Comparative metagenomic discovery of the dynamic cellulose-degrading process from a synergistic cellulolytic microbiota

Ming Yang · Jingjing Zhao · Yue Yuan · Xiaoyi Chen · Fan Yang · Xianzhen Li

Abstract To reveal the dynamic process of cellulose biodegradation and explore more potential cellulases, a microbiota (FPDM) with cellulose-degrading ability was cultivated, and different stages of filter paper degradation were compared. Ion chromatography and comparative metagenomic sequencing revealed that the diversity of FPDM enhanced as the hydrolysate length diversity increased. Sporocytophaga and Cohnella dynamically dominated the synergistic degradation of cellulose in early-intermediate and intermediate-final periods, respectively. Moreover, 363 declining shifting and 231 progressive shifting unannotated genes were speculated to participate in the catabolism of cellulose to cellodextrin/cellooligosaccharide and to cellobiose, respectively. Based on the dynamic changes in hydrolysates, community structure and gene abundance, a dynamic cellulose-degrading pathway of FPDM was predicted. Our work should provide a new perspective for subsequent identification of key cellulolytic strains and enzymes and clarification of the mechanism of cellulose biodegradation.

Keywords Comparative metagenome · Dynamic cellulose degradation process · Functional genes · Synergistic cellulolytic microbiota

Introduction
Currently, the development of cellulose-derived renewable biofuels is of great significance for recycling agricultural waste and relieving the shortage of fossil fuels (Rodionova et al. 2017). However, the biodegradation of cellulose is difficult due to its recalcitrant nature and the relatively low activity of currently available hydrolytic systems, which is considered to be a major obstacle to industrial-scale production of fuel from cellulose (Jayasekara and Ratnayake 2019). Therefore, it is urgent to further clarify the molecular mechanism of cellulose biodegradation and explore efficient cellulose hydrolysis systems.

Recently studies revealed that cellulose decomposition by pure cultures of the cellulolytic strain was
less efficient than in mixed cultures (Dumova and Kruglov 2009; Jiménez et al. 2014; Wongwilaiwalin et al. 2010). In natural environments, cellulolytic microbes work together with other microorganisms that utilize soluble sugars (cellobiose and glucose) to catabolize insoluble cellulosic matter (Deng and Wang 2016; Dumova and Kruglov 2009). The cellulolytic microbes could secrete several cellulases, including endoglucanase, cellobiohydrolase and beta-glucosidase (Juturu and Wu 2014), which catalyze the hydrolysis of the beta-1,4-glycosidic bonds in a cellulose molecule (Herrera et al. 2019; Lee et al. 2016). Catabolite repression of cellulases can be relieved due to the subsequent consumption of cellulose degradation products (Dumova and Kruglov 2009). Therefore, the de novo synthesis of cellulases and increased biomass of cellulolytic microbes drive continuous decomposition of cellulose. However, due to the diversity of natural cellulose substrates and habitats, these cellulolytic communities are biologically diverse, which hampers the elucidation of their structure and function (Wilson 2011).

As a result of the rapid development of next generation sequencing technology in recent years, metagenomic sequencing has become a powerful strategy for exploring the structure of cellulolytic microbiota and mining functional cellulose degradation related strains and enzymes (Knight et al. 2018). For instance, metagenome analysis of a rice straw-adapted microbial consortium enriched from compost revealed a distinctive set of microorganisms and carbohydrate-active enzyme (CAZyme) genes distributed within these bacteria phyla, suggesting a synergistic system efficient in catabolism of cellulosic components in the compost habitat (Wang et al. 2016). With the help of culture-independent “omics” techniques, the metabolic “blue-print” of a proficient cellulose-degrading microbiota enriched from soil sample was predicted, and the putative operation of cellulosome and free-enzyme within the community was demonstrated (Zhou et al. 2014). Despite the exciting findings from metagenomic researches of cellulolytic microbiota, there is still much we do not understand about the cellulose-degrading mechanism and the function of numerous putative proteins, which can be partly attribute to the lack of knowledge about the dynamic changes of the microbial community at different cellulose-degrading stages (Mande et al. 2012).

To further clarify the cellulose degrading mechanism and explore more potential cellulases, we cultured a stable soil-derived microbiota (FPDM) capable of efficiently degrading filter paper in the present study. The dynamic changes in filter paper hydrolysates and the microbiota composition at different degradation stages were respectively analyzed by ion chromatography (IC) and metagenomic shotgun sequencing, revealing the synergistic action of the core functional strains. Notably, based on short-time series expression miner (STEM) algorithm (Ernst and Bar-Joseph 2006), the function of some key functionally unannotated genes (FUGs) were predicted and a more distinct metabolic pathway involved in cellulose degradation were mapped.

Materials and methods

Enrichment of the filter paper-adapted microbiota from soil samples

Lignocellulose-rich soil samples were collected from woodland soil in the campus of Dalian Polytechnic University, China (38°58’59.2176” N, 121°31’55.1964” E). These soil samples were respectively suspended in sterilized water and dispersed onto a double-layered filter paper placed on a solid medium, then pressed with a flat-ended glass rod to ensure sufficient contact between the soil and filter paper and incubated at 30 °C. After 15-day fermentation, a piece of filter paper containing the cellulose-degrading microbiota was cut and transferred onto the same filter paper medium and incubated at 30 °C for 15 days. Then the second batch microbiota culture was transferred again onto the same new medium. This subculture process was repeated for 10 times, and a stable microbiota named FPDM with an efficient cellulolytic ability was established. Further fermentation was conducted in double-layered filter paper solid medium and liquid filter paper medium, respectively, and performed in biological triplicate. The double-layered filter paper solid medium applied a circular Whatman No. 1 filter paper (Whatman, Tokyo) as the only carbon source. The filter paper was presoaked with sterilized water and then placed on the agar medium with the inorganic components. The inorganic components of the solid media contained per liter: 0.5 g each of KCl, NaNO₃, and MgSO₄·7H₂O;
0.01 g of FeSO₄ 7H₂O; and 1 g of K₂HPO₄, pH 7.0. To prepare the liquid medium, 5 g per liter filter paper (cut into pieces 5 mm in length and width) was added into the same inorganic solution as the solid medium. The supernatant was collected from 100 mL microbiota FPDM culture after being cultivated in liquid medium at 30 °C with shaking (180 rev/min) for 3, 8, and 15 days for the subsequent IC analysis. The FPDM culture after being cultivated in solid medium for 3, 8, and 15 days was also sampled for the subsequent metagenomic analysis.

Ion chromatography

Considering the cellulose hydrolysates were difficult to be detected from the FPDM solid culture, hydrolysates from FPDM liquid culture were determined by high-performance anion-exchange chromatography using an ICS-5000 IC system (Dionex, USA) equipped with a pulsed amperometric detector and a CarboPac PA100 column. The samples were analyzed by gradient elution. A 1.0 mL/min of mobile phase flow rate with a 25 μL of injection volume was adjusted at 30 °C of column temperature.

