EU369622, unpublished). Unfortunately, corresponding database entries do not denote any species assignments. The second most abundant kind of hits could be affiliated to clones originating from a leachate of a closed municipal solid waste landfill (Huang et al., 2003). Corresponding sequences are related to *Methanoculleus olentangyi* and *M. oldenburgensis* belonging to the order *Methanomicrobiales* and to a clone that clusters with non-specified 16S-rDNA sequences close to the group of *Thermoplasmatales*. Dominance of species belonging to the genus *Methanoculleus* was also confirmed by identification of

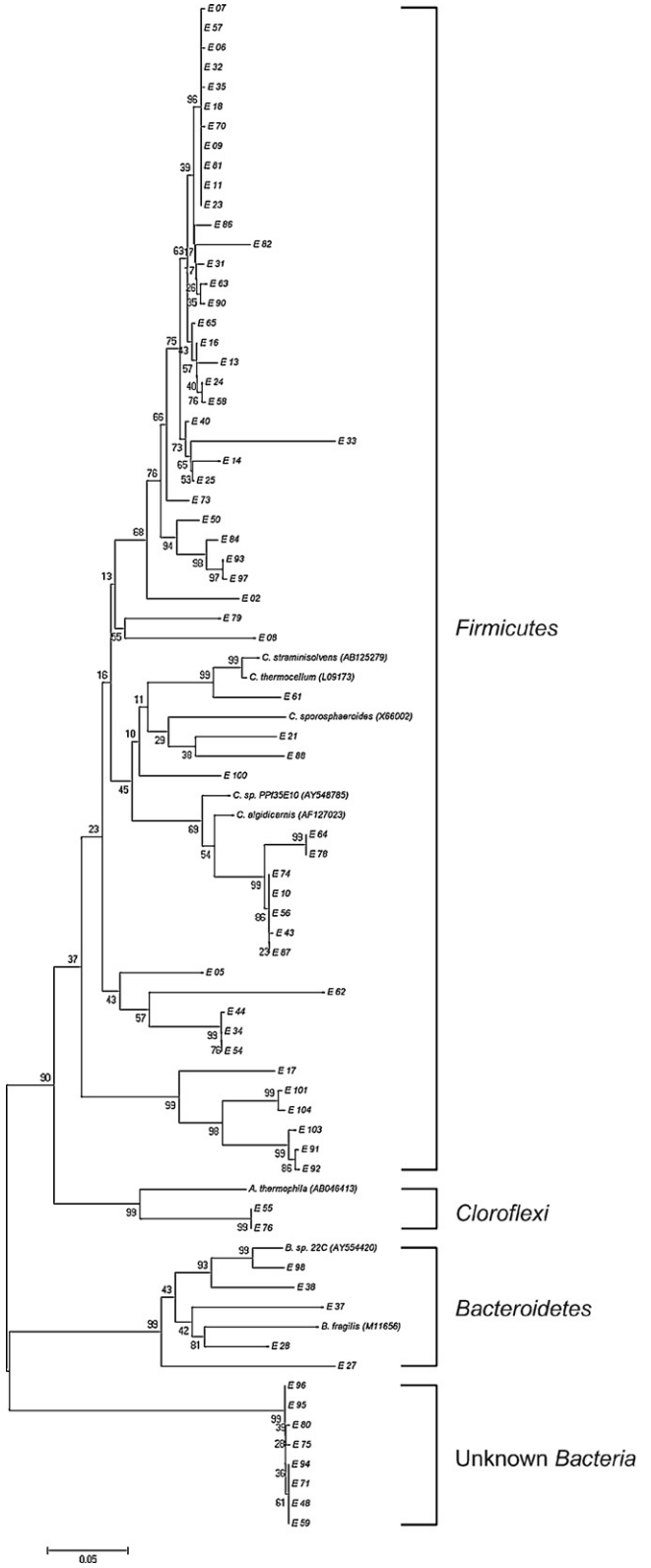

Fig. 2. Phylogenetic analysis of bacterial 16S-rDNA clone sequences by means of the neighbor-joining method using genetic distances defined by Jukes and Cantor. Gaps in alignments were removed by using the Gblocks program. The bar represents 5% sequence divergence. Numbers at the nodes represent bootstrap values (1000 resamplings). Clone Nos. are given next to the tree branches. Accession Nos. for reference sequences are given in addition to the species name. Further details on the applied phylogenetic analysis method and corresponding computer programs are outlined in Section 2.

four 16S-rDNA sequences that are very similar (99%) to a clone from a thermophilic anaerobic municipal solid waste (MSW) digestor that is closely related to the reference species *M. oldenburgensis* in the phylogenetic analysis of the corresponding study (Tang et al., 2004). This analysis also showed wide distribution of operational taxonomic units (OTUs) in the genus *Methanoculleus* and dominance of the *Methanoculleus* population in the analysed MSW digester, especially under micro-aeration (Tang et al., 2004). Likewise, two sequences are 99% identical to a clone obtained from a microbial community of a mesophilic anaerobic protein degradation process using bovine serum albumin as a substrate (Tang et al., 2005). The reference clone clusters with *Methanoculleus bourgensis*. Methanogens affiliated with the genera *Methanoculleus* and *Methanosaeta* were predominant in this protein degrading community (Tang et al., 2005). Further sequences from the analysed biogas plant correspond to a clone related to *Methanoculleus thermophilus* from fault-bordered aquifers in a Miocene formation in Japan (Shimizu et al., 2006) and to the 16S-rDNA sequence of the completely sequenced methanogen *Methanoculleus marisnigri* (Accession No. CP000562). RDP classifier analysis also confirmed that most of the archaeal sequences originate from strains belonging to the class *Methanomicrobia* (see Fig. 1B). Only five sequences were assigned to the order *Thermoplasmatales* and one sequence to 'Unclassified *Archaea*'. All *Methanomicrobia* sequences seem to stem from *Methanoculleus* species. Phylogeny of the *Methanoculleus* clones is visualised in the phylogenetic tree shown in Fig. 3. The community consists of two clusters, one representing clones closely related to *M. bourgensis* and *M. olentangyi*. These two species are subjective synonyms (Asakawa & Nagaoka, 2003). For the second cluster of clones it was not possible to find a closely related reference species. All other reference species of the *Methanoculleus* group are located outside of the aforementioned clusters. In conclusion, the methanogenic archaeal community seems to be dominated by species closely related to the *M. bourgensis*/*M. olentangyi* group.

