Microbial community structure in a full-scale anaerobic treatment plant during start-up and first year of operation revealed by high-throughput 16S rRNA gene amplicon sequencing

Else Marie Fykse, Tone Aarskaug, Elisabeth H Madslien, Marius Dybwad

PII: S0960-8524(16)31385-2
DOI: http://dx.doi.org/10.1016/j.biortech.2016.09.118
Reference: BITE 17141

To appear in: *Bioresource Technology*

Received Date: 4 July 2016
Revised Date: 26 September 2016
Accepted Date: 28 September 2016

Please cite this article as: Marie Fykse, E., Aarskaug, T., Madslien, E.H., Dybwad, M., Microbial community structure in a full-scale anaerobic treatment plant during start-up and first year of operation revealed by high-throughput 16S rRNA gene amplicon sequencing, *Bioresource Technology* (2016), doi: http://dx.doi.org/10.1016/j.biortech.2016.09.118

Microbial community structure in a full-scale anaerobic treatment plant during start-up and first year of operation revealed by high-throughput 16S rRNA gene amplicon sequencing

Else Marie Fykse*, Tone Aarskaug, Elisabeth H Madslien, and Marius Dybwad

Norwegian Defence Research Establishment (FFI)

P.O. box 27, N-2027 Kjeller, Norway

*Corresponding author

Else Marie Fykse

Norwegian Defence Research Establishment (FFI), P. O Box 25, N-2027 Kjeller, Norway

Abstract

High-throughput amplicon sequencing of six biomass samples from a full-scale anaerobic reactor at a Norwegian wood and pulp factory using Biothane Biobed Expanded Granular Sludge Bed (EGSB) technology during start-up and first year of operation was performed. A total of 106,166 16S rRNA gene sequences (V3-V5 region) were obtained. The number of operational taxonomic units (OTUs) ranged from 595 to 2472, and a total of 38 different phyla and 143 families were observed. The predominant phyla were Bacteroidetes, Chloroflexi, Firmicutes, Proteobacteria, and Spirochaetes. A more diverse microbial community was observed in the inoculum biomass coming from an Upflow Anaerobic Sludge Blanket (USAB) reactor, reflecting an adaptation of the inoculum diversity to the specific conditions of the new reactor. In addition, no taxa classified as obligate pathogens were identified and potentially opportunistic pathogens were absent or observed in low abundances. No *Legionella* bacteria were identified by traditional culture-based and molecular methods.

Keywords

Microbial community structure; anaerobic treatment plant; high-throughput sequencing; 16S rRNA gene

1 Introduction

The aeration ponds of the biological treatment plant at a wood and pulp factory operated by Borregaard Ind. Ltd., a world's leading supplier of lignin-based chemicals, were shut down in September 2008 due to high concentrations of *Legionella* spp. bacteria in the ponds (Blatny et al., 2008; Nygård et al., 2008; Olsen et al., 2010). Aeration ponds in activated sludge plants may provide a favorable growth environment for pathogenic microorganisms, including *Legionella* spp. and freshwater amoeba which can serve as host organisms for amoeba-resistant *Legionella* spp. (Fykse et al., 2014; Olsen et al., 2010). A concern when operating biological treatment plants is therefore the potential growth of pathogenic *Legionella* spp. and other microbial pathogens, and in particular the release of pathogens to the open environment through discharge of liquid effluent (e.g. to rivers) or aerosolization to the atmosphere (Blatny et al., 2011; Fykse et al., 2013).

In 2013, a new full-scale closed anaerobic treatment plant was established at Borregaard to reduce the potential for future releases of *Legionella* spp. and other potential human pathogens into the open atmosphere. The anaerobic treatment plant also has other advantages such as reduced discharge of organic material to the nearby river Glomma and allowing waste to be used for biogas production. The full-scale anaerobic reactor is based on the Biothane Biobed Expanded Granular Sludge Bed (EGSB) technology.

Anaerobic reactors contain highly complex microbial communities which play critical roles in the degradation processes which mainly proceed in four metabolic steps: hydrolysis, fermentation, acetogenesis, and methanogenesis (Zinder, 1984). Multiple types of bacteria, including *Eubacteria* and *Archaea*, contribute in a complex chain of events. During anaerobic fermentation, large organic molecules are broken down into hydrogen

and acetic acid, which can be used in methanogenic respiration (Sandoval Lozano et al., 2009). During the initial startup of an anaerobic reactor it is common to seed the reactor with biomass from another reactor to establish a microbial community that are able to quickly start digestion as a response to wastewater infeed (Ahring, 2003) to avoid buildup of harmful intermediate such as for example volatile fatty acids (VFA) (acidosis), which inhibits the methane production (Goux et al., 2015; Nelson et al., 2012). A link between microbial diversity and process robustness has been shown (Goux et al.,2015; Sundberg et al., 2013; Werner et al., 2010; Ziganshin et al., 2013) and knowledge about the microbial communities in anaerobic reactors may lead to increased awareness of factors that are important for efficient/optimal and stable operation as well as increased understanding of the potential for amplification and release of pathogenic microorganisms to the environment.

The microbial communities of anaerobic reactors have in the past been investigated using various methods, including: conventional microbiological methods (Roest et al., 2005), Sanger sequencing of 16S rRNA gene cloning libraries, denaturing gradient gel electrophoresis, and fluorescent *in situ* hybridization (Nelson et al., 2011; Pereira et al., 2006; Rivière et al., 2009). It is, however, assumed that only a small fraction of the microbial community present in an anaerobic reactor can be cultured under laboratory conditions (Amann et al. 1996). High throughput sequencing has therefore gained popularity in recent years as these techniques allow for a deeper and more convenient analysis of complex microbial communities without the need for time-consuming and labor-intensive Sanger sequencing and cloning processes or culture analysis. High-throughput sequencing has been used in several metagenomics studies to characterize the

microbial community in anaerobic reactors, both as high-throughput amplicon sequencing of the 16S rRNA gene (Li et al., 2013; Shu et al., 2015; Sundberg et al., 2013; Wong et al., 2013) and direct high-throughput shotgun sequencing (Guo et al., 2015). From these studies, sequences belonging to the phyla Proteobacteria, Bacteroidetes, Firmicutes and Chloroflexi were most abundantly identified. Deeper understanding of the microbial community structure in full-scale anaerobic reactors, including community dynamics and adaptation in response to environmental changes such as changes in wastewater composition, is important to optimize the performance of anaerobic treatment plants (Goux et al., 2015).

Few studies using high throughput sequencing to characterize the microbial community dynamics during the start-up phase of an anaerobic reactor have been performed (Goux et al., 2016; Solli et al., 2014). The main objective of the present study was to characterize the microbial community in a full-scale anaerobic reactor (EGSB) at a Norwegian wood and pulp factory using high throughput amplicon sequencing. The V3-V5 hypervariable region of the bacterial 16S rRNA gene was used to characterize the microbial community in the biomass used to seed the anaerobic reactor (inoculum) and comparison to the microbial community in the reactor and discharged effluent after six as well as 12 months of operation, respectively. Bioinformatic analysis was used to examine the alpha diversity and the taxonomy of the samples. Additionally, the secondary objective of this work was to investigate the potential presence (growth/amplification) of human pathogenic microorganisms with emphasis on *Legionella* spp. in the anaerobic treatment plant.