DNA preparation and metagenomic sequencing

Totally 1 g of cellulolytic microbiota FPDM samples adhere to the filter paper were collected for nucleic acid extraction. Metagenomic DNA was isolated using a PowerSoil Kit (MoBio, USA) following the manufacturer’s recommendations. Nanodrop-1000 (Thermo Scientific, USA) and gel electrophoresis were used to determine the DNA quality. High-quality DNA was then stored at −20 °C before library construction.

DNA libraries were constructed using a NEBNext® Ultra™ DNA Library Prep Kit for Illumina (New England Biolabs, USA) under the manufacturer’s protocols. First, 1 μg of each metagenomic DNA sample was fragmented to a size of ~350 base pairs (bp) using an ultrasonicator (Covaris, USA). Next, the fragmented DNA was end-polished, A-tailed, and ligated with Illumina adaptors for further polymerase chain reaction (PCR) amplification. The PCR products were then purified using the AMPure XP system (Beckman Coulter, USA); libraries were profiled for size distribution using an Agilent 2100 Bioanalyzer (Agilent Biotechnologies, USA) and quantified by quantitative PCR (Wang et al. 2018). Finally, the high-quality libraries were sequenced on an Illumina NovaSeq 6000 platform (Novogene, China) using the paired-end 150 bp sequencing strategy.

Metagenomic data preprocessing, assembly, and annotation

Raw reads that contained a >15 bp of adapter overlap, >40 low-quality bases (Q-value < 38) or >10% “N” bases were removed using SOAPnuke (Chen et al. 2018). The remaining clean reads from each sample were then mapped against the corresponding scaftigs using SOAPaligner (Li et al. 2009). Subsequently, all unmapped reads were mixed and assembled using SOAPdenovo2 with the same parameters. Open reading frames (ORFs) were predicted from scaftigs longer than 500 bp using MetaGeneMark (Sunagawa et al. 2015; Zhu et al. 2010). To obtain a nonredundant gene catalog, ORFs of more than 100 bp were clustered at 95% sequence identity using CD-HIT (Li and Godzik 2006; Sunagawa et al. 2015). SOAPaligner was used to align the clean reads with the nonredundant gene catalog, and only genes supported by at least three mapped reads were designated as unigenes for further analysis (Qin et al. 2010). The unigene abundance was calculated for each sample as the number of reads mapped to the unigene and normalized by unigene length (Villar et al. 2015). The unigenes were mapped against the National Center for Biotechnology Information (NCBI) nonredundant protein (NR) database (ftp://ftp.ncbi.nlm.nih.gov/blast/db/FASTA/) and the available fungi genomes from EnsemblFungi database (https://fungi.ensembl.org) for taxonomic annotation using blastp in DIAMOND (Buchfink et al. 2015), with a maximum e-value of 1e−5. MEGAN (Huson et al. 2011) was applied to determine the taxonomic level of each unigene using the lowest common ancestor-based algorithm. The abundance of each taxonomic group was estimated by summing the abundance of unigenes assigned to the same feature (Li et al. 2014). The relative abundance of sequence from eukaryotes, virus and unclassified organisms were also calculated, and showed that the sum of these
relative abundances is only 0.052%. Therefore, downstream analyses focused only on prokaryotic taxa, genes and genomes. The Kyoto Encyclopedia of Genes and Genomes (KEGG) database (Kanehisa et al. 2008) was used to evaluate the function and metabolic pathways of the unigenes via blastp in DIAMOND. To mine the carbohydrate-active genes, the unigenes were annotated and selected using a hidden Markov model-based database for automated CAZyme annotation (dbCAN) (Lombard et al. 2014; Yin et al. 2012). To analyze the shift pattern of relative abundance, the STEM algorithm was employed to cluster the relative abundance variant profiles of the genus-level taxa, functional genes and the CAZyme families respectively.

**Metagenomic contigs binning**

In order to recover high quality metagenome-assembled genomes (MAGs), all clean reads were mixed and assembled by SOAPdenovo2, and the resulting contigs were then binned using three different binning software tools with Binning module in MetaWRAP pipeline (v1.2.1) (Uritskiy et al. 2018), including MetaBAT2 (Kang et al. 2019), MaxBin 2.0 (Wu et al. 2016), and CONCOCT (Alneberg et al. 2014). The resulting bins were collectively processed to produce consolidated MAGs with both Bin_refinement (criterion: completeness > 70%; contamination < 10%) and Reassemble_bins modules in MetaWRAP. The bins resulting from the Bin_refinement module were qualified with Salmon using the Quant_bins module of MetaWRAP. Bins obtained from each relative step above were subjected to quality control by estimating completeness and contamination using CheckM (Parks et al. 2015). Gene prediction and functional annotation were carried out using Annotate_bins in MetaWRAP based on Prokka gene database (Seemann 2014). CAZyme annotation for each bin was performed using the same method to that for unigenes as described above. These MAGs were then uploaded to MiGA (Rodriguez-R et al. 2018) for a complementary analysis to determine the most likely taxonomic classification and novelty rank of the bins at genus-level resolution.

**Statistical analysis**

The alpha diversity was characterized by calculating the Shannon–Wiener and Simpson indices using Mothur (Schloss et al. 2009). The R package “agricolae” (http://tarwi.lamolina.edu.pe/~fmeniburu) was used to determine significant differences ($P < 0.05$) in alpha diversity indices and their change rates between treatments. The beta-diversity for Bray–Curtis dissimilarities was calculated using the “beta_diversity.py” program in QIIME (Caporaso et al. 2010), and visualized by nonmetric multidimensional scaling (NMDS), which was performed using the “WGCNA”, “stat”, and “ggplot2” packages in the R software. Bar plots and heat maps were created using the R packages “ggplot2” and “pheatmap”, respectively. Pairwise comparisons of the taxonomic and functional profiles between treatments were carried out using Metastats (White et al. 2009).

**Prediction of gene function and construction of cellulolytic metabolic pathway**

Based on the abundance variation patterns, FUGs were clustered together with annotated cellulolytic genes (ACGs) and were thus assigned a certain function similar to those of ACGs from the same pattern. Furthermore, a dynamic pathway of cellulose degradation was predicted for FPDM based on the results of IC, annotation and prediction of gene function, and taxonomical annotation. The detailed steps were as follows:

1. FUGs and ACGs that displayed no significant difference ($P > 0.05$) in relative abundance between treatments, had a low average relative abundance ($< 0.01\%$) in all treatments, or were not assigned at the genus level were discarded. Moreover, FUGs derived from cellulolytic bacteria were screened out for further analysis.

2. The remaining FUGs and ACGs were mixed together and clustered based on the abundance variation patterns using STEM algorithm. Abundance variation patterns with fewer than 10 ACGs were removed. FUGs clustered with a group of ACGs ($\geq 60\%$) that had only one certain function, were presumed to have the same function with them.
(3) To establish the cellulolytic pathway for FPDM, the degradation products determined by IC were first mapped to the cellulolytic pathway according to the order of metabolism. Subsequently, in combination with the pathways and reactions described in the KEGG database, ACGs and putative cellulolytic genes (PCGs) were mapped to the cellulolytic pathway of FPDM based on the annotated and putative functions, respectively.