3.4. Phylogenetic analysis of methanogen mcrA clone sequences

To investigate the phylogeny of the methanogenic *Archaea* with an independent marker, and to compare the results with those obtained with the 16S-rDNA analysis, a clone library based on *mcrA* amplicons was constructed and characterised. The *mcrA* gene encodes the α-subunit of methyl coenzyme-M reductase and also serves as phylogenetic marker gene for characterisation of methanogenic communities since it is present in all methanogenic *Archaea* analysed so far (Luton et al., 2002). Fragments of the *mcrA* gene were amplified by PCR using the specific primers ME1/ME2 and MLf/MLr (Juottonen et al., 2006) and subsequently cloned into the T-cloning vector pGEM-T Easy. RFLP analysis of 200 *mcrA* amplicon clones (109 clones obtained by using the ME1/ME2 primers and 91 clones amplified by means of the MLf/MLr primers) revealed 15 different restriction profile clusters as calculated by the computer program GelCompar II (Applied Math). Thirty representative clones were chosen from these clusters for sequence analysis. BLASTn analysis revealed that twenty-eight of the obtained *mcrA* amplicon sequences are similar to seventeen different uncultured archaeon clones amplified from a commercial, mesophilic CSTR (completely stirred tank reactor) biogas plant fed with cattle liquid manure and maize silage (see Table 1). Thus, the substrate of the CST reactor is very similar to the biomass that was fermented in the biogas reactor analysed in this study. Two further *mcrA* sequences are related to uncultured *Methanomicrobiaceae* archaeon clones from a community on rice roots. In addition to these similarities, all but three sequences are highly related (89–99% identity) to different *Methanoculleus bourgensis* strains including the type strain MS2 of this species (*M. bourgensis* DSM3045). *M. bourgensis* is able to

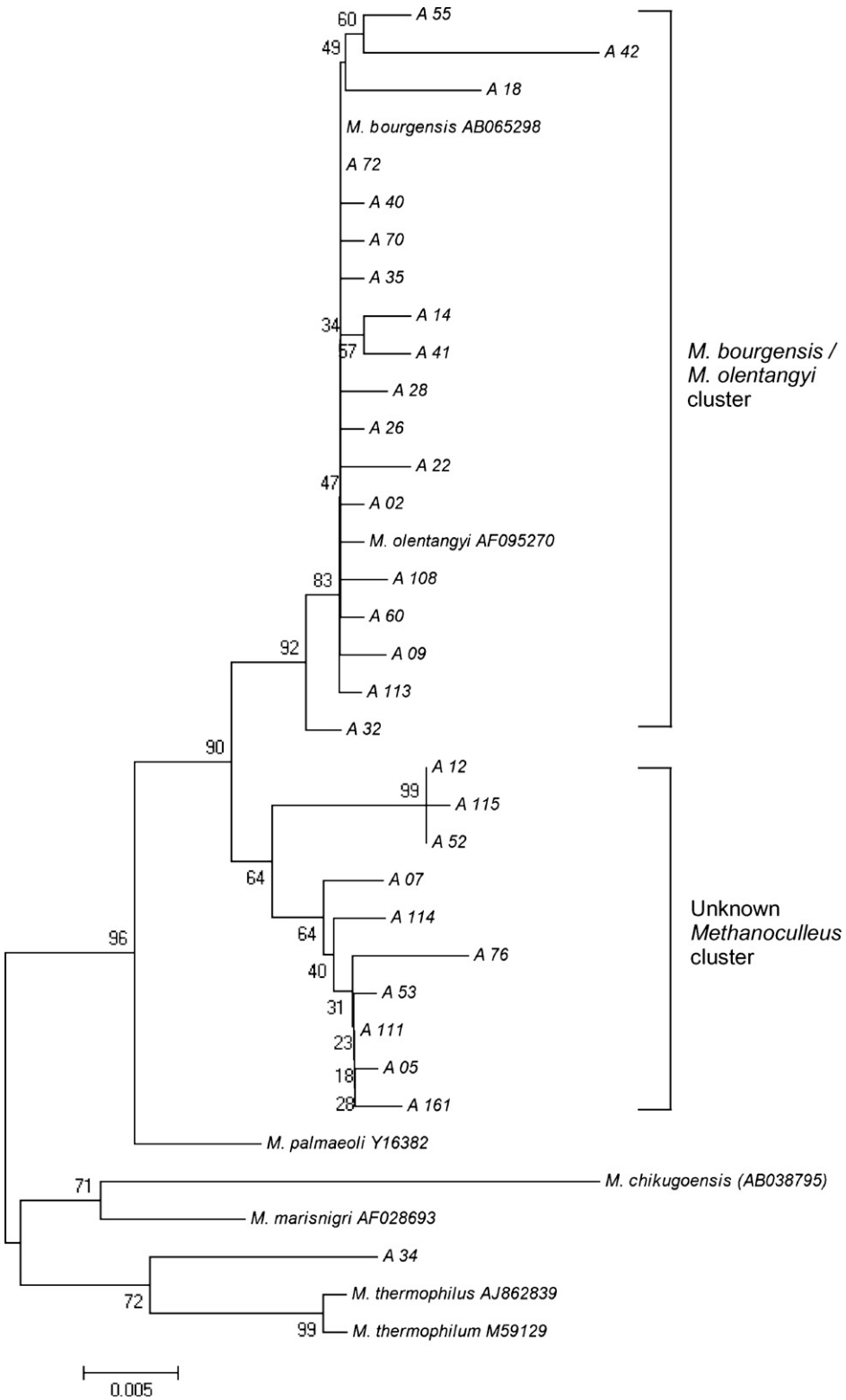

Fig. 3. Phylogenetic analysis of archaeal 16S-rDNA clone sequences by means of the neighbor-joining method using genetic distances defined by Jukes and Cantor. Gaps in alignments were removed by using the Gblocks program. The bar represents 0.5% sequence divergence. Numbers at the nodes represent bootstrap values (1000 resamplings). Clone Nos. are given next to the tree branches. Accession Nos. for reference sequences are given in addition to the species name. Further details on the applied phylogenetic analysis method and corresponding computer programs are outlined in Section 2.

