Journal of Bioscience and Bioengineering
VOL. xx No. xx, 1–9, 2017

www.elsevier.com/locate/jbiosc

Characterization of cellulolytic microbial consortium enriched on Napier grass using metagenomic approaches

Pattanop Kanokratana,[1,*] Sarunyou Wongwilaiwalin,[1] Wuttichai Mhuantong,[1] Sithichoke Tangphatsornruang,[2] Lily Eurwilaichitr,[1] and Verawat Champreda[1]

Enzyme Technology Laboratory, National Center for Genetic Engineering and Biotechnology, 113 Thailand Science Park, Pahonyothin Road, Klong Luang, Pathum Thani 12120, Thailand[1] *and Genomic Research Laboratory, National Center for Genetic Engineering and Biotechnology, 113 Thailand Science Park, Pahonyothin Road, Klong Luang, Pathum Thani 12120, Thailand*[2]

Received 21 May 2017; accepted 25 October 2017
Available online xxx

Energy grass is a promising substrate for production of biogas by anaerobic digestion. However, the conversion efficiency is limited by the enzymatically recalcitrant nature of cellulosic wastes. In this study, an active, structurally stable mesophilic lignocellulolytic degrading microbial consortium (Np-LMC) was constructed from forest compost soil microbiota by successive subcultivation on Napier grass under facultative anoxic conditions. According to tagged 16S rRNA gene amplicon sequencing, increasing abundance of facultative *Proteobacteria* was found in the middle of batch cycle which was then subsequently replaced by the cellulose degraders *Firmicutes* and *Bacteroidetes* along with decreasing CMCase, xylanase, and β-glucanase activity profiles in the supernatant after 5 days of incubation. Anaerobic/facultative bacteria *Dysgonomonas* and *Sedimentibacter* and aerobic bacteria *Comamonas* were the major genera found in Np-LMC. The consortium was active on degradation of the native and delignified grass. Direct shotgun sequencing of the consortium metagenome revealed relatively high abundance of genes encoding for various lignocellulose degrading enzymes in 23 glycosyl hydrolase (GH) families compared to previously reported cellulolytic microbial communities in mammalian digestive tracts. Enzymes attacking cellulose and hemicellulose were dominated by GH2, 3, 5, 9, 10, 26, 28 and 43 in addition to a variety of carbohydrate esterases (CE) and auxiliary activities (AA), reflecting adaptation of the enzyme systems to the native herbaceous substrate. The consortium identified here represents the microcosm specifically bred on energy grass, with potential for enhancing degradation of fibrous substrates in bioenergy industry.

[**Key words:** Carbohydrate-active enzymes; High-throughput sequencing; Lignocellulose degradation; Napier grass; Microbial consortium]

The global energy industry is undergoing a rapid transition from its dependence on the nonrenewable fossil resource to more sustainable carbon-neutral resources. Lignocellulosic plant biomass is considered a promising resource for production of renewable energy in the forms of biogas and liquid fuels together with its potential for conversion to commodity chemicals and bio-polymers in biorefinery. Napier grass (*Pennisetum purpureum*), also called Elephant grass or Giant King grass, is a C4 perennial grass with high biomass. Its potential for use as a feedstock for renewable energy production has assessed in several parts of the world (1–3). A recent feasibility study on utilizing Napier grass as an energy crop for biogas production by anaerobic digestion has been reported (4). Understanding of the complex natural bioconversion of Napier grass and also other lignocellulosic biomass is thus an essential starting step towards efficient utilization of this alternative crop for bioenergy production.

