Journal of Bioscience and Bioengineering
VOL. 120 No. 2, 174–180, 2015

www.elsevier.com/locate/jbiosc

Analysis of bacterial xylose isomerase gene diversity using gene-targeted metagenomics

Dini Nurdiani,[1,‡] Michihiro Ito,[2,3,‡] Toru Maruyama,[1,3] Takeshi Terahara,[1] Tetsushi Mori,[3,4] Shin Ugawa,[5] and Haruko Takeyama[1,2,3,*]

Department of Life Science and Medical Bioscience, Waseda University, 2-2 Wakamatsu-cho, Shinjuku-ku, Tokyo 162-8480, Japan,[1] Institute for Nanoscience and Nanotechnology, Waseda University, 513 Waseda-tsurumaki-cho, Shinjuku-ku, Tokyo 162-0041, Japan,[2] Core Research for Evolutionary Science and Technology Project (CREST), Japan Science and Technology Agency, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan,[3] International Center for Science and Engineering Programs (ICSEP), Waseda University, 3-4-1 Okubo, Shinjuku-ku, Tokyo 169-8555, Japan,[4] and Forestry and Forest Products Research Institute, 1 Matsunosato, Tsukuba, Ibaraki 305-8687, Japan[5]

Received 16 November 2014; accepted 22 December 2014
Available online 3 February 2015

Bacterial xylose isomerases (XI) are promising resources for efficient biofuel production from xylose in lignocellulosic biomass. Here, we investigated xylose isomerase gene (*xylA*) diversity in three soil metagenomes differing in plant vegetation and geographical location, using an amplicon pyrosequencing approach and two newly-designed primer sets. A total of 158,555 reads from three metagenomic DNA replicates for each soil sample were classified into 1127 phylotypes, detected in triplicate and defined by 90% amino acid identity. The phylotype coverage was estimated to be within the range of 84.0–92.7%. The *xylA* gene phylotypes obtained were phylogenetically distributed across the two known *xylA* groups. They shared 49–100% identities with their closest-related XI sequences in GenBank. Phylotypes demonstrating <90% identity with known XIs in the database accounted for 89% of the total *xylA* phylotypes. The differences among *xylA* members and compositions within each soil sample were significantly smaller than they were between different soils based on a UniFrac distance analysis, suggesting soil-specific *xylA* genotypes and taxonomic compositions. The differences among *xylA* members and their compositions in the soil were strongly correlated with 16S rRNA variation between soil samples, also assessed by amplicon pyrosequencing. This is the first report of *xylA* diversity in environmental samples assessed by amplicon pyrosequencing. Our data provide information regarding *xylA* diversity in nature, and can be a basis for the screening of novel *xylA* genotypes for practical applications.

© 2015, The Society for Biotechnology, Japan. All rights reserved.

[**Key words:** Xylose; Xylose isomerase; Metagenome; Next generation sequencing; Biodiversity; Lignocellulosic biomass; Bioethanol production]

Xylose, a five-carbon sugar abundant in hardwood and agricultural residues, is a potential resource for biofuel production. The economic fermentation of xylose using xylose-fermenting microorganisms would greatly facilitate the utilization of lignocellulose biomass (1). *Saccharomyces cerevisiae* is a well-known producer of ethanol from glucose. However, no wild *S. cerevisiae* reported to date can efficiently ferment xylose. Attempts have been made to use genetically modified *S. cerevisiae* to produce bioethanol from xylose (2–5).