Isolation, identification and functional analysis of dominant strains

The microbiota FPDM was cultured in liquid filter paper medium at 30 °C, and a tenfold serial dilution of the FPDM culture with sterile saline was spread on the solid medium plates consisting of 2% glucose, 0.2% yeast extract, 0.05% MgSO₄, and 1.5% agar at pH 7.0, and incubated at 30 °C for 8 days. Colonies were randomly chosen from the plates and purified by repeating the above-mentioned spread plate method for 5 times. The isolated strains were identified using full-length 16S rRNA sequence analysis as described previously (Hayat et al. 2012).

To confirm synergistic cellulose-degrading capability of two dominant strains, *Sporocytophaga* sp. CX and *Cohnella* sp. CX were simultaneously inoculated into liquid filter paper medium with 2% (v/v) inoculum volume for a single strain, and co-fermented at 30 °C for 3 days and 8 days, respectively. As a control, a single *Sporocytophaga* sp. CX strain was cultivated in liquid filter paper medium using filter paper and cellodextrin as the sole carbon source for 48 h. To measure the cellulose-degrading activity, enzyme assay of the culture supernatant was performed. The endo-1,4-beta-glucanase activity was measured by 3,5-dinitrosalicylic acid (DNS) method as described previously (Miller 1959), and carboxymethyl cellulose (CMC) was used as the substrate (Younesi et al. 2016). The cellobiosidase activity was assayed using p-nitrophenyl-beta-D-cellobioside (pNPC) (Song et al. 2016), microcrystalline cellulose (Avicel PH101) (Kołaczkowski et al. 2020) and phosphoric acid swollen cellulose (PASC) (Dos Santos et al. 2020) as substrate, respectively. The beta-glucosidase activity was tested with p-nitrophenyl-beta-D-glucopyranoside (pNPG) as substrate as reported before (Song et al. 2016). The cellobiose dehydrogenase activity was assayed according to Sulej et al. (2013). All the assays were performed in triplicate and the mean values ± standard deviation were presented.

Results and discussion

Sequencing data, genome assembly and annotation of microbiota FPDM

Overall, ~ 198.11 million high quality clean pair-end reads (mean ~ 22.01 million per sample) were generated from the nine microbiota FPDM samples, ~ 187.79 million (94.79%) of these reads were used to assemble for each corresponding sample and obtained a total of 82,261 scaffigs (> 500 bp) with a total length of 232.89 megabases (Mb) and an N50 length of 37.35 kilobases (Kb). Subsequently, the unused reads were mixed and assembled, which resulted in 3,473 scaffigs. Considering each assembly, an average of 29,436 ORFs per sample were predicted, and a total of 44,829 unigenes were obtained by clustering all these ORFs, of which 32,614 (72.81%) could be assigned a putative function based on NCBI NR database. In total, 172 CAZyme families and carbohydrate-binding modules (CBMs) were identified, and 1,956 (4.36%) unigenes were assigned to them (Table S1). Specially, almost all of these CAZymes and/or CBMs related genes were present in the members of Firmicutes, Bacteroidetes, Proteobacteria and Actinobacteria (Table S2).

To perform a comprehensive characterization of microbiota FPDM, the MAGs were retrieved from the metagenomic sequences using MetaWRAP pipeline, yielding six near-complete draft bacterial genomes with sizes from 3.41 to 8.06 Mb, GC contents from 36.44 to 69.45%, and estimated genome completeness over 96%. A total of 32,642 ORFs were predicted from these MAGs, of which 20,844 were functionally annotated based on Prokka gene database. The taxonomic assignments via determining average amino acid identity (AAI) based on the type material...
genomes and NCBI Genome database (Prokaryotes) of MiGA showed that the bins labeled from Bin.1 to Bin.6 were assigned to *Achromobacter* (Proteobacteria), *Herbaspirillum* (Proteobacteria), *Sporocytophaga* (Bacteroidetes), *Cohnella* (Firmicutes), *Microbacterium* (Actinobacteria) and *Paenibacillus* (Firmicutes), respectively (Table S3). CAZyme analysis showed that a lot of CAZyme genes were detected from these bins across the process of cellulose degradation, especially in Bin.3 and Bin.4 (Table S4).

The diversity of microbiota FPDM increased with the increment of filter paper hydrolysate complexity.

The filter paper degradation process by microbiota FPDM could be divided into three stages: the early stage (3-day cultivation, FPDM.3), the intermediate stage (8-day cultivation, FPDM.8) and the final stage (15-day cultivation, FPDM.15). During this process, the morphology of filter paper changed from the center light yellow strain lawn to all over spread sticky or even translucent strain lawn (Fig. 1). This observation was similar to that reported by Dumova and Kruglov (2009), who observed a slight degradation area in filter paper after the microbiota sample was cultured for 7 days. To further analyze the filter paper hydrolysates at different degradation stages, IC analysis was performed. As shown in Fig. 2, a large number of short-chain saccharide molecules with a degree of polymerization (DP) greater than 15 were detected in FPDM.3 culture. After 8 days of continuous degradation, the short-chain molecules were broken down into cello-oligosaccharides and cellobiose with DP values of 2–6. At FPDM.15 stage, the cello-oligosaccharides and cellobiose contents decreased, and the end product, glucose, was observed. All the results suggest that the product complexity, i.e., the length diversity of polymer chain in hydrolysates, increased during the early-intermediate period of cellulose degradation.

The disturbances in the diversity and structure of microbiota FPDM at different cellulose degradation stages were assessed using Shannon–Wiener and Simpson indices. As shown in Table 1, the species diversity of the microbiota FPDM significantly increased in the intermediate and final stages compared with the early stage (*P* < 0.05). In addition, the increasing rate of species diversity in the early-intermediate period was significantly faster than that in the intermediate-final period (*P* < 0.05). The 2-dimensional NMDS plot showed that the microbiota FPDM.3 was well separated from FPDM.8 and FPDM.15, which were relatively close to each other (Fig. 3), suggesting that the structure of FPDM was...
disturbed in the early-intermediate period but relatively stable in the intermediate-final period, which coincided well with the alpha diversity results. Considering the changes of filter paper hydrolysates during the degradation process, it could be deduced that the diversity of microbiota FPDM increased with the increment of the length diversity of polymer chain in hydrolysates. Therefore, we hypothesized that the complexity of filter paper hydrolysates affected the way strains interacted with greater cooperation according to the study of Deng and Wang (2016).

