



A genomic perspective on the potential of termite-associated *Cellulosimicrobium cellulans* MP1 as producer of plant biomass-acting enzymes and exopolysaccharides

Nguyen Thi-Hanh Vu^{1,2,*}, Tung Ngoc Quach^{1,2,*}, Xuan Thi-Thanh Dao^{3,4}, Ha Thanh Le³, Chi Phuong Le¹, Lam Tung Nguyen³, Lam Tung Le¹, Cuong Cao Ngo⁵, Ha Hoang¹, Ha Hoang Chu^{1,2} and Quyet-Tien Phi^{1,2}

¹Institute of Biotechnology, Vietnam Academy of Science and Technology, Hanoi, Vietnam, Hanoi, Vietnam

²Graduate University of Science and Technology, Vietnam Academy of Science and Technology, Hanoi, Vietnam

³School of Biotechnology and Food Technology, Hanoi University of Science and Technology, Hanoi, Vietnam

⁴Vinh University, Vinh, Vietnam

⁵Vietnam–Russia Tropical Center, Hanoi, Vietnam

*These authors contributed equally to this work.

ABSTRACT

Background. Lignocellulose is a renewable and enormous biomass resource, which can be degraded efficiently by a range of cocktails of carbohydrate-active enzymes secreted by termite gut symbiotic bacteria. There is an urgent need to find enzymes with novel characteristics for improving the conversion processes in the production of lignocellulosic-based products. Although various studies dedicated to the genus *Cellulosimicrobium* as gut symbiont, genetic potential related to plant biomass-acting enzymes and exopolysaccharides production has been fully untapped to date.

Methods. The cellulolytic bacterial strain MP1 was isolated from termite guts and identified to the species level by phenotypic, phylogenetic, and genomic analysis. To further explore genes related to cellulose and hemicellulose degradation, the draft genome of strain MP1 was obtained by using whole-genome sequencing, assembly, and annotation through the Illumina platform. Lignocellulose degrading enzymes and levan production in the liquid medium were also examined to shed light on bacterial activities.

Results. Among 65 isolates obtained, the strain MP1 was the most efficient cellulase producer with cellulase activity of 0.65 ± 0.02 IU/ml. The whole genome analysis depicted that strain MP1 consists of a circular chromosome that contained 4,580,223 bp with an average GC content of 73.9%. The genome comprises 23 contigs including 67 rRNA genes, three tRNA genes, a single tmRNA gene, and 4,046 protein-coding sequences. In support of the phenotypic identification, the 16S rRNA gene sequence, average nucleotide identity, and whole-genome-based taxonomic analysis demonstrated that the strain MP1 belongs to the species *Cellulosimicrobium cellulans*. A total of 30 genes related to the degradation of cellulases and hemicellulases were identified in the *C. cellulans* MP1 genome. Of note, the presence of *sacC1-levB-sacC2-ls* operon responsible

Submitted 11 March 2021

Accepted 1 July 2021

Published 28 July 2021

Corresponding authors

Ha Thanh Le, ha.lethanh@hust.edu.vn

Quyet-Tien Phi, tienpq@ibt.ac.vn

Academic editor

Craig Moyer

Additional Information and
Declarations can be found on
page 17

DOI 10.7717/peerj.11839

© Copyright
2021 Vu et al.

Distributed under
Creative Commons CC-BY 4.0

OPEN ACCESS

for levan and levan-type fructooligosaccharides biosynthesis was detected in strain MP1 genome, but not with closely related *C. cellulans* strains, proving this strain to be a potential candidate for further studies. Endoglucanases, exoglucanases, and xylanase were achieved by using cheaply available agro-residues such as rice bran and sugar cane bagasse. The maximum levan production by *C. cellulans* MP1 was 14.8 ± 1.2 g/l after 20 h of cultivation in media containing 200 g/l sucrose. To the best of our knowledge, the present study is the first genome-based analysis of a *Cellulosimicrobium* species which focuses on lignocellulosic enzymes and levan biosynthesis, illustrating that the *C. cellulans* MP1 has a great potential to be an efficient platform for basic research and industrial exploitation.

Subjects Biochemistry, Biotechnology, Genomics, Microbiology, Molecular Biology

Keywords Carbohydrate-active enzymes, *Cellulosimicrobium cellulans*, Levan, Lignocellulose, Termite guts, Whole-genome sequencing

INTRODUCTION

Termites are social insects contributing to nutrient recycling in terrestrial ecosystems and many vertebrate food chains (Wong *et al.*, 2014). Termites utilize microbes in the hindgut to hydrolyze efficiently cellulose of wood and lignocellulosic materials, into more easily digested sugars and short-chain fatty acids (Ohkuma, 2003; Pasari *et al.*, 2019). As demonstrated previously, termites were not able to digest the lignocellulosic biomass, a mixture of cellulose, hemicellulose, lignin, and pectin completely in the absence of symbiotic microorganisms such as bacteria, archaea (Tokuda & Watanabe, 2007; Wong *et al.*, 2014). Thus, termite gut served as a promising source of identifying novel cellulolytic enzymes as well as an excellent model for further investigating the symbiotic relationships between bacteria and their host.

Recently, bacterial symbionts are acquiring much attention from researchers over the world as potential sources for screening novel and highly efficient lignocellulose-degrading enzymes. They can be applied in many industries like biofuel, food, pulp and paper, and agriculture (Chutani & Sharma, 2016; Pandey, Edgard & Negi, 2016). Species belonging to genus *Bacillus*, *Brevibacillus*, *Cellulomonas*, *Streptomyces*, and *Paenibacillus* were well-studied as excellent cellulases, hemicellulases, xylanase and pectinases producers (Kamsani *et al.*, 2016; Ohkuma, 2003; Pasari *et al.*, 2019; Ventorino *et al.*, 2016). Cellulose hydrolysis is attributed to the synergistic activity of three different groups such as endoglucanase, exoglucanase, and β -glucosidase (Jäger & Büchs, 2012; Wi *et al.*, 2015). Typical hemicellulases are arabinoxylanases, mannanase, and xylanases involved in hemicellulose decomposition (Dashtban, Schraft & Qin, 2009). The whole-genome sequencing approach was used to promote rapid advances in the discovery of potent cellulase and hemicellulose enzymes. In regard to 20 identified CAZymes including 12 endoglucanases, two exoglucanases, and 6 β -glucosidases, halophilic bacterium *Parvularcula flava* NH6-79^T serves as cellulolytic enzymes producer (Abdul Karim *et al.*, 2020). *Micromonospora* sp. CP22 was reported to depolymerize the lignocellulosic biomass

based on 63 cellulolytic and hemi-cellulolytic CAZymes found in the genome (*Chen et al., 2020*). The genome *B. velezensis* LC1 was shown to contain 31 genes involved in lignocellulose degradation, and some of these genes were highly induced in presence of bamboo shoot powder (*Li et al., 2020*). This finding is especially significant given that few genomes of termite gut symbiotic bacteria are available.

Levan has received increasing scientific attention due to its application in the pharmaceutical, industrial, and food fields (*Belghith et al., 2012; Dahech et al., 2011; Yoo et al., 2004*). Indeed, levan is mainly composed of β -2,6 polyfructan with extensive branching through β -(2,1) linkages, that are mostly synthesized by bacterial enzymes (*Gojgic-Cvijovic et al., 2019; Mardo et al., 2017*). Japan, the US, and South Korea allowed food and pharmaceutical companies to manufacture levan as a functional food additive (*Verspreet et al., 2015*), while it is not commercially permitted in Europe (*Mardo et al., 2017*). Bacterial levan is synthesized in sucrose-rich environments through the action of levansucrase (EC 2.4.1.10) for energy reserve and biofilm formation (*Shih et al., 2005*). Under starvation conditions, accumulated levan was found to be converted into levan-type fructooligosaccharides (L-FOs) that are imported across the outer membrane (*Gray et al., 2021*). Until now, many bacteria were reported to produce levan such as *Bacillus*, *Erwinia*, *Pseudomonas*, *Microbacterium*, and *Zymomonas* (*Gojgic-Cvijovic et al., 2019*). The whole genome sequencing shed light on the full potential of levansucrase and levanase in levan-producing bacteria. As revealed in *B. subtilis*, *sacB* gene encoding for levansucrase catalyzes the synthesis of levan, which is then degraded mainly into levanbiose by the action of levanase such as YveA and YveB. In addition, the *sacB-yveB-yveA* levansucrase tricistronic operon is conserved across 12 complete genome sequences of *B. subtilis* (*Dogsa et al., 2013*). Interestingly, *sacB* gene was also found to be conserved in halophilic bacterium *Halomonas smyrnensis* AAD6R (*Diken et al., 2015*).

In previous studies, complete genome sequencing showed that genus *Cellulosimicrobium* is a rich source of glycosidases involved in plant-growth promoting and ginseng biotransformation abilities (*Eida et al., 2020; Zheng et al., 2017*). To the best of our knowledge, genomic analysis of *Cellulosimicrobium* has not been revealed to prove a better understanding of its genetic basis for other applications. In this study, we report for the first time, the identification and detailed genomic analysis of cellulose-degrading and levan-producing *Cellulosimicrobium cellulans* MP1 isolated from the termite gut. These findings provided a scientific basis for the further employment of strain MP1 and its potential genes in biotechnological processes.

MATERIALS & METHODS

Isolation of symbiotic cellulolytic bacteria

Drywood termites (*Cryptotermes domesticus*) were collected from rotten tree trunks and bagasse in Nghe An Province, Vietnam. Termites were surface-sterilized using 70% ethanol to remove contamination and then washed with sterile distilled water. The head of each termite was removed using forceps; the entire guts were removed, crushed with glass rods, and subsequently inoculated into 1 ml broth mineral medium M1 (NaNO₃ 2.5 g; NaCl

0.1 g; KH_2PO_4 2 g; MgSO_4 0.2 g; $\text{CaCl}_2 \cdot 6\text{H}_2\text{O}$ 0.1 g, pH 7.0 in a liter) containing 1% carboxymethylcellulose (CMC) or filter paper as a sole carbon source (Gupta, Samant & Sahu, 2012). These cultures were then incubated for 14 days in an incubating shaker at 30 °C. To isolate the cellulolytic bacteria, the growing cultures were spread on the M1 plate medium (K_2HPO_4 1 g; NaNO_3 2.5 g; KCl 2 g; peptone 2 g; MgSO_4 0.5 g; CMC 10 g; agar 15 g; pH 7.0 in a liter). All bacterial isolates were subsequently purified by re-streaking on the M2 agar plate. Confirmation of the cellulolytic ability of pure isolates was performed on the solid medium by covering the Petri dishes with Congo-red dye (Teather & Wood, 1982). The colonies showing yellow-colored halo zones by Congo-red staining were considered as positive cellulolytic bacteria and the clear zones were measured.

Physiological, biochemical, and 16S rRNA sequencing analysis

The shape and size of strain MP1 were determined by scanning electron microscope (SEM) JSM-5410 (JEOL, Tokyo, Japan). Gram staining was performed using the conventional methodology and confirmed using the KOH test (Powers, 1995). The effects of different conditions (ranges of pH, NaCl concentration and temperature, carbon and nitrogen sources) on the growth were investigated as previously described (Kamlage, 1996).

