Bioresource Technology 153 (2014) 388–392

Contents lists available at ScienceDirect

Bioresource Technology

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

Short Communication

Two amplicon sequencing strategies revealed different facets of the prokaryotic community associated with the anaerobic treatment of vinasses from ethanol distilleries

M.A. Martínez [a,b,*], H. Romero [c], N.I. Perotti [a,b]

[a] PROIMI-CONICET, Av. Belgrano y Pje. Caseros, SM de Tucumán 4000, Tucumán, Argentina
[b] Facultad de Ciencias Exactas y Tecnología, Universidad Nacional de Tucumán, Av. Independencia 1800, SM de Tucumán 4000, Tucumán, Argentina
[c] Laboratorio de Organización y Evolución del Genoma, Dpto. Ecología y Evolución, Facultad de Ciencias/CURE, Universidad de la República, Iguá 4225, Montevideo 11400, Uruguay

HIGHLIGHTS

- Diversity of a microbial consortium from anaerobic digestion of vinasses is studied.
- DNA amplified from the V4-16S region by two different sets of primers is compared.
- Each set of primers detected a limited diversity of the vinasses digesting consortia.
- One set of primers identified *Archaea* whilst the other set captured more diversity.

ARTICLE INFO

Article history:
Received 15 September 2013
Received in revised form 4 December 2013
Accepted 7 December 2013
Available online 14 December 2013

Keywords:
Anaerobic digestion
Vinasses
16S rRNA pyrosequencing
Community profiling

ABSTRACT

The prokaryotic consortium from a pilot-scale UASB reactor fed with vinasses from ethanol distilleries was evaluated by means of amplicon sequencing of the 16S rRNA gene. Two different sets of primers targeted to overlapping regions of the V4-16S region were used to gain a broad picture of such community and to perform a comparative analysis. From the two datasets obtained, prevalent phyla were *Firmicutes*, *Verrucomicrobia* and *Thermotogae*. Interestingly, one set of primers captured variability in both the bacterial and archaeal portions of the community, whilst the other one revealed a more diverse community structure, but only in the Bacteria domain. Although a certain level of agreement between the two strategies was observed, sharp differences indicate that different facets of the community were disclosed by each approach.

1. Introduction

In Northwest Argentina there are a growing number of large-scale distilleries integrated into sugar mills for alcohol production. Following molasses fermentation by yeasts, vinasses are a highly polluting waste typically produced in the range of 9–14/l of ethanol obtained and as a result, enormous amount of wastes frequently remains poorly treated in the region causing serious environmental problems. The anaerobic fermentation is an effective treatment technology since it improves both the economic and ecological results of biomass utilization (Dererie et al., 2011).

In the present work we examined the structure of the prokaryotic community performing the anaerobic digestion of a pilot scale up flow anaerobic sludge blanket type reactor (UASB) fed with vinasses by amplicon sequencing of the 16S rRNA gene, a methodology that has proven to be a valuable tool to get a glimpse of microbial communities in several environments (Liu et al., 2008). As a second goal, a comparison between two community profiles produced by two alternative strategies was performed. Indeed, two sets of primers were used to target different yet overlapping sequences of the hypervariable region V4 of the 16S rRNA gene to achieve a meaningful coverage of the prokaryotic diversity. The resulting data are thus comparable and provide an a priori similar taxonomic utility. Furthermore, among several dispersed regions of the bacterial 16S rRNA gene, the V4 region was found to be particularly informative for short reads sequencing studies (Cole et al., 2009; Bergmann et al., 2011; Ghyselinck et al., 2013).

To date, microbial consortia from a number of anaerobic digestion processes have been analyzed using next generation sequencing (NGS) platforms (Nettmann et al., 2010; Rademacher et al., 2012). However, wastes from ethanol distilleries have not been similarly examined. With the exception of a global 16S rRNA gene library of a laboratory scale reactor (España-Gamboa et al., 2012),

* Corresponding author at: PROIMI-CONICET, Av. Belgrano y Pje. Caseros, SM de Tucumán 4000, Tucumán, Argentina. Tel.: +54 3814344888.
E-mail addresses: ale.martinez.unt@gmail.com (M.A. Martínez), eletor@fcien.edu.uy (H. Romero), nperotti@herrera.unt.edu.ar (N.I. Perotti).