Different shift patterns of relative abundance reveal the synergistic action of strains in microbiota FPDM

To reveal the “jungle rule” in microbiota FPDM due to cooperation and competition for carbon sources in the process of cellulose degradation, the relative abundance of strains in microbiota FPDM and their shifts were analyzed. As shown in Fig. 4a, Bacteroidetes, Firmicutes and Proteobacteria were the main phyla; they accounted for over 96% of the abundance in FPDM samples from different cellulose degradation stages, which was consistent with previous studies (Berlemont et al. 2014; Rosnow et al. 2016). At the genus level, the genera with a total relative abundance greater than 0.5% included Sporocytophaga (40.82%), Cohnella (34.85%), Herbaspirillum (5.87%), Achrobacter (4.99%), Paenibacillus (2.00%), and Microbacterium (0.54%). As illustrated in Fig. 4b, there were 6 relative abundance-shifting patterns among microbial taxa in FPDM. These patterns could be roughly divided into two types: progressive type (pattern 8, 11, 12 and 15) and declining type (pattern 0 and 4). The declining shift microorganisms were deduced to be involved in the early stage of cellulose degradation (cellulose to cellodextrin). Notably, almost all declining taxa belonged to Bacteroidetes, with Sporocytophaga as the representative genus, which has been proved to glide on solid surfaces of cellulose substrates and initiate cellulose degradation (Berg et al. 1972; Liu et al. 2014; Taillefer et al. 2018). On the other hand, progressive shift taxa might play important roles in catabolizing metabolites such as cello-oligosaccharides, cellobiose and glucose that are produced in the intermediate-final degradation stage. Most of the progressive shift taxa were derived from Firmicutes and Proteobacteria. Cohnella, the most

| Table 1 Alpha diversity indices of different FPDM microbiota samples |
|------------------|------------------|------------------|
| Alpha diversitya | Shannon–Wiener   | Simpson          |
| FPDM.3           | 1.540 ± 0.023b   | 0.367 ± 0.012a   |
| FPDM.8           | 1.695 ± 0.012a   | 0.305 ± 0.003b   |
| FPDM.15          | 1.766 ± 0.008a   | 0.288 ± 0.002b   |

Normalized differential alpha diversityb

|                  | Shannon–Wiener (day−1) | Simpson (day−1) |
|------------------|------------------------|------------------|
| FPDM.3 vs FPDM.8 | 0.0311 ± 0.0047a        | - 0.0123 ± 0.0025a |
| FPDM.8 vs FPDM.15| 0.0102 ± 0.0014b        | - 0.0025 ± 0.0005b |

aFPDM.3, FPDM.8 and FPDM.15 refer to microbiota FPDM cultivated in fiber medium for 3, 8, and 15 days, respectively.
bNormalized differential alpha diversity is the difference between the latter and the former alpha diversity indices divided by the time lag.

Within a single column, statistical differences of the values between the treatments were labeled with different letters (P < 0.05); Values represent means ± standard error.
dominant genera from Firmicutes, has been reported to contain a set of genes encoding cellulases and possess the ability to utilize cellulose, hemicellulose, and cellobiose (Kämpfer et al. 2006; Eida et al. 2012; Khianngam et al. 2012; Hameed et al. 2013). Therefore, it could be deduced that these carbon source utilizing features of progressive shift and declining shift taxa in FPDM were consistent with IC results, which revealed that cellulose were processively hydrolyzed into short-chain saccharides (DP \[\leq 15\]), cello-oligosaccharides and cellobiose, and glucose during different degradation stages. Furthermore, all these progressive and declining microorganisms have been identified in cellulolytic microbiota from other samples, such as industrial samples, the gut of native insects (Manfredi et al. 2015), peat–straw (Dumova and Kruglov 2009) and cow manure (Qiao et al. 2019), however, these works focused on mining cellulolytic strains and cellulases, giving few information about the dynamic cooperation among stains during cellulose hydrolysis process.

To further understand the synergistic cellulose catabolism in microbiota FPDM, the relative abundances of CAZyme coding genes in different cellulose-degrading stages were compared and a heat map of the important CAZyme and CBM families was plotted (Fig. 5). Totally 7 relative abundance-shifting patterns among CAZyme coding genes could be observed. Similarly, these annotated protein families could be classified into progressive type (pattern 11 and 15) and declining type (pattern 0, 3 and 4). A large proportion of the declining shift families were cellulolytic enzymes, such as endo-1,4-beta-glucanase (enzyme commission (EC) 3.2.1.4, glycoside hydrolase family 5 (GH5), GH8–10, 44, 48, 74), and reducing end-acting 1,4-beta-cellobiosidase (EC3.2.1.176, GH48), etc., and some CBMs like CBM2, 3, and 16 etc. (Brulc et al. 2009; Cui et al. 2019; Dhaeseleer et al. 2013; Lombard et al. 2014; Zhang et al. 2018). Of the progressive shift families, several CAZyme families were observed, such as 6-phospho-beta-glucoosidase (EC3.2.1.86, GH1 and 4), beta-glucosidase (EC3.2.1.21, GH1, 3, 5, 9, 30, 39, and 116) and cellobiose dehydrogenase (EC1.1.99.18, auxiliary activity family 3 (AA3)) whose active mainly focus on the catabolism of cellulose products such as cello-oligosaccharide and cellobiose (Brulc et al. 2009; Hallberg et al. 2002; He et al. 2019; López-Mondejar et al. 2016a). Moreover, CBM5, 17, 28 and lytic polysaccharide monooxygenases (LPMOs, EC1.14.99.54 / EC1.14.99.56, AA10) which directly involved in the catabolism of cellulose were also detected in these progressive shift families (Cao et al. 2012; Medie et al. 2012; Taillefer et al. 2018; Talamantes et al. 2016).

At present, most researches explored cellulolytic enzymes mainly in annotated gene database produced by omic analysis (Wilhelm et al. 2018; Moraes et al. 2018; Cui et al. 2019; Tomazetto et al. 2020). In this work, 4,031 genes with high relative abundance (\(\geq 0.01\%\)) and clear species information were still functionally unannotated. To further explore the FUGs potentially related to cellulose catabolism, cluster analysis was performed using the STEM algorithm based on different abundance variation patterns of genes in different cellulose degradation stages. As shown in Table 2, four typical patterns (0, 4, 11 and 14), in which more than 10 cellulolytic genes were annotated, were selected for clustering. Annotated genes clustered to the declining pattern 0 and 4 usually encode endo-type cellulases (GH5, 8–10, 44, 48 and 74) and other non-hydrolytic CAZyme modules (CBM4 and 6) derived from cellulolytic bacteria, such as Sporocytophaga. Therefore, it could be speculated that 363 FUGs from pattern 0 and 4 might be related to early-intermediate stage of cellulose
degradation and could be categorized as putative cellulolytic gene 1 (PCG1). Most of the genes in progressive pattern 11 and 14 were mainly from Cohnella and Paenibacillus. Nearly all the annotated genes encoded beta-glucosidase (EC3.2.1.21, GH1, 3, 5, 9, 30, 39, and 116) and exo-beta-1,4-glucosidase (EC3.2.1.74, GH1, 3, 5, and 9), suggesting that the clustered 231 FUGs might code for enzymes that play important roles in intermediate-final stage of cellulose degradation. This type of FUGs was categorized as putative cellulolytic gene 2 (PCG2). These results showed high accordance with the synergism hypothesis that the hydrogen bonds of natural cellulose are destroyed by non-hydrolytic protein factors, forming disordered cellulose. The disordered cellulose is then hydrolyzed into dextrin and glucose under the action of endoglucanase, exoglucanase, beta-glucosidase and exo-beta-1,4-glucosidase (Juturu and Wu 2014; López-Mondejar et al. 2016b; Medie et al. 2012).