utilise hydrogen and carbon dioxide and formate as methanogenic substrates. Some secondary alcohols can also be metabolised for methanogenesis and acetate is required for growth (Asakawa and Nagaoka, 2003). The optimal growth temperature for *M. bourgensis* was determined to be 40 °C which is close to the temperature of the biogas fermenter from which the analysed sample was taken. Three *mcrA* sequences are 88% identical to *mcrA* of the completely sequenced methanogen *Methanoculleus marisnigri* JR1 (CP000562). In summary, the diversity within the methanogenic population in the analysed biogas plant fermenter seems to be restricted to the

Methanomirobiaceae. But it should be noted that there is diversity within this family. Some of the analysed archaeal *mcrA* amplicon sequences could not be assigned to characterised species and thus seem to represent so far unexplored microorganisms. The phylogenetic analysis based on *mcrA* sequences is shown in Fig. 4 and confirms the results obtained from the archaeal 16S-rDNA clone library. *Methanoculleus* species were previously identified in different anaerobic habitats including thermophilic municipal biogas plants (Tang et al., 2004; Weiss et al., 2008), a thermophilic upflow reactor treating distillery wastewater (Tang et al., 2007), mesophilic anaerobic chemostats (Tang et al., 2005) a thermophilic batch digestor inoculated with cattle manure (Chachkhiani et al., 2004) a biogas reactor fed with cattle dung (Rastogi et al., 2007) and a two-phase biogas reactor operated with plant biomass (Klocke et al., 2008). In some of the microbial communities residing in these environments, members belonging to the genus *Methanoculleus* were found to be dominant indicating that these are very well adapted to the conditions in the respective habitats.

3.5. Integrated analysis of 16S-rDNA clone library sequences and metagenome sequence data obtained by 454-pyrosequencing

Metagenomics, among other things, is aimed at obtaining an unbiased view of the community composition of a specific environment. This certainly cannot be achieved by sequencing of 16S-rDNA clone libraries for two reasons: (i) Amplification of 16S-rDNA sequences is biased by the choice of PCR primers that are designed on basis of known sequences stored in database. (ii) Many studies relying on the 16S-rDNA clone library approach only present limited numbers of clones insufficient to cover the whole complexity of the community. Currently, high-throughput sequencing technologies are prevalently applied in metagenomics to analyse community structures of different habitats. For example, the novel 454-pyrosequencing technology circumvents the need for clone library construction and thus should allow for a relatively unbiased view of the community structure when it is applied for whole-microbial-community DNA sequencing. However, 454-pyrosequencing currently only produces short read lengths which are not always sufficient to accurately characterise microbial communities (Wommack et al., 2008). Longer amplicon sizes of 16S-rDNA clone libraries usually allow more precise assignments of 16S-rDNA clones. Therefore a comparative analysis using 16S-rDNA clone library sequences and metagenome sequences obtained by 454-pyrosequencing was carried out for the analysed biogas-producing microbial community.

The same total community DNA preparation that was used as template DNA for the construction of the 16S-rDNA clone libraries in this study had been sequenced previously by applying the 454-pyrosequencing technology (Schlüter et al., 2008). This approach yielded 616,072 sequence reads with an average read length of 230 bases accounting for approximately 142 million bases sequence information. Reads encoding 16S-rRNAs were extracted from the metagenome data set by means of BLASTn on the basis of the Ribosomal Database Project (RDP database-release 10.5). A total of 2781 16S-rDNA sequences could be identified by this approach. All metagenome sequences were cut to the length that aligned to reference 16S-rDNA sequences. To analyse whether the metagenome 16S-rDNA sequence reads correspond to the PCR-amplified 16S-rDNA sequences of the clone libraries described above, BLASTn comparison of both data sets was performed. Results of this approach are shown in Table 2. 16S-rDNA metagenome reads could be assigned to 62 16S-rDNA amplicon sequences (out of a total of 109 sequences) with minimum alignment lengths of 60 bases and sequence identities of 98–100%. This analysis also disclosed abundance of distinct 16S-rDNA amplicon sequences. Due to the limited amount of 16S-rDNA amplicon sequences and the