2 Material and methods

2.1 Wastewater treatment plant and sample collection

The surveyed wastewater treatment plant was a full-scale anaerobic treatment plant operated by Borregaard located in Sarpsborg, Norway. The anaerobic reactor was based on Biothane Biobed EGSB technology. Two wastewater streams from Borregaard's production process, condensates from ethanol and alkacell production, are combined in a buffer tank before fed into a conditioning tank. The organic content of the ethanol condensate was as mainly acetic acid, ethanol, methanol, aldehydes, and amyl alcohols. The organic content of the alkacell condensate was identified as mainly methanol, ethanol, and terpenes. In the conditioning tank wastewater was mixed with recycled treated anaerobic effluent from a recirculation step to attain a homogenous substrate composition and to avoid overloading the microorganisms at the bottom of the reactor with substrate. The pH-value and the temperature of the raw wastewater were regulated in the pretreatment step in the conditioning tank. The reactor effluent was filtered with Hydrotech disc filters to remove suspended solids before discharge into the river Glomma (Case study Borregaard,)

To initiate the treatment plant, the reactor was inoculated with biomass delivered as 83 truckloads from a mesophilic Upflow Anaerobic Sludge Blanket (UASB) reactor (Norske Skog Saugbrugs, Halden, Norway). Biomass samples (500 ml) from nine random truckloads were collected and transported to the laboratory during startup of the anaerobic reactor (1-10 March 2013). The samples were homogenized and pooled together to make two samples (Inoculum 1, IN1, from truckloads 1, 11, 21, 31, and 41, and Inoculum 2, IN2, from truckloads 51, 61, 71, and 81) (Table 1). The samples were stored frozen (-20 °C) until further processing and DNA extraction. Biomass samples were also collected from the

anaerobic reactor (R) and the discharged effluent (E) after six (24 September 2013, sample id R6M and E6M) and 12 (20 May 2014, sample id R12M and E12M) months of operation (Table 1). These samples (R6M, E6M, R12M, and E12M) were processed and stored as described above for the IN1 and IN2 samples. Since the anaerobic reactor was fitted with an external conditioning tank providing continuous circulation, the reactor content was considered to be well-mixed and homogenous, and thus the collected biomass samples were considered to be representative.

2.2 DNA extraction

Total DNA was isolated from biomass samples using the Fast DNA Spin Kit for Soil (MP Biomedicals, Solon, OH, USA). Briefly, the samples were thawed and mixed before 400 µl was subjected to bead beating (1 min, 3200 rpm) in a Mini-Beadbeater-8 (BioSpec Products, city, country). DNA isolation was further performed as described by the kit manufacturer. Purified DNA samples were quantified using the PicoGreen dsDNA Assay Kit (Thermo Scientific, city, USA).

2.3 PCR amplification and 454-pyrosequencing

The purified DNA samples were sent to Eurofins MWG Operon (Ebersberg, Germany) for PCR amplification of the bacterial 16S rRNA gene and high-throughput amplicon sequencing. Briefly, the DNA samples were amplified in three independent PCR reactions with primers spanning the V3-V5 hypervariable region of the 16S rRNA gene. The sequence of the forward primer was (338F) 5'-ACTCCTACGGGAGGCAGCAG (Huse et al., 2008) and the reverse primer (907R) 5'-CCGTCAATTCMTTTRAGTTT (Muyzer et al., 1998). The PCR amplicons were pooled and sequenced on a Roche GS

Junior 454-pyrosequencing instrument (Roche diagnostics, Basel, Switzerland) using the titanium chemistry.

2.4 Data analysis

The raw sequence read sets were de-multiplexed and analyzed using BioNumerics (v7.6, Applied Maths, Belgium). The quality control process included barcode and primer trimming, chimera checking, and filtering of sequence reads containing homo-polymer stretches (>7 bases). Sequence reads failing to pass one of the quality criteria were removed. The quality-controlled sequence reads for each sample (106,166) were subjected to a diversity analysis and taxonomic classification in Bionumerics, which incorporates the Mothur analysis toolbox (Schloss et al., 2009). Briefly, the sequence read sets were aligned using the SILVA ribosomal database (SSU Ref NR 115 containing 479,726 classified entries, http://www.arb-silva.de) and clustered into operational taxonomic units (OTUs) at 97% sequence identity, followed by calculation of various diversity (richness, evenness, coverage) estimators including plotting of rarefaction and rank abundance curves and calculation of Good's sampling coverage as well as the Shannon diversity index, Inverse Simpson index, Shannon-based measure of evenness, Chao1 richness estimator, and Abundance Coverage-based Estimator (ACE) for each sample. Taxonomic classification of the sequence reads was also performed using the SILVA ribosomal database (SSU Ref NR 115). The classification results were used to calculate relative abundances at the phylum and family level for each sample, and these results were further used to describe and compare the observed microbial diversity and community structure.

2.5 *Legionella* analysis

The presence of *Legionella* spp. or *Legionella pneumophila* was investigated by real-time PCR and co-culture experiments with amoebae as described in Fykse et al., 2013, 2014, and growth analysis was performed according to the ISO11731 standard.

3 Results and discussion

3.1 Operational performance of the full-scale anaerobic reactor

This full scale anaerobic reactor was fed with two wastewater streams from the production processes at Borregaard Ind. Ltd. (condensates from ethanol production and alkacell production, respectively) as described in material and methods. The composition of the wastewater streams was kept relatively constant throughout the study period. In general, the ethanol production condensate contributed about 80% (Table 2) throughout the entire test period. On average the Chemical Oxygen Demand (COD) of the material fed into the reactor varied from 37000 to 59000 kg per day during the sampling period and the degradation efficiency was in average 70%. In general, the amount of volatile fatty acids measured (VFAs) in the anaerobic effluent was low (< 65 mg/L), which is positive for the reactor performance since VFAs intoxication can lead to loss in biogas production (Goux et al., 2015). The reactor temperature was kept around 36-37 °C, i.e. a typical mesophilic digestion process, while the pH was kept around 7. If necessary, NaOH was added to the conditioning tank to increase pH. The daily biogas production was varied from about 14000 to19000 m^3/day and the methane gas accounted for about 80 % after six months and 70 % after 12 months of operation. More detailed process parameters are shown in Table 2, which have been obtained from the records of the plant.

3.2 Composition of microbial community

In this work, a total of 106,202 raw sequence reads was generated by high-throughput amplicon sequencing of the 16S rRNA gene (V3-V5 hypervariable region) in six biomass samples collected during the start-up and first year of operation of an anaerobic wastewater treatment plant (Table 1). After quality filtering, which resulted in a <0.1% rejection rate, a total of 106,166 sequences with an average read length of 556 bp were retained for further analysis (Table 1). Alignment and clustering of the sequences showed that the number of OTUs based on a 97% sequence identity threshold ranged from 593 to 2472 for the six samples (IN1, IN2, E6M, R6M, E12M, and R12M), with the highest number in IN1 and IN2, i.e. the inoculum used to initiate the anaerobic reactor (Table 1). The number of OTUs after six and 12 months operation were 2-4 times lower compared to the inoculum, suggesting that the richness had decreased from start-up to six months of operation and then remained relatively stable from six to 12 months of operation. These results show that the OUT richness has been stabilized in the reactor and in the effluent. The rarefaction curves approached the plateau phase for all samples, suggesting that that the sequencing depth was sufficient to capture the majority of the bacterial diversity, including a large fraction of the rare diversity (Fig. 1). At the same time the rarefaction curves also showed that the observed richness was higher in the IN1 and IN2 samples compared to E6M, R6M, E12M, and R12M samples (Fig. 1). This finding was also supported by the calculated Shannon, Inverse Simpson, Chao1 and ACE estimators which were higher for IN1 and IN2 compared to the E6M, R6M, E12M, and R12M samples, further suggesting that the bacterial diversity was higher in the inoculum used to initiate the anaerobic reactor than in the reactor and the effluent after six and 12 months of operation

(Table 1). The calculated Good's sampling coverage for all samples suggested that the coverage was sufficient to capture most of the bacterial diversity, including a large fraction of the rare diversity (Table 1). The calculated sampling coverage for the samples E6M, R6M, E12M, and R12M were highly similar when compared to each other (0.97-0.98). However, the sampling coverage of IN1 and IN2 (0.93-0.94) was reduced when compared to the other samples. This indirectly supports the observation that the diversity of the bacterial community was higher in the inoculum samples. The rank abundance curves showed that the anaerobic reactor and discharged effluent samples after 12 months of operation (R12M and E12M, respectively) appeared to be dominated by a single highly abundant OTU when compared to the IN1, IN2, E6M, and R6M samples (Fig. S1). However, since the observation only applied to a single OTU, none of the diversity estimators, including the Shannon-based measure of evenness, seemed to be noticeably affected (Table 1).Such loss in diversity can potentially occur if the inoculum contained aerobic microbes, not surviving under anaerobic condition over time in the reactor. A decrease in the microbial diversity was also observed in a different study investigating the microbial community in a farm anaerobic reactor during the start-up phase (Goux et al., 2016).