![Fig. 5 Heat map of the relative abundance of cellulose-degrading related CAZyme and CBM families from microbiota FPDM. The tree on the left is clustered based on the change patterns of the relative abundances of CAZyme families. Each panel shows a z-value calculated by normalizing the relative abundance and translating it into color intensity (log scale), with reference to the color bar (top right). Each bracketed number followed CAZyme family refers to the change pattern profiles of the relative abundance (middle right). “*” refers to the relative abundance of CAZyme family more than 0.1%, and “**” refers to the relative abundance more than 1%. The different shapes after bracketed numbers refer to the cellulose-degrading related enzymes and the EC numbers (bottom right). Specifically, FPDM.3, FPDM.8 and FPDM.15 refer to microbiota FPDM cultivated in fiber medium for 3, 8, and 15 days respectively, and each time points sampled in biological triplicate with the postfix name of “1”, “2”, and “3”](image-url)
### Table 2  Top abundant genes belong to 4 different abundance disturbing patterns across treatment FPDM.3, FPDM.8, and FPDM.15

| Abundance disturbing pattern | Gene ID | EC Number | CAZy Family | Annotation or putative information |
|-----------------------------|---------|-----------|-------------|-----------------------------------|
| **Top 10 abundant genes belong to abundance disturbing pattern 0**                   |         |           |             |                                    |
| FPDM.8.2_21397              | NA      | NA        |             | Related to the initial and middle phase of cellulose degradation   |
| NOVO_MIX_965                | NA      | NA        |             | Related to the initial and middle phase of cellulose degradation   |
| FPDM.15.1_27227             | NA      | NA        |             | Related to the initial and middle phase of cellulose degradation   |
| FPDM.15.1_36486             | NA      | NA        |             | Related to the initial and middle phase of cellulose degradation   |
| FPDM.15.1_4658              | NA      | NA        |             | Related to the initial and middle phase of cellulose degradation   |
| FPDM.15.3_14026             | NA      | GH3       |             | Beta-glucosidase (EC3.2.1.21); exo-beta-1,4-glucosidase (EC3.2.1.74) |
|                             |         |           | CBM4/ GH10 | Binding of these modules has been demonstrated with Modules of approx. 150 residues found in bacterial enzymes. Binding of these modules has been demonstrated with xylan, beta-1,3-glucan, beta-1,3-1,4-glucan, beta-1,6-glucan and amorphous cellulose but not with crystalline cellulose/ Endo-1,4-beta-glucanase (EC3.2.1.4) |
|                             |         | CBM6      |             | The cellulose-binding function has been demonstrated in one case on amorphous cellulose and beta-1,4-xylan Some of these modules also bind beta-1,3-glucan, beta-1,3-1,4-glucan, and beta-1,4-glucan |
|                             | 3.2.1.8/ | CBM6      |             | The cellulose-binding function has been demonstrated in one case on amorphous cellulose and beta-1,4-xylan Some of these modules also bind beta-1,3-glucan, beta-1,3-1,4-glucan, and beta-1,4-glucan |
|                             | 3.2.1.32 |           |             | Endo-1,4-beta-glucanase (EC3.2.1.4); beta-glucosidase (EC3.2.1.21); exo-beta-1,4-glucosidase (EC3.2.1.74); cellulose 1,4-beta-cellobiosidase (non-reducing end) (EC3.2.1.91) |

| **Top 10 abundant genes belong to abundance disturbing pattern 4** |         |           |             |                                    |
| FPDM.15.1_17147             | NA      | NA        |             | Related to the initial and middle phase of cellulose degradation   |
| FPDM.8.1_9164                | NA      | NA        |             | Related to the initial and middle phase of cellulose degradation   |
| FPDM.15.3_5584               | NA      | NA        |             | Related to the initial and middle phase of cellulose degradation   |
| FPDM.15.2_11349              | NA      | NA        |             | Related to the initial and middle phase of cellulose degradation   |
| FPDM.8.3_12846              | NA      | CBM6      |             | Modules of ~120 residues. The cellulose-binding function has been demonstrated in one case on |
| FPDM.15.2_4723              | NA      | CBM6      |             |                                    |

**References**
- Nedashkovskaya and Kim (2015)
- Lladó et al. (2017)
- Taillefer et al. (2018)
Table 2 continued

| Entry                  | GH  | EC Description                                                                 | Organism          | Reference                |
|------------------------|-----|---------------------------------------------------------------------------------|-------------------|--------------------------|
| FPDM.15.1_16917        | 3.2.1.4 | GH9 amorphous cellulose and beta-1,4-xylan. Some of these modules also bind beta-1,3-glucan, beta-1,3-1,4-glucan, and beta-1,4-glucan | *Sporocytophaga* | Taillefer et al. (2018) |
| FPDM.15.1_32432        | 3.2.1.21 | GH3 Endo-1,4-beta-glucanase (EC3.2.1.21); exo-beta-1,4-glucosidase (EC3.2.1.74); cellulose 1,4-beta-cellobiosidase (non-reducing end) (EC3.2.1.91) | *Sporocytophaga* | Taillefer et al. (2018) |
| FPDM.15.1_29223        | 3.2.1.4 | GH5 Endo-1,4-beta-glucanase (EC3.2.1.21); exo-beta-1,4-glucosidase (EC3.2.1.74); cellulose 1,4-beta-cellobiosidase (non-reducing end) (EC3.2.1.91) | *Sporocytophaga* | Taillefer et al. (2018) |
| FPDM.15.1_20282        | 3.2.1.4 | GH8 Endo-1,4-beta-glucanase (EC3.2.1.4) | *Sporocytophaga* | Taillefer et al. (2018) |

Top 10 abundant genes belong to abundance disturbing pattern 11.

- **FPDM.3.1_13197** NA NA Related to the middle and last phase of cellulose degradation
- **FPDM.15.1_26209** NA NA Related to the middle and last phase of cellulose degradation
- **FPDM.3.3_14813** NA NA Related to the middle and last phase of cellulose degradation
- **FPDM.8.3_3197** NA NA Related to the middle and last phase of cellulose degradation
- **FPDM.3.3_6193** NA NA Related to the middle and last phase of cellulose degradation
- **FPDM.3.1_15556** 3.2.1.21 GH116 Beta-glucosidase (EC3.2.1.21); exo-beta-1,4-glucosidase (EC3.2.1.74) | *Cohnella* | Eida et al. (2012) |
- **FPDM.3.3_12658** 3.2.1.21 GH3 Beta-glucosidase (EC3.2.1.21); exo-beta-1,4-glucosidase (EC3.2.1.74) | *Cohnella* | Eida et al. (2012) |
- **FPDM.3.3_22593** 3.2.1.21 GH3 Beta-glucosidase (EC3.2.1.21); exo-beta-1,4-glucosidase (EC3.2.1.74) | *Cohnella* | Eida et al. (2012) |
- **FPDM.3.2_23567** 3.2.1.86 GH4 6-phospho-beta-glucosidase (EC3.2.1.86) | *Cohnella* | Eida et al. (2012) |
- **FPDM.3.1_15556** 3.2.1.21 GH116 Beta-glucosidase (EC3.2.1.21) | *Cohnella* | Eida et al. (2012) |
- **FPDM.3.3_21908** 3.2.1.86 GH4 6-phospho-beta-glucosidase (EC3.2.1.86) | *Cohnella* | Eida et al. (2012) |

Top 10 abundant genes belong to abundance disturbing pattern 14.