Enhancement of bio-digestibility of lignocellulosic plant biomass by biological processes is a promising strategy because of its low capital cost, and low consumption of energy and chemicals. The natural degradation of lignocellulosic biomass, with it complex molecular structure, requires synergistic and cooperative actions of microbes producing a variety of cellulolytic and hemicellulolytic enzymes (5). Symbiotic multi-species lignocellulose degrading consortia, ranging from dual-species systems to complex microcosms represent a good candidate for enhancing biomass degradation in several biotechnological processes, for example, in biogas production from fibrous wastes (6). Several mesophilic and thermophilic structurally stable lignocellulolytic microbial consortia were isolated previously by successive enrichment on different cellulosic substrates (7–9). These consortia have been reported to effectively degrade a variety of cellulosic materials, such as agricultural residues, filter paper, cotton, and pulp wastes (9–12). In an attempt to define the minimal requirements of a consortium and understand the metabolic interdependencies among different members, a simplified lignocellulolytic bacterial consortium was constructed by combining bacterial isolates obtained from an initial enriched consortium (7). Subsequent studies on this defined microcosm revealed the roles of microbes contributing towards consortium stability and cooperative activity on lignocellulose degradation (13,14). The ability to construct stable, simplified and defined consortia opens the door for industrial application. Although a defined consortium of microbes is desirable, the

* Corresponding author. Tel.: +66 2564 6700x3464; fax: +66 2564 6707.
E-mail address: pattanopk@biotec.or.th (P. Kanokratana).

1389-1723/$ – see front matter © 2017, The Society for Biotechnology, Japan. All rights reserved.
https://doi.org/10.1016/j.jbiosc.2017.10.014

ecological complexity of consortia makes isolation of the required species difficult. For this reason, culture-independent approaches employing molecular methods such as next generation sequencing have been used extensively to reveal diversity and metabolic capabilities of microbial communities residing in various environments active in lignocellulose decomposition, e.g., bagasse compost (15), peat soil (16), termite gut (17), switchgrass-adapted compost (18) and cow rumen (19).

In this work, a lignocellulolytic microbial consortium (Np-LMC) working optimally under mesophilic anoxic condition was constructed from forest compost soil by successive enrichment on Napier grass. The dynamic community profiles and metabolic potentials of the Np-LMC were investigated using culture-independent methods on a next generation sequencing platform. The work provides insights into the cooperative actions of microbes and enzymes in decomposition of energy grass and shows the potential of Np-LMC as a promising microbial agent for enhancing conversion efficiency of fibrous lignocellulosic substrates to bio-energy and bio-products of economic importance.

MATERIALS AND METHODS

Materials Decomposed forest litter, resulting from natural degradation of forest litter, collected from several locations in the Queen Sirikit Botanic Garden, Chiang Mai, Thailand was used as a microbial source for preparation of lignocellulolytic consortia. After removing fresh leaf litter at the top layer, the decomposed forest litter attached to soil surface was collected. The decomposed forest litter samples contained approximately 30% of forest soil. Napier grass was obtained from a Compressed Biogas Plant (Universal Absorbent and Chemicals PCL, Chiang Mai, Thailand). Rice straw was obtained from a rice crop field in Pathum Thani province, Thailand. The biomass was physically processed by a cutting mill to particles of diameter approximately 1 mm and used as a substrate for consortium. Pretreated Napier grass and rice straw were prepared by alkaline pretreatment (10% w/v NaOH with 3:1 liquid/solid ratio at 30°C for 24 h). The pretreated biomass was washed with tap water until neutral pH (pH 7–7.5) was obtained and dried at 60°C for 24 h. The carbohydrate compositions of the biomasses were determined using the standard NREL method (20). Native Napier grass contained 34.3 wt% cellulose, 19.8 wt% hemicellulose, 22.8 wt% lignin and 11.6 wt% ash. The rice straw contained 35.8 wt% cellulose, 21.5 wt% hemicelluloses, 24.4 wt% lignin, and 15.0 wt% ash. The lignocellulosic substrates were autoclaved twice at 121°C for 15 min before use.