Genomic DNA for 16S rRNA gene sequencing was prepared by phenol-chloroform extraction. The amplification of the 16S rRNA gene sequence of strain MP1 was performed by using the universal primer pair 27F (5' -TAACACATGCAAGTCGAACG-3') and 1429R (5'-GGTGTGACGGGCGGTGTGTA-3') (Phi et al., 2010). A sequence similarity search was carried out using the BLAST program (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). The phylogenetic tree was computed by using the neighbor-joining method with 1,000 bootstrap in MEGA version 6.0 (Tamura et al., 2013). Numbers at nodes indicate percentages of 1000 bootstrap re-samplings and *Bifidobacterium bifidum* DSM 20456 (S83624) was used as the outgroup branch. The 16S rDNA gene sequence of strain MP1 was deposited onto the GenBank (NCBI) under accession number MW534740.

Genome sequencing, de novo assembly, and annotation

For library construction, DNA was extracted from a pure culture of a single colony of strain MP1 using G-spinTM Total DNA Extraction Mini Kit according to the manufacturer's instructions. The quantity and quality of extracted DNA were evaluated by electrophoresis in 0.6% (w/v) agarose gel and NanoDrop spectrophotometer 2000 Thermo Scientific in order to construct a paired-end library. The constructed genome library was then sequenced using the Illumina platform (Illumina, California, USA). Quality control and read trimming were conducted using FastQC version 0.11.5 (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>) and Trimmomatic version 0.36 (Bolger, Lohse & Usadel, 2014). The de novo genome assembly was made with SPAdes v.3.13 (Bankevich et al., 2012), which was then analyzed for its completeness using the Benchmarking Universal Single-Copy Orthologous (BUSCO) version 3 (<https://gitlab.com/ezlab/busco>).

The draft genome assembled into contigs was annotated using Prokaryotic Genomes Annotation Pipeline (PGAP; http://www.ncbi.nlm.nih.gov/genome/annotation_prok/) at

NCBI. In addition, the CRISPRCasFinder was used to identify putative CRISPR loci and Cas cluster (Grissa, Vergnaud & Pourcel, 2007). Orthologous genes and Gene ontology (GO) were analyzed using clusters of orthologous genes (COGs) (Galperin et al., 2015) and InterProScan 5 (Jones et al., 2014), respectively. Virulence factors-encoding genes were identified using the Pathosystems Resource Integration Center (PATRIC) platform (Wattam et al., 2017). The graphical map of the circular genome was also generated using PATRIC. The draft genome sequence was deposited in the GenBank (NCBI) database under accession number: JAFGYF000000000.

Analysis of whole-genome similarity

To classify strain MP1 at the species level, whole-genome similarity including average nucleotide identity (ANI) calculation and digital DNA–DNA hybridization (dDDH) was performed. The ANI was calculated using the orthologous average nucleotide identity (OrthoANI) (Lee et al., 2016). The MP1 genome sequence data was uploaded to the Type (Strain) Genome Server (TYGS) for a whole-genome-based taxonomic analysis (<https://tygs.dsmz.de>). *In silico* dDDH and the branch lengths were and the Genome BLAST Distance Phylogeny evaluated using Genome-to-Genome Distance Calculator (GGDC) (Meier-Kolthoff et al., 2013).

Comparative genomics and prediction of Carbohydrate-active enzyme

A genome-wide comparison of COGs between the assembled genome of *Cellulosimicrobium cellulans* MP1 and three other *C. cellulans* available in GenBank, including J36 (NZ_JAGJ01000000.1), LMG16121 (NZ_CAOI01000000.1), and ZKA 48 (NZ_QUMZ01000000.1) was implemented using the OrthoVenn web server with default parameters (E -value $1e-5$ and inflation value 1.5) (Wang et al., 2015). The putative genes encoding CAZymes in the comparative genomes were predicted using the dbCAN2 meta server and classified by DIAMOND, HMMER, and Hotpep via CAZy and dbCAN databases, respectively. The top hits with e -value $<1E-17$, minimum homology rate $>50\%$, and coverage $>45\%$ were considered to be homologs.

Determination of extracellular enzymatic activities

To verify the production of endoglucanase, *C. cellulans* MP1 was cultivated in the TN medium (rice bran 20 g; soya peptone 10 g, casein 10 g, NaCl 1 g, pH 7.0 in a liter) at 37 °C with vigorous shaking. At different *time intervals*, samples were taken, followed by centrifugation at 4 °C, 12,000 rpm for 10 min to remove bacterial cells and insoluble materials from the culture broth. About 0.5 ml of the crude enzyme solution was added into 0.5 ml of 0.05 M sodium phosphate, pH 7.0 buffer containing 1% of CMC. The mixture was then incubated at 30 °C for 30 min, which was stopped by adding 1 ml of 3,5'-dinitrosalicylic acid (DNS) reagent followed by boiling the reaction mixture at 100 °C for 5 min (Miller, 1959). As for xylanase production, strain MP1 was cultivated on M1 medium (K_2HPO_4 1 g; KCl 2 g; NH_4Cl 2.5 g; yeast extract 2 g, $MgSO_4$ 0.5 g, sugar cane bagasse 5 g, pH 7.0 in a liter) at 37 °C for 4 days. At different *time intervals*, culture was centrifuged at 4 °C, 12,000 rpm for 10 min to obtain the crude enzyme. About 0.5 ml of the crude enzyme was mixed with 0.5 ml of the substrate solution and incubated at

50 °C for 5 min, followed by adding 1 ml DNS reagent. Xylanase activity was assayed using 1% birchwood xylan solution in acetate buffer, pH 5.0 as the substrate (Bailey, Biely & Poutanen, 1992). To quantify pectinase activity, the M2 medium (NH₄Cl 2.5 g; K₂HPO₄ 1 g; MgSO₄ 0.5 g; KCl 2 g; yeast extract 2 g, rice bran 10 g; lactose 10 g; pH 7.0 in a liter) was used. The pectinase assay was performed at 40 °C for 30 min as described previously (Mercimek Takci & Turkmen, 2016). The quantitative enzyme assays of endoglucanase, xylanase, and pectinase were performed according to standard IUPAC procedures and expressed as international units (IU) (Ghose, 1987). One unit (IU) of enzymatic activity is defined as the amount of enzyme that releases 1 μmol reducing sugars per ml of culture supernatant per minute under assay conditions.

Levan determination

To test the ability to synthesize levan, the *C. cellulans* MP1 was cultured overnight in batch culture medium (yeast extract 7 g, (NH₄)₂SO₄ 1.5 g, K₂HPO₄ 2.5 g, pH 7.0 in a liter) at 37 °C under vigorous agitation. The overnight culture was transferred to a new batch culture medium supplemented with 100 and 200 g/l sucrose and adjusted to an optical density at 600 nm of 0.1. The culture was centrifuged at 10,000 rpm for 10 min at different time intervals. Cell-free supernatant was used to determine levan produced during fermentation (Gojgic-Cvijovic et al., 2019). Levan was harvested by adding three volumes of ice-cold ethanol. The mixture was kept at 4 °C for 12 h, centrifuged at 12,000 rpm at 4 °C for 20 min, and washed with 75% ethanol to remove the residual sugars. The obtained precipitate was hydrolyzed with 0.1 M HCl at 100 °C for 1 h. After neutralizing with 2 M NaOH, levan content was determined according to Somogyi and Nelson (Somogyi, 1945).

RESULTS

Cellulose-degrading potential and identification of the strain MP1

Of the 65 isolates able to grow on CMC plates as the sole carbon source, a total of 8 bacterial isolates produced variable zones of CMC clearance after Congo-red staining. Among those, MP1 was selected due to showing the maximum zone of clearance (9 ± 1, 1 mm). In support of this result, enzyme assays for cellulase activity on CMC and filter paper were found to be the highest for MP1 with 0.66 ± 0.15 IU/ml after 48 h and 0.33 ± 0.05 FPU/ml after 72 h of cultivation, respectively (Table S1). These results indicated that isolate MP1 is a potent cellulolytic bacterium for further study.

The phenotypic examinations indicated that MP1 grew well on LB agar after 2 days of incubation at 37 °C, producing colonies that were circular, smooth, convex, and pale yellow in colour. MP1 cells were Gram-positive, non-spore-forming rods, and nonmotile. The strain MP1 could utilize D-galactose, D-sucrose, D-raffinose, and amygdalin as sole carbon and energy (Table S2). It gave a positive test for catalase, starch hydrolysis, β-galactosidase, and nitrate reduction, whereas negative for oxidase, gelatinase, urease, indole, H₂S, acetoin. The 16S rDNA gene sequence of MP1 was aligned with the similar nucleotide sequences in the GenBank database in which a phylogenetic tree was then constructed. The neighbor-joining phylogenetic tree showed the close relationship between MP1 and related *Cellulosimicrobium* species and the highest homology to *Cellulosimicrobium cellulans* DSM

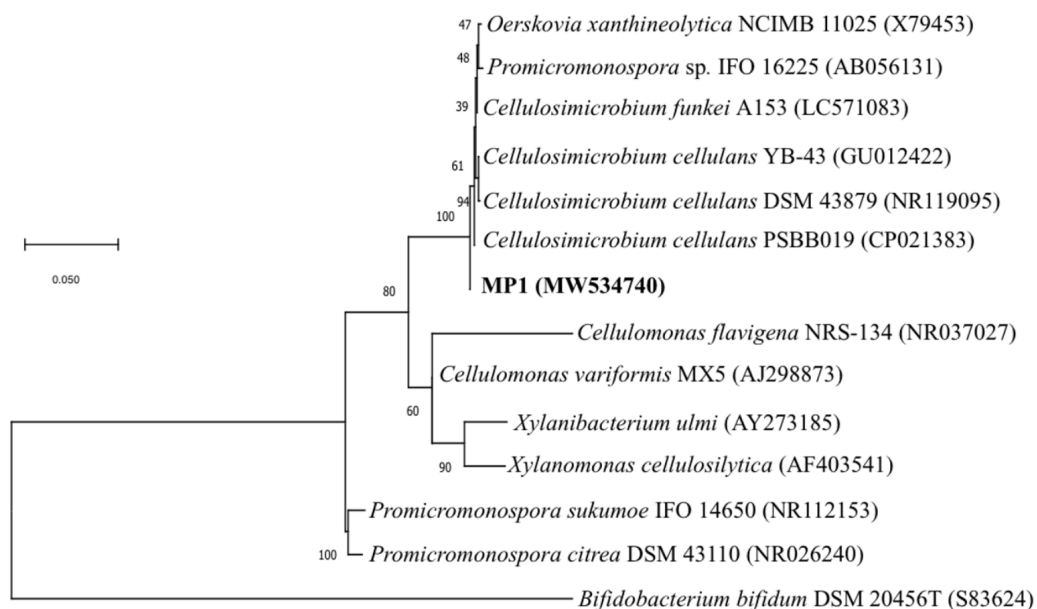


Figure 1 Identification of the strain MP1. The Maximum-likelihood phylogenetic tree based the 16S rRNA sequence of strain MP1 and representatives of reference type strains. Bar: 0.05 substitutions per site. [Full-size !\[\]\(fd7fe780e8fd8eece60268c87d0c3e04_img.jpg\) DOI: 10.7717/peerj.11839/fig-1](https://doi.org/10.7717/peerj.11839/fig-1)

43879 (99.5%) and *Cellulosimicrobium funkei* A153 (99.6%) (Fig. 1). Further, the OrthoANI software was performed to determine the OrthoANI value between isolate MP1 and five closely related *Cellulosimicrobium* species. It revealed that MP1 shared high similarity to *C. cellulans* DSM 43879 (88.71%), *C. funkei* JCM 14302 (91.35%), and *C. funkei* JCM NRBC 104118 (91.35%), and low nucleotide similarity was observed with *Promicromonospora sukumoe* SAI-064 that were out of distinct sub-clade (Fig. 2A). This result confirmed that this strain was not considered to be a novel species.

