Water Research 104 (2016) 1–10

Contents lists available at ScienceDirect

Water Research

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

Long-term successional dynamics of microbial association networks in anaerobic digestion processes

Linwei Wu [a], Yunfeng Yang [a, *], Si Chen [b], Mengxin Zhao [a], Zhenwei Zhu [b, 1], Sihang Yang [a], Yuanyuan Qu [c], Qiao Ma [c], Zhili He [d], Jizhong Zhou [a, d, e, **], Qiang He [b, f, ***]

[a] State Key Joint Laboratory of Environment Simulation and Pollution Control, School of Environment, Tsinghua University, Beijing 100084, China
[b] Department of Civil and Environmental Engineering, The University of Tennessee, Knoxville, TN 37996, USA
[c] Key Laboratory of Industrial Ecology and Environmental Engineering (Ministry of Education), School of Environmental Science and Technology, Dalian University of Technology, Dalian 116024, China
[d] Institute for Environmental Genomics and Department of Microbiology and Plant Biology, University of Oklahoma, Norman, OK 73019, USA
[e] Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[f] Institute for a Secure and Sustainable Environment, The University of Tennessee, Knoxville, TN 37996, USA

ARTICLE INFO

Article history:
Received 12 April 2016
Received in revised form
12 July 2016
Accepted 29 July 2016
Available online 2 August 2016

Keywords:
Anaerobic digestion
Microbial interactions
Process stability

ABSTRACT

It is of great interest to elucidate underlying mechanisms to maintain stability of anaerobic digestion, an important process in waste treatment. By operating triplicate anaerobic digesters continuously for two years, we found that microbial community composition shifted over time despite stable process performance. Using an association network analysis to evaluate microbial interactions, we detected a clear successional pattern, which exhibited increasing modularity but decreasing connectivity among microbial populations. Phylogenetic diversity was the most important factor associated with network topology, showing positive correlations with modularity but negative correlations with network complexity, suggesting induced niche differentiation over time. Positive, but not negative, correlation strength was significantly related ($p < 0.05$) to phylogeny. Furthermore, among populations exhibiting consistent positive correlations across networks, close phylogenetic linkages were evident (e.g. *Clostridiales* organisms). *Clostridiales* organisms were also identified as keystone populations in the networks (i.e., they had large effects on other species), suggestive of an important role in maintaining process stability. We conclude that microbial interaction dynamics of anaerobic digesters evolves over time during stable process performance.

1. Introduction

Widely applied in wastewater treatment and animal waste management, anaerobic digestion is an important microbial process in waste treatment and renewable energy recovery (Aydin et al., 2015a; Talbot et al., 2008; Zhang et al., 2011). Therefore, it is crucial to understand the ecology and function of microbial communities involved in anaerobic digestion. With the development and application of molecular microbial ecology techniques, progress has recently been made to characterize microbial community compositions in anaerobic digestion processes (Narihiro and Sekiguchi, 2007). For example, Aydin et al. (2015b, 2016) showed that changes in microbial community composition led to altered biodegradation capacity of organic waste and antibiotics during anaerobic digestion, which linked microbial community compositions to the function of anaerobic digesters. However, more studies are needed that focus on the potential interactions among microbial populations at the whole community level, which is expected to contribute more to system functions than individual populations (Ma et al., 2016).

Microorganisms live within complicated networks through a multitude of interactions, such as mutualism and competition

* Corresponding author.
** Corresponding author. State Key Joint Laboratory of Environment Simulation and Pollution Control, School of Environment, Tsinghua University, Beijing 100084, China.
*** Corresponding author. Department of Civil and Environmental Engineering, The University of Tennessee, Knoxville, TN 37996, USA.
E-mail addresses: yangyf@tsinghua.edu.cn (Y. Yang), jzhou@ou.edu (J. Zhou), qianghe@utk.edu (Q. He).
[1] Current affiliation: Loyola University Chicago, Institute of Environmental Sustainability, USA.

http://dx.doi.org/10.1016/j.watres.2016.07.072
0043-1354/

(Faust and Raes, 2012). However, most of those interactions among microbial populations cannot be directly observed, representing a great challenge for studying population interactions in microbial communities. Network analysis has been used to deduce potential interactions among microbial populations by uncovering strong, non-random associations (Faust et al., 2012). It has been applied to examine complex microbial communities in various habitats, such as oceans (Chow et al., 2014), soils (Barberán et al., 2012), human microbiomes (Faust et al., 2012) and bioreactors (Ju and Zhang, 2015). In addition, network analysis is capable of revealing changes in the topology of microbial networks (Deng et al., 2015; Zhou et al., 2010, 2011). Therefore, network analyses have been considered as powerful tools for studying population interactions in complex microbial communities (Lupatini et al., 2014).

Various approaches of network analyses have been developed and widely applied in functional genomics studies based on gene expression data, including differential equation-based network methods, Bayesian network analyses, and relevance/association network methods (Deng et al., 2012). Among them, the association network method based on co-occurrence/correlation is the most commonly used, owing to its computational simplicity and noise tolerance (Gardner and Faith, 2005). However, most studies employing association network analyses use arbitrary thresholds, thus compromising the constructed networks with subjectivity. To address this, a random matrix theory (RMT)-based approach was developed to objectively identify a threshold for network construction based on microarray data or high-throughput sequencing data (Luo et al., 2006, 2007). This approach was shown to be effective in identifying network interactions among microbial populations (Deng et al., 2012, 2015; Zhou et al., 2010, 2011).