Construction of lignocellulolytic microbial consortia The structurally stable Np-LMC was prepared according to the method modified from Haruta et al. (11) with forest compost soil as seed culture. One gram of the environmental sample was inoculated into a 50 ml screw-cap disposable tube containing 30 ml of PCS medium (0.1% yeast extract, 0.5% peptone, 0.5% $CaCO_3$, and 0.5% NaCl, pH 8.0) supplemented with 1% dried Napier grass as a biomass substrate and a filter paper strip (0.3 g) as an indicator for cellulase activity. The mixture was incubated at 40°C under static anoxic conditions until the filter paper had disintegrated (approximately 4–7 d), after which 1 ml of the culture was then transferred into fresh medium. The procedure was then repeated until a structurally stable cellulolytic microbial community was obtained as indicated by denaturing gradient gel electrophoresis (DGGE). DGGE profiles of bacterial 16S rRNA gene were compared based on band presence/absence (Jaccard coefficient) and then the dendrogram was constructed by unweighted pair group method with arithmetic mean (UPGMA) method. The consortium was maintained in PCS medium containing 20% glycerol with no cellulosic substrate at −80°C for long term storage. The structurally stable consortium was used for all subsequent experimental studies.

Evaluation of crude Np-LMC hydrolytic enzyme activities Activities of cellulases and hemicellulases in the culture medium were measured based on the amount of liberated reducing sugars using 3,5-dinitrosalicylic acid (DNS) method (21). The crude enzymes were clarified by centrifugation at 7500 $\times g$ for 15 min. Reactions (320 µl) contained the appropriate dilution of enzyme in either 50 mM sodium acetate buffer, pH 5.0 or 50 mM phosphate buffer, pH 8.0, and 1 % (w/v) of the corresponding substrate (Sigma–Aldrich, Singapore): carboxymethyl cellulose for CMCase activity, birchwood xylan for xylanase activity, and β-glucan for β-glucanase activity. The reaction was incubated at 50°C for 10 min and then stopped by the addition of 2 volume of DNS solution. The amount of reducing sugars was determined from the absorbance measurements at 540 nm and interpolated from a standard curve of the corresponding sugars. Assay reactions were performed in triplicate, and the mean was reported. The coefficients of variations were all ≤5%. One unit (U) was defined as the amount of enzyme which produced 1 µmol of reducing sugar in 1 min. The total protein of the culture medium was determined using the Bradford assay kit (Bio-Rad, Hercules, CA, USA) with bovine serum albumin as the standard.

Lignocellulosic substrate degradation by Np-LMC The efficiency of Np-LMC on degradation of selected lignocellulosic substrates was studied. One milliliter of stable Np-LMC was inoculated into 30 ml PCS medium supplemented with 1% w/v dried native or pretreated Napier grass or rice straw and then incubated under static conditions at 40°C. Remaining solid biomass was collected on day 3, 6, and 10 by filtration through a 1 mm sieve and the residual weight was then determined gravimetrically after drying at 60°C for 2 d with uninoculated medium as a control. The hydrolysis efficiency was evaluated based on residual weight of the biomass.

DGGE analysis of microbial structure and dynamics After successive cultivation of the consortia, the stability of the bacterial community was determined using DGGE. Total DNA of the consortia was extracted using an SDS-based DNA extraction method (22) with slight modification described by Kanokratana et al. (23). Partial 16S rRNA gene sequence was amplified from the purified DNA using a DGGE primer pair; 338 GC_F, which was attached to a GC clamp at the 5′-terminus (5′-CGCCCGCCGCGCGCGGCGGGCGGGGCGGGGGCACGGGGGGACTCCTACGGGAGGC A-3′; GC clamp sequence underlined) and 518R (5′- ATTACCGCGGCTGCTGG-3′). Amplification was performed using *Taq* DNA polymerase (Thermo Scientific, Espoo, Finland) on a MyCycler thermal cycler (Bio-Rad) for 30 cycles of denaturation at 94°C for 30 s, annealing at 55°C for 1 min, and extension at 72°C for 2 min. Amplified 16S rRNA fragments were analyzed by the DCode system (Bio-Rad). The DGGE conditions were 8% (w/v) polyacrylamide gradient gels, 1x TAE electrophoresis buffer, 40–60% denaturant gradient (100% is defined as 14.3 M urea with 40% formamide). Electrophoresis was performed at 60°C for 18 h at a constant voltage of 80 V. The gel was stained with ethidium bromide and visualized on a UV transilluminator.