S. cerevisiae lacks the ability to utilize xylose but can utilize and ferment its isomer D-xylulose (6). On the other hand, some bacteria and fungi possess xylose isomerases (XIs) (7), which directly convert xylose to D-xylulose. The introduction of a gene encoding XI (*xylA*) into *S. cerevisiae*, therefore, can theoretically result in recombinant yeast cells capable of fermenting xylose to ethanol (5,8). The *xylA* genes have been classified into two groups, group I and II, based on their length, amino acid sequence similarity, and divalent cation preference (9). Group I XIs have 380 to 390 amino acid residues while group II XIs have 440 to 460 amino acids. They share only 20–30 % amino acid identities with each other, but the active site residues are highly conserved among the two groups and no general difference in enzymatic properties has been known so far. Their taxonomic distribution are also different; group I XIs are mainly in the phyla *Actinobacteria* and *Deinococcus-Thermus*, while group II XIs are mainly in the phyla *Proteobacteria* and *Firmicutes*. Nevertheless, xylose-fermenting activity has remained insufficient in heterologous hosts with known *xylA* (5,10). Further exploration of bacterial *xylA* is required for improving XI activity. Estimates of the genetic diversity of *xylA* in nature will be helpful for the efficient screening of suitable genes in *S. cerevisiae*.

Because diverse microbes inhabit soil environments that contain lignocellulosic biomass, soil metagenomes could be a promising resource to explore potential bacterial *xylA* genes. The *xylA* gene diversity in soil may be affected by the following: (i) soil organic carbon and nutrient content (11), and (ii) xylose sugar and carbon content, which differ among plant species growing in soil (12,13), causing variation in plant litter decomposition rates. Thus, it can be expected that soils that differ in plant vegetation can contain *xylA* genes with different levels of genetic diversity.

* Corresponding author at: Department of Life Science and Medical Bioscience, School of Advanced Science and Engineering, Waseda University, 2-2 Wakamatsu-cho, Shinjuku-ku, Tokyo 162-8480, Japan. Tel.: +81 3 5369 7326; fax: +81 3 5369 7302.
E-mail address: haruko-takeyama@waseda.jp (H. Takeyama).
‡ The first two authors contributed equally to this work.

1389-1723/$ – see front matter © 2015, The Society for Biotechnology, Japan. All rights reserved.
http://dx.doi.org/10.1016/j.jbiosc.2014.12.022

High-throughput targeted metagenomics of amplified functional gene sequences based on next generation sequencing (NGS) has provided deep insights into their diversity, distribution, and ecological roles as well as into the relationships between microbial assemblages and particular ecological functions (14–18). Elucidating *xylA* diversity in soil by gene-targeted metagenomics will provide a basis for the screening of novel *xylA* genes useful for bioethanol production. Parachin and Gorwa-Grauslund (2011) reported the isolation of two novel *xylA* genes from a soil metagenomic library while they did not focus on the diversity and distribution of *xylA* genes in soil (10). In this study, we aimed to uncover the sequence diversity of *xylA* genes in three different soil metagenomes by *xylA*-targeted metagenomics. Although two degenerate primer sets have been reported to amplify *xylA* gene sequences (10), they were designed based on only 11 *xylA* gene sequences. Since our purpose was to obtain much more divergent *xylA* gene sequences, two sets of new degenerate primers were designed to explore group I and group II *xylA* diversity. The microbial community structures in the soil samples were also analyzed by amplicon pyrosequencing, and the association between *xylA* and 16S rRNA gene diversity was examined.

MATERIALS AND METHODS

Soil sampling Soil samples were collected from three sites: two were located at Tsukuba mountain (Ibaraki prefecture, Japan) (sites T1 and T2) and one was located at Shirakaba Lake (Nagano prefecture, Japan) (site S1). Soil sampling was conducted on May 2009 at S1 and June 2009 at T1 and T2. After carefully removing the surface plant litter, soil samples were collected measuring 50 (width) × 50 (length) × 5 (top soil depth) cm. The soil samples were placed into individual sterile plastic bags and stored at 4°C. Each soil sample was stored in 10 g aliquots at −20°C for subsequent DNA extraction.

The three sampling sites were chosen due to their differences in plant vegetation, to maximize the soil microbial composition and abundance. The T1 sampling site comprises *Criptomeria japonica* (needle-leaf tree) vegetation, while the T2 and S1 sites comprise *Fagus crenata* and *Betula platyphylla* plant vegetation, respectively, which are birch-leaf trees. Based on a reported analysis of plant litter composition (12), the holocellulose fraction in needle- and birch-leaf litter, as a whole, is fairly constant, while the xylan composition in birch-leaf litter is 2- to 3-fold higher than that of needle-leaf litter.

