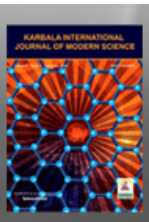

Karbala International Journal of Modern Science

Volume 12 | Issue 1 Article 3

Molecular analyses of Cellulosimicrobium cellulans Mosul isolated from crown galls of Olea europaea

Ghazwan Qasim Hasan
Department of Biology, College of Education for Pure Sciences, University of Mosul, Iraq,
dr.ghazwan@uomosul.edu.iq

Molecular analyses of Cellulosimicrobium cellulans Mosul isolated from crown galls of Olea europaea

Abstract

Olea europaea L. is farmed extensively for its oil in various parts of the world. Olives are abundant in bioactives and beneficial minerals that may have therapeutic and medicinal applications. *Cellulosimicrobium cellulans* belongs to the family Promicromonosporaceae and is classified as a Gram-positive bacillus. Since there is not much research on the topic, the goal of the current study was to analyze the whole genome sequence of this bacterium. The results showed that 99.8% of the raw values agreed well with the reference strain CP017660, with a GC content of 74.5% and a pairwise similarity of 98.1%. The *de novo* assembly analysis revealed that 3,250 contigs totaling 5,272,102 lengths were present, derived from 20% of the raw reads that were supplied. However, the findings showed that *C. cellulans* had 1772 encoding genes for proteins without functions and 2175 expressing genes for particular proteins with particular functions. The genome annotation result showed 41 completely characterized annotated features, including alanyl-tRNA synthetase, the aerotolerance protein BatB, argininosuccinate synthase, 3-hydroxyisobutyryl-CoA hydrolase, and others. The phylogenetic analysis based on the genomic DNA showed that the *C. cellulans* Mosul strain grouped in a clade that included closely associated strains, such as CP017660, CP041694, CP020857, CP072387, and CP017660 strains. Moreover, the heatmap matrix outcomes demonstrated a substantial relationship among *C. cellulans* and CP041694, CP072387, CP017660, CP020857, CP101475, and CP053419 strains. Lastly, the subsystem feature results showed that 134 genes encoding enzymes related to vitamins, pigments, cofactors, and prosthetic groups were present, along with 34 genes linked to defense, virulence, and disease.

Keywords

Whole genome sequencing; 16S rRNA; C. cellulans; O. europaea; Crown galls; Heatmap matrix; Genome annotation.

RESEARCH PAPER

Molecular Analyses of *Cellulosimicrobium cellulans* Mosul Isolated From Crown Galls of *Olea europaea*

Ghazwan Q. Hasan

Department of Biology, College of Education for Pure Sciences, University of Mosul, Iraq

Abstract

***Olea europaea* L. is farmed extensively for its oil in various parts of the world. Olives are abundant in bioactives and beneficial minerals that may have therapeutic and medicinal applications. *Cellulosimicrobium cellulans* belongs to the family Promicromonosporaceae and is classified as a Gram-positive bacillus. Since there is not much research on the topic, the goal of the current study was to analyze the whole genome sequence of this bacterium. The results showed that 99.8 % of the raw values agreed well with the reference strain CP017660, with a GC content of 74.5 % and a pairwise similarity of 98.1 %. The *de novo* assembly analysis revealed that 3250 contigs totaling 5,272,102 lengths were present, derived from 20 % of the raw reads that were supplied. However, the findings showed that *C. cellulans* had 1772 encoding genes for proteins without functions and 2175 expressing genes for particular proteins with particular functions. The genome annotation result showed 41 completely characterized annotated features, including alanyl-tRNA synthetase, the aerotolerance protein BatB, argininosuccinate synthase, 3-hydroxyisobutyryl-CoA hydrolase, and others. The phylogenetic analysis based on the genomic DNA showed that the *C. cellulans* Mosul strain grouped in a clade that included closely associated strains, such as CP017660, CP041694, CP020857, CP072387, and CP017660 strains. Moreover, the heatmap matrix outcomes demonstrated a substantial relationship among *C. cellulans* and CP041694, CP072387, CP017660, CP020857, CP101475, and CP053419 strains. Lastly, the subsystem feature results showed that 134 genes encoding enzymes related to vitamins, pigments, cofactors, and prosthetic groups were present, along with 34 genes linked to defense, virulence, and disease.**

Keywords: Whole genome sequencing, 16S rRNA, *C. cellulans*, *O. europaea*, Crown galls, Heatmap matrix, Genome annotation

1. Introduction

The little tree *Olea europaea*, which belongs to the Oleaceae family, is often found in western Asia, Africa, and the Mediterranean Basin [1]. Similar to many Mediterranean woody species that thrive in semi-arid environments, olive trees also tolerate soil water availability. Olive adaptability may be influenced by interactions with advantageous microbes [2]. Gram-positive is the defining feature of the genus *Cellulosimicrobum*, which mostly consists of three species: *Cellulosimicrobium cellulans*, *Cellulosimicrobium funkei*, and *Cellulosimicrobium terreum*. Of these, *C. cellulans* and *C. funkei* have been found to be pathogenic [3]. The short reads with an output of up to 6 Tb can be sequenced using Illumina's NovaSeq 6000 sequencing platform. This platform follows the standard Illumina sequencing procedure, which consists of library preparation, in situ amplification-based cluster creation, and synthesis-based sequencing. One of the main characteristics of the NovaSeq 6000 is its flexibility [4]. The genome sequence analysis of *C. cellulans* MP1 revealed 30 genes involved in the degradation of cellulases and hemicellulases. Also, the sequence involved 23 contigs, 4046 protein-coding genes, 1 mRNA gene, 3 tRNA genes, and 67 rRNA [5]. The Illumina Miseq analysis of the genome of *Cellulosimicrobium* sp. TH-20 strain revealed 3450 open reading frames, 9 rRNA, 51

Received 18 June 2025; revised 25 October 2025; accepted 30 October 2025.
Available online 21 November 2025

E-mail addresses: dr.ghazwan@uomosul.edu.iq, ghazwan.qasim@yahoo.com.

https://doi.org/10.33640/2405-609X.3438

tRNA, and 11 genes involved in the biotransformation of ginsenosides [6]. The 16S rRNA sequences of *Cellulosimicrobium composti* BIT-GX5T and *Cellulosimicrobium funkei* CT-R177 strains were 99.17 % and 99.86 % identical to *C. cellulans* LMG 16121T, respectively [7,8], also the phylogenetic analysis of the 16S showed present five housekeeping genes (recA, trpB, rpoB, gyrA, and atpD) [7].