Process stability is highly desirable during anaerobic digestion processes. Previous efforts have primarily focused on the roles of individual populations in process stability, especially on methanogens (Chen and He, 2015; Sekiguchi, 2006). Owing to the importance of microbial interactions in system functions (Ma et al., 2016), herein we evaluated microbial population interactions by performing network analysis of the microbial communities in anaerobic digesters operated continuously for two years. A clear successional pattern was identified, exhibiting increasing modularity but decreasing connectivity between populations over time. Furthermore, microbial phylogenetic diversity was found to be the most important factor associated with network topology, indicative of induced niche differentiation over time.

2. Material and methods

2.1. Anaerobic digester operation and biomass sampling

Triplicate mesophilic continuous anaerobic digesters, designated as C1, C2 and C3 hereafter, were established and operated with dairy waste as the substrate as previously described (Chen and He, 2015). All anaerobic digesters had a working volume of 3.6 L and were operated at a constant temperature of 35 °C. The hydraulic retention time was maintained at 20 days and the organic loading rate (OLR) was kept at 1.0 g volatile solids (VS)/L/day throughout the two-year study period. Process performance remained stable throughout the study period and biomass samples were collected from the digesters periodically, resulting in a total of 156 samples from 52 time points. All samples were stored at −80 °C before use. The detailed sampling points and process performance parameters are summarized in Supplementary Table S1.

2.2. Acquisition and processing of 16S rRNA gene sequences

DNA was extracted from biomass samples using previously described protocols (Ma et al., 2015). Briefly, biomass samples were suspended in 630 μL DNA-extraction buffer, followed by treatment with 60 μL of a lysozyme mixture (37 °C, 60 min), 60 μL of a protease mixture (37 °C, 30 min), and 80 μL of 20% sodium dodecyl sulfate (37 °C, 90 min). The treated sample suspension was subsequently extracted with phenol–chloroform–isoamyl alcohol (25:24:1) at 65 °C for 20 min and the supernatant was extracted using chloroform–isoamyl alcohol (24:1). DNA extract was then mixed with 0.6 volume of isopropanol and stored at 4 °C overnight. DNA was obtained by centrifugation followed by washing with 70% cold ethanol, drying and resuspension in nuclease-free water. DNA concentration and purity were analyzed with a NanoDrop spectrophotometer (NanoDrop Technologies Inc., Wilmington, DE, USA). The V4 region of microbial 16S rRNA genes was amplified by primer pairs (Wu et al., 2015), 515F (50-GTG CCA GCM GCC GCG GTA A-30) and 806R (50-GGA CTA CHV GGG TWT CTA AT-30). PCR was performed at 94 °C for 1 min; 35 cycles of 94 °C for 20 s, 53 °C for 25 s, and 68 °C for 45 s; and a final extension at 68 °C for 10 min using the AccuPrime High Fidelity Taq Polymerase (Invitrogen, Grand Island, NY, USA). PCR products were pooled and purified using the QIAquick Gel Extraction Kit (Qiagen, Valencia, CA, USA) and amplicon sequencing was performed with the Miseq Illumina platform at the Institute for Environmental Genomics (IEG), University of Oklahoma.

For sequencing data analysis, the primer sequences were trimmed from the paired-end sequences, which were then merged using FLASH. Merged sequences were processed to generate operational taxonomic units (OTUs) by UPARSE at the 97% sequence similarity threshold. Taxonomy was assigned with a confidence cutoff of 50% using the RDP classifier. Phylogenetic trees were then constructed from all representative sequences using the FastTree algorithm (Price et al., 2009). The phylogenetic distance between OTUs was then determined by their relatedness in the phylogenetic tree with function *cophenetic* in R *picante* package. The rRNA gene copy number for each OTU was estimated with the rrnDB database (Stoddard et al., 2014). The OTU matrices were rarefied to 11,558 sequences per sample. The abundance-weighted average rRNA gene copy number was then calculated for each sample.

2.3. TaqMan qPCR analysis

TagMan qPCR analyses were performed with triplicate biomass samples at 15 time points (Day 45, 73, 90, 111, 121, 132, 146, 167, 251, 289, 326, 347, 395, 453 and 501). Genus-specific TaqMan qPCR assays were used to quantify the populations of *Methanosarcina* and *Methanosaeta*. To determine the relative abundance of the both methanogens in the archaeal community, a domain-specific TaqMan qPCR assay was performed to quantify total archaeal populations. The characteristics of TaqMan primer/probe sets used in this study were summarized in Table S2, and the qPCR procedure was performed with a CFX96 Real-Time PCR Detection System (Bio-Rad, Hercules, California, USA) as previously described (Chen and He, 2015). In brief, the qPCR assays were performed in 25 μL reaction volumes with 15 pmol primers, 5 pmol probe, and Brilliant II QPCR Master Mix (Agilent, Santa Clara, California, USA). The thermal cycling was started by an incubation at 50 °C for 2 min and an initial denaturation at 95 °C for 10 min, followed by up to 45 cycles at 95 °C for 30 s and 60 °C (for all primer/probe sets) for 45 s.

2.4. Network construction

To construct a time-lag network, it is desirable to use a minimum of 12 samples with consistent time intervals between samples. We fit our time-series data to this criteria by categorizing samples into 9 operational intervals according to sampling time: 1)