The use of high-throughput sequencing has revolutionized the investigation of complex microbial communities in anaerobic wastewater treatment plant, by increasing the completeness (i.e. insight into the rare diversity) and reducing the costs and labor needs associated with diversity analyses. The increased completeness offered by high-throughput sequencing can be exemplified by comparing results from a microbial community

investigation in an anaerobic reactor using the traditional 16S rRNA gene cloning and Sanger sequencing methodology, which revealed only 69 OTUs (Rivière et al., 2009), to the results obtained using high-throughput 16S rRNA gene amplicon sequencing in this study, which revealed between 593 to 2472 OTUs (Table 1). The microbial diversity observed in this study supports work undertaken by others that also have used high-throughput 16S rRNA gene sequencing to study anaerobic reactors, and which reported between 700 to 9000 OTUs (Lee et al., 2012; Li et al., 2013; Shu et al., 2015; Werner et al., 2010). This study present microbial community data obtained using high-throughput sequencing of the inoculum biomass coming from a paper plant used to seed the reactor in a wood and pulp factory and comparing this to the microbial community structure after six and 12 months of operation. This study showed that the microbial richness of the inoculum was decreased after six months and then remained relatively stable from six to 12 months of operation of the reactor.

3.3 Taxonomic classification of the microbial communities

To better understand the microbial community structure in the anaerobic treatment plant, including changes over time, the sequence reads were taxonomically classified and the relative abundances calculated at the phylum and family level for each sample. The results are visualized in Fig. 2 and 3, respectively, and presented in full in supplementary material (Table S1). A total of 38 different phyla were observed in this study (Table S1). The number of observed phyla per sample was 30 in IN1, 32 in IN2, 31 in E6M, 27 in R6M, 36 in E12M, and 28 in R12M (Table S1). The results showed that the most abundant phyla in IN1 and IN2 were Bacteroidetes (30 and 34%, respectively) followed by Firmicutes (17.4 and 21.2%), Proteobacteria (17 and 19%) and Chloroflexi (9 and 9%)

(Table S1, Fig.2). Previous studies also identified these phyla to be the most abundant (Beale et al., 2016; Guo et al., 2015; Nelson et al., 2011; Shu et al., 2015; Sundberg et al., 2013; Ye and Zhang, 2013) in anaerobic reactors processing various wastewater such as sewage, food, pulp and paper. In reactors digesting sewage sludge the Spirochetes was one of the most prevalent phyla (Sundberg et al., 2013). Also in this study (waste water from a wood and pulp factory) the phylum Spirochetes was also identified as a major phylum (8 and 6.6%) (Fig. 2, Table S1). A total of 143 different families were observed and determined in six samples. The number of observed families per sample was 73 in IN1, 62 in IN2, 67 in E6M, 61 in R6M, 89 in E12M, and 61 in E12M (Table S1).The families represented at >2 % frequency (22 families) in at least one sample are shown in Fig. 3

A major observation was that the relative abundance of the phylum Bacteroidetes had decreased from 30-34 % in the inoculum to 5-7% after six months. However, after 12 months the relative abundance had again increased to >30% (Fig. 2). Bacteroidetes, a diverse group of Gram-negative bacteria, are commonly observed in anaerobic digestion processes (Beale et al., 2016; Guo et al., 2015; Nelson et al., 2011; Shu et al., 2015; Sundberg et al., 2013).The dominant class within the Bacteroidetes was Bacteroidia, whereas the Sphingobacteria accounted for < 10% (Table S1). In the inoculum (IN1, IN2) relative abundance of the family Prevotellaceae, belonging to the class Bacteroidia, was dominating (13 and 20%). However, after six and 12 months this family was almost absent and another family, the Porphyromonadaceae, was dominating ((Table S1, Fig. 3). A meta-analysis of microbial diversity in anaerobic reactors also identified Bacteroidetes as one of four dominating phyla (Nelson et al., 2011). The Bacteroidetes bacteria are supposed to

play a critical role in starch and papermaking wastewater treatment plants (Shu et al., 2015) containing saccharolytic chemoorganotrophic heterotroph bacterial groups. Bacteroidetes are supposed to be involved in breakdown of polysaccharides and fermentation processes (Garrity et al., 2005).

The phylum Firmicutes containing syntrophic bacteria (Morris et al., 2012), which in general plays an important role in degradation of VFAs, are commonly detected in anaerobic digesters. Degradation of VFAs is important to avoid overloading the reactor with VFAs, which inhibit methane production (Goux et al., 2015). The predominant class of Firmicutes identified in this study was Clostridia (>90%) (Table S1). Clostridia are a highly versatile class of Gram-positive anaerobic endospore-forming bacteria, including several species capable of degrading proteins, lipids, and polymeric carbohydrates and they are capable to degrade biomass and fermentation of the wastewater (Lynd et al., 2002). The relative abundance of Clostridia accounted for 15-23 % of the sequences in all samples except in the E6M sample where the Clostridia accounted for about 30 % (Table S1). Three families belonging to the order Clostridiales were dominating, the Ruminococcaceae in all samples, the Christensenellaceae in E6M (16 %) and the Lachnospiraceae in R6M (20 %) (Fig.3), dominated by the genus *Acetitomaculum an* acetogenic bacteria (Fig. S2, Table S1). Clostridia have also been identified as a major class in microbial communities in anaerobic reactors in previous studies (Beale et al., 2016; Guo et al., 2015; Nelson et al., 2011; Shu et al., 2015; Sundberg et al., 2013). The class also contains highly pathogenic bacteria such as bacteria belonging to the family Clostridiaceae and genus *Clostridium*. However, the relative abundance of this family in this study was extremely low (<0.2%)

(Table S1). A similar abundance of clostridia in all samples indicate that a stable microbial community has been established in the reactor. In general, methanogenic microorganisms can degrade primarily only one-carbon compounds. Therefore, acetate, propionate, ethanol, and their higher homologs have to be fermented to one-carbon compounds by syntrophic fermentations before methanogenesis by the Archeae.

Proteobacteria is another dominating phylum evenly distributed in all samples analyzed in this study. This phylum is one of the four dominating phyla previously identified in anaerobic treatment plants (Beale et al., 2016; Guo et al., 2015; Nelson et al., 2011; Shu et al., 2015; Sundberg et al., 2013). Relative abundance of this phylum accounted for 19, 17, 5, 11, 12 and 15 % respectively, in all samples (Table S1, Fig. 2), confirming its ubiquitous presence in such environments. All five classes within the phylum Proteobacteria were observed; however, the relative abundance of Alphaproteobacteria was very low (<0.4%) (Table S1), and the Betaproteobacteria was mainly represented in the inoculum (Table S1). The dominating Proteobacteria was the class Deltaproteobacteria accounting for in average 30 % of the Proteobacteria. Together, Delta, Epsilon and Gammaproteobacteria accounted for > 90 % of the Proteobacteria in the six and 12 months samples (Table S1). The distribution of the Deltaproteobacteria at family level was different in the various samples. Desulfomicrobiaceae, a family within the order Desulfovibrionales, were dominating after six and 12 months operation (Table S1, Fig. 3). This is a family of bacteria defined by a wide range of morphological and chemotaxonomic properties capable of syntrophic growth. In general, they are strictly anaerobic, having a respiratory type or fermentative type of metabolism. Members are either mesophilic or

moderately thermophilic sulfate-reducing bacteria (Kuever et al., 2005). In the IN1 and IN2 samples the family Syntrophaceae and Bacteriovoracaceae were the dominating families (Table S1, Fig. 3). The Syntrophaceae are strictly anaerobic bacteria. Campylobacteraceae, a family belonging to the Epsilonproteobacteria, were present in the 12 months samples (2 and 7% in the effluent and reactor, respectively), and in the other samples only as minor families (Table S1, Fig. 3). Bacteria within the Campylobacteraceae family were classified to the genus *Arcobacter* (Fig, S2, Table S1), which has a wide environmental distribution, e.g. water, animals, and food products, and consist of more than 19 identified species (Hsu and Lee, 2015). One of these species, *Arcobacter butzleri,* is considered an opportunistic emerging pathogen and has been linked to enteritis (Prouzet-Mauléon et al., 2006). The highest abundance of *Arcobacter* was identified in the reactor (7%), while a lower abundance (2%) was measured in the discharged effluent. The family Pseudomonadaceae, which might contain potential opportunistic pathogenic species such as *Pseudomonas aeruginosa*, was identified at very low abundance in the reactor (average <0.1%) and were absent from the discharged effluent after 12 months (Table S1). The relative abundance of (< 0.2 %) bacteria in the family Enterobacteriales indicating a low recovery rate from the reactor and the effluent of enteric bacteria that include many species of common human enteric pathogens. Bacteria in the order Burkholderiales, which contain potential human pathogenic species, were observed at low abundance in the inoculum (3.3%) and present at extremely low abundance in the six and 12 months samples (0.2% in E12M) (Table S1).