Scientists worldwide are currently focusing on bacterial symbionts as possible sources for detecting new and extremely effective lignocellulose-degrading enzymes, which can be utilized in a wide range of sectors, including agriculture, pulp and paper, food, and biofuel [5]. The synergistic activity of three distinct groups, including β-glucosidase, endoglucanase, and exoglucanase, is responsible for cellulose hydrolysis [9]. While the common hemicellulases that are involved in hemicellulose disintegration are mannanase, arabinofuranosidases, xylanases, and other enzymes [10]. Among 99 enzymes, glycoside hydrolases, which break down the majority of plant biomass, including hemicellulose and cellulose, were projected to be the most prevalent subfamily found in *C. cellulans* [5]. The genome analyses of the *Cellulosimicrobium* genus showed an abundant provider of glycosidases implicated in ginseng biological transformation capabilities and plant growth promotion [11]. The current study was conducted in order to investigate the complete genome analysis of *C. cellulans*.

2. Materials and methods

2.1. Isolation of C. cellulans

The crown galls isolated from the stems of *O. europaea* were thoroughly crushed with a ceramic mortar and then immersed in ordinary saline to maintain bacterial viability. The samples were incubated overnight at 28 °C. The nutrient agar medium was made by dissolving 28 g of nutritional agar in 1 L of distilled water, per the instructions provided by the Neogen/UK (Spanish) manufacturer. The streak plate method was applied to spread a tiny amount of the inoculum onto petri dishes containing NA medium using an inoculation loop [12]. The microbes then were isolated by consecutive streaks. Once bacterial cells were placed on agar plates, streaking initiated a gradient dilution on the agar surfaces. The dense growth of bacteria occurred in the portion of the medium with the fewest bacterial cells as a result of this range of dilution. To guarantee purity, single colonies were isolated and picked up using sterile loops prior to being disseminated into a fresh NA medium. The *C. cellulans* colonies were allowed to grow for 72 h at 28 °C [13].

2.2. DNA extraction

The genomic DNA was extracted from a pure culture of *C. cellulans* using the cetyltrimethylammonium bromide (CTAB) method, as mentioned by Ref. [14], with some changes, including the incubation of preheated CTAB buffer for 1 h, then spinning down of the deposited DNA. with some modifications (the precipitated DNA was spun down at 735 g for 5 min after the prepared CTAB buffer was incubated for 45 min.

2.3. High-throughput sequencing and data analysis

The TruSeq DNA library preparation kits (Illumina, San Diego, CA, USA) were used to create the library at the DNA Link company/Republic of Korea. Following the manufacturer's recommendations, the DNA was sequenced using the Illumina NovaSeq6000 platform, a 2 × 150 bp reads technique, and the WGS (PCR-Free 550 prep). To remove low-quality reads and create clean, high-quality reads, the DNA raw reads were first run with Trimmomatic v 0.40 and then through the Geneious Prime BBDuk program [15,16]. The total number of FASTQ reads was multiplied by the read length (151 bp) to get the genome sequencing coverage, which was then divided by the entire genome size of the bacterium [16].

2.4. Map to reference

Geneious mapper (medium-low sensitivity) was used to map DNA data to the reference sequences, with up to five iterations. The whole reads were mapped against the CP017660 strain, the primary causative agent, *C. cellulans*. The findings were presented in a report that included statistics in assembled readings and the overall number of reads used [15].

2.5. Pairwise alignment

The entire length of the bacterium sequence was aligned with a closely related sequence, applying Geneious alignments in Geneious prime. Pairwise identity and genome variation were reported.

2.6. Phylogenetic tree analysis of genomic DNA

The whole-genome-based phylogenetic tree of thirteen strains of *C. cellulans* and the most closely

related strains was inferred using the REALPHY tool https://realphy.unibas.ch/realphy/ [17]. The genome in FASTA format was uploaded to the server with the default settings. The heatmap matrix of the examined strains was also analyzed based on the outcomes of the tree alignments.

2.7. Phylogenetic tree of 16S ribosomal RNA

Neighbor joining method was applied to construct the phylogeny tree. For alignment and manual optimization, the Geneious Prime V. 11 [15]; http://www.geneious.com was utilized. Sequences of all aligned lengths were then extracted using ClustalW alignment. General Time Reversible (GTR) has been used to recreate the evolutionary tree. The phylogeny inference was accomplished using Bayesian inference of phylogeny (MrBayes 3.2.6) [18]. The phylogenetic analysis included twelve strains.

2.8. De novo assembly

De novo assembly of the entire raw reads was done using Geneious software assembler. Geneious software paired the bi-directional reads before subjecting them to the assembler, then 20 % of the reads were taken to create over 1000 contigs of varying lengths, each of which was constructed of several overlapping assembled reads with a consensus sequence.

2.9. Genome annotation

The assembled genome was annotated with the Rapid Annotation using Subsystem Technology (RAST) server [19]. The SEED tool [20] was used for predicting functional genes in subsystem categories.

3. Results

3.1. C. cellulans genomic analysis

The BBduk tool was used to obtain the raw reads of the *C. cellulans* genome. The outcomes revealed that 7440 reads were removed, and 9051 reads were trimmed to improve the product quality, where 85.97 % of the total reads were covered by K-mer (Table 1).

3.2. Whole genome reads mapping of C. cellulans

The results showed that 99.8 % of the data matched the reference genome CP017660 with 98.1 % pairwise identity, and 74.5 % GC content (Table 2, Fig. 1).

3.3. De novo assembly of the data

The results showed that 3250 contigs with a total length of 5,272,102 bp were obtained from 20 % of applied raw reads (Table 3).

3.4. General genome features of C. cellulans

The results showed that *C. cellulans* contains 2175 encoding genes for specific proteins with specific functions and 1772 encoding genes for proteins without functions. These proteins were categorized with coverage of 90.15, 44.89, and 93.49 % for hypothetical characteristics, local protein family features, and protein-encoding features, respectively (Table 4).

3.5. C. cellulans genome annotation

The genome was annotated using the RAST tool, and 41 fully classified annotated features were identified, including 3-hydroxyisobutyryl-CoA hydrolase, Aerotolerance protein BatB, Alanyl-tRNA synthetase, argininosuccinate acid synthase, and other features (Table 5).