DNA extraction Metagenomic DNA was extracted from the aliquot of the homogenate of each soil sample using ISOIL (Nippongene, Tokyo, Japan) according to the manufacturer's instructions. The final DNA concentration was measured spectrophotometrically yielding DNA concentrations averaging between 4.6 and 5.2 μg/g soil. Each soil DNA sample was extracted in triplicate, and each metagenomic DNA (nine total samples) was used as template DNA for PCR amplification of partial *xylA* sequences.

PCR amplification of *xylA* and 16S rRNA gene To amplify *xylA* genes from soil metagenomes, two degenerate primer sets (named A and B) were designed based on conserved regions of 112 known amino acid XI sequences collected from the NCBI database. Primer set A consisted of xyl1 (5′-TGGGGNGGNCGNGARGGNA-3′) and xyl2 (5′-RAAYTSRTCNGTRTCCCARCC-3′) (Fig. 1). The two oligonucleotides were designed manually against the conserved amino acid regions WGGREGY and GWDTDEF, respectively. Primer set B consisted of xyl30F (5′-TGTGTTTTGG GGCGGNMKNGANGG-3′) and xyl30.4R (5′-GTTATGGCCCGCCADNKKNKCRTG-3′) (Fig. 1). The primers were designed against the conserved amino acid regions VFWGGREG and HEQMAGHN, respectively, using the CODEHOP program (19).

Partial fragments of *xylA* genes were amplified from the respective soil metagenomic DNA samples with the following primer sets containing 454 pyrosequencing adaptors (underlined), common MID sequence tags for 454 pyrosequencing (indicated in bold), and the degenerate primer sets describe above. The primer set designated pyro-set A comprised the forward and reverse primers 5′-CCATCTCATCCCTGCGTGCTCCGACTCAG-[common MID]-TGGGGNGGNCGNGARGGNTA-3′ and 5′-CCTATCCCCTGTGTGCCTTGGCAGTCTCAGRAAYTSRTCNGTRTCCCARCC-3′, and the primer set designated pyro-set B comprised the forward and reverse primers 5′-CCATCTCATCCCTGCGTGCTCCGACTCAG-[common MID]-TGTGTTTTGGGGCGGNMKNGANGG-3′ and 5′-CCTATCCCCTGTGTGCCTTGGCAGTCTCAGGTTATGGCCCGCCADNKKNKCRTG-3′. PCR amplification with pyro-set A was performed in a 25 μL mixture (total volume) containing 10 ng soil DNA, 2.5 μL 10-fold reaction Ex Taq buffer, 0.2 mM dNTP, 4 μM of each primer, and 1.25 U Ex Taq HS polymerase (TaKaRa-Bio Inc., Otsu, Japan) with the following PCR conditions: 2 min of initial denaturation at 94°C, 25 cycles of 30 s denaturation at 94°C, 30 s annealing at 68°C, and 30 s extension at 72°C, followed by a 7 min final extension at 72°C. For amplification using pyro-set B, a second round of PCR was conducted with the PCR products amplified by primer set B as the template. The amplification using set B and pyro-set B was performed in a 25 μL mixture (total volume) containing 10 ng soil DNA, 4 μM of each primer, and 12.5 μL Premix Ex Taq HS polymerase (TaKaRa-Bio Inc.) with the following PCR conditions: 2 min of initial denaturation at 94°C, 25 cycles of 30 s denaturation at 94°C, 30 s annealing at 58°C, and 15 s extension at 72°C, followed by a 7 min final extension at 72°C. PCR products were then analyzed by agarose gel electrophoresis and purified using a Gel Extraction Kit (Qiagen GmbH, Hilden, Germany) according to the manufacturer's instructions.