- **FPDM.3.3_12658** 3.2.1.21 GH3 Beta-glucosidase (EC3.2.1.21); exo-beta-1,4-glucosidase (EC3.2.1.74) | *Cohnella* | Eida et al. (2012) |
- **FPDM.3.3_6193** NA NA Related to the middle and last phase of cellulose degradation
- **FPDM.3.3_15298** NA NA Related to the middle and last phase of cellulose degradation
- **FPDM.3.1_13589** NA NA Related to the middle and last phase of cellulose degradation
- **FPDM.15.1_24753** NA NA Related to the middle and last phase of cellulose degradation
- **FPDM.3.2_10241** NA NA Related to the middle and last phase of cellulose degradation
- **FPDM.3.1_11213** 3.2.1.21 GH3 Beta-glucosidase (EC3.2.1.21); exo-beta-1,4-glucosidase | *Cohnella* | Eida et al. (2012) |
Dynamic shifts in taxa and CAZymes make the cellulose catabolism pathway more distinct

Currently, extensive efforts have been made for clarifying the cellulose catabolism pathway in cellulytic microbial consortia (Raman et al. 2011; Fosses et al. 2017). Zhou et al. (2014) predicted how a structured and cooperative microbial community functions in biomass conversion process by analyses of reconstructed bacterial draft genomes from a soil-derived cellulytic-degrading microbial community. Based on microbial draft genomes, Cui et al. (2019) predicted a mutualistic interaction among the dominant microbes regarding the cellulolytic process in the niche. However, all these works built the cellulolytic pathway at a certain culture stage of the microbiota. Few works focused on the dynamically synergistic action at both the gene and genera levels throughout the cellulose degradation process.

Using temporal metagenomic analysis, the microbial composition that thrives in hypersaline environments in the tropics has been investigated (Couto-Rodriguez and Montalvo-Rodriguez 2019), which sparked our interest in constructing a more clear cellulose catabolism pathway based on the dynamic analysis of the filter paper hydrolysates, the community structure and composition, and the clustered functional genes of microbiota FPDM. As shown in Fig. 6, two dominant genera, Sporocytophaga and Cohnella dynamically participated in the synergistic degradation process and had different functional tendencies in different stages. In the early-intermediate stage, cellulose was first hydrolyzed into short-chain cellulose, celloextrin and celllobiose by endo-1,4-beta-glucanase (EC3.2.1.4, GH5, 8–10, 44, 48, and 74), cellulose 1,4-beta-cellobiosidase (nonreducing end) (EC3.2.1.91, GH5 and 9), cellulose 1,4-beta-cellobiosidase (reducing end) (EC3.2.1.176, GH48) and PCG1 secreted from progressive shift microorganisms (e.g. Sporocytophaga). In the intermediate-final stages, beta-glucosidase (EC3.2.1.21, GH1, 3, 5, 9, 30, 39, and 116), exo-beta-1,4-glucosidase

Table 2 continued

| FPDM.3.3_11198 | NA | CBM6 | (EC3.2.1.74) Modules of ~120 residues. The cellulose-binding function has been demonstrated in one case on amorphous cellulose and beta-1,4-xylan. Some of these modules also bind beta-1,3-glucan, beta-1,3-1,4-glucan, and beta-1,4-glucan. | Cohnella | Eida et al. (2012) |
| FPDM.3.3_16369 | NA | CBM6 | (EC3.2.1.74) Modules of ~120 residues. The cellulose-binding function has been demonstrated in one case on amorphous cellulose and beta-1,4-xylan. Some of these modules also bind beta-1,3-glucan, beta-1,3-1,4-glucan, and beta-1,4-glucan. | Cohnella | Eida et al. (2012) |
| FPDM.15.1_12031 | 3.2.1.21 | GH3 | Beta-glucosidase (EC3.2.1.21); exo-beta-1,4-glucosidase (EC3.2.1.74) | Cohnella | Eida et al. (2012) |
| FPDM.3.3_18618 | 3.2.1.21 | GH3 | Beta-glucosidase (EC3.2.1.21); exo-beta-1,4-glucosidase (EC3.2.1.74) | Cohnella | Eida et al. (2012) |

*Was the variation pattern of gene abundance clustered by STEM algorithm, the significant patterns of 0, 4, 11 and 14 with more than 10 annotated cellulolytic genes were displayed, and the top 5 of relative abundance of putative cellulolytic genes and candidate cellulytic genes were listed, respectively

*Represents the species that carries the cellulose-degrading related genes detected in microbiota FPDM, including cellulases and exoglucanases

*Represents the species that carries the celllobiose-degrading related genes detected in microbiota FPDM, including beta-glucosidases and phosphocellobiase

Putative cellulolytic genes were highlighted in gray
(EC3.2.1.74, GH1, 3, 5, and 9) and PCGs secreted from both progressive and declining shift microorganisms continued to hydrolyze short-chain cellulose and cellodextrin (DP > 15) into shorter cellobio-oligosaccharides and cellobiose. Then, cellobio-oligosaccharides and cellobiose were completely hydrolyzed to glucose and glucose 6-phosphate by beta-glucosidase (EC3.2.1.21, GH1, 3, 5, 9, 30, 39, and 116), exo-beta-1,4-glucosidase (EC3.2.1.74, GH1, 3, 5, and 9), phosphocellulioase (EC3.2.1.86, GH1 and 4) and PCG2 produced from declining shift microorganisms (e.g. Cohnella).

To further identify the biological function of two dominant genera, Sporocytophaga sp. CX and Cohnella sp. CX were isolated from microbiota FPDM and co-cultured using filter paper as the substrate. The activities of endo-1,4-beta-glucanases (EC3.2.1.4, GH5, 8–10, 44, 48 and 74) (250 ± 32 U/L) could be detected in supernatant of Sporocytophaga sp. CX, and Cohnella sp. CX showed the activities of beta-glucosidases (EC3.2.1.21, GH1, 3, 5, 9, 30, 39, and 116, and EC3.2.1.74, GH1, 3, 5, and 9) (14.6 ± 0.3 U/L) and cellobiose dehydrogenases (EC1.1.99.18, AA3) (39.0 ± 3.3 U/L) (Fig. 6). Activities of other cellulases were not detected probably due to the low secretion level. As shown in Fig. 7, when Sporocytophaga sp. CX was cultivated for 8 days, filter paper was partly degraded. However, when this strain was co-fermented with Cohnella sp. CX, a completely hydrolysis of filter paper could be observed. All these results confirmed the synergistic work of the two strains deduced by our metagenomic analysis.
Conclusions

The cellulolytic function of microbiota FPDM was clarified by the dynamic analysis of hydrolysates and metagenomes at different cellulose-degrading stages. The biodiversity of FPDM gradually enhanced as the length diversity of polymer chain in cellulose hydrolysates increased. FPDM-mediated cellulose degradation was accomplished by the collaboration and competition of two types of taxa, represented by Sporocytophaga and Cohnella respectively. Based on the deduced cellulolytic function of 432 annotated and 594 unannotated genes, a dynamic cellulose-degrading pathway of FPDM dominated by two core genera was predicted. Our work systematically depicted the synergistic and dynamic action of different taxa and their enzymes in microbiota during cellulose-degrading process. Significantly, based on the temporal analysis of cellulose-degrading microbiota samples, functional strains and proteins belong to the same shift pattern could be clustered, which should be valuable for mining key cellulolytic strains and enzymes and rebuilding more efficient cellulose-degrading system (Fig. 8).