3.6. Phylogenetic tree analysis based on genomic DNA and 16S rRNA gene of C. cellulans

The phylogenetic analysis showed that isolate *C. cellulans* Mosul was positioned in a clade involved closely related strains of CP101475 (strain TL2; China: Shandong) from poultry farm *Cellulosimicrobium cellulans* strain TL2 chromosome, complete genome - Nucleotide - NCBI, CP041694 (strain NEB113; USA) *Cellulosimicrobium cellulans* strain NEB113 chromosome, complete genome - Nucleotide - NCBI, CP053419 (72-3: China) from domestic animals *Cellulosimicrobium* sp. 72-3 chromosome, complete genome - Nucleotide - NCBI, CP020857 (TH-20: China; Fusong) from ginseng soil *Cellulosimicrobium* sp. TH-20 chromosome, complete genome - Nucleotide - NCBI, CP072387 (strain ORNL-0100 1; USA) from Human oral cavity *Cellulosimicrobium cellulans* strain ORNL-0100 chromosome, complete gen - Nucleotide - NCBI, and CP017660 (Saudia Arabia: Jizan) that collected from *Panicum turgidum* *Cellulosimicrobium* sp. JZ28, complete genome - Nucleotide - NCBI [11] also see (Fig. 2). However, the evolutionary analysis of 16S rRNA showed clustering of *C. cellulans* Mosul

Table 1. Total raw reads of C. cellulans genomic before and after trimming (producing clean reads) by the BBduk tool in Geneious Prime software.

Bacteria (total DNAseq)	No. of reads before trimming	No. of reads after trimming	Q trimmed	K trimmed	Total removed	quality-trimming both ends
C. cellulans	52,254,426	52,246,986	9051 reads (0.02 %)	44,923,358 reads (85.97 %)	7440 reads (0.01 %)	Q6.0

The result showed that only 7440 reads have been totally removed and 9051 reads trimmed for their quality, while K-mer trimming covered 44,923,358 reads.

Table 2. Mapping the whole reads of C. cellulans against CP017660 (Cellulosimicrobium sp).

Accession number	Reference bacteria	No. of assembled reads/bp	Consensus sequence length/bp	Pairwise Identity with the ref. virus	Coverage ratio	GC content
CP017660	*Cellulosimicrobium* sp.	48,664,758	4,139,600	98.1 %	99.8 %	74.5 %

The result showed that mapping has covered 99.8 % of the reference genome with 98.1 % pairwise identity and GC content of 74.5 %.

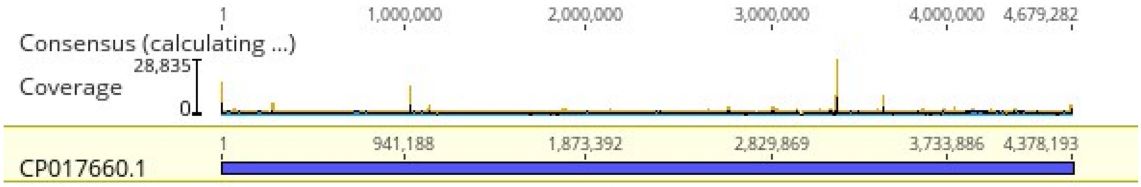

Fig. 1. Assembled reads to the reference sequence of Cellulosimicrobium sp. Mapping to reference shows that the alignment between ref. bacteria (CP017660; Saudia Arabia: Jizan that collected from P. turgidum) and the examined strain C. cellulans is 98.1 % similar.

(PV276936) among a clade of two strains of CP053419 (72-3: China) from domestic animals, and CP017660 (Saudia Arabia: Jizan) that collected from *P. turgidum*, and a clade of nine strains as shown in (Fig. 3). The evolutionary analysis based on 16S rRNA of *C. cellulans* Mosul is consistent with [11] and https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_013127315.1/

3.7. Heatmap matrix of C. cellulans

The heatmap visualization displayed a greater relationship (high intensity) between *C. cellulans* and CP101475, CP072387, CP017660, CP041694, CP020857, and CP053419 strains, and a moderate relationship (medium intensity) with CP052757, CP059447, CP159290, and CP021383, and a limited relationship (low intensity) with both CP048210 and CP042174 (Fig. 4).

3.8. Subsystem feature counts of C. cellulans

The subsystem result revealed that 34 genes were linked to defense, illness, and virulence (Fig. 5). Enzymes linked to pigments, co-factors, vitamins, and prosthetic groups are encoded by 134 genes.

3.9. Plant growth promotion and protection

The findings indicated that thiazole synthase, succinate dehydrogenase (SDH), sulfur acceptor protein, and sulfur carrier protein are among the proteins that may be linked to the promotion of plant growth (Table 5).

3.10. Cell division proteins

Three membrane proteins, FtsQ, FtsL, and cell division trigger factor, that are expected to be

Table 3. De novo assembly of the raw reads data.

Statistics	Unused Reads	All Contigs	Contigs ≥ 100 bp	Contigs ≥ 1000 bp
Number of	20,770	3250	3175	195
Min Length (bp)	10	24	101	1005
Median Length (bp)		260	262	16,121
Mean Length (bp)	146	1622	1658	22,876
Max Length (bp)	151	121,247	121,247	121,247
N50 Length (bp)		30,443	30,443	38,782
Number of contigs ≥ N50		50	50	38
Length Sum (bp)	3,034,486	5,272,102	5,267,309	4,461,004

The 20 % of the raw reads data was applied to obtain 3250 contigs with total length 5,272,102 bp.

Table 4. General genome quality features of C. cellulans generated using RAST server.

Feature	Value
Genome ID	157920.49
Genome name	*C. cellulans*
Reference genome	1710.12
Coarse consistency (%)	98.2
Fine consistency (%)	97.2
Completeness (%	100
Contamination (%)	0.8
Evaluation Group	R100 (Cellulomonas sp. JC225)
Contig count	1
DNA size (bp)	4378193
Contigs N50 (bp)	4378193
Contigs L50	1
Overpresent Roles	25
Underpresent Roles	14
Predicted Roles	1470
Completeness Roles	368
Total Distinct Roles	2018
Protein-Encoding Genes with Functional Assignment	2175
Protein-Encoding Genes without Functional Assignment	1772
% Protein-Encoding Feature Coverage	90.15
% Features that are Hypothetical	44.89
% Features that are in Local Protein Families	93.49

General genome quality features of *C. cellulans* generated showed 2175 protein-encoding genes with functional assignment, and 1772 protein-encoding genes without functional assignment. Also, proteins were classified as 90.15 % protein-encoding feature coverage, 44.89 % features that are hypothetical; and 93.49 % features that are in local protein families.

involved in cell division were found in the *C. cellulans* genome (Table 5).

