Archives of Microbiology (2023) 205:278
https://doi.org/10.1007/s00203-023-03617-6

ORIGINAL PAPER

Draft genome sequence of *Hahella* sp. CR1 and its ability in producing cellulases for saccharifying agricultural biomass

Melvin Chun Yun Tan[1] · Muhammad Ramziuddin Zakaria[1] · Kok Jun Liew[1] · Chun Shiong Chong[1]

Received: 21 April 2023 / Revised: 20 June 2023 / Accepted: 23 June 2023 / Published online: 7 July 2023

Abstract

Hahella is a genus that has not been well-studied, with only two identified species. The potential of this genus to produce cellulases is yet to be fully explored. The present study isolated *Hahella* sp. CR1 from mangrove soil in Tanjung Piai National Park, Malaysia, and performed whole genome sequencing (WGS) using NovaSeq 6000. The final assembled genome consists of 62 contigs, 7,106,771 bp, a GC ratio of 53.5%, and encoded for 6,397 genes. The CR1 strain exhibited the highest similarity with *Hahella* sp. HN01 compared to other available genomes, where the ANI, dDDH, AAI, and POCP were 97.04%, 75.2%, 97.95%, and 91.0%, respectively. In addition, the CAZymes analysis identified 88 GTs, 54 GHs, 11 CEs, 7 AAs, 2 PLs, and 48 CBMs in the genome of strain CR1. Among these proteins, 11 are related to cellulose degradation. The cellulases produced from strain CR1 were characterized and demonstrated optimal activity at 60 ℃, pH 7.0, and 15% (w/v) sodium chloride. The enzyme was activated by K^+, Fe^{2+}, Mg^{2+}, Co^{2+}, and Tween 40. Furthermore, cellulases from strain CR1 improved the saccharification efficiency of a commercial cellulase blend on the tested agricultural wastes, including empty fruit bunch, coconut husk, and sugarcane bagasse. This study provides new insights into the cellulases produced by strain CR1 and their potential to be used in lignocellulosic biomass pre-treatment.

Keywords Halophile · Cellulolytic · Whole genome sequencing · Glycosyl hydrolase · Saccharification

Introduction

Lignocellulose is abundantly present in the environment, primarily from the living biomass in forest ecosystems and wastes from human activities. Approximately 1.3 billion tons of lignocellulosic biomass is generated annually from wastes of agricultural plantations and industries that produce commercial products, such as paper, fabric, food, and beverages (Mujtaba et al. 2023). Various approaches in separating lignocellulose components (cellulose, hemicelluloses, and lignin) have been investigated to convert by-products efficiently into valuable materials (Nanda et al. 2014; Okolie et al. 2021). One of the techniques used in this process is biological pre-treatment involving biological agents, particularly enzyme-producing bacteria. This method is advantageous due to the low energy consumption and operational cost compared to other physicochemical methods (Wu et al. 2022a, b).

Cellulases are key enzymes in cellulose degradation, a main component in the lignocellulosic biomass. This enzyme is categorized into endoglucanase, exoglucanase, and β-glucosidase, based on the catalytic mode of action in breaking down β-1,4-glycosidic linkages of cellulose. Cellulases function synergistically in attacking the polymeric cellulose at different levels of crystallinity and molecular bonds (Paul et al. 2021). For instance, endoglucanase cleaves cellulose chains randomly within the amorphous region, while exoglucanase cuts at the reducing and non-reducing ends of the polymeric chain and releases short-chain oligosaccharides and cellobioses (Houfani et al. 2020; Prawisut et al. 2020). Subsequently, β-glucosidases complete the hydrolysis by breaking down these short polymers into glucose (Aytaş et al. 2023). Additionally, these cellulases have distinct catalytic domains classified as glycoside hydrolases (GH) in the carbohydrate-active enzymes (CAZymes) database (Sidar et al. 2020). Endoglucanases and exoglucanases are primarily related to GH5, GH6, GH7, GH9, GH12, and GH45

Communicated by Wen-Jun Li.

✉ Chun Shiong Chong
cschong@utm.my

[1] Department of Biosciences, Faculty of Science, Universiti Teknologi Malaysia, 81310 Skudai, Johor, Malaysia

families, while β-glucosidases are associated with GH1, GH2, and GH3 families (Andlar et al. 2018; Carbonaro et al. 2023). Some cellulases also comprise carbohydrate-binding modules (CBM) that enhance enzymatic activity by targeting specific substrates (Sidar et al. 2020).

Cellulase-producing microorganisms, such as bacteria and fungi, can be isolated from various sources, such as soil (Naresh et al. 2019), compost (Li et al. 2023), insect guts (Dar et al. 2019; Xie et al. 2023), and animal rumens (Astuti et al. 2022). Despite advancements in cellulase studies, challenges remain in establishing stable and economic cellulase systems with a pre-treatment step that ensures high sugar yield before the bioconversion (Dar et al. 2019). Therefore, bioprospecting for potential cellulolytic bacterial strains from new sources remains relevant. Mangrove ecosystems, for instance, are a potential source for mining various novel enzymes that are functionally robust in extreme conditions (Behera et al. 2017; Mamangkey et al. 2021). Mangrove swamps in tropical and subtropical regions are coastal areas flooded with seawater in intervals (Liu et al. 2019). Mangrove soil contains a significant proportion of water-soluble sugars derived from lignocellulose components, comprising around half of its organic matter (Behera et al. 2017). Due to its salinity and nutrient richness, this environment provides an ideal habitat for a wide range of marine and freshwater microorganisms, including those possessing cellulase degradation properties (Palit et al. 2022).