Another main observation of this work is that the abundance of bacteria classified to the phylum Chloroflexi, was increased from 9 % in the IN1 and IN2 samples to

approximately 20 % after six and 12 months of operation, indicating an adaption of the inoculum microbial community to the condition in the reactor after six and 12 months caused by the wastewater fed into the reactor (Table S1, Fig. 2). The dominating family was the Anaerolineaceae, which accounted for about 40 % in the E6M sample (Table S1, Fig.3). The Chloroflexi is taxa of bacteria usually associated with biological treatment of pulp material and has been identified through multiple studies investigating the microbial diversity in digested sludge from municipal wastewater treatment plants (Chouari et al., 2005; Rivière et al., 2009; Yamada and Sekiguchi, 2009).

Bacteria within the phylum Spirochaetes are frequently detected in anaerobic digestion systems that treat municipal sludge, sewage sludge, industrial wastewater, livestock wastewater, and even synthetic organic matters (Lee et al., 2012; Sundberg et al., 2013). However, their proportions are reported to be quite varied and abundance from 4 to 47% has been reported (Lee et al., 2012). The anaerobic species of Spirochaetes, both facultative and obligate anaerobes, grow abundantly whenever cellobiose, which is a product when cellulose is degraded by other organisms, is available to them and the bacteria have been associated with glucose and acetate uptake (Leschine et al., 2006). The relative abundance of Spirochaetes in this study varied between 6-8% in all samples except for the R6M sample where the abundance accounted for more than 13% (Table S1, Fig. 2). For the phylum Spirochaetes, the individual taxonomic classification indicated that taxa with known human pathogenic members such as *Leptospira*, *Treponema* and *Borrelia* species were not present (Table S1).

The most abundant archaea family represented in this study was the Methanosaetaceae important for methanogenesis. The observed abundance was very low (<0.2%) (Table S1). An explanation is that amplicon sequencing of bacterial and archaea 16S rRNA genes are biased because of a potential low efficiency of the 16 S rRNA amplification of archaea compared to bacteria (Baker et al., 2003). Other studies using next generation sequencing to study the diversity of anaerobic reactors showed that the sequences affiliated to Archaea in general constituted less than 10 % (Sundberg et al., 2013; Guo et al., 2015). However, in some reactors the abundance was higher (Sundberg et al., 2013). Another reason for the low abundance can be that pyrosequencing will detect only those fragments present above a certain threshold, which is dependent on the sequencing depth of the sample.

In this study, no *Legionella* bacteria were identified in the inoculation biomass or in the six and 12 months samples using microbial growth method, co-culture with amoebae and real-time PCR (results not shown).

3.4 Differences in microbial community structure

In this study a more diverse microbial population was observed in the IN1 and IN2 samplescompared to the samples collected after six and 12 months operation of the reactor. The substrate or/and process parameters of the reactor have promoted a shift in the microbial population (Table S1, Fig. 2, 3, S2). The E6M, R6M, E12M and R12M samples were dominated by few families belonging to the phyla Bacteroidetes and Chloroflexi, and the classes Clostridia and Deltaproteobacteria (Table S1, Fig.2, 3). The increased presence in the six and 12 months samples of strictly anaerobic bacteria such as the

Deltaproteobacteria, capable of syntrophic growth (Desulfovibrionales) reflects an environment in the reactor for effective methane production. Several factors may influence the community structure in an anaerobic treatment plant, including the type and rate of wastewater feed material, reactor oxygen status (i.e. degree of anaerobicity), pH, temperature, and several others (Nelson et al., 2012; Sundberg et al., 2013). In this work the pH (7) and temperature (35-37 °C) were kept constant during the study period (12 months) and the composition of the infeed to the reactor was also reasonable stable with ethanol production condensates from wood and pulp being the major wastewater infeed. A main observation in this study is that the abundance of Chloroflexi bacteria in the six and 12 months samples has increased whereas the abundance of Bacteroidetes decreased by more than 20 % in the six months samples compared to the inoculum. The general high abundance of Clostridia in all samples indicates a stable anaerobic environment in the reactor throughout the study period. The low level of VFAs indicates that the reactor performs well with optimal processes for anaerobic degradation of the organic waste from the wood and pulp factory for biogas production. High abundances of Bacteroidetes and Firmicutes are in general associated with stable operational conditions in a reactor. Therefore, the increase in Bacteroidetes to more than 30 % after 12 months might play an important role for stabilization of the reactor and production of precursor molecules for biogas production. Chloroflexi species have been identified in digesters fermenting different sugars into acetate and other short chain fatty acids precursors for methane production (Yamada and Sekiguchi, 2009). The increased abundance of Chloroflexi in the six and 12 months samples indicates a microbial community in the anaerobic reactor beneficial for biogas and methane production reflected by the high and efficient methane

production in this reactor. In general, this work supports previous studies showing that microbial composition of a reactor is mainly governed by the substrate differences and the process temperature of the reactor (Sundberg et al., 2013).

4 Conclusions

High-throughput amplicon sequencing of six biomass samples from a full-scale anaerobic treatment plant during start-up and first year of operation was performed in this work. A more diverse microbial community was observed in the inoculum biomass compared to the six and 12 months samples. The dominant bacterial communities were affiliated to the phyla Bacteroidetes, Chloroflexi, Firmicutes, Proteobacteria and the Spirochaetes. The observed community structure changes probably reflect an adaptation of the inoculum microbial community to the specific anaerobic conditions of the new reactor, highly suitable for methane production. No obligate pathogens or *Legionella* spp. were identified in this study.

Acknowledgements

The work was funded by Borregaard ind. Ltd. and by the Norwegian Defence Research Establishment (FFI). The authors acknowledge Simen Settem Simonsen, Rune Gunnar Bergh and Viggo Waagen at Borregaard for valuable help and support including providing the samples and the operational performance data of the reactor.

References

1. Ahring, B. K., 2003. Perspectives for anaerobic digestion. Adv. Biochem. Eng. Biotechnol. 81, 1-30.

2. Amann, R., Ludwig, W., Schleifer, K.-H., 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiol. Rev. 59, 143-169.

3. Baker, G. C., Smith, J. J., Cowan, D. A., 2003. Review and re-analysis of domain-specific 16S primers. J. Microbiol. Method. 55, 541-555.

4. Beale, D. J, Karpe, A.V., McLeod, J. D., Gondalia, S.V., Muster, T. H., Othman, M. Z., Palombo, E. A., Joshi, D..,2016. An 'omics' approach towards the characterisation of laboratory scale anaerobic digesters treating municipal sewage sludge. Water Res. 88, 346-357. doi: 10.1016/j.watres.2015.10.029.

5. Blatny, J. M., Reif, B. A .P., Skogan, G., Andreassen, Ø., Høiby, E. A., Ask , E., Waagen, V., Aanonsen, D., Aaberge, I. S., and Caugant, D. A,.2008. Tracking Airborne Legionella spp. and Legionella pneumophila at a biological treatment plant. Environ Sci Technol. 42, 7360-73687.

6. Blatny, J. M., Ho, J., Skogan, G., Fykse, E. M., Aarskaug ,T., Waagen, V., 2011. Airborne Legionella bacteria from pulp waste treatment plant: aerosol particles characterized as aggregates and their potential hazard. Aerobiologia. 27. 147-162.