4. Discussion

A variety of amino acids and their derivatives, such as levan polysaccharide and yeast-lytic enzymes such as endo-β-1,3-glucanases, are produced by *C. cellulans* and are utilized in the digestion of yeast cell walls [21,22]. *Cellulomonas cellulans*, *Oerskovia xanthineolytica*, and *Arthrobacter luteus* are aliases for *C. cellulans*, an actinomycete that excretes lytic enzyme complexes from yeast cell walls such as endo-β-1,3-glucanases EC 3.2.1.39 and 3.2.1.6 [21,22]. The genome mapping of *C. cellulans* showed a similar genome size to some extent to *Cellulosimicrobium* sp. JZ28, which was 4,139,600 bp in *C. cellulans* and 4,378,193 bp in JZ28, and the GC content was also identical (74.5 %) with [11]. The genome features result has a considerable amount of similarity to Ref. [11] in terms of the number of proteins that they contain. These proteins were categorized as 90.15 %, 44.89 %, and 93.49 % protein-encoding feature coverage, features that are hypothetical, and features that are in local protein families, respectively (Table 4). The genomes of *C. cellulans* have a median of about 4000 protein-coding genes, a circular chromosome with a length of 4.2–4.6 million base

Table 5. Genome annotation of C. cellulans by the RAST server.

Role	Annotated count
3-hydroxyisobutyryl-CoA hydrolase (EC 3.1.2.4)	1
Aerotolerance protein BatB	1
Aerotolerance protein BatC	0
Alanyl-tRNA synthetase (EC 6.1.1.7)	2
Allophanate hydrolase 2 subunit 1 (EC 3.5.1.54)	0
Argininosuccinate synthase (EC 6.3.4.5)	2
Aspartyl-tRNA (Asn) synthetase (EC 6.1.1.23)	2
ATP-dependent Clp protease ATP-binding subunit ClpX	2
ATP-dependent Clp protease, ATP-binding subunit ClpC	2
ATP synthase epsilon chain (EC 3.6.3.14)	3
ATP synthase F0 sector subunit a (EC 3.6.3.14)	2
Cell division protein FtsL	0
Cell division protein FtsQ	0
Cell division trigger factor (EC 5.2.1.8)	2
Chorismate mutase II (EC 5.4.99.5)	1
Dephospho-CoA kinase (EC 2.7.1.24)	2
Flagellar assembly protein FliH	0
Flagellar basal body-associated protein FliL	0
Flagellar biosynthesis protein FliO	0
Flagellar biosynthesis protein FliS	0
Flagellar M-ring protein FliF	2
Flagellar protein FliJ	0
Iron-dependent repressor IdeR/DtxR	2
Polyphosphate glucokinase (EC 2.7.1.63)	2
Prolyl-tRNA synthetase (EC 6.1.1.15), bacterial type	2
PTS system, N-acetylglucosamine-specific IIC component	1
RNA polymerase sigma factor for flagellar operon	1
Succinate dehydrogenase cytochrome *b*-556 subunit	0
Succinate dehydrogenase hydrophobic membrane anchor protein	1
Sulfur acceptor protein => iron-sulfur cluster assembly SufE	1
Sulfur carrier protein ThiS	0
Thiazole synthase (EC 2.8.1.10)	1
Tricarboxylate transport sensor protein TctE	1
Tricarboxylate transport transcriptional regulator TctD	0
Ubiquinone biosynthesis regulatory protein kinase UbiB	1
Urease accessory protein UreD	0
Urease accessory protein UreF	2
Uricase (urate oxidase) (EC 1.7.3.3)	2
Uroporphyrinogen-III synthase (EC 4.2.1.75)	0

The genome was annotated using the RAST server, and genome report showed 41 fully characterized annotated features over the whole sequence of the strain.

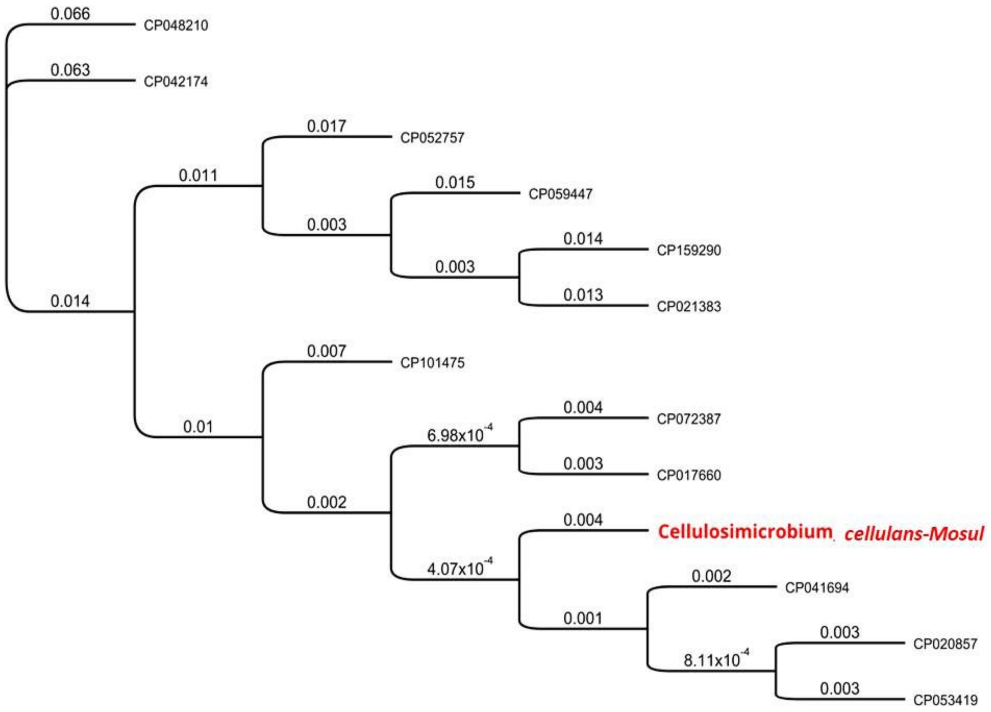

Fig. 2. Phylogenetic tree of genomic DNA of C. cellulans. The tree shows that C. cellulans Mosul is closely related to CP053419, CP020857, CP072387, and CP017660 strains.

