Bioresource Technology 216 (2016) 260–266

Contents lists available at ScienceDirect

Bioresource Technology

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

Deeper insight into the structure of the anaerobic digestion microbial community; the biogas microbiome database is expanded with 157 new genomes

Laura Treu [a], Panagiotis G. Kougias [a,*], Stefano Campanaro [b], Ilaria Bassani [a], Irini Angelidaki [a]

[a] Department of Environmental Engineering, Technical University of Denmark, Kgs. Lyngby DK-2800, Denmark
[b] Department of Biology, University of Padua, Via U. Bassi 58/b, 35121 Padova, Italy

HIGHLIGHTS

- 236 genome bins were extracted from metagenomic assembly in biogas reactors.
- Existence of a core essential microbial group in biogas production system.
- Expansion of the biogas microbiome database with 157 new genomes.
- Comparison of metagenomes unveiled differences in phylogenetic distribution.

ARTICLE INFO

Article history:
Received 7 March 2016
Received in revised form 2 May 2016
Accepted 6 May 2016
Available online 24 May 2016

Keywords:
Biogas
Microbiome
Metagenome
Assembly
Binning

ABSTRACT

This research aimed to better characterize the biogas microbiome by means of high throughput metagenomic sequencing and to elucidate the core microbial consortium existing in biogas reactors independently from the operational conditions. Assembly of shotgun reads followed by an established binning strategy resulted in the highest, up to now, extraction of microbial genomes involved in biogas producing systems. From the 236 extracted genome bins, it was remarkably found that the vast majority of them could only be characterized at high taxonomic levels. This result confirms that the biogas microbiome is comprised by a consortium of unknown species. A comparative analysis between the genome bins of the current study and those extracted from a previous metagenomic assembly demonstrated a similar phylogenetic distribution of the main taxa. Finally, this analysis led to the identification of a subset of common microbes that could be considered as the core essential group in biogas production.

1. Introduction

Biogas production is a striking technology for sustainable generation of renewable energy. The produced biogas is derived as a result of the anaerobic decomposition of organic matter via a biological process mediated mainly by a complex consortium of bacteria and archaea (Luo et al., 2015). Despite the fact that this technology is well established, considering the proliferation of the biogas plants worldwide, fundamental aspects related to the microbiology of the process is still unclear.

In the cited literature the composition of biogas-producing microbial communities has been generally determined via construction of 16S-rRNA clone libraries and subsequent analysis of 16S-rRNA amplicons (De Francisci et al., 2015; Kröber et al., 2009; Luo et al., 2015). The taxonomic assignment of the microbial species was commonly based on sequence similarity search against reference 16S rRNA sequences deposited in public databases. Even in the shotgun sequencing studies, most of the reference genomes used for profiling the composition of the microbial communities are isolates from various environments different from the anaerobic digestion system, while it is known that only a small fraction of microorganisms have been cultivated (Albertsen et al., 2013; Hugenholtz, 2002). Therefore, even if the phylogeny of these genomes is related to the ones found in biogas communities, it is uncertain whether they serve the same function during the anaerobic digestion (AD) process. Moreover, by profiling only phylogenetic marker genes, such as the 16S rRNA gene, it is impossible to acquire insights on the community's functional capabilities (Langille et al., 2013), and thus fundamental information regarding essential roles of predominantly uncultivated microbes (e.g. symbiotic or competitive behavior) in the formation of a collective network are limited (Tyson et al., 2004). Another aspect of

* Corresponding author at: Department of Environmental Engineering, Technical University of Denmark, Bld 113, 2800 Lyngby, Denmark.
E-mail address: panak@env.dtu.dk (P.G. Kougias).

http://dx.doi.org/10.1016/j.biortech.2016.05.081
0960-8524/

particular attention is the definition of a core microbiome in biogas production. Riviere et al. (2009) demonstrated that in sludge digesters there is a fraction of phylotypes that are always present constituting the common prokaryotic community, while another fraction of phylotypes are site specific. Nevertheless, such information are lacking in cases of biogas reactors treating agricultural and industrial residues. It is imperative to extend the analysis of the core microbiome at genomic level in AD systems elucidating the genome structure of the stable taxa and of those specific of different operational conditions.

The advancement of sequencing technologies and bioinformatic tools allow nowadays a deep characterization of complex communities, such as the one of the biogas microbiome. Therefore, in the cited literature the number of metagenomic analyses, even without performing assembly or binning processes, is increasing. Currently, most of the metagenomic studies on anaerobic digesters determined the functional properties of the microorganisms using non-assembled short reads (Eikmeyer et al., 2013), or in others works the gene finding was achieved using few number of short scaffolds (Schlüter et al., 2008; Stolze et al., 2015; Wirth et al., 2012). Bremges et al. (2015) assembled the metagenome of a single agricultural production-scale biogas facility and managed to reconstruct most of the genes involved in methane metabolism.