7. Case study Borregaard; Pulp & Paper | Case Study. http://technomaps.veoliawatertechnologies.com/processes/lib/_casestudies_/3260,Case-Study-Borregaard2.pdf.

8. 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.

9. Fykse, E. M., Aarskaug, T., Thrane, I., Blatny, J. M., 2013. *Legionella* and non-*Legionella* bacteria in a biological treatment plant. Can. J. Microbiol. 59, 102-109 (doi: 10.1139/cjm-2012-0166).

10. Fykse, E. M., Aarskaug, T., Blatny, J. M., 2014. Detection of *Legionella pneumophila* in a biological treatment plant by co-cultivation with *Acanthamoeba castellanii*. Current Environ. Engineer. 1, 1-9. 10.2174/2212717801666140707161649.

11. Garrity, G., Bell, J. A., Lilburn, T. G. (eds). 2005. Bergey's Manual of Systematic Bacteriology, Release 5.0, Springer, New York, INc., NY.

12. Goux, X ., Calusinska, M., Lemaigre, S., Marynowska, M., Klocke, M ., Udelhoven, T ., Benizri, E., Delfosse, P., 2015. Microbial community dynamics in replicate anaerobic digesters exposed sequentially to increasing organic loading rate, acidosis, and process recovery. Biotechnol. Biofuels. 8, 122 DOI: 10.1186/s13068-015-0309-9

13. Goux , X., Calusinska, M., Fossépré, M., Benizri , E., Delfosse, P. ,2016. Start-up phase of an anaerobic full-scale farm reactor – Appearance of mesophilic anaerobic conditions and establishment of the methanogenic microbial community. Bioresource Technol. 212, 217–226.

14. Guo, J., Peng, Y., Ni, B. J. Han, X., Fan, L., Yuan, Z., 2015. Dissecting microbial community structure and methane-producing pathways of a full-scale anaerobic reactor digesting activated sludge from wastewater treatment by metagenomic sequencing. Microb. Cell. Fact. 14, 14:33. doi: 10.1186/s12934-015-0218-4.

15. Hsu,T.-T. D., Lee, J., 2015. Global Distribution and Prevalence of Arcobacter in Food and Water. Zoonoses and Public Health. 62, 579–589.

16. Huse, S.M., Dethlefsen, L., Huber, J. A., Welch, M. D., Relman, D. A., Sogin, M. L., 2008. Correction: Exploring Microbial Diversity and Taxonomy Using SSU rRNA Hypervariable Tag Sequencing. PLoS Genet. 4(12): 10.1371/annotation/3d8a6578-ce56-45aa-bc71-05078355b851.

17. Kuever, J., Rainey, F. A., Widdel, F., 2005. Genus III. Desulfothermus gen. nov. In Bergey's Manual of Systematic Bacteriology, 2nd edn, vol. 2, The Proteobacteria, Part C, The Alpha-,Beta-, Delta-, and Epsilonproteobacteria,. D. J. Brenner, D. J., Krieg, N. R., Staley, J. T., Garrity, G. M (Eds). pp. 955–956, Springer, New York.

18. Lee, S.-H., Kang, H.-J., Lee, Y. H., Lee, T. J., Han, K., Choi, Y., Park, H.-D., 2012. Monitoring bacterial community structure and variability in time scale in full-scale anaerobic digesters. J Environ. Monit. 14, 1893-1905.

19. Leschine, S., Paster, B., Canale-Parola, E., 2006. Free-living saccharolytic spirochetes: The genus *Spirochaeta*. Prokaryotes. 7, 195-210.

20. Li, A., Chu, Y., Wang, X., Ren, L,, Yu, J., Liu X., Yan, J., Zhang, L., Wu, S., Li, S., 2013. A pyrosequencing-based metagenomic study of methane-producing microbial community in solid-state biogas reactor. Biotechnol. Biofuels. 6, 3.

21. Lynd, L.R., Weimer, P.J., arthrobactervan Zyl, W.H., Pretorius, I.S., 2002. Microbial cellulose utilization: Fundamentals and biotechnology. Microbiol. Mol. Biol. Rev. 66, 506-577.

22. Morris, B. E. L., Herbst, F. A., Bastida, F. Seifert, J., von Bergen, M., Richnow, H. H., Suflita, J. M., 2012. Microbial interactions during residual oil and n-fatty acid

metabolism by a methanogenic consortium. Environ. Miocrobiol. Rep. 4, 297-306. DOI: 10.1111/j.1758-2229.2012.00333.x

23. Muyzer, G., Brinkhoff T, Nuebel U, Santegoeds C, Schaefer H, Waver, C., 1998. Denaturing gradient gel electrophoresis (DGGE) in microbial ecology. In: Akkermans, A. D. L., van Elsas, J. D., de Bruijn, F. J. (Eds) Molecular Microbial Ecology Manual. Kluwer Academic Publishers; p. 1-27. Dordrecht, The Netherlands.

24. Nygård, K., Werner-Johansen, Ø., Rønsen, S., Caugant, D. A., Simonsen, Ø., Kanestrøm, A., Ask, E., Ringstad, J., Ødegård, R., Jensen, T., Krogh, T., Høiby, E. A., Ragnildstveit, E., Aaberge, I. S., Aavidtsland, P., 2008. An outbreak of Legionnaires' disease caused by long-distance spread from an air scrubber in Sarpsborg, Norway. Clin. Infect. Dis. 46, 61-69.

25. Nelson, M. C., Morrison, M., Yu, Z., 2011. A meta-analysis of the microbial diversity observed in anaerobic digesters. Bioresour. Technol.102, 3730-3739. doi: 10.1016/j.biortech.2010.11.119.

26. Nelson, M. C., Morrison, M., Schanbacher, F., Yu, Z., 2012.Shifts in microbial community structure of granular and liquid biomass in response to changes to infeed and digester design in anaerobic digesters receiving food-processing wastes. Bioresour. Technol.107, 135-43. doi: 10.1016/j.biortech.2011.12.070

27. Olsen, J.S., Aaskaug, T., Thrane, .I, Pourcel, C., Ask, E., Johansen, G., Waagen, V., Blatny, J. M., 2010. Alternative routes for dissemination of legionella pneumophilia causing three outbreaks in Norway. Environ. Sci .Technol. 44, 8712-8717.

28. Pereira, M. A., Roest, K., Stams, A. J. M., Mota, M., Alves, M., Akkermans, A. D., 2006. Molecular monitoring of microbial diversity in expanded granular sludge bed (EGSB) reactors treating oleic acid. FEMS Microbiol. Ecolol. 41, 95-103.

29. Prouzet-Mauléon,V., Labadi, L., Bouges, N., Ménard, A., Mégraud, F., 2006. *Arcobacter butzleri*: underestimated enteropathogen. Emerg. Infect. Dis. 12, 307-309.

30. Rivière D, Desvignes V, Pelletier E, Chaussonnerie S, Guermazi S, Weissenbach J, Li T, Camacho P, Sghir A., 2009. Towards the definition of a core of microorganisms involved in anaerobic digestion of sludge. ISME, 3:700-714.

31. Roest, K., Heilig, H. G., Smidt, H., de Vo, W. M., Stams, A. J., Akkermans, AD., 2005. Community analysis of a full-scale anaerobic bioreactor treating paper mill wastewater. Syst. Appl. Microbiol. 28, 175-185.

32. Sandoval Lozano, C. J., Vergara Mendoza, M., Carreño de Arango, M., Castillo Monroy, E. F., 2009. Microbiological characterization and specific methanogenic activity of anaerobe sludges used in urban solid waste treatment. J. Waste Management 29, 704–11.

33. Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., Lesniewski, R. A., Oakley, B. B., Parks, D. H., Robinson, C. J., Sahl, J. W., Stres, B., Thallinger, G. G., Van Horn, D. J., and Weber, C. F. 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl. Environ. Microbiol.75, 7537–7541.

34. Shu, D., He, Y., Yue, H., Wang, Q. 2015. Microbial structures and community functions of anaerobic sludge in six full-scale wastewater treatment plants as revealed by 454-high-throughput pyrosequencing. Bioresource technology, 186, 163-172.