To make identification more accurate at the species level, the whole-genome-based taxonomic analysis conducted by the Type Strain Genome Server (TYGS) platform suggested that *Cellulosimicrobium* sp. MP1 was closest to *C. cellulans* DSM 43879 with digital DNA-DNA hybridization (dDDH) values and differences in guanine-cytosine (GC) content of 57.5% (formula d_6) and 0.56%, respectively (Fig. 2B). As shown previously, the differences between *C. funkei* and *C. cellulans* in phenotypic characteristics are motility and ability to utilize raffinose (Hamada et al., 2016; Yoon et al., 2007). Based on the phenotypic characteristics and genome-wide comparison, this strain was identified as *Cellulosimicrobium cellulans* MP1. This bacterium was deposited at VAST-Culture Collection of Microorganisms (VCCM) with the accession number VCCM 14150.

Genome sequence and general features of the *C. cellulans* strain MP1

Briefly, the standard short insert library yielded 443,986,169 bases resulting in 4,487,842 mapped reads with about 92.2-fold sequencing depth. The draft genome of strain MP1 (4,580,223 bp with GC content of 73.9%) was assembled into a ring chromosome, comprising of 23 contigs and no plasmid was detected (Fig. 3). The genome was

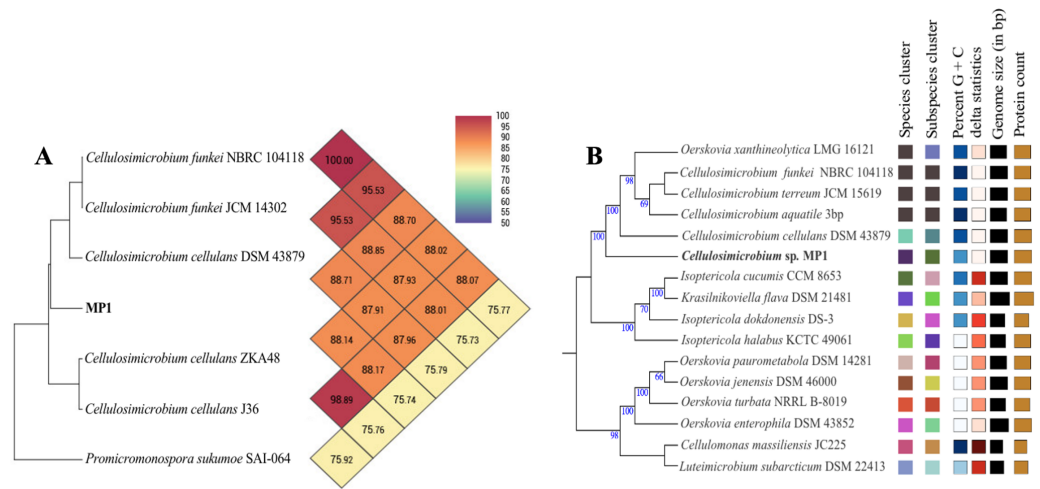


Figure 2 Phylogenomic classification of *Cellulosimicrobium* sp. MP1 based on genome analysis. (A) Heatmap of OrthoANI values for *Cellulosimicrobium* sp. MP1 and five closely related species. (B) Genome Basic Local Alignment Search Tool (BLAST) distance phylogenies (GBDP) using Type Strain Genome Server (TYGS) platform. Branch lengths are scaled in terms of GBDP distance formula d_5 . Numbers above branches are GBDP pseudo-bootstrap support values > 60% from 100 replications, with an average branch support of 84.3%. Tree was rooted at midpoint.

Full-size DOI: 10.7717/peerj.11839/fig-2

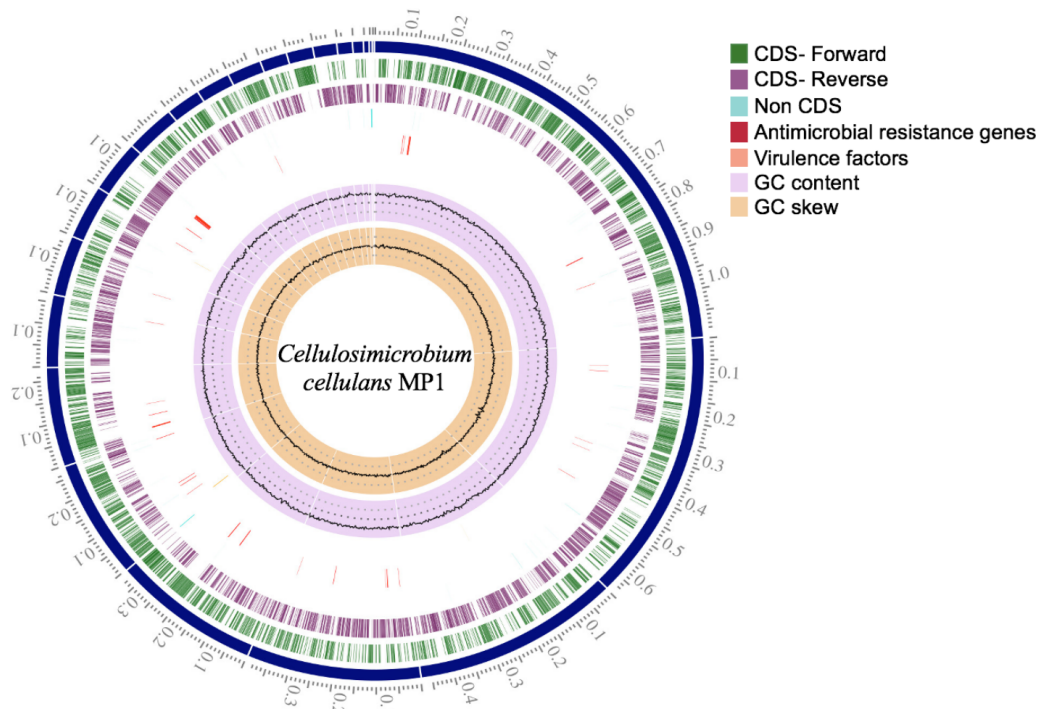


Figure 3 Circular genome map of *C. cellulans* MP1.

Full-size DOI: 10.7717/peerj.11839/fig-3

Table 1 Genomic features of *C. cellulans* MP1.

Features	Chromosome
Genome size (bp)	4,580,223
No. of contigs	23
No. of plasmids	0
G + C content (%)	73.9
Genes (total)	4,088
CDSs (coding)	3,964
rRNAs	61
tRNAs	55
ncRNAs	3
Pseudogenes	63
CRISPRS	0
GenBank accession number	JAFGYF000000000

predicted to have 4,088 genes assigned for 3,964 protein-coding sequences (CDS), 61 rRNA sequences, 55 tRNA sequences, 3 ncRNA sequences (Table 1). Moreover, a total of three virulence factors present in *C. cellulans* MP1 included dihydroxy-acid dehydratase, CarD-like transcriptional factor, and two calmodulin-like proteins (Fig. 3).

A total of 79.5% (3,216 out of 4,046) of the protein-coding sequences were assigned to 21 out of 25 COG functional categories (Fig. 4A). Transcription (K: 365 protein-coding sequences), carbohydrate transport and metabolism (G: 307), amino acid transport and metabolism (E: 224), energy production and conversion (C: 199), and inorganic ion transport and metabolism (P: 180) were found to be the largest categories. By contrast, the least represent groups included cell cycle control, cell division, chromosome partitioning (D: 33), chromatin structure and dynamics (B: 2), and cell motility (N: 1).

GO analysis was used to provide a deeper understanding of the functional catalogs of strain MP1. A total of 3,165 genes were assigned to 43 subclasses, including 13 subclasses of the cellular component (CC) class, 10 subclasses of the molecular function (MF) class, and 20 subclasses of the biological process (BP) class (Fig. 4B). In detail, the CC class occupied the most genes (1638 genes; 45.3%), followed by the BP (1424 genes; 39.4%), and MF (553 genes; 15.3%) class. The most abundant pathways were cell (GO:0005623; 508 genes), cell part (GO:0044464; 508 genes), and membrane (GO:0016020; 340 genes), which could be considered as the main functional groups of genes belonging to the CC class. Within the BP class, the three most prevalent molecular functions were metabolic process (GO:0008152; 378 genes), cellular process (GO:0009987; 379 genes), and growth (GO:0040007; 281 genes) (Fig. 4B). Among all subclasses belonging to the MF class, catalytic activity (GO:0003824) and binding (GO:0005488) contributed the most genes, with 311 and 176, respectively.