16S rRNA genes were amplified with a universal primer set (27F/338R) (20) accompanied by 454 pyrosequencing adaptors (underlined) and the common MID tag sequences for each metagenome sample (as indicated in bold). The primer sequences were as follows: pyro-27F (5′-CCATCTCATCCCTGCGTGTCTCCGACTCAG-[common MID]-AGAGTTTGATCMTGGCTCAG-3′) as the forward primer and pyro-338R (5′-CCTATCCCCTGTGTGCCTTGGCAGTCTCAGTGCTGCCTCCCGTAGGAGT-3′) as the reverse primer. For each sample, PCR amplification of 16S rRNA gene was performed in a 25 μL mixture (total volume) containing 10 ng soil metagenome DNA, 12.5 μL PrimeSTAR Max DNA polymerase (Takara) and 0.2 μM of forward and reverse primers with the following PCR conditions: 5 min of initial denaturation at 98°C, 25 cycles of 10 s denaturation at 98°C, 15 s annealing at 49°C, and 5 s extension at 72°C,

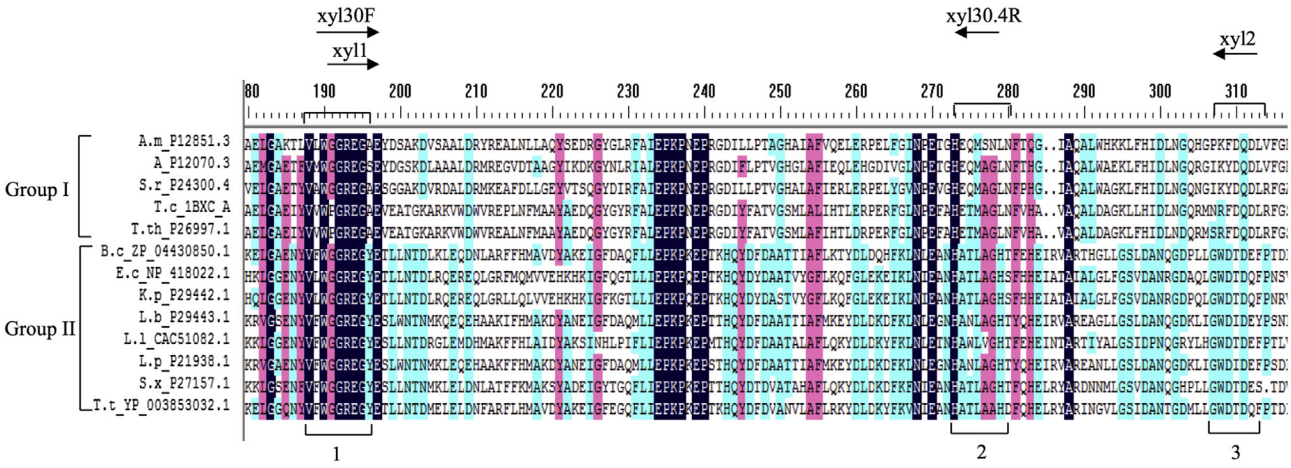

FIG. 1. Multiple sequence alignment of known *xylA* genes collected from the NCBI database. Species names are followed by accession numbers; A.m, *Actinoplanes missouriensis*; A, *Arthrobacter* sp.; S.r, *Streptomyces rubiginosus*; T.c, *Thermus caldophilus*; T.th, *Thermus thermophilus*; B.c, *Bacillus cereus*; E.c, *Escherichia coli*; K.p, *Klebsiella pneumoniae*; L.b, *Lactobacillus brevis*; L.l, *Lactococcus lactis*; L.p, *Lactobacillus pentosus*; S.x, *Staphylococcus xylosus*; T.t, *Thermoanaerobacterium thermosaccharolyticum*). The colors in the alignment indicate conservation among the sequences; dark blue, completely conserved; pink, ≥80% conserved; green, ≥50% conserved. Degenerate primer design for *xylA* gene amplification was based on the conserved regions. Primer set A was designed based on conserved regions 1 and 3 with a fragment length of approximately 380 bp, while primer set B was designed based on conserved regions 1 and 2 with a fragment length of approximately 290 bp. The arrows represent the primer regions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Download English Version:

https://daneshyari.com/en/article/20249

Download Persian Version:

https://daneshyari.com/article/20249

Daneshyari.com