35. Solli, L., Håvelsrud, O. E., Horn, S. J., Rike, A. G., 2014. A metagenomic study of the microbial communities in four parallel biogas reactors. Biotechnol. for Biofuels, 7, 146. DOI: 10.1186/s13068-014-0146-2.

36. Sundberg, C., Al-Soud, W.A., Larsson, M., Alm, E., Yekta, S. S., Svensson, B. H. Sørensen, S. J., Karlsson, A., 2013. 454 pyrosequencing analyses of bacterial and archaeal richness in 21 full-scale biogas digesters. FEMS Microbiol. Ecol. 85, 612–626. DOI: 10.1111/1574-6941.12148

37. Werner, J.J. Knights, D., Garcia, M.L., Scalfone, N.B., Smith, S., Yarasheski, K., Cummings, T.A., Beers, A.R., Knight, R., Angenent, L.T., 2010. Bacterial community structures are unique and resilient in full-scale bioenergy systems. PNAS 108, 4158–4163.

38. Wong, MT., Zhang, D., Li, J., Hui, RKH., Tun, HM., Brar, MS., Park, TJ., Chen, YG., Leung, FC., 2013. Towards a metagenomic understanding on enhanced biomethane production from waste activated sludge after pH 10 pretreatment. BIOTECHNOLOGY FOR BIOFUELS, Volume: 6, Article Number: 38, DOI: 10.1186/1754-6834-6-38.

39. Yamada, T., Sekiguchi,Y.,2009. Cultivation of Uncultured Chloroflexi Subphyla: Significance and Ecophysiology of Formerly Uncultured Chloroflexi 'Subphylum I' with Natural and Biotechnological Relevance. Microbes and Environ. 24, 205-216. http://doi.org/10.1264/jsme2.ME09151S.

40. Ye, L., Zhang, T. 2013. Bacterial communities in different sections of a municipal wastewater treatment plant revealed by 16S rDNA 454 pyrosequencing. Appl. Microbiol. Biotechnol., 97, 2681–2690.

41. Ziganshin AM, Liebetrau J, Proeter J, Kleinsteuber S., 2013. Microbial community structure and dynamics during anaerobic digestion of various agricultural waste materials. Appl Microbiol Biot. 97, 5161–5174.

42. Zinder, S. H., 1984. Microbiology of anaerobic conversion of organic wastes to methane: recent developments. ASM News 50, 294-298.

Figure Captions

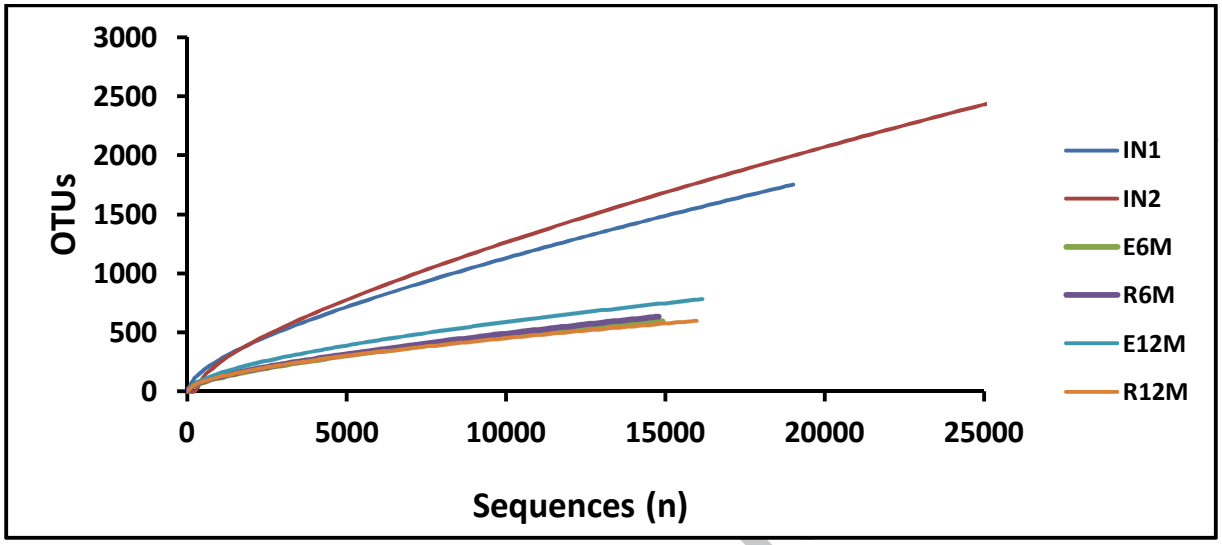

Fig. 1. Observed bacterial community richness (rarefaction curves) for six biomass samples from the full-scale anaerobic treatment plant during startup and first year of operation. Operational taxonomic units (OTUs) were defined at 97% sequence identity. The rarefaction curves for all samples approached the plateau phase, suggesting that the sequencing depth was sufficient to capture the majority of the diversity including a large fraction of the rare diversity. The rarefaction curves show that the observed richness is higher in the IN1 and IN2 samples compared to E6M, R6M, E12M, and R12M samples. Inoculum sample 1 (IN1), Inoculum sample 2 (IN2), Discharged reactor effluent after six months of operation (E6M), Reactor content after six months of operation (R6M), Discharged reactor effluent after 12 months of operation (E12M), Reactor content after 12 months of operation (R12M).

(Color in print)