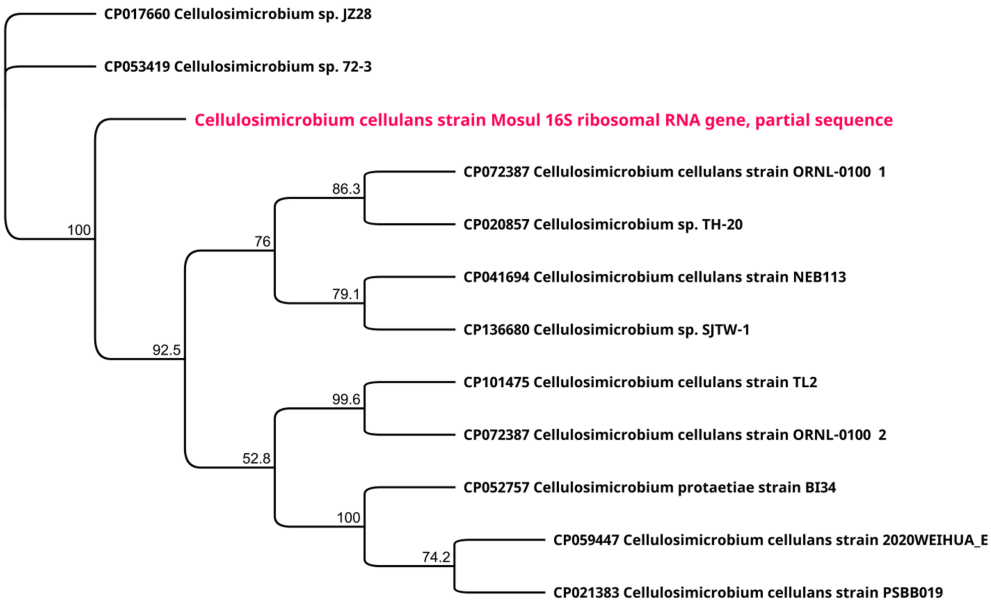

Fig. 3. Evolutionary analysis of C. cellulans based on 16S rRNA gene (PV276936). The tree shows that isolate C. cellulans Mosul was featured between CP053419 and CP017660, and a clade of nine strains.

pairs, and a high G + C concentration. They are known to have genes for the breakdown of carbohydrates, such as cellulases and hemicelluloses, and to have genes for denitrification and oxygen/sulfur cycling, which are adaptations to different environments [5]. The genus *Cellulosimicrobium* exhibits considerable genetic variation, according to full-genome phylogeny. *C. cellulans* may comprise several genomospecies, which would contradict existing taxonomic categories. Using whole-genome sequencing data, these phylogenomic techniques classify closely related species,

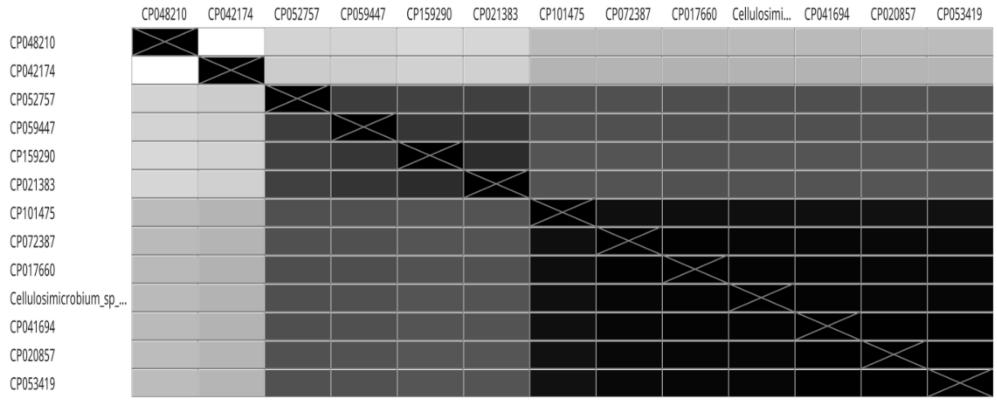

Fig. 4. Heatmap matrix of the examines strain of C. cellulans and some related isolates based on the multiple alignments.

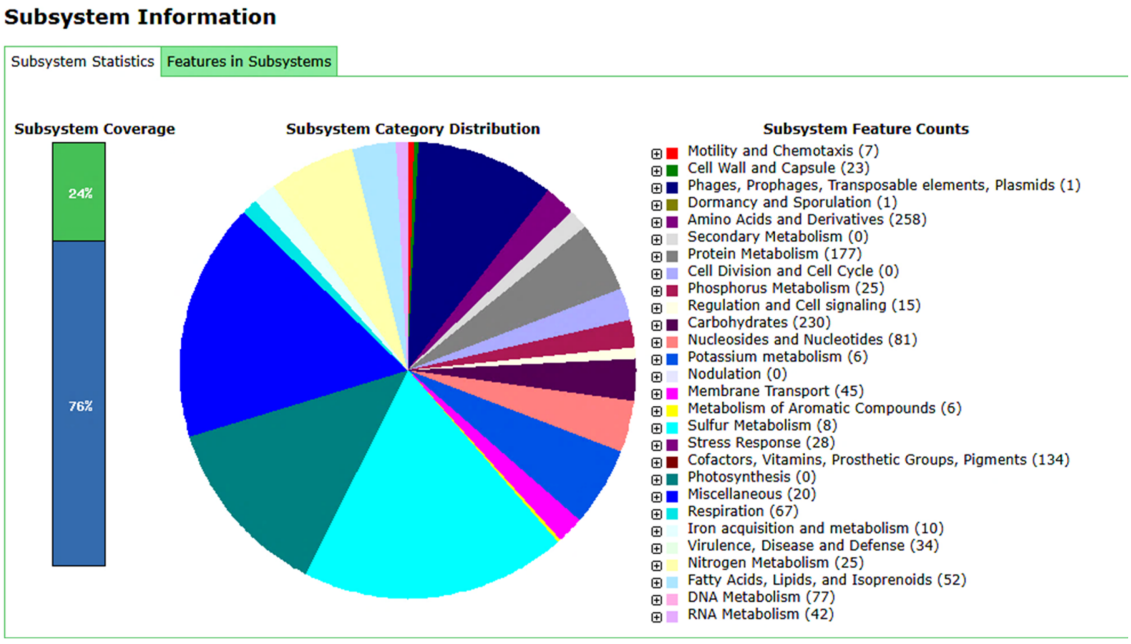

Fig. 5. Subsystem category distribution statistics of C. cellulans. Each subsystem feature count is displayed in a pie chart, and SEED viewer is used to demonstrate the subsystem coverage. The percent of proteins included in the subsystems is shown by the green bar of the subsystem coverage, whilst the percent of proteins excluded from the subsystems is represented by the blue bar.