Fig. 7 The degradation process of filter paper by dominant strains isolated from microbiota FPDM. The strains were cultivated in liquid culture medium containing round filter paper at 30 °C. Sporocytophaga sp. CX was fermented for 3 days (a) and 8 days (c), respectively. Meanwhile, Sporocytophaga sp. CX and Cohnella sp. CX were co-fermented for 3 days (b) and 8 days (d), respectively.
Acknowledgments FY, XL, and XC were supported by grants from the National Natural Sciences Foundation of China (31671796, 31771907, and 31801469), and the Liaoning BaiQianWan Talents Program is also gratefully acknowledged.

Author contributions MY: designed the strategy for metagenomic analysis, developed the scripts and wrote the manuscript. JZ: contributed with reagents, materials, and analyses, performed the experiments, and wrote the manuscript. YY, XC: contributed with reagents, materials, and performed the experiments. FY, XL: drafted, wrote and revised the manuscript. All authors agreed with the submitted version of the paper. All authors read and approved the final manuscript.

Funding This work was supported by the National Natural Sciences Foundation of China (31671796, 31771907, and 31801469), and the Liaoning BaiQianWan Talents Program is also gratefully acknowledged.

Data availability The metagenomic reads generated in this study have been submitted to NCBI Sequence Read Archive (SRA) under the accession number SRP255666. A previous version of this paper was published as preprint in Research Square (https://doi.org/10.21203/rs.3.rs-22654/v1) (Yang et al. 2020).

Compliance with ethical standards Conflict of interest The authors declare that they have no conflict of interest.

References

Alneberg J, Bjarnason BS, de Bruijn I et al (2014) Binning metagenomic contigs by coverage and composition. Nat Methods 11:1144–1146. https://doi.org/10.1038/nmeth.3103

Berg B, Hofsten B, Pettersson G (1972) Electronmicroscopic observations on the degradation of cellulose fibres by Cellibivinio fulvus and Sporocytophaga myxococcoides. J Appl Bacteriol 35:215–219. https://doi.org/10.1111/j.1365-2672.1972.tb03692.x

Berlemont R, Allison SD, Weihe C et al (2014) Cellulytic potential under environmental changes in microbial communities from grassland litter. Front Microbiol 5:639. https://doi.org/10.3389/fmicb.2014.00639

Berlemon R, Allison SD, Weihe C et al (2014) Cellulytic potential under environmental changes in microbial communities from grassland litter. Front Microbiol 5:639. https://doi.org/10.3389/fmicb.2014.00639

Buchfink B, Xie C, Huson DH (2015) Fast and sensitive protein alignment using DIAMOND. Nat Methods 12:59–60. https://doi.org/10.1038/nmeth.3176

Cao R, Jin Y, Xu D (2012) Recognition of cello-oligosaccharides by CBM17 from Clostridium cellulovorans: molecular dynamics simulation. J Phys Chem B 116:6087–6096. https://doi.org/10.1021/jp3010647

Caporaso JG, Kuczynski J, Stombaugh J et al (2010) QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7:335–336. https://doi.org/10.1038/nmeth.f.303

Chen Y, Chen Y, Shi C et al (2018) SOAPnuker: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. GigaScience 7:1–6. https://doi.org/10.1093/gigascience/gix120

Couto-Rodriguez RL, Montalvo-Rodriguez R (2019) Temporal analysis of the microbial community from the crystallizer ponds in Cabo Rojo, Puerto Rico, using metagenomics. Genes 10:422. https://doi.org/10.3390/genes10060422

Cui J, Mai G, Wang Z et al (2019) Metagenomic insights into a cellulose-rich niche reveal microbial cooperation in cellulose degradation. Front Microbiol 10:618. https://doi.org/10.3389/fmicb.2019.00618

Deng YJ, Wang SY (2016) Synergistic growth in bacteria depends on substrate complexity. J Microbiol 54:23–30. https://doi.org/10.1007/s12275-016-5461-9
Dhaeseleer P, Gladden JM, Allgaier M et al (2013) Proteogenomic analysis of a thermophilic bacterial consortium adapted to deconstruct switchgrass. PLoS ONE 8:e68465. https://doi.org/10.1371/journal.pone.0068465

Dos Santos PC, de Oliveira MAS, Seixas FAV et al (2020) A novel cellbiohydrolase I (CBBH) from Penicillium digitatum: production, purification, and characterization. Appl Biochem Biotechnol 192:257–282. https://doi.org/10.1007/s12010-020-03307-9

Dunova VA, Kruglov YY (2009) A cellulase-decomposing bacterial association. Microbiology 78:234–239. https://doi.org/10.1134/S0026261709020155

Eida MF, Nagaoka T, Wasaki J et al (2012) Isolation and characterization of the cellulase-secretome produced by the Antarctic bacterium Flavobacterium sp. AUG42. Microbes Environ 27:226–233. https://doi.org/10.1266/jsme2.ME11299

Ernst J, Bar-Joseph Z (2006) STEM: a tool for the analysis of short time series gene expression data. BMC Bioinformatics 7:191. https://doi.org/10.1186/1471-2105-7-191

Fosses A, Mateó M, Franche N et al (2017) A seven-genome cluster in Rumuniclostridium cellulolyticum is essential for signalization, uptake and catabolism of the degradation products of cellulose hydrolysis. Biotechnol Biofuels 10:250. https://doi.org/10.1186/s13068-017-0933-7

Hallberg BM, Henriksson G, Pettersson G et al (2002) Crystal structure of the flavoprotein domain of the extracellular flavocytochrome celllobiose dehydrogenase. J Mol Biol 315:421–434. https://doi.org/10.1006/jmbi.2001.5246

Hameed A, Hung M, Lin S et al (2013) Cohnella formosensis sp. nov., a xylanolytic bacterium isolated from the rhizosphere of Medicago sativa L. Int J Syst Evol Microbiol 63:2806–2812. https://doi.org/10.1099/ijs.0.045831-0

Hayat R, Sheiridl RA, Ifikhar-ul-Hassan M et al (2012) Characterization and identification of compost bacteria based on 16S rRNA gene sequencing. Ann Microbiol 63:905–912. https://doi.org/10.1007/s13213-012-0542-4