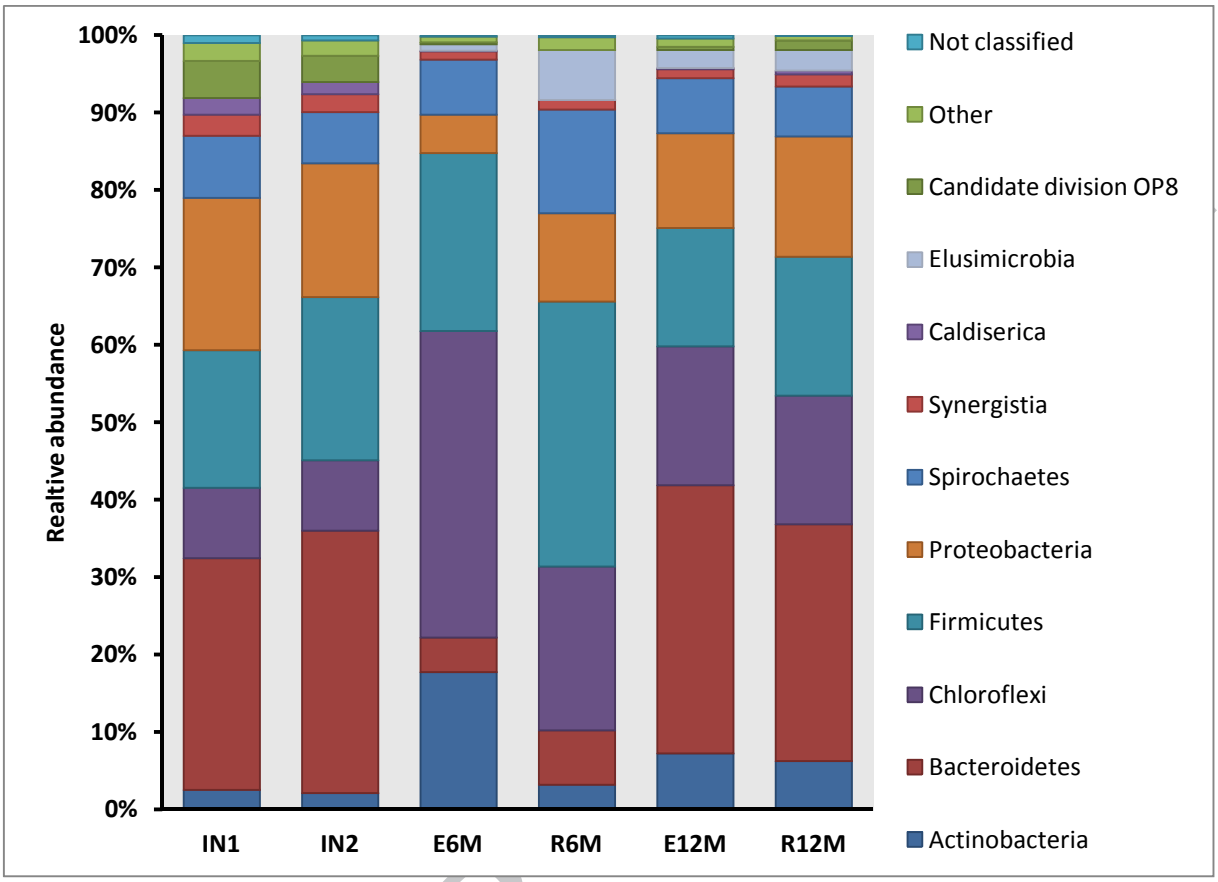

Fig.2. Relative abundance of major bacterial phyla, present as ≥1% of the sequence reads in at least one sample, in six biomass samples from the full-scale anaerobic treatment plant during start-up and first year of operation. Inoculum sample 1 (IN1), Inoculum sample 2 (IN2), Discharged reactor effluent after six months of operation (E6M), Reactor content after six months of operation (R6M), Discharged reactor effluent after 12 months of operation (E12M), Reactor content after 12 months of operation (R12M).

(Color in print)

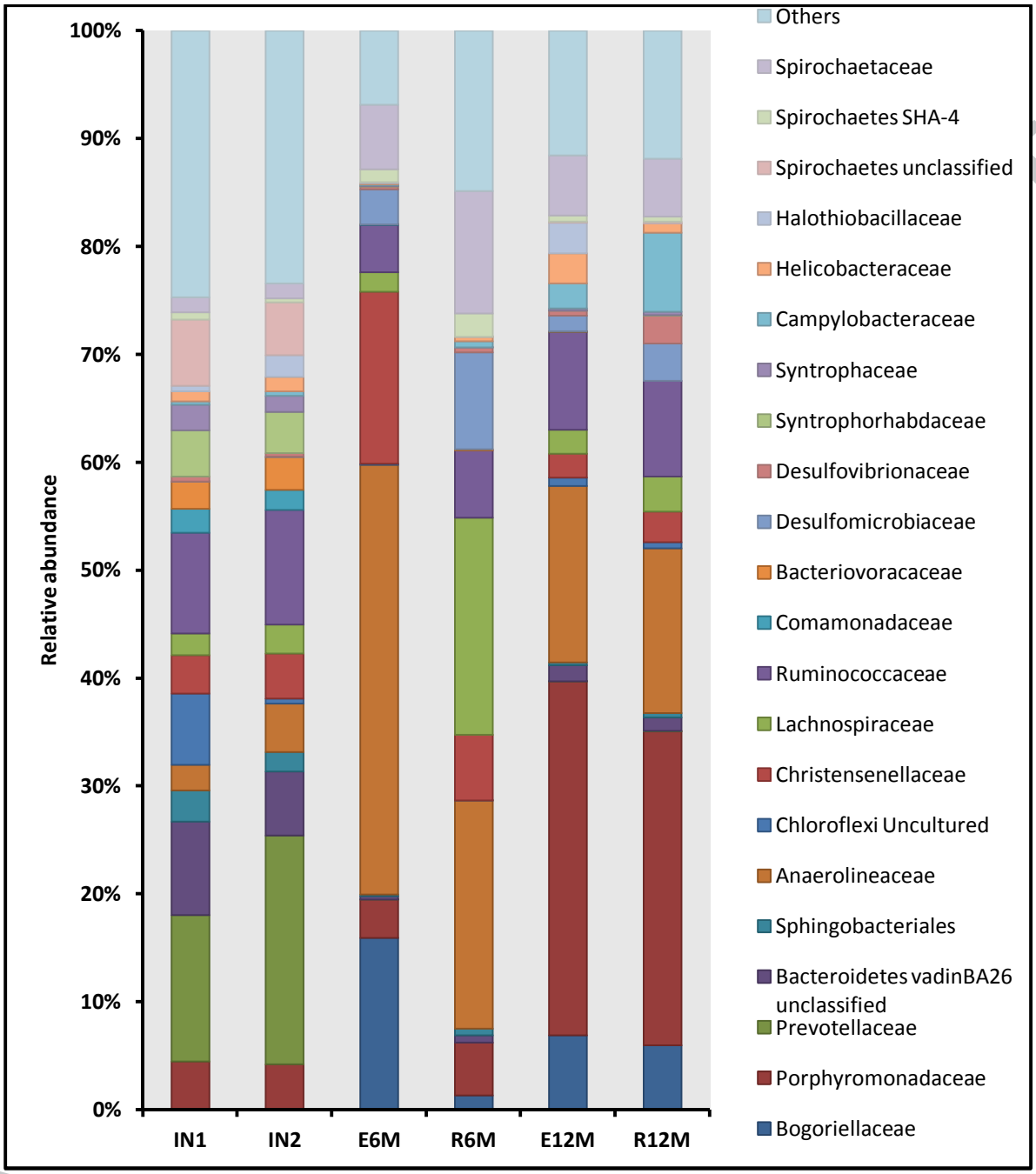

Fig. 3. Relative abundance of major bacterial families, present as ≥2% of the sequence reads in at least one sample, in six biomass samples from the full-scale anaerobic treatment plant during startup and first year of operation. Inoculum sample 1 (IN1), Inoculum sample 2 (IN2), Discharged reactor effluent after six months of operation (E6M), Reactor content after six months of operation (R6M), Discharged reactor effluent after 12 months of operation (E12M), Reactor content after 12 months of operation (R12M). (Color in print)

Tables

Table 1. Number of sequence reads after quality filtering and operational taxonomic units (OTUs) at 97% sequence identity and calculated diversity estimators for six biomass samples collected from the full-scale anaerobic treatment plant during startup and first year of operation

Sample[1]	Total # of quality-filtered sequence reads	OTUs	Good's sampling coverage	Shannon diversity index	Shannon-based measure of evenness	Inverse Simpson index	Chao1 richness estimator	Abundance coverage-based estimator (ACE)
IN1	19017	1751	0.94	5.30	0.69	62.49	6354	12387
IN2	25280	2472	0.93	5.36	0.66	49.91	9048	18592
E6M	14912	593	0.98	3.43	0.52	11.96	1617	2811
R6M	14786	630	0.97	3.64	0.55	14.11	1990	4034
E12M	16174	782	0.97	3.68	0.54	8.33	2049	3246
R12M	15997	595	0.98	3.49	0.53	9.77	1684	2843

[1] Inoculum sample 1 (IN1), Inoculum sample 2 (IN2), Discharged reactor effluent after six months of operation (E6M), Reactor content after six months of operation (R6M), Discharged reactor effluent after 12 months of operation (E12M), Reactor content after 12 months of operation (R12M).

Table 2. The average process parameters of the anaerobic reactor in this study. These parameters are: organic load (ORL), COD value, contaminants removal efficiency, biogas production, volatile fatty acids (VFAs), temperature and pH.

Reactor parameters	**Average reactor parameters 30 days before sampling**	
Period	**18.08.2013 – 17.09.2013**	**20.04.2014 – 20.05.2014**
OLR applied[1]	0.37	0.33
Ethanol condensate (m3/day)	3717 ± 565	6252 ± 258
Alcacell condensate (m3/day)	964 ± 286	2082± 201
COD loading (kg/day)	37000 ± 8000	59000 ± 5000
COD removal efficiency (%)	71 ± 6	70 ± 6
Biogas production (m3/day)[2]	14066 ± 2869	19455 ± 1743
Methane content (v/v %)	83	74
VFAs (mg/L CH3COOH)	21 ± 22	62 ± 19
Temp (°C)	35.96 ± 0.44	36.29 ± 0.95
pH	7.03 ± 0.13	7.02 ± 0.17

1) OLR: (tCOD/t OTS/day); ton COD loaded per day/OTS (Organic Total Solids) per day
2) m3/day, biogas production normalized at 0 °C and 1 atm.