Diversity analysis of Np-LMC using tagged 16S rRNA gene amplicon sequencing The purified metagenomic DNA was used as the template for amplification of the partial 16S rRNA gene using universal bacterial primers E785F (5′-GGATTAGATACCCTGGTAGTCC-3′) and E1081R (5′-CTCACGRCACGAGCTGACG-3′) encompassing the 5 and 6 hypervariable regions in prokaryotic 16S rRNA gene attached with tagged barcode sequences (24). Polymerase chain reactions were performed using Phusion DNA polymerase (Thermo Scientific, Espoo, Finland) on a MyCyclerthermocycler for 25 cycles of denaturation at 98°C for 10 s, annealing at 66°C for 30 s, and extension at 72°C for 30 s. The amplicons were purified using Qiagen PCR purification kit (Qiagen, Hilden, Germany) and were quantified using NanoDrop ND-1000 Spectrophotometer (ThermoScientific, Wilmington, DE, USA). The sequences were determined using an ION PGM platform (Life Technologies, CA, USA) following the manufacturer's recommended protocols. The sequencing dataset was initially cleaned by removing low quality score reads with a cutoff of 20 for Phred quality score. Next, the remaining reads were classified into groups based on the tagged sequences and trimmed of the tag and primer sequences using quantitative insights into microbial ecology (QIIME) (25). Chimeric sequences were identified and removed using UCHIME (26). Taxonomic classification was assigned using the RDP classifier tool (27) to determine approximate phylogeny with 95% confidence threshold and further classified using BLASTN against the NCBI 16S rRNA database for phylogenetic assignment at the genus level using an expected E-value cutoff of 1e-06. Operational taxonomic units (OTUs) were determined at sequence similarity levels of 0.03, 0.05, and 0.15 by the furthest-neighbor method of MOTHUR (28). The diversity indexes and nonparametric diversity were calculated to measure and compare diversity among the dataset. Good's coverage (G), an estimator for sampling completeness, was calculated as $G = 1 - (n/N)$, where n is the number of singleton phylotypes and N is the total number of sequences in the sample. The 16s rRNA sequences were deposited in the NCBI Short Read Archive (www.ncbi.nlm.nih.gov/Trace/sra) with the accession number SRP063648.

Shotgun sequencing of Np-LMC metagenome The Np-LMC metagenome DNA library was sequenced on an ION Proton sequencing system (Life Technologies) following the manufacturer's protocols. The metagenome sequence reads obtained are available in the NCBI Sequence Read Archive (www.ncbi.nlm.nih.gov/Trace/sra) with the accession number SRP063648. The sequencing dataset was initially cleaned by removing reads with low quality scores and short reads with cutoff of 20 and 100 for Phred quality score and length, respectively. rRNA sequences were identified by BLASTN against the NCBI 16S rRNA database using an E-value cutoff of 1e-5. Non ribosomal-hit sequences were annotated using a composition-based approach as implemented in Metagenome Composition Vector (MetaCV) version 2.3.0 (29) against a reference database of 2059 prokaryote genomes. Outputs from MetaCV were filtered by a quality score of 20 (as recommended by Lie et al. (29) for read lengths of 100 nucleotides). Three carbohydrate-active enzyme classes from the CAZy database (30), namely glycoside hydrolases (GHs), carbohydrate esterase (CEs), and the recently introduced auxiliary activities (AAs) were identified by searching the sequence dataset against a library of modules derived from all entries in the CAZy database. All BLAST searches were performed using an expected value cutoff of 1e − 10. Heat maps were plotted via library "gplots" in R (31) using hits per million of reads as an input.

Please cite this article in press as: Kanokratana, P., et al., Characterization of cellulolytic microbial consortium enriched on Napier grass using metagenomic approaches, J. Biosci. Bioeng., (2017), https://doi.org/10.1016/j.jbiosc.2017.10.014