revealing unique species groups and occasionally casting doubt on conventional classifications based on the 16S rRNA gene. Research employing whole-genome data or conserved single-copy genes offers a more precise and high-resolution picture of the evolutionary relationships of *Cellulosimicrobium* species [23]. Hydrolytic enzymes produced by *Cellulosimicrobium* sp. may help them to survive in harsh environments and compete with other natural bacteria [24,25]. The succinate dehydrogenase (SDH) plays physiological roles in plants through its effects on stomata function, fungal defense, root elongation, and photosynthesis. [26]. The thiamin-biosynthesis process involves the sulfur-carrier protein that is expressed by this gene. Thiazole-phosphate synthase (ThiG) catalyzes the formation of the sulfur carrier protein (ThiS) [27]. The inclusion of the -S-C=N group as a toxophoric unit in the molecular structure has been linked to the importance of thiazoles' antibacterial qualities [27,28]. Bacteria cannot survive without the presence of disome proteins. One divisome protein that has been suggested to act as a scaffold protein during divisome construction is the FtsQBL complex [29]. It implies that mobility to the mid-cell could only occur after complex formation and that the FtsQ, FtsB, and FtsL connections may exist in cells before translocation to this region

[30]. Additionally, it generates cyclo-(dehydroala-L-Leu), anthranilic acid, and other cyclic dipeptides as metabolites [31]. Using spectroscopic techniques, these chemicals were separated from bacterial cultures and identified; anthranilic acid had strong antifungal efficacy against plant infections [31]. Although there is a study on peptides from marine sources, no known natural products from *C. cellulans* or other marine sources are known to be "celluloids" and were classified as straight-chain peptides. Rather, studies have identified a new class of linear peptides called cellulamides from the actinomycetota genus *Cellulosimicrobium*, which includes *C. cellulans*. The cellulamide A peptides are a novel discovery from the species *C. cellulans*, which is linked to marine organisms [8]. The creation of enzymes that degrade plant cell walls, yeast, and even tough organic pollutants like benzo(a)pyrene is part of the diverse metabolic protein of *C. cellulans* [32]. It also uses important metabolic processes such as glycolysis, the TCA cycle, and several others for biosynthesis and energy production [23]. *C. cellulans* uses a variety of carbohydrates, such as lactose, maltose, sucrose, ribose, glucose, and rhamnose, those present in cellulose and other compounds obtained from plants [33]. It also has enzymes that are capable of hydrolyzing complex polysaccharides, like glucans, which are present in the cell walls of yeast [34]. Moreover, the *C. cellulans* probably uses a broad range of vitamins and cofactors for its metabolic functions [35]. Some general information can be deduced from related bacteria and metabolic pathways, even though the search results presented do not reveal many specifics regarding the precise cofactors and vitamins that this species needs. It has been shown that *C. cellulans* may degrade plant cellulosimicrobium, a nitrogen metabolism-related activity. Specifically, strains of *C. cellulans* have been found to utilize several types of nitrogen sources and break down cellulose, xylan, and even the nitrogen-containing compound quinclorac (QNC) [36–38]. These bacteria have also been connected to the promotion of plant growth and biocontrol due to their ability to fix nitrogen and participate in other enzymatic activities. The ability of *C. cellulans* to decompose complex materials, such as plant biomass, is well documented. Because it contains a variety of enzymes, including cellulases and hemicellulases, it can degrade cellulose, hemicellulose, and other components of plant cell walls [36,37,39]. Furthermore, C. cellulans possesses a unique ability to degrade significant pollutants such as the polycyclic aromatic hydrocarbon (PAH) benzo(a) pyrene [32].

5. Conclusions

The results showed that 99.8 % of the raw values agreed well with the reference strain CP017660, with a pairwise similarity of 98.1 % and a GC content of 74.5 %. Based on 20 % of the raw reads, the *de novo* assembly analysis showed that 3250 contigs with a total length of 5,272,102 were present. The *C. cellulans* has 3947 protein-encoding genes. Alanyl-tRNA synthetase, aerotolerance protein BatB, argininosuccinate synthase, and others were among the 41 fully defined annotated features that were revealed by the genome annotation result. The phylogenetic analysis of the genomic DNA of *C. cellulans* Mosul revealed that this strain was grouped in a clade of most closely related strains, including CP017660, CP041694, CP020857, CP072387, and CP017660. Furthermore, the results of the heatmap matrix showed a strong correlation between *C. cellulans* Mosul and the previously mentioned strains. Finally, the subsystem feature results revealed 34 genes associated with defense, virulence, and disease, as well as 134 genes encoding enzymes related to vitamins, pigments, cofactors, and prosthetic groups.

Funding

I would like to inform you that this work has been done privately (self-funding).

Ethical statement

This work does not include human and animal experiments.

Conflict of interest

There is no conflict of interest.

Acknowledgements

The author would like to thank the University of Mosul/Iraq for its continued support; without its help, this work would never have been possible. The author also would like to thank Dr Osama Alisawi/ College of Agriculture, University of Kufa for his friendly assistance.

References

[1] M. Shafiq, M.Z. Iqbal, M. Athar, The beneficial effects of olive tree (*Olea europaea* L.) in the nutritional, pharmaceutical and industrial application: a review, J. Plant. Dev. 31 (2024) 247–266, https://doi.org/10.47743/jpd.2024.31.1.950.

[2] M. Cardoni, J. Mercado-Blanco, Confronting stresses affecting olive cultivation from the holobiont perspective, Front. Plant Sci. 14 (2023) 1261754, https://doi.org/10.3389/fpls.2023.1261754.