0960-8524/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.biortech.2013.12.030

this is the first pyrosequencing approach to evaluate the microbial consortium growing with vinasses as the sole substrate, by two complementary approaches.

2. Methods

2.1. Reactor operation and wastewater description

Sludge samples were obtained from an anaerobic digester type UASB with a working volume of 180 l. The reactor was configured for continuous operation with recirculation, thus modifying it to an expanded granular reactor with an ascending rate of 1 m h^{-1}. The inoculum used for setting up the processes was municipal household waste. The digester's working parameters were temperature 30 °C and pH 6.55. Operation was stable at a pH value of 6.55, and occasional acidification caused by volatile fatty acid (VFA) accumulation was controlled by adjusting the influent pumping. (For a detailed description of the reactor set up see the Supplementary material.)

2.2. Total community DNA preparation and amplicon sequencing of metagenomic DNA

The activated sludge samples were taken from three sampling points of the reactor (at 1.00, 2.00, and 2.50 m) and were mixed immediately after collection. After centrifugation for 15 min at 7000×*g*, the sludge pellets were washed three times with PBS (Phosphate Buffered Saline) plus EDTA 10 mM and stored at −20 °C until processing. Aliquots of the pellets were used for total community DNA extraction using the MO BIO PowerSoil® DNA Isolation Kit. DNA concentration and integrity were estimated spectrophotometrically and by gel electrophoresis. The V4 hypervariable region of the bacterial 16S rRNA gene was amplified using two sets of primers (named RK and RDP) modified for 454 pyrosequencing (Table S1). PCR amplifications and sequencing services were carried out on a Genome Sequencer FLX (Roche Applied Science) at the INDEAR genome sequencing facility (Argentina).

2.3. Sequence analyses, graphs and other statistical analyses

Analyses were performed using QIIME software package (Caporaso et al., 2010) with slight variations of the standard pipeline, using default parameters and options of the scripts unless specifically stated. In short, reads were assigned to each strategy (RK or RDP) and then were clustered into Operational Taxonomic Units (OTUs). Representative sequences were aligned using PyNast (Caporaso et al., 2010) and chimeric sequences were removed with ChimeraSlayer. Alpha diversity metrics Chao, Shannon and PD_distance, and rarefaction curves were calculated as implemented in QIIME. Taxonomical characterization of each OTU was performed with the 'RDP classifier' using the version of the Ribosomal Database Project database (Cole et al., 2009). Phylogenetic inference was performed using PhyML (Guindon et al., 2010). (For a detailed version see the Supplementary material.)

3. Results and discussion

3.1. Microbial consortium according to RDP and RK amplicon sequencing schemes

A strategy using different library construction schemes with two sets of primers was applied in order to select the methodologies to explore the prokaryotic community from an anaerobic digester fed with vinasses from ethanol distilleries. One set, RDP, is well defined by the online matching tool RDP probe match (Table S1). The second set, RK, was selected since it has shown to achieve good coverage with no significant bias against several bacterial and archaeal groups (Bates et al., 2011; Bergmann et al., 2011). The sequences amplified covered overlapping regions of the V4-16S rRNA region and produced reads with an average length of 220 and 270 bp, where RDP amplified sequences are nested within RK ones (Table S1). A total of 3048 V4-16S rRNA gene sequences were obtained after quality trimming, filtering and chimera detection. Of these, 1803 reads resulted from amplicon sequencing using RDP primers and 1246 from the RK primers. Fig. 1a and b show prevalent phyla detected in the UASB sampled. The RDP primers produced sequences that matched almost exclusively to the Bacteria domain (with the exception of 3 reads). Instead, RK primers revealed a population comprising 82% of *Bacteria* and 17.8% of *Archaea*, as was observed in other microbial communities from anaerobic digesters (Schlüter et al., 2008; Rademacher et al., 2012) and consistent with the decreased organic content of the treated waste obtained (data not shown).