In our previous work, it was demonstrated that by assembling the shotgun metagenome sequences and following a binning strategy, it was possible to dissect the bioma of multiple thermophilic biogas reactors treating manure-based substrates (Campanaro et al., 2016). In this approach, *de novo* assembly procedure can be applied to analyze complex microbial communities generating a large set of scaffolds, which can be subsequently classified in single biological entities with a procedure named binning. This classification can be performed with different strategies, but the most innovative is based on the rationale that in different environmental conditions one bacterial species can be present at different relative abundances, consequently scaffolds belonging to the same genome change their coverage concertedly and they can be attributed to the same microbe (Albertsen et al., 2013; Nielsen et al., 2014). The results led to the identification of 106 microbial genomes (Genome Bins, GBs), and a conservative estimation indicated the presence of more than 450 microorganisms in the biogas microbial community. This estimate was derived considering that approximately 70% of the assembly could not be assigned to a specific GB. Moreover, this argument was further reinforced as during the assembly process, the reads belonging to the least abundant microorganisms were discarded.

This study is a continuation of the previous work aiming to further elucidate the biogas microbial community by enriching the biogas microbiome database with reference genomes present in anaerobic digesters. The samples were obtained from mesophilic and thermophilic continuous reactors used to upgrade and enhance biogas production via hydrogen assisted methanogenesis. The microbial community found in the current study was compared with the corresponding one of our previous assembly. This allowed the determination of similarities and differences among the microbiota and the identification of a potential existence of common microbes that can serve as the core essential group for biogas production.

2. Materials and methods

2.1. Reactor configuration and sample collection

Samples were obtained from the secondary reactor of a serial configuration operating either in mesophilic (35 ± 1 °C) or thermophilic (55 ± 1 °C) conditions. The collection of the samples was performed once the reactors were operating under steady state conditions (i.e. after a period of 3 Hydraulic Retention Times) before and after H_2 addition to ensure representative process conditions and microbial community stability. As the upgrading process occurred in the secondary reactor of the serial configuration, only samples from the secondary stage were analyzed. Each configuration was comprised by two Continuously Stirred Tank Reactors (CSTR) connected in series with volume ratio between the primary/secondary reactor equal to 0.75. For the mesophilic conditions, the Hydraulic Retention Time (HRT) of the primary and secondary reactors were 25 and 33 days, respectively, while the corresponding HRT for the thermophilic setup were 15 and 20 days, respectively. The primary reactor of each set was serving as conventional biogas producing digester fed with cattle manure. The characteristics of the manure used as substrate are given in Table 1. The digestate of the primary reactor along with external H_2 gas were introduced to the secondary reactor in order to upgrade the biogas quality by coupling the CO_2 contained in the biogas with the injected H_2. The H_2 flow rate and a detailed description of the reactor operation are described by Bassani et al. (2015).

2.2. DNA extraction and high throughput sequencing

Initially, each sample was filtered using a 100 μm nylon cell strainer filter in order to remove all the fibrous residues of animal nutrition present in the digested manure. Subsequently, the samples were centrifuges at 2500*g* for 10 min in order to recover ~2 g of pellet. Genomic DNA was extracted using RNA PowerSoil® DNA Elution Accessory Kit (MO BIO Laboratories, Carlsbad, CA). NanoDrop (ThermoFisher Scientific, Waltham, MA) and Qbit fluorimeter (Life Technologies, Carlsbad, CA) were used to evaluate the quality and quantity of the extracted DNA. Metagenome sequencing was performed using Illumina NextSeq 500 desktop system and Nextera XT kit (Illumina, San Diego, CA) for library preparation (150+150 bp).

2.3. Metagenomic assembly and binning process

Trimmomatic software was used to filter the raw reads in FASTQ format and to remove the adaptors (Bolger et al., 2014). Overlapped paired-ends were merged using Flash (Magoč and Salzberg, 2011) using standard parameters, except from the maximum overlap parameter, which was set to 150 bases. Assembly and binning strategy was performed using a previously established method (Campanaro et al., 2016); all the perl scripts used for binning were obtained from "http://www.biogasmicrobiome.com/" (binning process v1). For the metagenome assembly both paired-end reads and single-end reads (both those merged using Flash

Table 1
Chemical composition of cattle manure used in the experiment.

Parameter	Unit	Values
pH	–	7.44 ± 0.01
Total solids (TS)	g/L	47.40 ± 1.86
Volatile solids (VS)	g/L	34.56 ± 1.40
Total Kjeldahl Nitrogen (TKN)	g-N/L	3.03 ± 0.10
Ammonium Nitrogen (NH_4^+)	g-N/L	2.07 ± 0.01
Total Volatile fatty acids (VFA)	mg/L	6831 ± 477
Acetate	mg/L	4151 ± 394
Propionate	mg/L	1421 ± 67
iso-butyrate	mg/L	142 ± 1
Butyrate	mg/L	793 ± 16
iso-valerate	mg/L	224 ± 1
Valerate	mg/L	88 ± 1
n-hexanoate	mg/L	12 ± 1