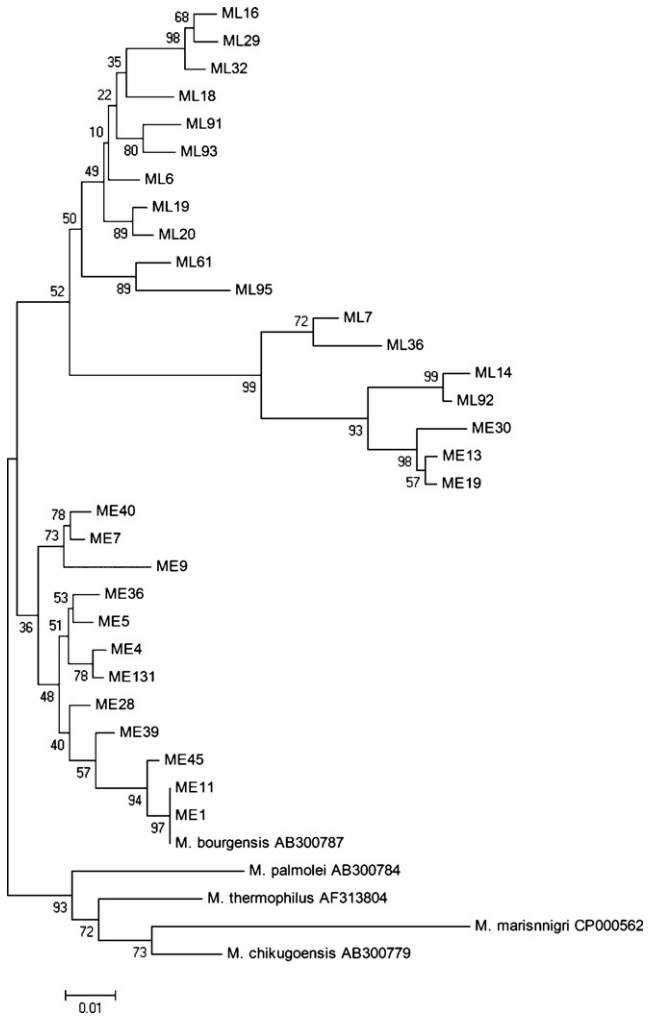

Fig. 4. Phylogenetic analysis of archaeal *mcrA* clone sequences by means of the neighbor-joining method using genetic distances defined by Jukes and Cantor. Gaps in alignments were removed by using Gblocks. The bar represents 1% sequence divergence. Numbers at the nodes represent bootstrap values (1000 resamplings). Clone Nos. are given next to the tree branches. Accession Nos. for reference sequences are given in addition to the species name. Further details on the applied phylogenetic analysis method and corresponding computer programs are outlined in Section 2.

selection procedure conducted to identify representative clones, frequency of occurrence estimations are not possible for the original 16S-rDNA clone library data set. The most abundant amplicon sequence is E79 which was assigned to the class *Clostridia*. A very similar sequence was previously identified in a clone library from a thermophilic anaerobic solid waste bioreactor (unpublished). Other abundant 16S-rDNA amplicon sequences are related to clones from a thermophilic anaerobic solid waste digestor (unpublished), to a *Bacteroidetes* clone from a packed-bed reactor degrading organic solid waste (Sasaki et al., 2007), to non-*Planctomycetes* OTUs encountered in a spring source sediment (Elshahed et al., 2007), and to a clone from a thermophilic digestor of biowastes (unpublished). Interestingly, 46 16S-rDNA amplicon sequences (42% of the sequenced 16S-rDNA amplicons) found no highly similar counterparts among the 16S-rDNA metagenome reads indicating that these clones represent rare species in the analysed fermentation sample. This result clearly demonstrates that the metagenome sequencing approach was not carried out to saturation. Moreover, of the 2781 metagenome reads representing 16S-rDNAs, 2114 (76%) could not be assigned to any of the 16S-rDNA amplicon sequences. Of these 2114 reads, 2050 were assigned to an OTU according to their nearest neighbor in the RDP reference database. The remaining 64 reads did

Table 2
BLAST-based analysis of 16S-rDNA metagenome reads obtained by 454-pyrosequencing versus 16S-rDNA clone library sequences.

16S-rDNA amplicon[a]	Phylum[b]	Class[b]	Order[b]	Genus[b]	Number of hits (95–98%)[c]	Number of hits (98–100%)[d]
E_79	*Firmicutes*	*Clostridia*			11	58
E_54	*Firmicutes*	*Clostridia*			2	46
E_81	*Firmicutes*	*Clostridia*			0	35
E_90	*Firmicutes*	*Clostridia*	*Clostridiales*		1	35
E_38	*Bacteroidetes*	*Bacteroidetes*	*Bacteroidales*		5	30
E_86	*Firmicutes*	*Clostridia*	*Clostridiales*		2	28
E_96					0	28
E_58	*Firmicutes*	*Clostridia*	*Clostridiales*		7	25
E_24	*Firmicutes*	*Clostridia*	*Clostridiales*		1	18
E_42	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae Incertae Sedis*	3	17
E_40	*Firmicutes*	*Clostridia*	*Clostridiales*		0	16
E_70	*Firmicutes*	*Clostridia*	*Clostridiales*		0	15
E_57	*Firmicutes*	*Clostridia*	*Clostridiales*		1	14
E_08	*Firmicutes*	*Clostridia*	*Clostridiales*		0	13
E_94					0	13
E_65	*Firmicutes*	*Clostridia*	*Clostridiales*		0	12
E_87	*Firmicutes*	*Clostridia*	*Clostridiales*	*Anaerobacter*	4	12
A_108	*Euryarchaeota*	*Methanomicrobia*	*Methanomicrobiales*	*Methanoculleus*	0	11
A_161	*Euryarchaeota*	*Methanomicrobia*	*Methanomicrobiales*	*Methanoculleus*	1	10
E_31	*Firmicutes*	*Clostridia*	*Clostridiales*		1	9
A_113	*Euryarchaeota*	*Methanomicrobia*	*Methanomicrobiales*	*Methanoculleus*	0	8
E_101	*Firmicutes*				1	8
E_44	*Firmicutes*	*Clostridia*			0	8
E_63	*Firmicutes*	*Clostridia*	*Clostridiales*		0	8
E_104					2	7
E_25	*Firmicutes*	*Clostridia*	*Clostridiales*		1	7
E_97	*Firmicutes*	*Clostridia*	*Clostridiales*		2	7
A_22	*Euryarchaeota*	*Methanomicrobia*	*Methanomicrobiales*	*Methanoculleus*	0	6
E_100	*Firmicutes*	*Clostridia*	*Clostridiales*		3	6
E_75	*Firmicutes*	*Clostridia*	*Clostridiales*		0	6
E_05	*Firmicutes*	*Clostridia*			3	5
E_103	*Bacteroidetes*	*Bacteroidetes*	*Bacteroidales*		0	5
E_61	*Firmicutes*	*Clostridia*	*Clostridiales*		2	5
E_16	*Firmicutes*	*Clostridia*	*Clostridiales*		0	4
E_37	*Bacteroidetes*	*Bacteroidetes*	*Bacteroidales*	*Alkaliflexus*	0	4
E_73	*Bacteroidetes*	*Bacteroidetes*	*Bacteroidales*		1	4
E_95					0	4
A_111	*Euryarchaeota*	*Methanomicrobia*	*Methanomicrobiales*	*Methanoculleus*	0	3
E_17	*Firmicutes*	*Erysipelotrichi*	*Erysipelotrichales*		0	3
E_82	*Firmicutes*	*Clostridia*			0	3
E_91	*Firmicutes*				0	3
E_92	*Firmicutes*				0	3
A_114	*Euryarchaeota*	*Methanomicrobia*	*Methanomicrobiales*	*Methanoculleus*	0	2
A_115	*Euryarchaeota*	*Methanomicrobia*	*Methanomicrobiales*	*Methanoculleus*	0	2
E_02	*Firmicutes*	*Clostridia*	*Clostridiales*		2	2
E_27	*Bacteroidetes*				1	2
E_28	*Bacteroidetes*	*Bacteroidetes*	*Bacteroidales*		4	2
E_32	*Firmicutes*	*Clostridia*			1	2
E_50	*Firmicutes*	*Clostridia*	*Clostridiales*		4	2
E_84	*Firmicutes*	*Clostridia*	*Clostridiales*	*Anaerobacter*	0	2
E_93	*Firmicutes*	*Clostridia*			2	2
A_52	*Euryarchaeota*	*Methanomicrobia*	*Methanomicrobiales*	*Methanoculleus*	0	1
A_80	*Euryarchaeota*	*Thermoplasmata*	*Thermoplasmatales*		0	1
E_13	*Firmicutes*	*Clostridia*	*Clostridiales*		0	1
E_19	*Euryarchaeota*	*Thermoplasmata*	*Thermoplasmatales*		0	1
E_21	*Firmicutes*	*Clostridia*	*Clostridiales*		0	1
E_43	*Firmicutes*	*Clostridia*	*Clostridiales*	*Clostridium*	0	1
E_56	*Firmicutes*	*Clostridia*	*Clostridiales*	*Clostridium*	1	1
E_62	*Firmicutes*				4	1
E_64	*Firmicutes*	*Clostridia*	*Clostridiales*	*Anaerobacter*	0	1
E_88	*Firmicutes*	*Clostridia*	*Clostridiales*		0	1
E_98	*Bacteroidetes*	*Bacteroidetes*	*Bacteroidales*		2	1
E_71					1	0