Most bacterial groups were differentially detected by each strategy. However, for both datasets, prevalent phyla were *Firmicutes*, *Verrucomicrobia* and *Thermotogae*. On the other hand, *Synergistetes*, *Choloroflexi*. *Bacteroidetes* and *Proteobacteria* occurred less frequently. Minor populations, grouped as other bacteria, included *Spirochaetes*, *Panctomycetes* and MPV-15 (⩽1%) as well as *Acidobacteria*, *Amatimonadetes*, *Cyanobacteria*, *Deferribacteres*, *Elusimicrobia* and *Lentisphaerae* (Fig. 1a and b). Most of the reads were assigned at the class rank; only a small number of reads could be classified at the genus rank, which is a limitation commonly reported for 16S rRNA based NGS studies (Liu et al., 2008; Rademacher et al., 2012).

3.2. Bacterial representatives involved in the anaerobic vinasses digestion

Firmicutes was noted as the dominant phylum when total diversity is considered. However, this phylum was unequally detected, comprising 40.3% and 12.4% of assignments in the RDP and RK-based reads, respectively (Fig. 1a and b). *Clostridiales* were mostly identified while *Bacilli* were barely represented by the order *Lactobacillales*, thus seem to play a less important role in this anaerobic consortium.

An unexpectedly high abundance of the phylum *Verrucomicrobia* was observed. Its members were similarly detected by both amplicon sequencing schemes applied (36.1% and 35.0% from the RK and RDP-based reads, respectively). *Verrucomicrobia* are rarely cultured yet are commonly detected in metagenomic libraries from various environments. Among well-studied members of this phylum, different fermentation patterns leading to acetate over of propionate production were observed when *Opitutus terrae* was grown in a mixed culture with a hydrogenotrophic partner (Chin and Janssen, 2002). Accordingly, the *Verrucomicrobia* representatives might contribute to the acidogenic stages in the UASB sampled here, and some members might also establish a syntrophic association with *Archaea* for acetate oxidation. Moreover, Bergmann et al. (2011) described a higher frequency of *Verrucomicrobia* in soils analyzed using the RK primers. These observations might also explain the high abundance of *Verrucomicrobia* observed.

Thermotogae and *Synergistetes* phyla exhibited abundance rates of 11.9% and 6.7%, and 10.4% and 1.2% from the RK and RDP-based reads, respectively (Fig. 1a and b). The thermotogales comprised 8.63% of the total number of sequences. All of them were clustered in only one OUT and taxonomically related *Kosmotoga* spp. (Fig. 2). This could simply be explained by its dominance, but also by the fact that there are a limited number of cultivated species from this taxon (Ben Hania et al., 2013). The thermotogales detected might