Genome-wide comparison of COGs and Carbohydrate-active enzymes

Given that microbial evolution is due to vertical descent from a single ancestral gene leading to orthologous genes in different species, it is necessary to explore gene function, gene structure, and molecular evolution by using a genome-wide comparison of COGs in

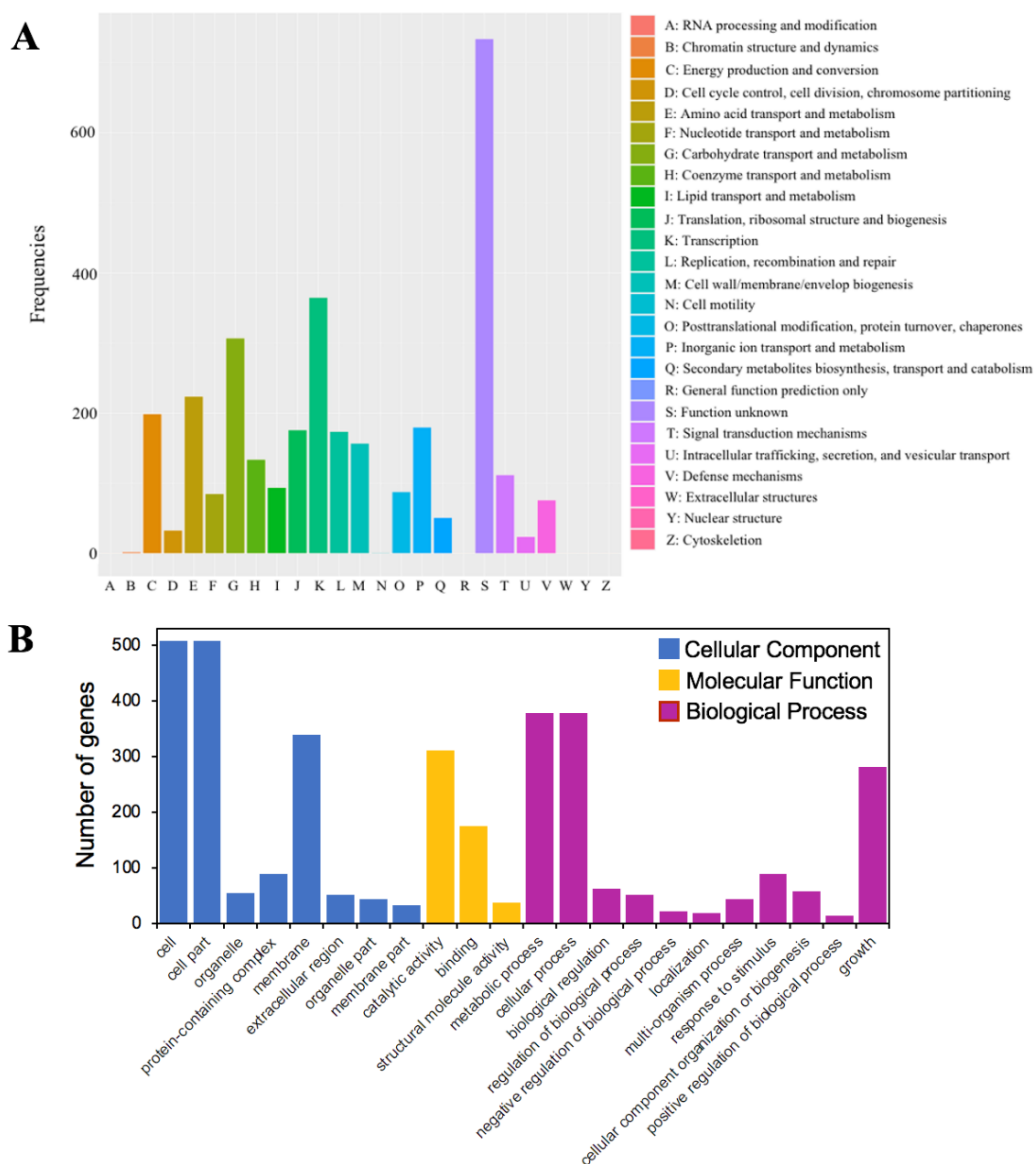


Figure 4 The functional annotations of *C. cellulans* MP1. (A) Cluster of orthologous gene (COG) classification. (B) Gene ontology (GO) functional classification.

Full-size DOI: [10.7717/peerj.11839/fig-4](https://doi.org/10.7717/peerj.11839/fig-4)

different strains. The COGs of *C. cellulans* MP1 were compared with four other strains, including J36, LMG16121, and ZKA48. It showed that *C. cellulans* MP1 comprised of 3405 COGs and 794 singletons. *C. cellulans* J36 and ZKA48 included 3604 and 3605 COGs, respectively, whereas the lowest COGs were observed in LMG16121 (Fig. 5A). Surprisingly, strain MP1 had the largest number of singletons ($n = 794$). Venn diagram denoted that a total of 2,539 COGs were commonly shared by all four strains of *C. cellulans*. Among the unique COGs observed in all strains, strain MP1 had the largest number of 20 (Fig. 5A). To support this result, the REALPHY phylogeny builder web tool was used to compare *C. cellulans* genomes. A maximum-likelihood phylogenetic tree showed that MP1 and

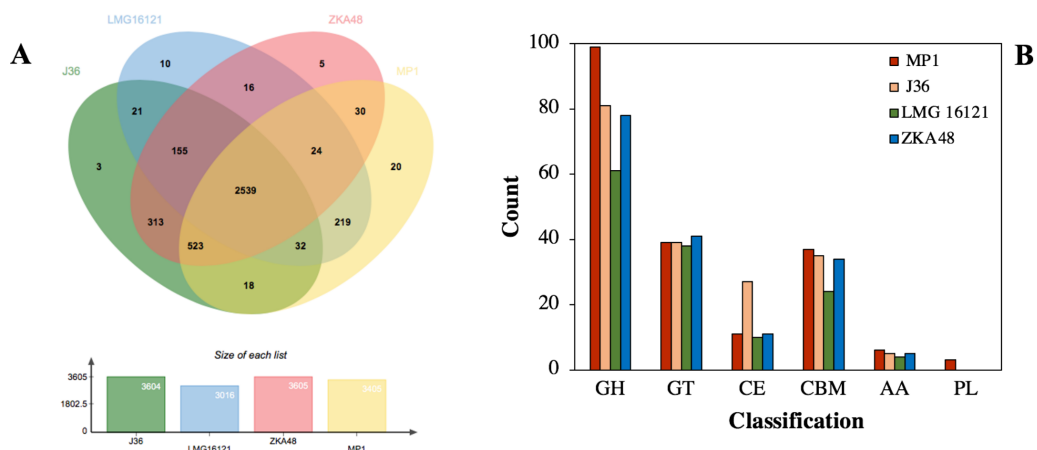


Figure 5 Comparative genomes between *C. cellulans* MP1 and three other *C. cellulans* strains. (A) Venn diagram represents the numbers of unique and shared orthologous genes of each strain. (B) Comparative genomic analysis of CAZymes across *C. cellulans* strains.

Full-size [DOI: 10.7717/peerj.11839/fig-5](https://doi.org/10.7717/peerj.11839/fig-5)

LMG16121 formed a monophyletic clade, suggesting a strong evolutionary relationship (Fig. S1).

Using dbCAN 2 meta server, the putative genes encoded for CAZymes present in *C. cellulans* MP1 were screened to find out the genes responsible for cellulose and hemicellulose degradation. After removing the sequences that did not meet the filtering criteria, a total of 195 predicted CAZymes was identified corresponding to 4.8% of the total of 3,216 protein-coding sequences observed in this strain. Glycoside hydrolases (GHs) involved in the degradation of the most plant biomass such as cellulose and hemicellulose were predicted to be the most abundant subfamily with 99 enzymes. Next, 39 glycosyltransferases (GTs), 37 carbohydrate-binding modules (CBMs), 11 carbohydrate esterases (CEs), 6 enzymes with auxiliary activities (AAs), and 3 polysaccharide lyases (PLs) were detected (Fig. 5B).

The genome of *C. cellulans* MP1 was compared to closely related *C. cellulans* such as *C. cellulans* J36, *C. cellulans* LMG16121, *C. cellulans* ZKA48. Generally, *C. cellulans* MP1 had significantly more CAZyme domains found than other *C. cellulans* strains. Despite forming a monophyletic clade with strain LMG16121, MP1 was 56 CAZyme domains higher than LMG16121. Meanwhile, with the exception of LMG16121, the family numbers of GT, CBM, and AA family numbers were the same in all compared strains, suggesting the coexistence of these genes in breaking down cellulosic biomass (Fig. 5B).

Mining of plant biomass-acting enzymes in *C. cellulans* MP1 gene pool

CAZyme analysis suggested that 30 cellulose-related sequences were detected in the genome of *C. cellulans* MP1. The major families related to the degradation of cellulose are GH6, GH9, GH48, GH10, GH16, GH1, GH3, GH13, and GH64. Based on the annotation, 5 endoglucanases (two GH6 and three GH9), 3 exoglucanases (GH6, GH10, and GH48), and lichenase (GH16) were revealed in the genome of strain MP1 (Table 2). A total of five

Table 2 List of predicted cellulolytic and hemicellulolytic enzymes present in the genome of *C. cellulans* MP1.

Classification	Locus tag	Predicted function
Cellulose-related	Orf_454, Orf_1616, Orf_2130, Orf_2755, Orf_3244	Endoglucanase [EC 3.2.1.4]
	Orf_1607, Orf_1610, Orf_1611	Exoglucanase [EC 3.2.1.91]
	Orf_2289	Lichenase [EC 3.2.1.73]
	Orf_2294, Orf_2388, Orf_2464, Orf_2606, Orf_3385, Orf_401, Orf_403, Orf_404, Orf_922	β -glucosidase [EC 3.2.1.21]
	Orf_2704, Orf_2893	Oligo-1,6-glucosidase [EC 3.2.1.10]
	Orf_2898	Maltodextrin glucosidase [EC 3.2.1.20]
	Orf_802	Glucan endo-1,3- β -glucosidase [EC 3.2.1.39]
	Orf_3703	β -galactosidase [EC 3.2.1.23]
	Orf_2449	Trehalose-6-phosphate hydrolase [EC 3.2.1.93]
	Orf_102	Pectate trisaccharide-lyase [EC 4.2.2.22]
	Orf_1326	Pectate lyase [EC 4.2.2.2]
	Orf_19, Orf_21	Levanase [EC 3.2.1.80]
	Orf_20	Levanbiose-producing levanase [EC 3.2.1.64]
	Orf_22	Levansucrase [EC 2.4.1.10]
Hemicellulose-related	Orf_905	Mannan endo-1,4- β -mannosidase [EC 3.2.1.78]
	Orf_3988, Orf_4004	Bifunctional β -xylosidase/ α -arabinosidase [EC 3.2.1.37; EC 3.2.1.55]
	Orf_4003	Arabinoxylan arabinofuranohydrolase [EC 3.2.1.55]
	Orf_35, Orf_2605	α -xylosidase [EC 3.2.1.177]
	Orf_24, Orf_2083, Orf_3386, Orf_3999, Orf_4000, Orf_4002	α -L-arabinofuranosidase [EC 3.2.1.55]
	Orf_34, Orf_3650, Orf_4001	Non-reducing end β -L-arabinofuranosidase [EC 3.1.1.185]
	Orf_3698	Exo- α -(1->6)-L-arabinofuranosidase [EC 3.2.1.-]
	Orf_1000, Orf_2117, Orf_3772	Endo-1,4- β -xylanase [EC 3.2.1.8]
	Orf_266, Orf_2247	α -galactosidase [EC 3.2.1.22]

out of the eight annotated endoglucanases and exoglucanases (*orf_454*, *orf_1616*, *orf_3244*, *orf_1607*, *orf_1611*) contained a CBM2 domain appended to them, which have been known to bind to crystalline cellulose, insoluble chitin, and xylan (McLean *et al.*, 2002). CBM2 was the major CBMs present in the genome of strain MP1. In the β -glucosidases family, three GH1 and six GH3 were considered as other important members for cellulose degradation and three GH13 had α -glucosidase activity. It is interesting to note that two pectate lyases such as PL4 (*orf_102*) and PL1 (*orf_1326*) belonging to the large class of PL were identified (Table 2). Given that pectin is involved in providing structural support for plant such as cell walls, wall integrity, and cell–cell cohesion, pectate lyases (EC 4.2.2.2) belonging to pectinase catalyze the eliminative cleavage of α -1,4-glycosidic bonds between C₄ and C₅ of pectin or pectic acid, producing unsaturated methyloligogalacturonates (Abbott, Gilbert & Boraston, 2010; Hugouvieux-Cotte-Pattat, Condemine & Shevchik, 2014). Pectate lyases are important symbiotic bacteria through facilitating their growth in presence of highly pectinolytic bacteria and under pectin-rich environments (Hugouvieux-Cotte-Pattat, Condemine & Shevchik, 2014).