[a] 16S-rDNA amplicon sequences of the clone libraries. Clone numbers beginning with an 'E' designate clones of the bacterial clone library, whereas clone numbers beginning with an 'A' designate clones of the archaeal clone library.

[b] 16S-rDNA amplicon sequences were assigned to taxonomic ranks (phylum, class, order and genus) by means of the Ribosomal Database Project (RDP) Classifier (release 10.5). Amplicon sequences E_96, E_94, E_104, E_95, E_29 and E_71 could only be assigned to the superkingdom *Bacteria*. Only assignments with a confidence level above 80% are listed.

[c] 16S-rDNA metagenome reads obtained by 454-pyrosequencing were assigned to 16S-rDNA clone library sequences by means of BLASTn (minimum alignment length: 60 bases, sequence identity: 95–98%).

[d] 16S-rDNA metagenome reads obtained by 454-pyrosequencing were assigned to 16S-rDNA clone library sequences by means of BLASTn (minimum alignment length: 60 bases, sequence identity: 98–100%).

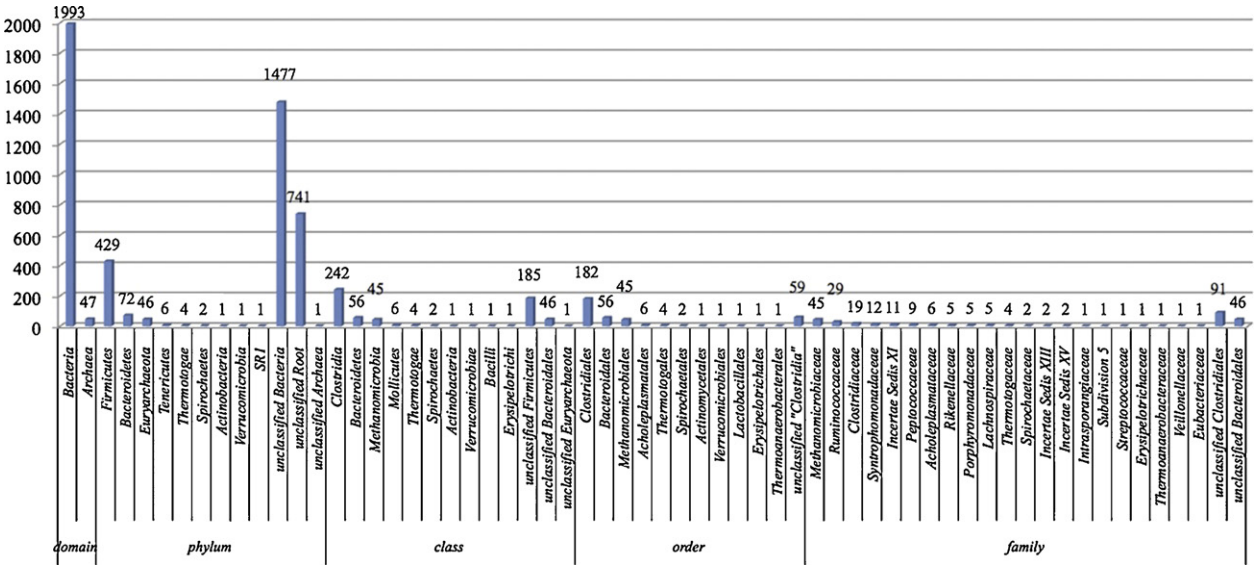

Fig. 5. Analysis of metagenome sequence reads containing 16S-rDNA sequence information by means of the Ribosomal Database Project (RDP) Classifier. 16S-rDNA sequences were assigned to the following taxa: domain, phylum, class, order, and family. Only assignments with confidence values above 80% were considered for this chart. The term '*Incertae Sedis*' indicates that the placement of these taxa within the taxonomic hierarchy is not clear yet.

not produce significant BLAST-hits to any of the reference sequences and hence were excluded. In total, 251 OTUs were observed. The largest three OTUs comprise more than 220 metagenome reads each. These results indicate that the clone library approach missed many taxonomic groups present in the biogas-producing microbial community, including three abundant groups to which more than 26% of all metagenome 16S-rRNA reads were assigned. This result was expected since the metagenome data sets comprises 616,072 sequence reads and thus allows for a much deeper coverage of the intrinsic community.