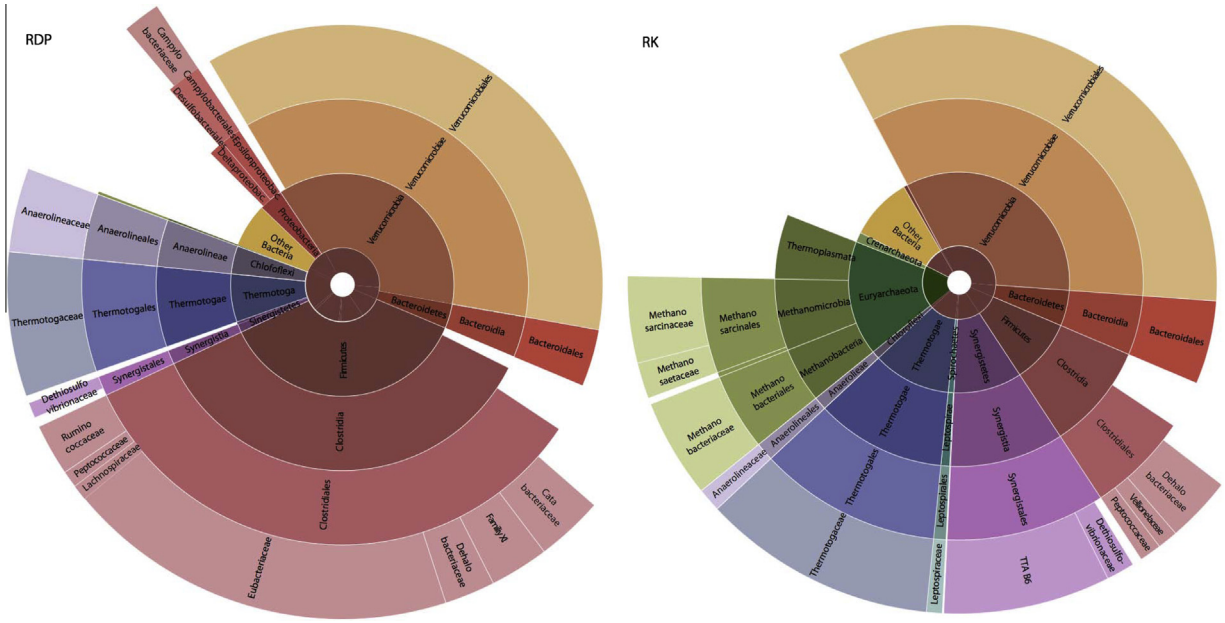

Fig. 1. Sunburst graphs of both amplicon sequencing results with abundances at family level for RDP (a) and RK-based (b) amplicon sequencing schemes. The order *Methanomicrobiales* (1.7%) is indicated as a minor group within *Methanomicrobia*.

be associated with syntrophic acetate oxidation, as was suggested in other methanogenic consortia (Sasaki et al., 2011).

It was also noticeable the presence of *Chloroflexi* with 1.2% and 3.9% (RK and RDP reads, respectively; Fig. 1a and b). Some authors have proposed that members of this phylum act as scavengers of organic compounds derived from bacterial cells in anaerobic reactors (Kindaichi et al., 2012). Considering this, the relative abundance of this group may reflect the need for a high turnover of the microbial consortium growing on vinasses from ethanol distilleries as substrate.

3.3. Archaeal sequences

The *Euryarcheota* (17.3% of total assigned reads) and *Chrenarcheota* phyla (1.5%) were almost exclusively detected with the RK-based approach (Fig. 1b). The *Thermoplasmatales* order was a prevailing representative (31.1% of all assigned archaeal reads), thus their role in this consortium might be of significant relevance and needs to be further assessed. Moreover, it has been recently proposed that uncultured *Thermoplasmatales* may represent a new order of methanogenic *Archaea*, *Methanoplasmatales* (Paul et al., 2012).

The order *Methanobacteriales* (26.5% of all archaeal reads), included *Methanobacterium* species, which utilizes H_2/CO_2 or formate to produce methane (Whitman et al., 2006). Other hydrogenotrophic methanogens belonging to the order *Methanomicrobiales* were less abundant (1.7% of RK-reads, Fig. 1b). In contrast, it has been reported that the *Methanomicrobiales* were dominant in five out of six agricultural biogas plants operated under mesophilic conditions (Schlüter et al., 2008; Nettmann et al., 2010). Nevertheless, prevalent taxa detected in the UASB reactor sampled here might suggest the presence of other candidates as hydrogenotrophic partners for a wide range of acidogenic bacteria fitting to the major phyla described in this work.

Aceticlastic methanogens belonging to the order *Methanosarcinales* were detected including both, the *Methanosaetaceae* and *Methanosarcinaceae* families (9.7% and 23.1% of all assigned archaeal reads, respectively; Fig. 1b). While *Methanosaeta* spp. are obligate aceticlastic methanogens, *Methanosarcina* contains species with extensive metabolic capabilities (Whitman et al., 2006) and might contribute to the need for a considerable metabolic plasticity in the anaerobic digestion of vinasses from ethanol distilleries.

3.4. Comparison of the RDP and RK amplicon sequencing diversity profiles

This approach can be seen as a comparison between two methodological replicates, since the sample and DNA extraction were the same, but the primers were different. Sequences were clustered into 96 and 203 Operational Taxonomic Units (OTUs) for RK and RDP respectively, using 97% sequence identity threshold. A more diverse picture of the community was captured by the RDP primer set as indicated by the alpha-diversity measures (Table S2) and rarefaction curves (Fig. S1).