[3] M. Rivero, J. Alonso, M.F. Ramón, G. Nancy, P. Ana, I. Marín, A. Navascués, R. Juanbeltz, Infections due to *Cellulosimicrobium* species: case report and literature review, BMC Infect Dis 19 (2019) 816, https://doi.org/10.1186/s12879-019-4440-2.
[4] A. Modi, S. Vai, D. Caramelli, M. Lari, M, The illumina sequencing protocol and the NovaSeq 6000 system, in: Bacterial Pangenomics: Methods and Protocols, Springer US, New York, NY. 2021, p. 1542, https://doi.org/10.1007/978-1-0716-1099-2_2.
[5] N.T.H. Vu, T.N. Quach, X.T.T. Dao, H.T. Le, C.P. Le, L. T. Nguyen, L.T. Le, C.C. Ngo, H. Hoang, H.H. Chu, Q.T. Phi, A genomic perspective on the potential of termite-associated *Cellulosimicrobium cellulans* MP1 as producer of plant biomass-acting enzymes and exopolysaccharides, PeerJ 9 (2021) e11839, https://doi.org/10.7717/peerj.11839.
[6] F. Zheng, W. Zhang, X. Chu, Y. Dai, J. Li, H. Zhao, L. Wen, H. Yue, S. Yu, Genome sequencing of strain *Cellulosimicrobium* sp. TH-20 with ginseng biotransformation ability, 3 Biotech 7 (2017) 1–10, https://doi.org/10.1007/s13205-017-0850-2.
[7] L. Hu, M. Xia, X. Gao, Y.X. Huo, Y. Yang, *Cellulosimicrobium composti* sp. nov., a thermophilic bacterium isolated from compost, Int. J. Syst. Evol. Microbiol. 71 (2021) 004905, https://doi.org/10.1099/ijsem.0.004905.
[8] M. Girão, J. Murillo-Alba, J. Martín, I. Pérez-Victoria, R. B. Leite, R. Urbatzka, P.N. Leão, M.F. Carvalho, F. Reyes, Cellulamides: a new family of marine-sourced linear peptides from the underexplored *cellulosimicrobium* genus, Mar. Drugs. 22 (2024) 268, https://doi.org/10.3390/md22060268.
[9] V. Rani, S. Mohanram, R. Tiwari, L. Nain, A. Arora, Beta-glucosidase: key enzyme in determining efficiency of cellulase and biomass hydrolysis, J. Bioprocess Biotech. 5 (2014) 197, https://doi.org/10.4172/2155-9821.1000197.
[10] D. Shallom, Y. Shoham, Microbial hemicellulases, Curr. Opin. Microbiol 6 (2003) 219–228, https://doi.org/10.1016/S1369-5274(03)00056-0.
[11] A.A. Eida, S. Bougouffa, I. Alam, M.M. Saad, H. Hirt, Complete genome sequence of the endophytic bacterium *Cellulosimicrobium* sp. JZ28 isolated from the root endosphere of the perennial desert tussock grass *Panicum turgidum*, Arch. Microbiol. 202 (2020) 1563–1569, https://doi.org/10.1007/s00203-020-01859-2.
[12] D.S. Katz, The Streak Plate Protocol, American Society for Microbiology, Washington, DC, USA. 2008, pp. 1–14 [Google Scholar].
[13] E. Osdaghi, A.J. Young, R.M. Harveson, Bacterial wilt of dry beans caused by *Curtobacterium flaccumfaciens* pv. *flaccumfaciens*: a new threat from an old enemy, Mol. Plant Pathol. 21 (2020) 605–621, https://doi.org/10.1111/mpp.12926.
[14] P. Tiels, E. Baranova, K. Piens, C. De Visscher, G. Pynaert, W. Nerinckx, S. Jan, F. Franck, H. Paco, T. Simon, G. Steven, V.H. Annelies, V. Albena, V. Wouter, R. Han, C. Nico Callewaert, A bacterial glycosidase enables mannose-6-phosphate modification and improved cellular uptake of yeast-produced recombinant human lysosomal enzymes, Nat. Biotechnol. 30 (2012) 1225–1231, https://doi.org/10.1038/nbt.2427.
[15] M. Kearse, R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, T. Tobias, A. Bruce, M. Peter, D. Alexei, Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data, Bioinformatics 28 (2012) 1647–1649, https://doi.org/10.1093/bioinformatics/bts199.
[16] B.D. Wingfield, L. De Vos, A.M. Wilson, T.A. Duong, N. Vaghefi, A. Botes, N.K. Ravindra, C. Ramesh, P. Barsha, A. Habibu, J.B. Martin, C. ShuaiFei, D.M. Pieterr, L. FeiFei, N. Sudhir, S. Shagun, T.S. Emma, S. Hiroyuki, A.T. Kalonji, A.V.D.N. Magriet, J.W. Michael J. Wingfield, IMA Genome-F16: draft genome assemblies of *Fusarium marasasianum, Huntiella abstrusa*, two *Immersiporthe knoxdaviesiana* isolates, *Macrophomina pseudophaseolina, Macrophomina phaseolina, Naganishia randhawae*, and *Pseudocercospora cruenta*, IMA Fungus 13 (2022) 3, https://doi.org/10.1186/s43008-022-00089-z.
[17] F. Bertels, O.K. Silander, M. Pachkov, P.B. Rainey, E. Van Nimwegen, Automated reconstruction of whole-genome phylogenies from short-sequence reads, Mol. Biol. Evol. 31 (2014) 1077–1088, https://doi.org/10.1093/molbev/msu088.
[18] F. Ronquist, M. Teslenko, P. van der Mark, D.L. Ayres, A. Darling, S. Höhna, B. Larget, L. Liu, M.A. Suchard, H. P. Huelsenbeck, MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space, Syst. Biol. 61 (2012) 539–542, https://doi.org/10.1093/sysbio/sys029.
[19] R.K. Aziz, D. Bartels, A.A. Best, M. DeJongh, T. Disz, R. A. Edwards, K. Formsma, S. Gerdes, E.M. Glass, M. Kubal, M. Folker, J.O. Gary, O. Robert, L.O. Andrei, A.O. Ross, K. M. Leslie, P. Daniel, P. Tobias, P. Bruce, D.P. Gordon, R. Claudia, S. Rick, V. Olga, V. Veronika, W. Andreas, Z. Olga, The RAST server: rapid annotations using subsystems technology, BMC Genomics 9 (2008) 1–15, https://doi.org/10.1186/1471-2164-9-75.
[20] R. Overbeek, R. Olson, G.D. Pusch, G.J. Olsen, J.J. Davis, T. Disz, R.A. Edwards, S. Gerdes, B. Parrello, M. Shukla, V. Veronika, R.W. Alice, X. Fangfang, S. Rick Stevens, The SEED and the Rapid annotation of microbial genomes using subsystems technology (RAST), Nucleic Acids Res. 42 (2014) D206–D214, https://doi.org/10.1093/nar/gkt1226.
[21] P. Ferrer, Revisiting the *Cellulosimicrobium cellulans* yeast-lytic β-1, 3-glucanases toolbox: a review, Microb. Cell Fact. 5 (2006) (2006) 10, https://doi.org/10.1186/1475-2859-5-10.
[22] Y. Tanabe, M. Oda, Molecular characterization of endo-1, 3-β-glucanase from *Cellulosimicrobium cellulans*: effects of carbohydrate-binding module on enzymatic function and stability, Biochim. Biophys. Acta 1814 (2011) 1713–1719, https://doi.org/10.1016/j.bbapap.2011.09.004.
[23] A. Sharma, J.A. Gilbert, R. Lal, (Meta) genomic insights into the pathogenome of *Cellulosimicrobium cellulans*, Sci. Rep. 6 (2016) 25527, https://doi.org/10.1038/srep25527.
[24] E. Nabti, L. Bensidhoum, N. Tabli, D. Dahel, A. Weiss, M. Rothballer, M. Schmid, A. Hartmann, A, growth stimulation of barley and biocontrol effect on plant pathogenic fungi by a *Cellulosimicrobium* sp. strain isolated from salt-affected rhizosphere soil in northwestern Algeria, Eur. J. Soil Biol. 61 (2014) 20–26, https://doi.org/10.1016/j.ejsobi.2013.12.008.
[25] Y. Yuan, Y. Hu, C. Hu, J. Leng, H. Chen, X. Zhao, J. Gao, Y. Zhou, Overexpression and characterization of a glycoside hydrolase family 1 enzyme from *Cellulosimicrobium cellulans* sp. 21 and its application for minor ginsenosides production, J. Mol. Catal. B Enzym. 120 (2015) 60–67, https://doi.org/10.1016/j.molcatb.2015.06.015.
[26] S. Huang, A.H. Millar, Succinate dehydrogenase: the complex roles of a simple enzyme, Curr. Opin. Plant Biol. 16 (2013) 344–349, https://doi.org/10.1016/j.pbi.2013.02.007.
[27] P.C. Dorrestein, H. Zhai, F.W. McLafferty, T.P. Begley, The biosynthesis of the thiazole phosphate moiety of thiamin: the sulfur transfer mediated by the sulfur carrier protein ThiS, Chem. Biol. 11 (2004) 1373–1381, https://doi.org/10.1016/j.chembiol.2004.08.009.
[28] J. Hu, Z. Su, B. Dong, D. Wang, X. Liu, H. Meng, H. Zhou, Characterization of a Bacillus subtilis S-16 thiazole-synthesis-related gene thiS knockout and antimicrobial activity analysis, Curr. Issues Mol. Biol. 45 (2023) 4600–4611, https://doi.org/10.3390/cimb45060292.
[29] Y. Choi, J. Kim, H.J. Yoon, K.S. Jin, S. Ryu, H.H. Lee, Structural insights into the FtsQ/FtsB/FtsL complex, a key component of the divisome, Sci Rep 8 (2018) 18061, https://doi.org/10.1038/s41598-018-36001-2.
[30] N. Buddelmeijer, J. Beckwith, A complex of the Escherichia coli cell division proteins FtsL, FtsB and FtsQ forms independently of its localization to the septal region, Mol Microbiol 52 (2004) 1315–1327, https://doi.org/10.1111/j.1365-2958.2004.04044.x.