All 16S-rDNA metagenome reads were taxonomically classified by means of the Ribosomal Database Project (RDP) Classifier. Fig. 5 shows that only 73% of all 16S-rDNA reads (2781) could be taxonomically assigned to the domains *Bacteria* or *Archaea* which most probably is due to short read lengths and the fact that many reads had not found closely related sequences in the RDP database. Moreover, most reads could only be assigned to higher taxonomic ranks suggesting that most members of the biogas-producing community are still unexplored. The Classifier analysis illustrates the composition of the community. The most abundant phyla are the *Firmicutes*, *Bacteroidetes* and *Euryarchaeota*. At the rank 'order' the *Clostridiales*, *Bacteroidales* and *Methanomicrobiales* are dominant. Within the *Euryarchaeota*, other taxa than the *Methanomicrobiales* seem to be underrepresented. The apparent diversity within the bacterial community of the analysed fermentation sample was not unexpected since many species have to perform tasks in different substrate hydrolysis steps and downstream fermentative pathways such as acidogenesis and acetogenesis. In summary, the relatively large 16S-rDNA amplicon sequences facilitate more precise assignments to corresponding reference sequences stored in databases, whereas the short metagenome reads cover the diversity of the community and allow for quantitative estimations.

3.6. Concluding remarks

Sequencing and analysis of 16S-rDNA library clones was the accepted standard for assessing the microbial community composition of environments of interest for a long period of time. Although this approach is biased by the choice of PCR primers for amplification of 16S-rDNA and cloning efficiencies, it allows accurate assignment of 16S-rDNA sequences to corresponding database entries which is mainly supported by relatively large amplicon sequence lengths of ca. 1300–1500 bp. However, in case of the 16S-rDNA amplicon sequences from the biogas plant sample, it was observed that almost no sequences could be assigned to lower taxonomic ranks. Obviously, in most cases database 16S-rDNA sequences with valid taxonomic assignments are not closely related to sequences from the analysed libraries which leads to the conclusion that a significant proportion of the species residing in the biogas fermentation sample are still unexplored. In future experiments, cloned 16S-rDNA amplicons will therefore represent valuable sequence tags, facilitating identification of isolated biogas community members, for example by making use of defined 16S-rDNA clones as hybridisation probes. Moreover, analysis of 16S-rDNA clone library sequences provides insights into the phylogenetic structure of the community since they can be used to calculate evolutionary distances between equivalent sequence segments. For example, phylogenetic analysis of the archaeal clone sequences from the biogas community revealed the fine structure of the methanogen sub-community which is mainly composed of strains closely related to *Methanoculleus bourgensis* and so far non-characterised species of the genus *Methanoculleus*. In recent years, metagenome studies are often conducted by applying high-throughput sequencing technologies such as the 454-pyrosequencing method. Massively parallel shotgun sequencing of environmental DNA resulting in hundred thousands of single reads is very well suited for estimating the biodiversity of complex microbial communities (Manichanh et al., 2008). However, short read lengths are not always sufficient to allow for BLASTx detection of more-distant homologs in databases (Wommack et al., 2008). This also seems to be the case for metagenome sequence reads obtained from the biogas total community DNA, since a major proportion of the reads produced no significant BLAST hits to the NCBI nucleotide sequence database (Krause et al., 2008b; Schlüter et al., 2008). Apparently, nucleotide sequence databases only insufficiently represent genetic information from biogas-producing microbial communities. Nevertheless, analysis of short sequence reads resulting from 454-pyrosequencing provide a good estimate of the biodiversity of the biogas community, mainly because of the quantity of the analysed data set. Moreover, integrated analysis of 16S-rDNA clone library sequences and shotgun metagenome reads by means of BLAST enabled frequency of

abundance estimations for 16S-rDNA clone library sequences. So far, valid phylogenetic analyses considering phylogenetic distances of shotgun metagenome reads is problematic, since these reads usually do not represent equivalent large segments of phylogenetically evaluable marker genes with indisputable orthology. In conclusion, clone library sequences enable high resolution phylogenetic analysis of abundant taxonomic units, whereas metagenome sequence reads are more appropriate to describe the diversity of the community, but these sequences cannot be used for self-contained phylogenetics. Accordingly, an integrated analysis approach, considering clone library sequences and shotgun metagenome reads combines advantages and options of both sequence data types.

Acknowledgements

The Bioinformatics Resource Facility (BRF) team is acknowledged for their support on running bioinformatics software at the Center for Biotechnology (CeBiTec Bielefeld University). We also thank Irene Krahn for excellent technical assistance.

References

Angelidaki, I., Ellegaard, L., Ahring, B.K., 2003. Applications of the anaerobic digestion process. Adv. Biochem. Eng. Biotechnol. 82, 1–33.

Asakawa, S., Nagaoka, K., 2003. *Methanoculleus bourgensis*, *Methanoculleus olentangyi* and *Methanoculleus oldenburgensis* are subjective synonyms. Int. J. Syst. Evol. Microbiol. 53, 1551–1552.

Bayer, E.A., Lamed, R., Himmel, M.E., 2007. The potential of cellulases and cellulosomes for cellulosic waste management. Curr. Opin. Biotechnol. 18, 237–245.

Bayer, E.A., Shimon, L.J., Shoham, Y., Lamed, R., 1998a. Cellulosomes-structure and ultrastructure. J. Struct. Biol. 124, 221–234.

Bayer, E.A., Chanzy, H., Lamed, R., Shoham, Y., 1998b. Cellulose, cellulases and cellulosomes. Curr. Opin. Struct. Biol. 8, 548–557.

Bayer, E.A., Belaich, J.P., Shoham, Y., Lamed, R., 2004. The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides. Annu. Rev. Microbiol. 58, 521–554.

Chachkhiani, M., Dabert, P., Abzianidze, T., Partskhaladze, G., Tsiklauri, L., Dudauri, T., Godon, J.J., 2004. 16S rDNA characterisation of bacterial and archaeal communities during start-up of anaerobic thermophilic digestion of cattle manure. Bioresour. Technol. 93, 227–232.