Both approaches shared 48 OTUs (Fig. S2). Despite this low number of shared OTUs, when all the reads that each OTU represent are taken into account, 941 sequences from the RK set and 1318 from the RDP set were included in the intersection subset (76% and 73% of the total number of sequences, respectively). This is indicative that many of the non-shared OTUs were representing a rather small number of sequences. To gain more insight into the nature of these results, the analyses were repeated using 90% and 80% of sequence identity threshold. At the 90% identity level, the sequences clustered in 61 different RK-based OTUs and in 139 RDP-based OTUs, sharing 42 of them. Finally, at the 80% identity level, 26 and 36 OTUs were obtained (RK and RDP sets, respectively), sharing 19 of them and the number of sequences belonging to the shared OTUs increased reaching almost 90% for both primer sets. Then, the Bray–Curtis (BC) dissimilarity indexes between the two strategies were computed using the different cutoff sequence identity levels. BC index had a value of 0.54 for the 97% cutoff, 0.51 for the 90% cutoff and 0.51 using the 80% cutoff.

The same approach using now the taxonomic assignments and the corresponding partitions at different taxonomic levels was also applied. In this case, the results of the BC indexes were 0.66, 0.45 and 0.44, at Family, Class and Phylum taxonomic ranks, respectively. Interestingly, lowering the identity threshold for OTUs classification does not increase substantially the congruency of both strategies, suggesting that most of this differentially captured