For hemicellulose degrading genes mining, a total of 21 annotated proteins were deduced to involve in hemicellulose degradation. The major families responsible for the breakdown of hemicellulose were GH43, GH31, GH127, GH51, GH10, GH36, and GH4, and the total number of enzymes, including all families, was 21 (Table 2). Among them, the most abundant enzymes were attributed to GH43 (9 enzymes), followed by GH124 (3 enzymes) and GH31 (2 enzymes), indicating that strain MP1 might have great potentials for degradation of hemicellulosic backbones or debranding hemicellulose. Hemicellulose-related genes were predicted as mannan endo-1,4- β -mannosidase, arabinoxylan arabinofuranohydrolase, α -xylosidase, α -L-arabinofuranosidase, β -L-arabinofuranosidase, endo-1,4- β -xylanase, α -galactosidase, and bifunctional β -xylosidase/ α -arabinosidase (Table 2). Only two enzymes, mannan endo-1,4- β -mannosidase and α -galactosidase, have the ability to hydrolyse parts of mannan, the second major component in hemicellulose. Two endo-1,4- β -xylanase (*orf_100* and *orf_3772*) and one bifunctional β -xylosidase/ α -arabinosidase (*orf_3988*) were coupled with either CBM2 or CBM9. Surprisingly, genes encoding arabinofuranosidase (*orf_3999*, *orf_4000*, *orf_4001* and *orf_4002*), arabinoxylan arabinofuranohydrolase (*orf_4003*), and bifunctional β -xylosidase/ α -arabinosidase (*orf_4004*) are clustered in an operon. However, AA enzymes were not found to be related to cellulose and hemicellulose degradation.

To confirm the ability to produce endoglucanase, xylanase, and pectinase of strain MP1, the quantitative enzyme assay was monitored across different incubation periods (24 h, 48 h, 72 h, 96 h, 120 h) (Fig. 6). Using rice bran as an inducible substrate for endoglucanase production, the lowest endoglucanase activity of 0.15 ± 0.01 IU/ml was observed at 24 h, followed by an increase of activity within 48 h to 96 h. The highest endoglucanase production of about 3.17 ± 0.10 IU/ml was achieved after 72 h of incubation. The reduced activity of endoglucanase was notified as 1.40 ± 0.17 IU/ml at 120 h. Xylanase activity of strain MP1 was monitored on M1 medium supplemented with sugar cane bagasse at 37 °C for 120 h. The result was a steep increase within 0 to 72 h of incubation times. The highest enzyme production was recorded during the stationary phase reaching maximum (1.84 ± 0.08 IU/ml) at 96 h, followed by a slight decrease after 120 h (Fig. 6C). Regarding to pectinase, the maximum pectinase yield was 0.16 ± 0.01 IU/ml after 48 h of incubation. Increasing incubation time was subjected to a significant decrease of pectinase activity. Despite the fact that a number of studies involved in the yields of endoglucanase, xylanase, and pectinase have been performed, it is hard to compare due to the effect of production conditions, substrate and assay conditions, and the way of defining the units.

Levan exopolysaccharide biosynthesis and degradation

Apart from cellulose-related genes, we were interested in the gene encoding for levansucrase (*orf_22*) since the ability of the genus *Cellulosimicrobium* to produce levan has not been studied yet. Levansucrase (EC 2.4.1.10) belonging to GH32 participates in synthesizing levan using high sucrose concentration as the main substrate (Raga-Carbajal et al., 2018). The *C. cellulans* MP1 *ls* gene consists of 1,851 nucleotides encoding a protein of 616 amino acids with the predicted molecular mass of 66.6 kDa. Alignment of the amino acid sequence of *C. cellulans* MP1 *Ls* shows the highest level of identity with *Gluconacetobacter*

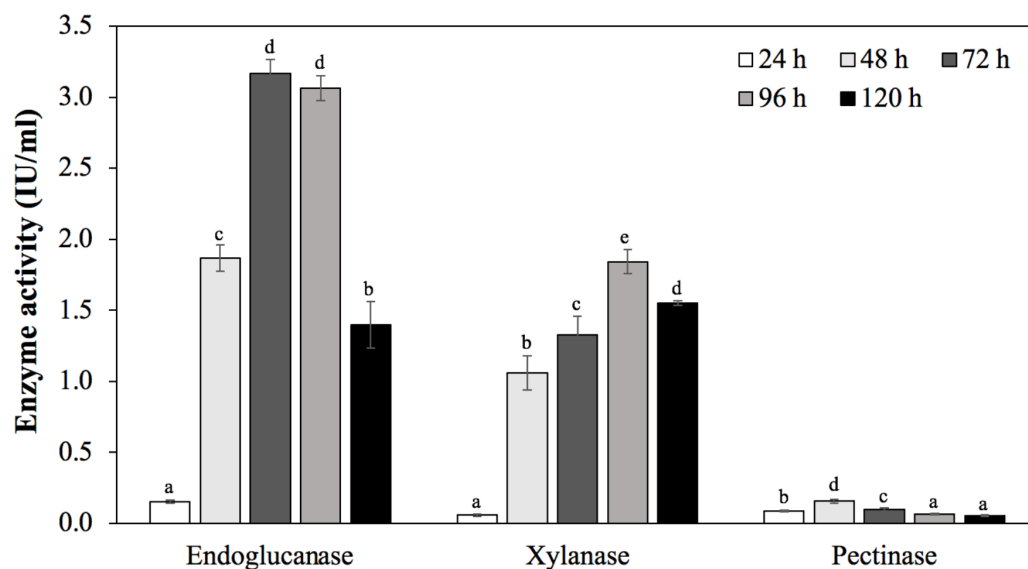


Figure 6 Enzymatic activities of *C. cellulans* MP1 observed in different incubation periods. Mean values with different letters a-d are significantly different according to the Fisher LCD test ($P < 0.05$).

Full-size [DOI: 10.7717/peerj.11839/fig-6](https://doi.org/10.7717/peerj.11839/fig-6)

diazotrophicus (61%) and *Microbacterium saccharophilum* (58%) (Fig. S2). In addition, levansucrase is clustered alongside levanbiose-producing levanase (*levB-orf_20*) and levanases (*sacC 1-orf_19* and *sacC 2-orf_21*) in an operon (Fig. 7A). This result suggested that levan produced by *C. cellulans* MP1 may be hydrolyzed into L-FOs such as levanbiose and levanases. It is noteworthy that the *sacC1-levB-sacC2-ls* operon was not present in the compared *C. cellulans* strains

The levan production in media with different initial sucrose concentrations was monitored as shown in Fig. 7B. Levan production depended on cultivation time and sucrose concentration. In the medium with 100 g/l sucrose, the maximum levan concentration was achieved at 20 h (9.9 ± 0.11 g/l), followed by a significant decrease after 24 h (5.5 ± 0.02 g/l). When sucrose concentration was increased to 200 g/l, the highest yield of levan was 14.8 ± 1.2 at 20 h. It was noteworthy that levan was accumulated in the stationary phase.

DISCUSSION

More recently, the exploitation of new plant biomass-acting enzymes and exopolysaccharides with special characteristics has become important since the technological use of agro-industrial residues is increasing. They are exploited favourably in food, paper, cosmetic and pharmaceutical industries (Angelin & Kavitha, 2020; Walia et al., 2017; Zheng et al., 2017). In this context, termites are thought to rely on the gut microbiome to digest wood and other types of plant biomass consisting mainly of cellulose as well as hemicellulose (Calusinska et al., 2020), making them an ideal source to search for new enzymes. Many cellulolytic bacteria from termite have been extensively investigated in the past, despite difficulties in isolation and cultivation (Kamsani et al., 2016). However, despite the use of high-throughput next-generation sequencing, potent biomass-acting enzymes have yet

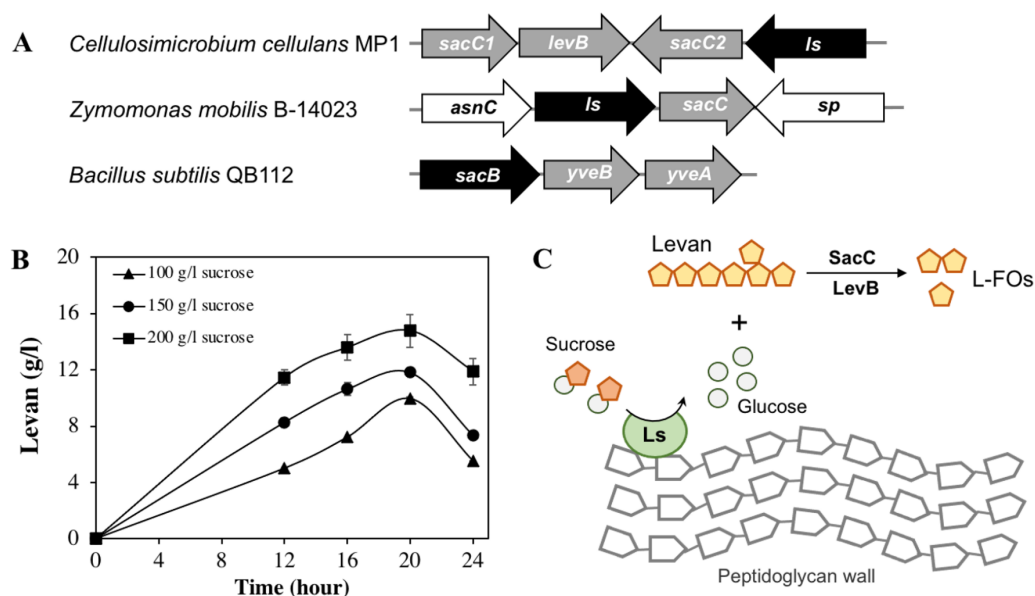


Figure 7 Levan biosynthesis in *C. cellulans* MP1. (A) Genetic organization of levan operon in *C. cellulans* MP1, *Zymomonas mobilis* B-14023 and *Bacillus subtilis* QB112. The *ls* and *sacC-levB* homologous genes are denoted in black and gray, respectively. (B) Time course of levan production. (C) Proposed mechanism for levan and levan-type fructooligosaccharide (L-FOs) biosynthesis.

Full-size [DOI: 10.7717/peerj.11839/fig-7](https://doi.org/10.7717/peerj.11839/fig-7)

to be revealed. Our work is the first genome analysis systematically elucidating enzymes related to the decomposition of cellulose and hemicellulose and levan production of termite-associated *C. cellulans*. The findings provide valuable genome information for biotechnological applications.