Supplementary material

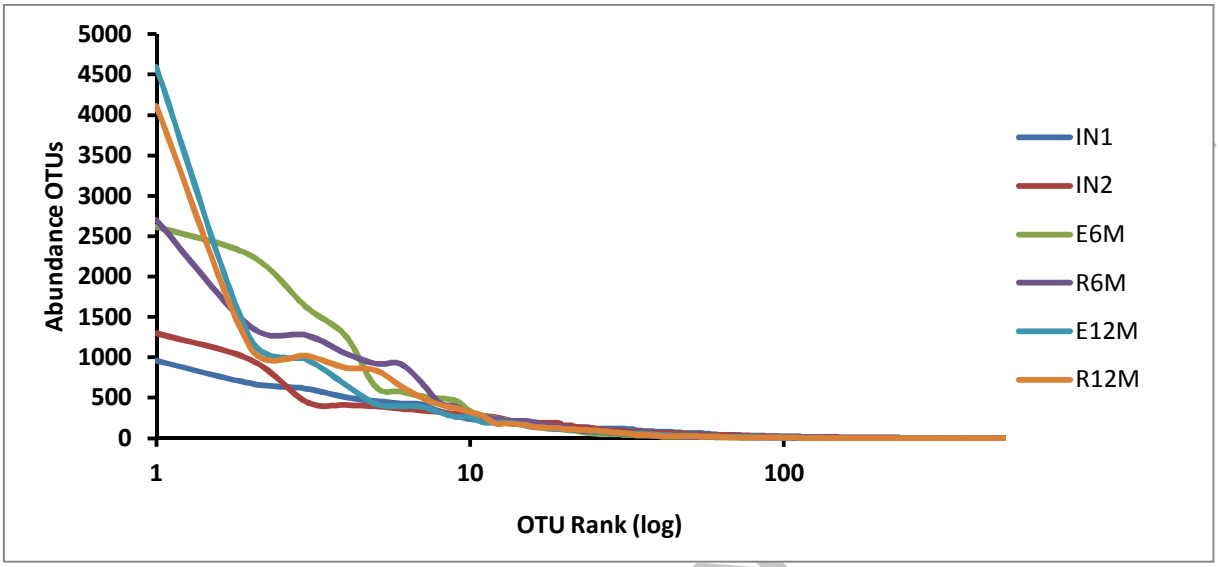

Fig. S1. Rank Abundance plot for six biomass samples from the full-scale anaerobic treatment plant during startup and first year of operation. Operational taxonomic units (OTUs) were defined at 97% sequence identity. The rank abundance curves show that the anaerobic reactor and discharged effluent samples after 12 months of operation (R12M and E12M, respectively) appeared to be dominated by a single highly abundant OTU when compared to the IN1, IN2, E6M, and R6M samples Inoculum sample 1 (IN1), Inoculum sample 2 (IN2), Discharged reactor effluent after six months of operation (E6M), Reactor content after six months of operation (R6M), Discharged reactor effluent after 12 months of operation (E12M), Reactor content after 12 months of operation (R12M).

(Online color)

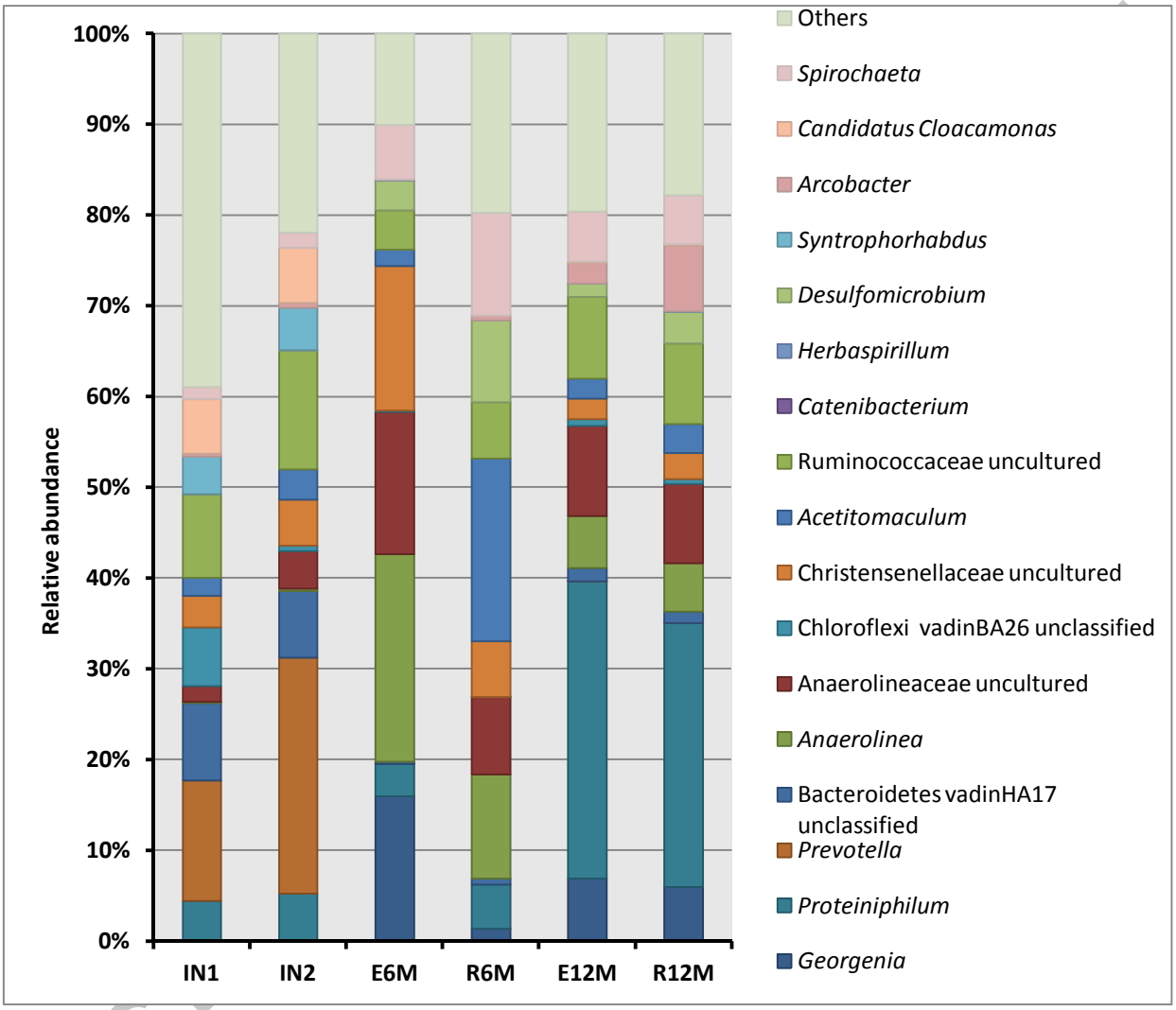

Fig. S2. Relative abundance of major bacterial genera, present as ≥3% of the sequence reads in at least one sample, in six biomass samples from the full-scale anaerobic treatment plant during startup and first year of operation. Inoculum sample 1 (IN1), Inoculum sample 2 (IN2), Discharged reactor effluent after six months of operation (E6M), Reactor content after six months of operation (R6M), Discharged reactor effluent after 12 months of operation (E12M), Reactor content after 12 months of operation (R12M).

(online color)

Table S1. Relative number of reads (%) per sample in six biomass samples from the full-scale anaerobic treatment plant during startup and first year of operation. The reads are assigned to the tax paths listed (according to SILVA tax, rel. 115). Inoculum sample 1 (IN1), Inoculum sample 2 (IN2), Discharged reactor effluent after six months of operation (E6M), Reactor content after six months of operation (R6M), Discharged reactor effluent after 12 months of operation (E12M), Reactor content after 12 months of operation (R12M).

Highlights

- An anaerobic reactor was studied during start-up and first year of operation
- Microbial community were investigated by high throughput amplicon sequencing
- Abundant phyla were Bacteroidetes, Chloroflexi, Firmicutes, Proteobacteria, and Spirochaetes
- The diversity of the inoculum was reduced after six and 12 months of operation