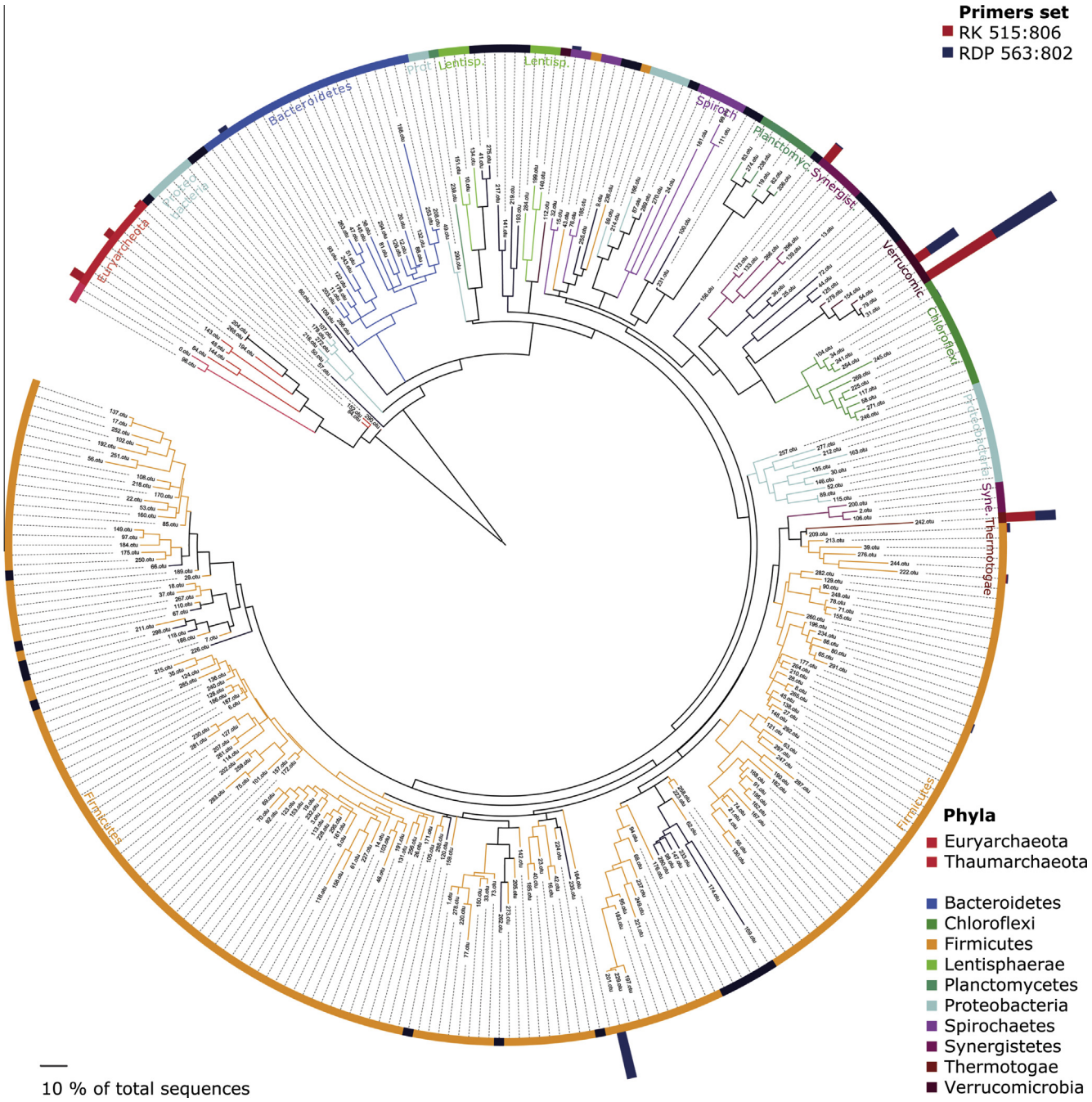

Fig. 2. Phylogenetic tree of the 299 OTUs detected for RDP (blue) and RK (red) amplicon sequencing strategies combined. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

diversity is the product of low abundance and highly divergent sequences. Then, BC indexes disregarding OTUs with only one sequence were recalculated; for the descending cutoff values 97%, 90% and 80% were 0.36, 0.34, and 0.28, respectively. Putting together these results, it could be observed that while moving towards a more coarse-grained analysis, the general picture captured by RK and RDP strategies tend to be more similar, as indicated by both classification methods, i.e. the OTU based and the taxonomy based. Nevertheless, minor differences appear in the step 97%–90% for the OTU based approach and in the step from class to phylum in the taxonomy based approach.

Notably, the amplicon sequencing RK-based approach captures variability in both the bacterial and archaeal fractions of the community, fact that might be an alternative to the use of specific primers for the examination of *Archaea*. On the other hand, the diversity indexes estimated showed that the RDP-based approach consistently displayed higher levels of diversity than the RK one (Table S2 and Fig. S2), particularly within *Firmicutes*, by detecting 'rare' sequences (i.e. low abundance and highly divergent sequences). It is important to bear in mind that the primers targeted the same region, V4-16S rRNA, being the RDP segment nested within the RK one. Therefore the differences observed are not likely to be the product of the intrinsic information content of the sequences but from differential amplification of the community DNA by each primer set.

The taxonomic profile obtained and their metabolic role was briefly discussed here since rarefaction curves were non-asymptotic (Fig. S1), which suggest that the richness in the sample

require additional sequencing efforts. Additionally, active members of methanogenic communities are selected by different substrates and operating conditions, allowing to microbial taxa that are lower in abundance may emerge to become active and prevalent in a particular niche. The DNA pyrosequencing based analyses do not reflect them accurately, thus further comparative studies over 16S genes in the DNA as well as in the RNA can overcome such limitations (Lu et al., 2013).

4. Conclusion

The prokaryotic consortium from a pilot-scale UASB reactor was evaluated by means of two different amplicon sequencing schemes directed to the V4-16S rRNA domain, thus contributing to the studies that examine alternative methodological approaches based on high-throughput amplicon sequencing of total DNA. As a result, 299 OTUs were detected (using 97% identity threshold), where dominant phyla growing on vinasses as substrate were *Firmicutes*, *Verrucomicrobia* and *Thermotogae*. Interestingly, one of the approaches captured the archaeal representatives of the community, where *Thermoplasmatales* were as much abundant as *Methanobacteriales* and *Methanosarcinales*.