Despite having potential in many fields, the genus *Cellulosimicrobium* remains poorly investigated. To date, the genus *Cellulosimicrobium* consists 7 species (six validly and one non-validly published names): *C. variabile* (Bakalidou et al., 2002), *C. funkei* (Brown et al., 2006), *C. cellulans* (Schumann, Weiss & Stackebrandt, 2001), *C. terreum* (Yoon et al., 2007), *C. marium* (Hamada et al., 2016), *C. aquatile* (Sultanpuram et al., 2015), and *C. arenosum* (Oh et al., 2018). Only *C. variabile* was reportedly isolated from the hindgut of termites (Qin et al., 2018; Sharma, Gilbert & Lal, 2016). Since *C. cellulans* MP1 produced highly active cellulolytic activity on CMC and filter paper, we subsequently analyzed its whole genome for the presence of lignocellulose-degrading CAZymes. Of note, the genome possesses 17 cellulolytic and 21 hemicellulolytic enzymes, in which a total of 11 enzymes featuring CBM domains (CBM2, CBM4, CBM6, CBM9, CBM35) were found. CBM domains play an important role in enhancing the substrate-binding capability of the enzymes, and some CBMs are able to enhance the thermostability of the enzyme (Gilbert, Knox & Boraston, 2013). CBM2 binds to various GH families including GH6 (*orf_454*, *orf_3244*), GH9 (*orf_1616*), GH10 (*orf_1000*, *orf_1611*), GH43 (*orf_3988*), and GH48 (*orf_1607*), which indicates the ability of these enzymes to bind to and support catalytic domains to hydrolyze crystalline cellulose and xylan. CBM9 domain was found only in endo-1,4- β -xylanase (*orf_3772*), which is responsible mainly for attacking the xylan backbone and

is similar to *Clostridium stercorarium* Xyn10B and *Thermotoga maritima* Xyn10A (Lee & Lee, 2014). By contrast, endo-1,4- β -xylanase (*orf_2117*) had no CBM domain, indicating that other substrate-binding regions might replace the function of CBM. The detection of endoglucanase, exoglucanase, and xylanase activities is in agreement with the presence of the aforementioned genes. It is interesting to note that the strain MP1 cultivated in LB medium containing 0.5% CMC had the lowest endoglucanase but activity of both enzymes increased at least 5.7 fold in TM3 medium supplemented with 2% rice bran, a low-cost agro-residue (Table S3). This result was considered as a promising strategy for reducing the cost of enzyme production and increasing enzyme efficiency.

Pectate trisaccharide-lyase (*orf_102*) and pectate lyase (*orf_1326*), which are involved in plant tissue maceration and modification of the cell wall structure (Atanasova et al., 2018), were identified based on the genomic analysis of strain MP1. Due to the ability to cleave pectin using a β -elimination mechanism, pectate lyases are produced either by bacteria living in close proximity with plants or by symbiotic gut bacteria (Biz et al., 2014; Jayani, Saxena & Gupta, 2005). Pectate lyases are important enzymes for industrial applications such as the beverage industry, pulp processing, waste treatment, leading to increasing attention of researchers over the world (Jayani, Saxena & Gupta, 2005; Zhao et al., 2018). Despite 72 genome sequences available at the time of writing and an array of extracellular enzymes found in *Cellulosimicrobium*, pectate lyases have not been identified yet. *C. cellulans* MP1 seemed to have acquired 2 different pectate lyases to establish itself in the termite gut. Additionally, pectinase produced by strain MP1 was low activity due to unoptimized medium. The production of pectinase can be optimized by using the response surface methodology as well as the recombinant enzyme, which is an interesting subject for further studies.

Interestingly, the presence of genes related to the levan exopolysaccharide in the genome of *C. cellulans* MP1 was predictable. Levan is one of the two main types of fructan biopolymers, produced from sucrose through the action of levansucrase (GH68) (Feng et al., 2015). Many studies showed that the GH68 is present in a number of genera, including *Gluconobacter*, *Gluconacetobacter*, *Komagataeibacter*, *Asaia*, *Neoasaia*, *Bacillus*, and *Kozakia* (Jakob et al., 2019); however, it has not been reported in the genus *Cellulosimicrobium*. Our finding revealed that the *sacC1-levB-sacC2-ls* operon is not conserved across well-studied bacteria and that the *ls* gene is encoded for active levansucrase, catalyzing the synthesis of higher-molecular-weight levan in presence of sucrose, which may serve as carbon storage for *C. cellulans* MP1. The molecules are hydrolyzed to L-FOs by levan-degrading enzymes including *sacC12* and *levB* (Fig. 7C). Some studies indicate that levansucrase LevB hydrolyzes levan to generate levanbiose predominantly and SacC is reported to be active against levan and inulin leading to the formation of free fructose (Débarbouillé et al., 1991; Raga-Carbajal et al., 2018). In *B. subtilis*, anti-terminator, SacY is phosphorylated in presence of the excess sucrose, resulting in upregulation of conserved levansucrase *sacB* and two endolevanase *yveAB* (Pereira, Petit-Glatron & Chambert, 2001). Another study demonstrated that activation of *sacB* also is regulated by pleiotropic regulatory genes *degS/degU*, *degQ* and *degR* (Débarbouillé et al., 1991). While no apparent homologs of SacY have been identified in the chromosome of strain MP1, 4 transcriptional factors *degA*

(*orf_114, orf_207, orf_1005, orf_2448*) and *degU* (*orf_296, orf_527, orf_1182, orf_2409*) are present. Known that deletion of *sacC* in *Zymomonas mobilis* enhanced levan production 15.5 g/L to 21.2 g/L (*Senthilkumar et al., 2004*), further investigation on the regulation mechanisms of this operon is an interesting subject for improving levan production.

Interestingly, Ls shares the 2 conserved cysteines (at positions 383 and 439) and an overall sequence identity of 57–62% with orthologs of gram-negative bacteria such as *Microbacterium saccharophilum*, *Burkholderia vietnamiensis*, *Gluconacetobacter diazotrophicus* (*Fig. S2*). This result was in contrast to most gram-positive levansucrases that lack a pair of conserved cysteine residues (*Jakob et al., 2019*). In *G. diazotrophicus* levansucrase, Cys339-Cys395 intramolecular disulfide bond plays an important role in protein folding and stability. Serin substitution for either Cys339 or Cys395 abolished sucrose hydrolysis activity and levan-forming activity via preventing the extended loop between β -strands IIIB and IIIC with the insertion located between blades III and IV (*Martínez-Fleites et al., 2005*). Further crystal structure analysis is required to reveal the conformational changes of *C. cellulans* Ls upon reduction and oxidation conditions.

CONCLUSIONS

This study highlighted the ability of strain MP1 to degrade cellulose and hemicellulose and produce levan. Out of 65 isolated termite gut symbiotic bacteria, isolate MP1, identified as *C. cellulans*, exhibited the highest specific cellulase activity. The genome of *C. cellulans* MP1 is one of 17 genomes of *C. cellulans* that are released onto the NCBI genome database, but it is the first sequence that has been reported in detail from a biotechnological perspective. Both genomic and experimental evidence proved that *C. cellulans* MP1 possesses 30 cellulose and 21 hemicellulose-related sequences, which were functionally redundant for endoglucanases, endoxylanase, β -glucosidases, xylanase, β -xylosidases, arabinofuranosidase, and pectate lyase. Moreover, the *sacC1-levB-sacC2-ls* operon involved in levan and L-FOs production was pronounced for the first time, which could be a selective advantage during host-adaptation and colonization. These findings not only enrich the genome database but also provide a valuable source of information to continue research into the potential applications of *C. cellulans* MP1, including its possible use in the biofuel, pulp and paper, and pharmaceutical industries.

ACKNOWLEDGEMENTS

We are grateful to Emerson Addison, Central Michigan University for critically reading the manuscript.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This study was financially supported by the Institute of Biotechnology, Vietnam Academy of Science and Technology (VAST) under grant number: CN40SH.01/18-21. The funders

had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:

The Institute of Biotechnology, Vietnam Academy of Science and Technology (VAST): CN40SH.01/18-21.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Nguyen Thi-Hanh Vu conceived and designed the experiments, performed the experiments, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Tung Ngoc Quach performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Xuan Thi-Thanh Dao and Lam Tung Le performed the experiments, prepared figures and/or tables, and approved the final draft.
- Ha Thanh Le conceived and designed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Chi Phuong Le analyzed the data, prepared figures and/or tables, and approved the final draft.
- Lam Tung Nguyen, Cuong Cao Ngo and Ha Hoang performed the experiments, analyzed the data, prepared figures and/or tables, and approved the final draft.
- Ha Hoang Chu and Quyet-Tien Phi conceived and designed the experiments, analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:

The raw data is available in the [Supplemental Files](#).

The 16S rDNA gene ([MW534740](#)) and genomic sequences ([JAFGYF000000000](#)) of the strain *Cellulosimicrobium cellulans* MP1 are available at NCBI GenBank.

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.11839#supplemental-information>.

REFERENCES

- Abbott DW, Gilbert HJ, Boraston AB. 2010.** The active site of oligogalacturonate lyase provides unique insights into cytoplasmic oligogalacturonate beta-elimination. *Journal of Biological Chemistry* **285**:39029–39038 DOI [10.1074/jbc.M110.153981](https://doi.org/10.1074/jbc.M110.153981).