Chouari, R., Le Paslier, D., Daegelen, P., Ginestet, P., Weissenbach, J., Sghir, A., 2005. Novel predominant archaeal and bacterial groups revealed by molecular analysis of an anaerobic sludge digester. Environ. Microbiol. 7, 1104–1115.

Cole, J.R., Chai, B., Farris, R.J., Wang, Q., Kulam-Syed-Mohideen, A.S., McGarrell, D.M., Bandela, A.M., Cardenas, E., Garrity, G.M., Tiedje, J.M., 2007. The ribosomal database project (RDP-II): introducing myRDP space and quality controlled public data. Nucleic Acids Res. 35, D169–172.

Cole, J.R., Wang, Q., Cardenas, E., Fish, J., Chai, B., Farris, R.J., Kulam-Syed-Mohideen, A.S., McGarrell, D.M., Marsh, T., Garrity, G.M., Tiedje, J.M., 2008. The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. Nucleic Acids Res..

Elshahed, M.S., Youssef, N.H., Luo, Q., Najar, F.Z., Roe, B.A., Sisk, T.M., Buhring, S.I., Hinrichs, K.U., Krumholz, L.R., 2007. Phylogenetic and metabolic diversity of *Planctomycetes* from anaerobic, sulfide- and sulfur-rich Zodletone Spring, Oklahoma. Appl. Environ. Microbiol. 73, 4707–4716.

Friedrich, M.W., 2005. Methyl-coenzyme M reductase genes: unique functional markers for methanogenic and anaerobic methane-oxidizing *Archaea*. Methods Enzymol. 397, 428–442.

Huang, L.N., Zhu, S., Zhou, H., Qu, L.H., 2005. Molecular phylogenetic diversity of bacteria associated with the leachate of a closed municipal solid waste landfill. FEMS Microbiol. Lett. 242, 297–303.

Huang, L.N., Zhou, H., Chen, Y.Q., Luo, S., Lan, C.Y., Qu, L.H., 2002. Diversity and structure of the archaeal community in the leachate of a full-scale recirculating landfill as examined by direct 16S rRNA gene sequence retrieval. FEMS Microbiol. Lett. 214, 235–240.

Huang, L.N., Chen, Y.Q., Zhou, H., Luo, S., Lan, C.Y., Qu, L.H., 2003. Characterization of methanogenic *Archaea* in the leachate of a closed municipal solid waste landfill. FEMS Microbiol. Ecol. 46, 171–177.

Juottonen, H., Galand, P.E., Yrjala, K., 2006. Detection of methanogenic *Archaea* in peat: comparison of PCR primers targeting the *mcrA* gene. Res. Microbiol. 157, 914–921.

Kleerebezem, R., van Loosdrecht, M.C., 2007. Mixed culture biotechnology for bioenergy production. Curr. Opin. Biotechnol. 18, 207–212.

Klocke, M., Mähnert, P., Mundt, K., Souidi, K., Linke, B., 2007. Microbial community analysis of a biogas-producing completely stirred tank reactor fed continuously with fodder beet silage as mono-substrate. Syst. Appl. Microbiol. 30, 139–151.

Klocke, M., Nettmann, E., Bergmann, I., Mundt, K., Souidi, K., Mumme, J., Linke, B., 2008. Characterization of the methanogenic *Archaea* within two-phase biogas reactor systems operated with plant biomass. Syst. Appl. Microbiol. 31, 190–205.

Krause, L., Diaz, N.N., Goesmann, A., Kelley, S., Nattkemper, T.W., Rohwer, F., Edwards, R.A., Stoye, J., 2008a. Phylogenetic classification of short environmental DNA fragments. Nucleic Acids Res. 36, 2230–2239.

Krause, L., Diaz, N.N., Edwards, R.A., Gartemann, K.H., Krömeke, H., Neuweger, H., Pühler, A., Runte, K.J., Schlüter, A., Stoye, J., Szczepanowski, R., Tauch, A., Goesmann, A., 2008b. Taxonomic composition and gene content of a methane-producing microbial community isolated from a biogas reactor. J. Biotechnol. 136, 91–101.

Lawson, P., Dainty, R.H., Kristiansen, N., Berg, J., Collins, M.D., 1994. Characterization of a psychrotrophic *Clostridium* causing spoilage in vacuum-packed cooked pork: description of *Clostridium algidicarnis* sp. nov. Lett. Appl. Microbiol. 19, 153–157.

Ley, R.E., Hamady, M., Lozupone, C., Turnbaugh, P.J., Ramey, R.R., Bircher, J.S., Schlegel, M.L., Tucker, T.A., Schrenzel, M.D., Knight, R., Gordon, J.I., 2008. Evolution of mammals and their gut microbes. Science 320, 1647–1651.

Liu, W.T., Chan, O.C., Fang, H.H.P., 2002. Characterization of microbial community in granular sludge treating brewery wastewater. Water Res. 36, 1767–1775.

Lueders, T., Chin, K.J., Conrad, R., Friedrich, M., 2001. Molecular analyses of methyl-coenzyme M reductase alpha-subunit (*mcrA*) genes in rice field soil and enrichment cultures reveal the methanogenic phenotype of a novel archaeal lineage. Environ. Microbiol. 3, 194–204.

Luton, P.E., Wayne, J.M., Sharp, R.J., Riley, P.W., 2002. The *mcrA* gene as an alternative to 16S rRNA in the phylogenetic analysis of methanogen populations in landfill. Microbiology 148, 3521–3530.

Madkour, M., Mayer, F., 2003. Structural organization of the intact bacterial cellulosome as revealed by electron microscopy. Cell. Biol. Int. 27, 831–836.

Mai, V., Greenwald, B., Morris Jr., J.G., Raufman, J.P., Stine, O.C., 2006. Effect of bowel preparation and colonoscopy on post-procedure intestinal microbiota composition. Gut 55, 1822–1823.