He B, Jin S, Cao J et al (2019) Metatranscriptomics of the Hu sheep rumen microbiome reveals novel cellulolytic bacteria. Biotechnol Biofuels 12:153. https://doi.org/10.1186/s13068-020-01779-9

Herrera LM, Brana V, Franco Fraguas L et al (2019) Characterization of the cellulase-secretome produced by the Antarctic bacterium Flavobacterium sp. AUG42. Microbiol Res 223–225:13–21. https://doi.org/10.1016/j.micres.2019.03.009

Huson DH, Mitra S, Ruscheweyh HJ et al (2011) Integrative analysis of environmental sequences using MEGAN4. Genome Res 21:1552–1560. https://doi.org/10.1101/gr.120618.111

Jayasekara S, Ratnayake R (2019) Microbial cellulases: an overview and applications. In: Pascual AR, Martin MEE (eds) Cellulose. IntechOpen, London, pp 1–18. https://doi.org/10.5772/intechopen.84531

Jiménez DJ, Korenblum E, van Elsas JD (2014) Novel multispecies microbial consortia involved in lignocellulose and 5-hydroxymethylfurfural bioconversion. Appl Microbiol Biotechnol 98:2789–2803. https://doi.org/10.1007/s00253-013-5253-7

Juturu V, Wu JC (2014) Microbial cellulosomes: engineering, production and applications. Renew Sust Energ Rev 33:188–203. https://doi.org/10.1016/j.rser.2014.01.077

Kämpfer P, Rossello-Mora R, Falsen E et al (2006) Cohnella thermotolerans gen. nov., sp. nov., and classification of ‘Paenibacillus hongkongensis’ as Cohnella hongkongensis sp. nov. Int J Syst Evol Microbiol 56:781–786. https://doi.org/10.1099/ijs.0.63985-0

Kanehisa M, Araki M, Goto S et al (2008) KEGG for linking genomes to life and the environment. Nucl Acids Res 36:D480–D484. https://doi.org/10.1093/nar/gkm882

Kang DY, Li F, Kirton E et al (2019) MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. PeerJ 7:e7359. https://doi.org/10.7717/peerj.7359

Khianngam S, Tanasupawat S, Akaracharanya A et al (2012) Cohnella cellulosilytica sp. nov., isolated from buffalo faeces. Int J Syst Evol Microbiol 62:1921–1925. https://doi.org/10.1099/ijs.0.032607-0

Knight R, Vrbanac A, Taylor BC et al (2018) Best practices for analysing microbiomes. Nat Rev Microbiol 16:410–422. https://doi.org/10.1038/s41579-018-0029-9

Kolaczkowski BM, Schaller KS, Sorensen TH et al (2020) Removal of N-linked glycans in cellulobiohydrolase Cel7A from Trichoderma reesei reveals higher activity and binding affinity on crystalline cellulose. Biotechnol Biofuels 13:1–13. https://doi.org/10.1186/s13068-020-01779-9

Lee CK, Jang MY, Park HR et al (2016) Cloning and characterization of xylanase in cellulolytic Bacillus sp. strain JMY1 isolated from forest soil. Appl Biol Chem 59:415–423. https://doi.org/10.1097/doi.org/10.1093/bioinformatics/btp336

Li W, Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22:1658–1659. https://doi.org/10.1093/bioinformatics/bt1158

Li R, Yu C, Li Y et al (2009) SOAP2: an improved ultrafast tool for short read alignment. Bioinformatics 25:1966–1967. https://doi.org/10.1093/bioinformatics/btp336

Li J, Jia H, Cai X et al (2014) An integrated catalog of reference genes in the human gut microbiome. Nat Biotechnol 32:834–841. https://doi.org/10.1038/nbt.2942

Liu L, Gao P, Chen G et al (2014) Draft genome sequence of Cohnella cellulosilytica sp. nov., isolated from forest sawdust and coffee residue composts. Microbes Environ 29:1–11. https://doi.org/10.1266/jsme2.ME11299

Li W, Sasaki T, Masunaga S et al (2013) Cellulose metabolism, uptake and catabolism of the degradation products of cellulose hydrolysis. Biotechnol Biofuels 13:1–13. https://doi.org/10.1186/s13068-020-01779-9

Llorente-Forner L, Juturu V, Wu JC (2014) Microbial cellulosomes: engineering, production and applications. Renew Sust Energ Rev 33:188–203. https://doi.org/10.1016/j.rser.2014.01.077

López-Mondejar R, Zühlke D, Becher D et al (2016a) Cellulose and hemicellulose decomposition by forest soil bacteria proceeds by the action of structurally variable enzymatic systems. Sci Rep 6:25279. https://doi.org/10.1038/srep25279
Wilhelm RC, Singh R, Eltis LD et al (2018) Bacterial contributions to delignification and lignocellulose degradation in forest soils with metagenomic and quantitative stable isotope probing. ISME J 13:413–429. https://doi.org/10.1038/s41396-018-0279-6

Wilson DB (2011) Microbial diversity of cellulose hydrolysis. Curr Opin Microbiol 14:259–263. https://doi.org/10.1016/j.mib.2011.04.004

Wongwilaiwalin S, Rattanachomsri U, Laothanachareon T et al (2010) Analysis of a thermophilic lignocellulose degrading microbial consortium and multi-species lignocellulolytic enzyme system. Enzyme Microb Technol 47:283–290. https://doi.org/10.1016/j.enzmictec.2010.07.013

Wu YW, Simmons BA, Singer SW (2016) MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. Bioinformatics 32:605–607. https://doi.org/10.1093/bioinformatics/btv638

Yang M, Zhao J, Yuan Y et al (2020) Comparative metagenomic discovery of the dynamic cellulose-degrading process from a synergistic cellulolytic microbiota. Res Sq. https://doi.org/10.21203/rs.3.rs-22654/v1

Yin Y, Mao X, Yang J et al (2012) dbCAN: a web resource for automated carbohydrate-active enzyme annotation. Nucl Acids Res 40:W445–W451. https://doi.org/10.1093/nar/gks479

Younesi FS, Pazhang M, Najavand S et al (2016) Deleting the Ig-like domain of Alicyclobacillus acidocaldarius endoglucanase Cel9A causes a simultaneous increase in the activity and stability. Mol Biotechnol 58:12–21. https://doi.org/10.1007/s12033-015-9900-3

Zhang KD, Li W, Wang YF et al (2018) Processive degradation of crystalline cellulose by a multimodular endoglucanase via a wirewalking mode. Biomacromol 19:1686–1696. https://doi.org/10.1021/acs.biomac.8b00340

Zhou Y, Pope PB, Li S et al (2014) Omics-based interpretation of synergism in a soil-derived cellulose-degrading microbial community. Sci Rep 4:5288. https://doi.org/10.1038/srep05288

Zhu W, Lomsadze A, Borodovsky M (2010) Ab initio gene identification in metagenomic sequences. Nucl Acids Res 38:e132. https://doi.org/10.1093/nar/gkq275

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.