- Abdul Karim MH, Lam MQ, Chen SJ, Yahya A, Shahir S, Shamsir MS, Chong CS.** 2020. Draft genome sequence of *Parvularcula flava* strain NH6-79 T, revealing its role as a cellulolytic enzymes producer. *Archives of Microbiology* **202**:2591–2597 DOI [10.1007/s00203-020-01967-z](https://doi.org/10.1007/s00203-020-01967-z).
- Angelin J, Kavitha M.** 2020. Exopolysaccharides from probiotic bacteria and their health potential. *International Journal of Biological Macromolecules* **162**:853–865 DOI [10.1016/j.ijbiomac.2020.06.190](https://doi.org/10.1016/j.ijbiomac.2020.06.190).
- Atanasova L, Dubey M, Grujić M, Gudmundsson M, Lorenz C, Sandgren M, Kubicek CP, Jensen DF, Karlsson M.** 2018. Evolution and functional characterization of pectate lyase PEL12, a member of a highly expanded *Clonostachys rosea* polysaccharide lyase 1 family. *BMC Microbiology* **18**:178 DOI [10.1186/s12866-018-1310-9](https://doi.org/10.1186/s12866-018-1310-9).
- Bailey MJ, Biely P, Poutanen K.** 1992. Interlaboratory testing of methods for assay of xylanase activity. *Journal of Biotechnology* **23**:257–270 DOI [10.1016/0168-1656\(92\)90074-J](https://doi.org/10.1016/0168-1656(92)90074-J).
- Bakalidou A, Kämpfer P, Berchtold M, Kuhnigk T, Wenzel M, König H.** 2002. *Cellulosimicrobium variable* sp. nov. a cellulolytic bacterium from the hindgut of the termite *Mastotermes darwiniensis*. *International Journal of Systematic and Evolutionary Microbiology* **52**:1185–1192 DOI [10.1099/00207713-52-4-1185](https://doi.org/10.1099/00207713-52-4-1185).
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA.** 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology: A Journal of Computational Molecular Cell Biology* **19**:455–477 DOI [10.1089/cmb.2012.0021](https://doi.org/10.1089/cmb.2012.0021).
- Belghith KS, Dahech I, Belghith H, Mejdoub H.** 2012. Microbial production of levansucrase for synthesis of fructooligosaccharides and levan. *International Journal of Biological Macromolecules* **50**:451–458 DOI [10.1016/j.ijbiomac.2011.12.033](https://doi.org/10.1016/j.ijbiomac.2011.12.033).
- Biz A, Farias FC, Motter FA, De Paula DH, Richard P, Krieger N, Mitchell DA.** 2014. Pectinase activity determination: an early deceleration in the release of reducing sugars throws a spanner in the works! *PLOS ONE* **9**:e109529 DOI [10.1371/journal.pone.0109529](https://doi.org/10.1371/journal.pone.0109529).
- Bolger AM, Lohse M, Usadel B.** 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**:2114–2120 DOI [10.1093/bioinformatics/btu170](https://doi.org/10.1093/bioinformatics/btu170).
- Brown JM, Steigerwalt AG, Morey RE, Daneshvar MI, Romero L-J, McNeil MM.** 2006. Characterization of clinical isolates previously identified as *Oerskovia turbata*: proposal of *Cellulosimicrobium funkei* sp. nov. and emended description of the genus *Cellulosimicrobium*. *International Journal of Systematic and Evolutionary Microbiology* **56**:801–804 DOI [10.1099/ijs.0.63882-0](https://doi.org/10.1099/ijs.0.63882-0).
- Calusinska M, Marynowska M, Bertucci M, Untereiner B, Klimek D, Goux X, Sillam-Dussès D, Gawron P, Halder R, Wilmes P, Ferrer P, Gerin P, Roisin Y, Delfosse P.** 2020. Integrative omics analysis of the termite gut system adaptation to *Miscanthus* diet identifies lignocellulose degradation enzymes. *Communications Biology* **3**:275 DOI [10.1038/s42003-020-1004-3](https://doi.org/10.1038/s42003-020-1004-3).

- Chen SJ, Lam MQ, Thevarajoo S, Abd Manan F, Yahya A, Chong CS. 2020.** Genome analysis of cellulose and hemicellulose degrading *Micromonospora* sp. CP22. *3 Biotech* **10**:160 DOI [10.1007/s13205-020-2148-z](https://doi.org/10.1007/s13205-020-2148-z).
- Chutani P, Sharma KK. 2016.** Concomitant production of xylanases and cellulases from *Trichoderma longibrachiatum* MDU-6 selected for the deinking of paper waste. *Bioprocess and Biosystems Engineering* **39**:747–758 DOI [10.1007/s00449-016-1555-3](https://doi.org/10.1007/s00449-016-1555-3).
- Dahech I, Belghith KS, Hamden K, Feki A, Belghith H, Mejdoub H. 2011.** Antidiabetic activity of levan polysaccharide in alloxan-induced diabetic rats. *International Journal of Biological Macromolecules* **49**:742–746 DOI [10.1016/j.ijbiomac.2011.07.007](https://doi.org/10.1016/j.ijbiomac.2011.07.007).
- Dashtban M, Schraft H, Qin W. 2009.** Fungal bioconversion of lignocellulosic residues; opportunities & perspectives. *International Journal of Biological Sciences* **5**:578–595 DOI [10.7150/ijbs.5.578](https://doi.org/10.7150/ijbs.5.578).
- Débarbouillé M, Martin-Verstraete I, Arnaud M, Klier A, Rapoport G. 1991.** Positive and negative regulation controlling expression of the *sac* genes in *Bacillus subtilis*. *Research in Microbiology* **142**:757–764 DOI [10.1016/0923-2508\(91\)90052-c](https://doi.org/10.1016/0923-2508(91)90052-c).
- Diken E, Ozer T, Arikan M, Emrence Z, Oner ET, Ustek D, Arga KY. 2015.** Genomic analysis reveals the biotechnological and industrial potential of levan producing halophilic extremophile, *Halomonas smyrnensis* AAD6T. *Springerplus* **4**:393 DOI [10.1186/s40064-015-1184-3](https://doi.org/10.1186/s40064-015-1184-3).
- Dogsa I, Brloznic M, Stopar D, Mandic-Mulec I. 2013.** Exopolymer diversity and the role of levan in *Bacillus subtilis* biofilms. *PLOS ONE* **8**:e62044 DOI [10.1371/journal.pone.0062044](https://doi.org/10.1371/journal.pone.0062044).
- Eida AA, Bougouffa S, Alam I, Saad MM, Hirt H. 2020.** Complete genome sequence of the endophytic bacterium *Cellulosimicrobium* sp. JZ28 isolated from the root endosphere of the perennial desert tussock grass *Panicum turgidum*. *Archives of Microbiology* **202**:1563–1569 DOI [10.1007/s00203-020-01859-2](https://doi.org/10.1007/s00203-020-01859-2).
- Feng J, Gu Y, Quan Y, Zhang W, Cao M, Gao W, Song C, Yang C, Wang S. 2015.** Recruiting a new strategy to improve levan production in *Bacillus amyloliquefaciens*. *Scientific Reports* **5**:13814 DOI [10.1038/srep13814](https://doi.org/10.1038/srep13814).
- Galperin MY, Makarova KS, Wolf YI, Koonin EV. 2015.** Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucleic Acids Research* **43**:D261–D269 DOI [10.1093/nar/gku1223](https://doi.org/10.1093/nar/gku1223).
- Ghose TK. 1987.** Measurement of cellulase activities. *Pure and Applied Chemistry* **59**:257–268 DOI [10.1351/pac198759020257](https://doi.org/10.1351/pac198759020257).
- Gilbert HJ, Knox JP, Boraston AB. 2013.** Advances in understanding the molecular basis of plant cell wall polysaccharide recognition by carbohydrate-binding modules. *Current Opinion in Structural Biology* **23**:669–677 DOI [10.1016/j.sbi.2013.05.005](https://doi.org/10.1016/j.sbi.2013.05.005).
- Gojgic-Cvijovic GD, Jakovljevic DM, Loncarevic BD, Todorovic NM, Pergal MV, Ciric J, Loos K, Beskoski VP, Vrvic MM. 2019.** Production of levan by *Bacillus licheniformis* NS032 in sugar beet molasses-based medium. *International Journal of Biological Macromolecules* **121**:142–151 DOI [10.1016/j.ijbiomac.2018.10.019](https://doi.org/10.1016/j.ijbiomac.2018.10.019).
- Gray DA, White JBR, Oluwole AO, Rath P, Glenwright AJ, Mazur A, Zahn M, Baslé A, Morland C, Evans SL, Cartmell A, Robinson CV, Hiller S, Ranson NA, Bolam DN,**

- Van den Berg B.** 2021. Insights into SusCD-mediated glycan import by a prominent gut symbiont. *Nature Communications* **12**:44 DOI [10.1038/s41467-020-20285](https://doi.org/10.1038/s41467-020-20285).
- Grissa I, Vergnaud G, Pourcel C.** 2007. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Research* **35**:W52–W57 DOI [10.1093/nar/gkm360](https://doi.org/10.1093/nar/gkm360).
- Gupta P, Samant K, Sahu A.** 2012. Isolation of cellulose-degrading bacteria and determination of their cellulolytic potential. *International Journal of Microbiology* **2012**:578925 DOI [10.1155/2012/578925](https://doi.org/10.1155/2012/578925).
- Hamada M, Shibata C, Tamura T, Nurkanto A, Ratnakomala S, Lisdiyanti P, Suzuki K-I.** 2016. Cellulosimicrobium marinum sp. nov. an actinobacterium isolated from sea sediment. *Archives of Microbiology* **198**:439–444 DOI [10.1007/s00203-016-1204](https://doi.org/10.1007/s00203-016-1204).
- Hugouvieux-Cotte-Pattat N, Condemine G, Shevchik VE.** 2014. Bacterial pectate lyases, structural and functional diversity. *Environmental Microbiology Reports* **6**:427–440 DOI [10.1111/1758-2229.12166](https://doi.org/10.1111/1758-2229.12166).
- Jäger G, Büchs J.** 2012. Biocatalytic conversion of lignocellulose to platform chemicals. *Biotechnology Journal* **7**:1122–1136 DOI [10.1002/biot.201200033](https://doi.org/10.1002/biot.201200033).
- Jakob F, Quintero Y, Musacchio A, Estrada-delos Santos P, Hernández L, Vogel RF.** 2019. Acetic acid bacteria encode two levansucrase types of different ecological relationship. *Environmental Microbiology* **21**:4151–4165 DOI [10.1111/1462-2920.14768](https://doi.org/10.1111/1462-2920.14768).
- Jayani RS, Saxena S, Gupta R.** 2005. Microbial pectinolytic enzymes: a review. *Process Biochemistry* **40**:2931–2944 DOI [10.1016/j.procbio.2005.03.026](https://doi.org/10.1016/j.procbio.2005.03.026).
- Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong SY, Lopez R, Hunter S.** 2014. InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**:1236–1240 DOI [10.1093/bioinformatics/btu031](https://doi.org/10.1093/bioinformatics/btu031).
- Kamlage B.** 1996. Methods for general and molecular bacteriology. *Food / Nahrung* **40**:103–103 DOI [10.1002/food.19960400226](https://doi.org/10.1002/food.19960400226).
- Kamsani N, Salleh MM, Yahya A, Chong CS.** 2016. Production of lignocellulolytic enzymes by microorganisms isolated from *Bulbitermes* sp. termite gut in solid-state fermentation. *Waste and Biomass Valorization* **7**:357–371 DOI [10.1007/s12649-015-9453-5](https://doi.org/10.1007/s12649-015-9453-5).
- Lee SH, Lee YE.** 2014. Cloning and characterization of a multidomain GH10 xy-lanase from *Paenibacillus* sp. DG-22. *Journal of Microbiology and Biotechnology* **24**:1525–1535 DOI [10.4014/jmb.1407.07077](https://doi.org/10.4014/jmb.1407.07077).
- Lee I, Ouk Kim Y, Park SC, Chun J.** 2016. OrthoANI: an improved algorithm and software for calculating average nucleotide identity. *International Journal of Systematic and Evolutionary Microbiology* **66**:1100–1103 DOI [10.1099/ijsem.0.000760](https://doi.org/10.1099/ijsem.0.000760).
- Li Y, Lei L, Zheng L, Xiao X, Tang H, Luo C.** 2020. Genome sequencing of gut symbiotic *Bacillus velezensis* LC1 for bioethanol production from bamboo shoots. *Biotechnology for Biofuels* **13**:34 DOI [10.1186/s13068-020-1671-9](https://doi.org/10.1186/s13068-020-1671-9).
- Mardo K, Visnapuu T, Vija H, Aasamets A, Viigand K, Alamäe T.** 2017. A highly active endo-levanase BT1760 of a dominant mammalian gut commensal bacteroides

