Journal of Biotechnology 136 (2008) 77–90

Contents lists available at ScienceDirect

Journal of Biotechnology

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

The metagenome of a biogas-producing microbial community of a production-scale biogas plant fermenter analysed by the 454-pyrosequencing technology

Andreas Schlüter[a,*], Thomas Bekel[b], Naryttza N. Diaz[b], Michael Dondrup[b], Rudolf Eichenlaub[c], Karl-Heinz Gartemann[c], Irene Krahn[a], Lutz Krause[b,g], Holger Krömeke[b], Olaf Kruse[d], Jan H. Mussgnug[d], Heiko Neuweger[b], Karsten Niehaus[e], Alfred Pühler[a], Kai J. Runte[b], Rafael Szczepanowski[a], Andreas Tauch[a], Alexandra Tilker[a], Prisca Viehöver[f], Alexander Goesmann[b]

[a] *Lehrstuhl für Genetik, Universität Bielefeld, D-33594 Bielefeld, Germany*
[b] *Centrum für Biotechnologie (CeBiTec), Universität Bielefeld, D-33594 Bielefeld, Germany*
[c] *Lehrstuhl für Gentechnologie/Mikrobiologie, Universität Bielefeld, D-33594 Bielefeld, Germany*
[d] *Algenbiotechnologie, Universität Bielefeld, D-33594 Bielefeld, Germany*
[e] *Proteom und Metabolomforschung, Universität Bielefeld, D-33594 Bielefeld, Germany*
[f] *Lehrstuhl für Genomforschung, Universität Bielefeld, D-33594 Bielefeld, Germany*
[g] *Nestlé Research Center, BioAnalytical Science Department, CH-1000 Lausanne 26, Switzerland*

ARTICLE INFO

Article history:
Received 12 March 2008
Received in revised form 16 April 2008
Accepted 8 May 2008

Keywords:
Biogas fermentation
Methanogenesis
Methane production
454-Pyrosequencing
Metagenome

ABSTRACT

Composition and gene content of a biogas-producing microbial community from a production-scale biogas plant fed with renewable primary products was analysed by means of a metagenomic approach applying the ultrafast 454-pyrosequencing technology. Sequencing of isolated total community DNA on a Genome Sequencer FLX System resulted in 616,072 reads with an average read length of 230 bases accounting for 141,664,289 bases sequence information. Assignment of obtained single reads to COG (Clusters of Orthologous Groups of proteins) categories revealed a genetic profile characteristic for an anaerobic microbial consortium conducting fermentative metabolic pathways. Assembly of single reads resulted in the formation of 8752 contigs larger than 500 bases in size. Contigs longer than 10 kb mainly encode house-keeping proteins, e.g. DNA polymerase, recombinase, DNA ligase, sigma factor RpoD and genes involved in sugar and amino acid metabolism. A significant portion of contigs was allocated to the genome sequence of the archaeal methanogen *Methanoculleus marisnigri* JR1. Mapping of single reads to the *M. marisnigri* JR1 genome revealed that approximately 64% of the reference genome including methanogenesis gene regions are deeply covered. These results suggest that species related to those of the genus *Methanoculleus* play a dominant role in methanogenesis in the analysed fermentation sample. Moreover, assignment of numerous contig sequences to clostridial genomes including gene regions for cellulolytic functions indicates that clostridia are important for hydrolysis of cellulosic plant biomass in the biogas fermenter under study. Metagenome sequence data from a biogas-producing microbial community residing in a fermenter of a biogas plant provide the basis for a rational approach to improve the biotechnological process of biogas production.

1. Introduction

Renewable resources for energy production come more and more into public focus because of problems caused by the predictable shortage of fossil fuels in the next decades and by global warming due to CO_2 release from burning of fossil fuels. These problems can partly be circumvented by the production of biogas from plant or waste material in a biological process (Angelidaki and Ellegaard, 2003; Daniels, 1992; Weiland, 2003; Yadvika et al., 2004). Anaerobic degradation of plant biomass carried out in biogas plants can be subdivided into different metabolic steps. First, plant compounds including cell wall material such as cellulose and xylan are hydrolysed and converted into mono-, di- and oligosaccharides (Bayer et al., 2004; Cirne et al., 2007; Lynd et al., 2002). This hydrolysis step is conducted

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

0168-1656/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jbiotec.2008.05.008

mainly by cellulolytic *Clostridia* and *Bacilli*, but is often inefficient under anaerobic conditions. Sugar intermediates are fermented to organic acids (acidogenesis) which in turn are converted to acetate, CO_2 and H_2 by bacteria performing secondary fermentations (Drake et al., 1997, 2002; Myint et al., 2007; Shin and Youn, 2005). The final methanogenesis step is conducted by *Archaea* which are restricted to a limited spectrum of input substrates (acetate, CO_2 and H_2, some C_1 compounds like formate and alcohols) that can be used for methane formation (Deppenmeier et al., 1996). Hydrolysis, acidogenesis, and acetogenesis are conducted by members of the *Eubacteria*. Several biochemical reactions are thermodynamically only possible in close interaction of at least two different bacterial partners (e.g. syntrophic H_2 feeding) (Schink, 1997, 2006). The enzymology of methanogenic pathways has been analysed in detail for model systems (Blaut, 1994; Deppenmeier, 2002; Ferry, 1992, 1999; Reeve, 1992; Reeve et al., 1997; Schnürer et al., 1999). However, the composition and interactions within a biogas-producing microbial community, and the contribution of a specific bacterium to the overall process are mainly unknown. Moreover, the influence of physicochemical parameters on population structure and efficiency of biogas formation is still under investigation (Karakashev et al., 2005; Shigematsu et al., 2004, 2006). Thus, a rational approach to improve the performance of biogas plants is impossible at the moment.