Acknowledgements

Authors thank MINCyT for Grant PFIP-ESPRO 2010-164/10, Argentina and the Faculty programme of the Asociación de Universidades del Grupo Montevideo.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.biortech.2013.12.030.

References

Bates, S.T., Berg-Lyons, D., Caporaso, J.G., et al., 2011. Examining the global distribution of dominant archaeal populations in soil. ISME J. 5, 908–917.

Ben Hania, W., Postec, A., Aüllo, T., et al., 2013. *Mesotoga infera* sp. nov., a novel mesophilic member of the order thermotogales, isolated from an underground gas storage in France. Int. J. Syst. Evol. Microbiol. 63, 3003–3008.

Bergmann, G.T., Bates, S.T., Eilers, K., et al., 2011. The under-recognized dominance of *Verrucomicrobia* in soil bacterial communities. Soil Biol. Biochem. 43, 1450–1455.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., et al., 2010. QIIME allows analysis of high-throughput community sequencing data. Nat. Methods 7, 335–336.

Chin, K.J., Janssen, P.H., 2002. Propionate formation by *Opitutus terrae* in pure culture and in mixed culture with a hydrogenotrophic methanogen and implications for carbon fluxes in anoxic rice paddy soil. Appl. Environ. Microbiol. 68, 2089–2092.

Cole, J.R., Wang, Q., Cardenas, E., et al., 2009. The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. Nucleic Acids Res. 37, D141–D145 (Data issue).

Dererie, D.Y., Trobro, S., Momeni, M.H., Hansson, H., Blomqvist, J., 2011. Improved bio-energy yields via sequential ethanol fermentation and biogas digestion of steam exploded oat straw. Bioresour. Technol. 102, 4449–4455.

España-Gamboa, E., Mijangos-Cortés, J.O., Hernández-Zárate, G., et al., 2012. Methane production by treating vinasses from hydrous ethanol using a modified UASB reactor. Biotechnol. Biofuels 5, 82.

Ghyselinck, J., Pfeiffer, S., Heylen, K., et al., 2013. The effect of primer choice and short read sequences on the outcome of 16S rRNA gene based diversity studies. PLoS One 8 (8), e71360.

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59, 307–321.

Kindaichi, T., Yuri, S., Ozaki, N., Ohashi, A., 2012. Ecophysiological role and function of uncultured *Chloroflexi* in an anammox reactor. Water Sci. Technol. 66, 2556–2561.

Liu, Z., DeSantis, T.Z., Andersen, G.L., Knight, R., 2008. Accurate taxonomy assignments from 16S rRNA sequences produced by highly parallel pyrosequencers. Nucleic Acids Res. 36 (18), e120.

Lu, X., Rao, S., Shen, Z., Lee, P.K., 2013. Substrate induced emergence of different active bacterial and archaeal assemblages during biomethane production. Bioresour. Technol. 148, 517–524.

Nettmann, E., Bergmann, I., Pramschüfer, S., Mundt, K., Plogsties, V., Herrmann, C., Klocke, M., 2010. Polyphasic analyses of methanogenic archaeal communities in agricultural biogas plants. Appl. Environ. Microbiol. 76, 2540–2548.

Paul, K., Nonoh, J.O., Mikulski, L., Brune, A., 2012. "*Methanoplasmatales*". *Thermoplasmatales*-related *Archaea* in termite guts and other environments are the seventh order of methanogens. Appl. Environ. Microbiol. 78, 8245–8253.

Rademacher, A., Zakrzewski, M., Schlüter, A., et al., 2012. Characterization of microbial biofilms in a thermophilic biogas system by high-throughput metagenome sequencing. FEMS Microbiol. Ecol. 79, 785–799.

Sasaki, D., Hori, T., Haruta, S., Ueno, Y., Ishii, M., Igarashi, Y., 2011. Methanogenic pathway and community structure in a thermophilic anaerobic digestion process of organic solid waste. J. Biosci. Bioeng. 111, 41–46.

Schlüter, A., Bekel, T., Diaz, N.N., et al., 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.

Whitman, W.B., Bowen, T.L., Boone, D.R., 2006. The methanogenic bacteria. Prokaryotes 3, 165–207.