[31] S.Y. Park, S.H. Shim, Characterization of metabolites from cultures of *Cellulosimicrobium cellulans*, J. Korean Soc. Appl. Biol. Chem. 57 (2014) 481–484, https://doi.org/10.1007/s13765-014-4118-9.
[32] W. Qin, F. Fan, Y. Zhu, X. Huang, A. Ding, X. Liu, J. Dou, Anaerobic biodegradation of benzo (a) pyrene by a novel *Cellulosimicrobium cellulans* CWS2 isolated from polycyclic aromatic hydrocarbon-contaminated soil, Braz. J. Microbiol. 49 (2018) 258–268, https://doi.org/10.1016/j.bjm.2017.04.014.
[33] R. Antony, K.P. Krishnan, S. Thomas, W.P. Abraham, M. Thamban, Phenotypic and molecular identification of *Cellulosimicrobium cellulans* isolated from antarctic snow, Antonie Van Leeuwenhoek 96 (2009) 627–634, https://doi.org/10.1007/s10482-009-9377-9.
[34] Y. Fu, L. Cheng, Y. Meng, S. Li, X. Zhao, Y. Du, H. Yin, *Cellulosimicrobium cellulans* strain E4-5 enzymatic hydrolysis of curdlan for production of (1→ 3)-linked β-D-glucan oligosaccharides, Carbohydrate Polymers 134 (2015) 740–744, https://doi.org/10.1016/j.carbpol.2015.08.019.
[35] M. Verce, L. De Vuyst, S. Weckx, Shotgun metagenomics of a water kefir fermentation ecosystem reveals a novel *Oenococcus* species, Front. Microbiol. 10 (2019) 479, https://doi.org/10.3389/fmicb.2019.00479.
[36] J.M. Song, D.Z. Wei, Production and characterization of cellulases and xylanases of *Cellulosimicrobium cellulans* grown in pretreated and extracted bagasse and minimal nutrient medium M9, Biomass Bioenergy 34 (2010) 1930–1934, https://doi.org/10.1016/j.biombioe.2010.08.010.
[37] T.Y. Dou, J. Chen, Y.F. Hao, X. Qi, Effects of different carbon sources on enzyme production and ultrastructure of *Cellulosimicrobium cellulans*, Curr. Microbiol. 76 (2019) 355–360, https://doi.org/10.1007/s00284-019-01633-8.
[38] S. Huang, J. Pan, M. Tuwang, H. Li, C. Ma, M. Chen, X. Lin, Isolation, screening, and degradation characteristics of a Quinclorac-Degrading bacterium, strain D, and its potential for bioremediation of rice fields polluted by quinclorac, Microbiol. Spectr. 9 (2021) e00398-21, https://doi.org/10.1128/Spectrum.00398-21.
[39] A. Walia, P. Mehta, S. Guleria, C.K. Shirkot, Improvement for enhanced xylanase production by *Cellulosimicrobium cellulans* CKMX1 using central composite design of response surface methodology, 3 Biotech. 5 (2015) 1053–1066, https://doi.org/10.1007/s13205-015-0309-2.