- thetiaotaomicron cleaves not only various bacterial levans, but also levan of Timothy grass. *PLOS ONE* 12:e0169989 DOI 10.1371/journal.pone.0169989.
- Martínez-Fleites C, Ortiz-Lombardía M, Pons T, Tarbouriech N, Taylor EJ, Arrieta JG, Hernández L, Davies GJ. 2005.** Crystal structure of levansucrase from the Gram-negative bacterium *Gluconacetobacter diazotrophicus*. *The Biochemical Journal* 390:19–27 DOI 10.1042/BJ20050324.
- McLean BW, Boraston AB, Brouwer D, Sanaie N, Fyfe CA, Warren RA, Kilburn DG, Haynes CA. 2002.** Carbohydrate-binding modules recognize fine sub-structures of cellulose. *The Journal of Biological Chemistry* 277:50245–50254 DOI 10.1074/jbc.M204433200.
- Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. 2013.** Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 14:60 DOI 10.1186/1471-2105-14-60.
- Mercimek Takcı HA, Turkmen FU. 2016.** Extracellular pectinase production and purification from a newly isolated *Bacillus subtilis* strain. *International Journal of Food Properties* 19:2443–2450 DOI 10.1080/10942912.2015.1123270.
- Miller GL. 1959.** Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Analytical Chemistry* 31:426–428 DOI 10.1021/ac60147a030.
- Oh M, Kim JH, Yoon JH, Schumann P, Kim W. 2018.** Cellulosimicrobium arenosum sp. nov. Isolated from Marine Sediment Sand. *Current Microbiology* 75:901–906 DOI 10.1007/s00284-018-1464-6.
- Ohkuma M. 2003.** Termite symbiotic systems: efficient bio-recycling of lignocellulose. *Applied Microbiology and Biotechnology* 61:1–9 DOI 10.1007/s00253-002-1189-z.
- Pandey AK, Edgard G, Negi S. 2016.** Optimization of concomitant production of cellulase and xylanase from *Rhizopus oryzae* SN5 through EVOP-factorial design technique and application in Sorghum Stover based bioethanol production. *Renewable Energy* 98:51–56 DOI 10.1016/j.renene.2016.05.071.
- Pasari N, Gupta M, Eqbal D, Yazdani SS. 2019.** Genome analysis of *Paenibacillus polymyxa* A18 gives insights into the features associated with its adaptation to the termite gut environment. *Scientific Reports* 9:6091 DOI 10.1038/s41598-019-42572-5.
- Pereira Y, Petit-Glatron MF, Chambert R. 2001.** *yveB*, Encoding endolevanase LevB, is part of the *sacB-yveB-yveA* levansucrase tricistronic operon in *Bacillus subtilis*. *Microbiology* 147:3413–3419 DOI 10.1099/00221287-147-12-3413.
- Phi QT, Park YM, Seul KJ, Ryu CM, Park SH, Kim JG, Ghim SY. 2010.** Assessment of root-associated *Paenibacillus polymyxa* groups on growth promotion and induced systemic resistance in pepper. *Journal of Microbiology and Biotechnology* 20:1605–1613.
- Powers EM. 1995.** Efficacy of the Ryu nonstaining KOH technique for rapidly determining gram reactions of food-borne and waterborne bacteria and yeasts. *Applied and Environmental Microbiology* 61:3756–3758 DOI 10.1128/AEM.61.10.3756-3758.1995.

- Qin W, Fan F, Zhu Y, Huang X, Ding A, Liu X, Dou J. 2018.** Anaerobic biodegradation of benzo(a)pyrene by a novel *Cellulosimicrobium cellulans* CWS2 isolated from polycyclic aromatic hydrocarbon-contaminated soil. *Brazilian Journal of Microbiology* **49**:258–268 DOI [10.1016/j.bjrm.2017.04.014](https://doi.org/10.1016/j.bjrm.2017.04.014).
- Raga-Carbajal E, López-Munguía A, Alvarez L, Olvera C. 2018.** Understanding the transfer reaction network behind the non-processive synthesis of low molecular weight levan catalyzed by *Bacillus subtilis* levansucrase. *Scientific Reports* **8**:15035 DOI [10.1038/s41598-018-32872-7](https://doi.org/10.1038/s41598-018-32872-7).
- Schumann P, Weiss N, Stackebrandt E. 2001.** Reclassification of *Cellulomonas cellulans* (Stackebrandt and Keddie 1986) as *Cellulosimicrobium cellulans* gen. nov. comb. nov. *International Journal of Systematic and Evolutionary Microbiology* **51**:1007–1010 DOI [10.1099/00207713-51-3-1007](https://doi.org/10.1099/00207713-51-3-1007).
- Senthilkumar V, Rameshkumar N, Busby SJ, Gunasekaran P. 2004.** Disruption of the *Zymomonas mobilis* extracellular sucrose gene (*sacC*) improves levan production. *Journal of Applied Microbiology* **96**:671–676 DOI [10.1111/j.1365-2672.2003.02169.x](https://doi.org/10.1111/j.1365-2672.2003.02169.x).
- Sharma A, Gilbert JA, Lal R. 2016.** (Meta)genomic insights into the pathogenome of *Cellulosimicrobium cellulans*. *Scientific Reports* **6**:25527 DOI [10.1038/srep25527](https://doi.org/10.1038/srep25527).
- Shih IL, Yu YT, Shieh CJ, Hsieh CY. 2005.** Selective production and characterization of levan by *Bacillus subtilis* (Natto) Takahashi. *Journal of Agricultural and Food Chemistry* **53**:8211–8215 DOI [10.1021/jf058084o](https://doi.org/10.1021/jf058084o).
- Somogyi M. 1945.** A new reagent for the determination of sugars. *Journal of Biological Chemistry* **160**:61–68 DOI [10.1016/S0021-9258\(18\)43097-9](https://doi.org/10.1016/S0021-9258(18)43097-9).
- Sultanpuram VR, Mothe T, Chintalapati S, Chintalapati VR. 2015.** *Cellulosimicrobium aquatile* sp. nov., isolated from Panagal reservoir, Nalgonda, India. *Antonie Van Leeuwenhoek* **108**:1357–1364 DOI [10.1007/s10482-015-0588-y](https://doi.org/10.1007/s10482-015-0588-y).
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013.** MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution* **30**:2725–2729 DOI [10.1093/molbev/mst197](https://doi.org/10.1093/molbev/mst197).
- Teather RM, Wood PJ. 1982.** Use of Congo red-polysaccharide interactions in enumeration and characterization of cellulolytic bacteria from the bovine rumen. *Applied and Environmental Microbiology* **43**:777–780 DOI [10.1128/AEM.43.4.777-780.1982](https://doi.org/10.1128/AEM.43.4.777-780.1982).
- Tokuda G, Watanabe H. 2007.** Hidden cellulases in termites: revision of an old hypothesis. *Biology Letters* **3**:336–339 DOI [10.1098/rsbl.2007.0073](https://doi.org/10.1098/rsbl.2007.0073).
- Ventorino V, Ionata E, Birolo L, Montella S, Marcolongo L, De Chiaro A, Espresso F, Faraco V, Pepe O. 2016.** Lignocellulose-adapted endo-cellulase producing *Streptomyces* strains for bioconversion of cellulose-based materials. *Frontiers in Microbiology* **7**:2061 DOI [10.3389/fmicb.2016.02061](https://doi.org/10.3389/fmicb.2016.02061).
- Verspreet J, Dornez E, VandenEnde W, Delcour JA, Courtin CM. 2015.** Cereal grain fructans: structure, variability and potential health effects. *Trends in Food Science & Technology* **43**:32–42 DOI [10.1016/j.tifs.2015.01.006](https://doi.org/10.1016/j.tifs.2015.01.006).
- Walia A, Guleria S, Mehta P, Chauhan A, Parkash J. 2017.** Microbial xylanases and their industrial application in pulp and paper biobleaching: a review. *3 Biotech* **7**:11 DOI [10.1007/s13205-016-0584-6](https://doi.org/10.1007/s13205-016-0584-6).

- Wang Y, Coleman-Derr D, Chen G, Gu YQ. 2015.** OrthoVenn: a web server for genome wide comparison and annotation of orthologous clusters across multiple species. *Nucleic Acids Research* **43**:W78–W84 DOI [10.1093/nar/gkv487](https://doi.org/10.1093/nar/gkv487).
- Wattam AR, Davis JJ, Assaf R, Boisvert S, Brettin T, Bun C, Conrad N, Dietrich EM, Disz T, Gabbard JL, Gerdes S, Henry CS, Kenyon RW, Machi D, Mao C, Nordberg EK, Olsen GJ, Murphy-Olson DE, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Vonstein V, Warren A, Xia F, Yoo H, Stevens RL. 2017.** Improvements to PATRIC, the all-bacterial bioinformatics database and analysis resource center. *Nucleic Acids Research* **45**:D535 DOI [10.1093/nar/gkw1017](https://doi.org/10.1093/nar/gkw1017).
- Wi SG, Cho EJ, Lee D-S, Lee SJ, Lee YJ, Bae H-J. 2015.** Lignocellulose conversion for biofuel: a new pretreatment greatly improves downstream biocatalytic hydrolysis of various lignocellulosic materials. *Biotechnology for Biofuels* **8**:228 DOI [10.1186/s13068-015-0419-4](https://doi.org/10.1186/s13068-015-0419-4).
- Wong LJ, H'Ng PS, Wong SY, Lee SH, Lum WC, Chai EW, Wong WZ, Chin KL. 2014.** Termite digestomes as a potential source of symbiotic microbiota for lignocelluloses degradation: a review. *Pakistan Journal of Biological Sciences* **17**:956–963 DOI [10.3923/pjbs.2014.956.963](https://doi.org/10.3923/pjbs.2014.956.963).
- Yoo SH, Yoon EJ, Cha J, Lee HG. 2004.** Antitumor activity of levan polysaccharides from selected microorganisms. *International Journal of Biological Macromolecules* **34**:37–41 DOI [10.1016/j.ijbiomac.2004.01.002](https://doi.org/10.1016/j.ijbiomac.2004.01.002).
- Yoon JH, Kang SJ, Schumann P, Oh TK. 2007.** Cellulosimicrobium terreum sp. nov. isolated from soil. *International Journal of Systematic and Evolutionary Microbiology* **57**:2493–2497 DOI [10.1099/ijms.0.64889-0](https://doi.org/10.1099/ijms.0.64889-0).
- Zhao Y, Yuan Y, Zhang X, Li Y, Li Q, Zhou Y, Gao J. 2018.** Screening of a novel polysaccharide lyase family 10 pectate lyase from *Paenibacillus polymyxa* KF-1: cloning, expression and characterization. *Molecules* **23**:2774 DOI [10.3390/molecules23112774](https://doi.org/10.3390/molecules23112774).
- Zheng F, Zhang W, Chu X, Dai Y, Li J, Zhao H, Wen L, Yue H, Yu S. 2017.** Genome sequencing of strain *Cellulosimicrobium* sp. TH-20 with ginseng biotransformation ability. *3 Biotech* **7**:237 DOI [10.1007/s13205-017-0850-2](https://doi.org/10.1007/s13205-017-0850-2).