The composition of biogas-producing microbial communities commonly is determined *via* construction of 16S-rDNA clone libraries and subsequent sequencing of 16S-rDNA amplicons (Huang et al., 2002; Klocke et al., 2007; McHugh et al., 2003; Mladenovska et al., 2003). Moreover, Polymerase Chain Reaction Single Strand Conformation Polymorphism (PCR-SSCP) followed by sequencing of obtained DNA-molecules was also used to elucidate community structures in biogas reactors (Chachkhiani et al., 2004). Another valuable marker for the analysis of methanogenic communities is the *mcrA* gene encoding a key-enzyme of methanogenesis, namely the α-subunit of methyl-coenzyme M reductase (MCR). Many methanogenic communities were analysed by using the *mcrA* gene as a phylogenetic marker (Lueders et al., 2001; Luton et al., 2002; Friedrich, 2005; Juottonen et al., 2006; Rastogi et al., 2007).

Development of second-generation ultrafast sequencing technologies such as 454-pyrosequencing led to the realisation of cost-effective large-scale environmental shotgun sequencing projects. Metagenomics became a versatile approach for exploration of different habitats for the structure, gene content and function of the respective autochthonous microbial communities. The number of metagenome projects using ultrafast sequencing techniques is constantly increasing (Angly et al., 2006; Edwards et al., 2006; Gill et al., 2006; Turnbaugh et al., 2006). Bioinformatics for the interpretation of metagenomic data has coordinately been improved (Raes et al., 2007). Recently, a novel gene finding algorithm that allows for exploitation of the limited information contained in the 250 nucleotides reads generated by 454-pyrosequencing for the prediction of coding sequences was developed (Krause et al., 2006). Moreover, design of bioinformatics strategies and tools for metagenomic data processing facilitates insights in community structures and gene content of microbial consortia from different habitats (Krause et al., 2008a,b).

Here, insight into the metagenome of a biogas-producing microbial community residing in the main fermenter of a production-scale biogas plant is presented. Obtained nucleotide sequence data were analysed at the single read and contig level for their genetic information content by applying different bioinformatics approaches.

2. Materials and methods

2.1. Total community DNA preparation from a fermentation sample of a biogas reactor

A fermentation sample was taken from the first biogas fermenter of an agricultural biogas plant located in Bielefeld-Jöllenbeck (Germany) in August 2007. The sample was stored in entirely filled, screw capped bottles and transferred to the laboratory. The analysed 500 kW biogas plant consists of two fermenters and a storage reservoir and was continuously fed with maize silage (63%), green rye (35%) and low amounts of chicken manure (appr. 2%). The substrate was fermented at appr. 41 °C at a pH-value of 7.7. The retention period of the substrate was 40–60 days. Further data for the analysed bioreactor were: volatile organic acids, 7739 mg acid L^{-1}; total anorganic carbon, 15,159 mg $CaCO_3$ L^{-1} and 2628 mg acetic acid L^{-1}. The biogas plant started to operate in December 2005.

First microscopic analysis of the fermentation sample was carried out within 2 h upon sampling. Samples were diluted with three parts of sterile tap-water. The diluted fermentation sludge was stained for 20 min by the addition of 2 μg/ml 4′,6′-diamidino-2-phenylindole hydrochloride (DAPI). Bacteria were visualised by a Nikon eclips 80i epi-fluorescence microscope equipped with an Plan Apo 60× (na 1.2) objective and DAPI filter settings (EX 340-380; DM 400; BA 435-485). Photos were taken using a Nikon camera in automatic mode.

A 20 g aliquot of the fermentation sample was used for total community DNA preparation by applying a CTAB (cetyltrimethylammonium bromide) containing DNA extraction buffer as described previously (Entcheva et al., 2001; Henne et al., 1999; Zhou et al., 1996). The obtained DNA pellet was resuspended in 8 ml TE buffer. One milliliter of the total genomic DNA preparation was purified on 10 MicroSpin S-400 HR sephacryl columns (GE Healthcare, München, Germany). After this final purification step ten DNA-eluates were pooled and subjected to precipitation using 40 μl NaCl (5 M) and 2 ml ethanol (−20 °C). After centrifugation (11,500 rpm, 10 min) the DNA-pellet was resuspended in 100 μl TE buffer. DNA concentration was estimated by means of the NanoDrop 2000 instrument (NanoDrop Technologies, Wilmington, USA) and analysed by gelelectrophoresis. The applied method yielded a highly pure genomic DNA ($A_{260}/A_{280} = 1.8$) with a concentration of 444 ng/μl.

2.2. Sequencing of the biogas reactor total community DNA preparation on a Genome Sequencer FLX System

Sequencing of the genomic DNA derived from the biogas reactor sample was done by applying the whole-genome-shotgun sequencing approach on the Genome Sequencer FLX System (Roche Applied Science, Mannheim, Germany). Approximately 5 μg of the DNA-preparation were used to generate a whole-genome-shotgun library according to the protocol supplied by the manufacturer. After titration, 3.5 DNA-copies per bead were used for the main sequencing run. After emulsion PCR and subsequent bead recovery, 900,000 DNA-beads were loaded onto each half of the PicoTiterPlate and subjected to sequencing. Obtained metagenome sequence reads are available *via* the link: ftp://ftp.cebitec.uni-bielefeld.de/pub/supplements/SchlueterEtAl_Metagenome_JournalBiotech_2008.zip.

The Genome Sequencer *De Novo* Assembler Software (Roche Applied Science, Mannheim, Germany) was used for assembly of the obtained nucleotide sequence reads.