Manichanh, C., Chapple, C.E., Frangeul, L., Gloux, K., Guigo, R., Dore, J., 2008. A comparison of random sequence reads versus 16S rDNA sequences for estimating the biodiversity of a metagenomic library. Nucleic Acids Res. 36, 5180–5188.

McHugh, S., Carton, M., Mahony, T., O'Flaherty, V., 2003. Methanogenic population structure in a variety of anaerobic bioreactors. FEMS Microbiol. Lett. 219, 297–304.

Mladenovska, Z., Dabrowski, S., Ahring, B.K., 2003. Anaerobic digestion of manure and mixture of manure with lipids: biogas reactor performance and microbial community analysis. Water Sci. Technol. 48, 271–278.

Ohmiya, K., Sakka, K., Kimura, T., 2005. Anaerobic bacterial degradation for the effective utilization of biomass. Biotechnol. Bioprocess. Eng. 10, 482–493.

Pelletier, E., Kreimeyer, A., Bocs, S., Rouy, Z., Gyapay, G., Chouari, R., Riviere, D., Ganesan, A., Daegelen, P., Sghir, A., Cohen, G.N., Medigue, C., Weissenbach, J., Le Paslier, D., 2008. Candidatus *Cloacamonas acidaminovorans*": genome sequence reconstruction provides a first glimpse of a new bacterial division. J. Bacteriol. 190, 2572–2579.

Rastogi, G., Ranade, D.R., Yeole, T.Y., Patole, M.S., Shouche, Y.S., 2007. Investigation of methanogen population structure in biogas reactor by molecular characterization of methyl-coenzyme M reductase A (*mcrA*) genes. Bioresour. Technol..

Sasaki, K., Haruta, S., Ueno, Y., Ishii, M., Igarashi, Y., 2007. Microbial population in the biomass adhering to supporting material in a packed-bed reactor degrading organic solid waste. Appl. Microbiol. Biotechnol. 75, 941–952.

Schlüter, A., Krahn, I., Kollin, F., Bönemann, G., Stiens, M., Szczepanowski, R., Schneiker, S., Pühler, A., 2007. IncP-1-beta plasmid pGNB1 isolated from a bacterial community from a wastewater treatment plant mediates decolorization of triphenylmethane dyes. Appl. Environ. Microbiol. 73, 6345–6350.

Schlüter, A., Bekel, T., Diaz, N.N., Dondrup, M., Eichenlaub, R., Gartemann, K.H., Krahn, I., Krause, L., Krömeke, H., Kruse, O., Mussgnug, J.H., Neuweger, H., Niehaus, K., Pühler, A., Runte, K.J., Szczepanowski, R., Tauch, A., Tilker, A., Viehöver, P., Goesmann, A., 2008. The metagenome of a biogas-producing microbial community of a production-scale biogas plant fermenter analysed by the 454-pyrosequencing technology. J. Biotechnol. 136, 77–90.

Shigematsu, T., Era, S., Mizuno, Y., Ninomiya, K., Kamegawa, Y., Morimura, S., Kida, K., 2006. Microbial community of a mesophilic propionate-degrading methanogenic consortium in chemostat cultivation analyzed based on 16S rRNA and acetate kinase genes. Appl. Microbiol. Biotechnol. 72, 401–415.

Shimizu, S., Akiyama, M., Ishijima, Y., Hama, K., Kunimaru, T., Naganuma, T., 2006. Molecular characterization of microbial communities in fault-bordered aquifers in the Miocene formation of northernmost Japan. Geobiology 4, 203–213.

Shiratori, H., Ikeno, H., Ayame, S., Kataoka, N., Miya, A., Hosono, K., Beppu, T., Ueda, K., 2006. Isolation and characterization of a new *Clostridium* sp. that performs effective cellulosic waste digestion in a thermophilic methanogenic bioreactor. Appl. Environ. Microbiol. 72, 3702–3709.

Tang, Y., Shigematsu, T., Morimura, S., Kida, K., 2005. Microbial community analysis of mesophilic anaerobic protein degradation process using bovine serum albumin (BSA)-fed continuous cultivation. J. Biosci. Bioeng. 99, 150–164.

Tang, Y., Shigematsu, T., Ikbal, Morimura, S., Kida, K., 2004. The effects of microaeration on the phylogenetic diversity of microorganisms in a thermophilic anaerobic municipal solid-waste digester. Water Res. 38, 2537–2550.

Tang, Y.Q., Fujimura, Y., Shigematsu, T., Morimura, S., Kida, K., 2007. Anaerobic treatment performance and microbial population of thermophilic upflow anaerobic filter reactor treating awamori distillery wastewater. J. Biosci. Bioeng. 104, 281–287.

Wang, Q., Garrity, G.M., Tiedje, J.M., Cole, J.R., 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl. Environ. Microbiol. 73, 5261–5267.

Weidner, S., Arnold, W., Puhler, A., 1996. Diversity of uncultured microorganisms associated with the seagrass *Halophila stipulacea* estimated by restriction frag-

ment length polymorphism analysis of PCR-amplified 16S rRNA genes. Appl. Environ. Microbiol. 62, 766–771.
Weiland, P., 2003. Production and energetic use of biogas from energy crops and wastes in Germany. Appl. Biochem. Biotechnol. 109, 263–274.
Weiss, A., Jerome, V., Freitag, R., Mayer, H.K., 2008. Diversity of the resident microbiota in a thermophilic municipal biogas plant. Appl. Microbiol. Biotechnol. 81, 163–173.
Wommack, K.E., Bhavsar, J., Ravel, J., 2008. Metagenomics: read length matters. Appl. Environ. Microbiol. 74, 1453–1463.
Yadvika, Santosh, Sreekrishnan, T.R., Kohli, S., Rana, V., 2004. Enhancement of biogas production from solid substrates using different techniques—a review. Bioresour. Technol. 95, 1–10.
Zverlov, V.V., Kellermann, J., Schwarz, W.H., 2005. Functional subgenomics of *Clostridium thermocellum* cellulosomal genes: identification of the major catalytic components in the extracellular complex and detection of three new enzymes. Proteomics 5, 3646–3653.