The present study isolated *Hahella* sp. CR1 halophilic marine bacterium of the family *Hahellaceae* from the mangrove soil in Tanjung Piai National Park, Malaysia. Currently, six genome sequences belong to the genus *Hahella* available in the NCBI database, but only two species have been published (*Hahella chejuensis* KCTC 2396 ^{T} and *Hahella ganghwensis* DSM 17046 ^{T}) (Lee et al. 2001; Baik et al. 2005). Despite the availability of genomic data, the cellulases and other hydrolytic enzymes produced by these bacteria and their potential applications in lignocellulosic bioconversion remained understudied. Therefore, this study aimed to analyze the sequenced genome of *Hahella* sp. CR1 and investigate its potential to be developed into cellulase systems applied in lignocellulosic biomass pre-treatment.

Materials and methods

Bacterial isolation & identification

Mangrove soil was collected at the terrestrial site with a sampling depth of 15.0 cm during the low tide period (November 2020) in Tanjung Piai National Park, Johor, Malaysia (1°16′01.6″ N, 103°30′27.6″ E) (CJB permit no.: G No. 887005). The fresh mangrove soil sample was serially diluted immediately after collection, spread plated onto Difco Marine Broth 2216 agar (MA) (BD, USA) and incubated at 30 ℃ up to seven days. The strain CR1 single colony was streaked repeatedly to obtain a pure colony. Subsequently, the 16S ribosomal ribonucleic acid (rRNA) gene was amplified using the polymerase chain reaction (PCR) technique with 27F (5′-AGAGTTTGATCMTGGCTCAG-3′) and 1525R (5′-AAGGAGGTGWTCCARCC-3′) primers (Weisburg et al. 1991). The amplified gene was Sanger-sequenced and nucleotide BLAST against the National Center for Biotechnology Information (NCBI) 16S rRNA sequences database (https://blast.ncbi.nlm.nih.gov/Blast.cgi) to identify the bacterium.

Morphological and biochemical characterization

The morphological characteristics of strain CR1 colonies cultured on MA were observed after 48 h of incubation at 30 ℃, including the size, form, elevation, margin, surface, pigmentation, and optical characteristics (Holt et al. 1994). The strain CR1 sample was then subjected to Gram staining, as recommended by Tripathi et al. (2021), to identify the morphological features under a light microscope. Subsequently, the catalase and oxidase tests were performed on strain CR1 (Reiner 2010; Shields and Cathcart 2010). The hydrolysis test was conducted by streaking strain CR1 on bile esculin agar, and MA supplemented with 1% (w/v) of carboxy-methyl cellulose (CMC), casein, starch, xylan, locust bean gum (LBG), azure-B, and Tween (20, 40, 60, and 80). The agar plates were incubated at 30 ℃ for up to seven days, and the clearance zone around the colonies was recorded. The CMC-supplemented agar was stained with a 0.1% (w/v) Congo red solution and de-stained with a 1 M sodium chloride (NaCl) solution before observation. Agar plates supplemented with starch, xylan, and LBG were stained with Lugol's iodine solution.

Further detection and characterization of strain CR1 were performed using API® 20 E (bioMerieux, France). The API® 20 E strip containing 20 micro-tubes of dehydrated substrates was inoculated with the prepared bacterial suspension and incubated at 30 °C for two days. The tests include β-galactosidase, arginine di-hydrolase, lysine decarboxylase, ornithine decarboxylase, urease, tryptophan deaminase, tryptophanase, gelatinase, citrate utilization, hydrogen sulfide production, and Voges–Proskauer test (glucose, mannose, inositol, sorbitol, rhamnose, sucrose, melibiose, amygdalin, and arabinose fermentation). The color changes in the strip were observed after the incubation period (Holmes et al. 1978; Thaochan et al. 2010). Furthermore, the antibiotic resistance of strain CR1 was evaluated using the Kirby–Bauer disk diffusion method (Wayne 2015; Henciya et al. 2020) using selected antibiotics. First, the fresh bacterial culture of strain CR1 was incubated overnight in marine broth (MB). After incubation, the bacterial culture

(100 µL) was spread onto MA with a sterile glass rod. The prepared antibiotic disks were placed on the agar surface and incubated at 30 ℃ for three days. The inhibition zone was observed, and the diameter was recorded according to the Clinical and. Laboratory Standards (CLSI) susceptibility test interpretive criteria.

Genome extraction, assembly, and annotation

The genomic DNA of strain CR1 was extracted from the culture with 5×10^6 colony-forming units (CFU) using Quick-DNA™ Miniprep Kit (ZYMO Research, USA) and DNA Clean & Concentrator™-25 (ZYMO Research, USA), according to the manufacturer's protocol. The quality of the extracted genomic DNA was determined qualitatively using 1% (w/w) gel electrophoresis and quantitatively using NanoDrop™ 1000 Spectrophotometer (Thermo Scientific, USA). The sequencing library of the genomic DNA was prepared using Nextera™ DNA Sample Preparation Kits (Illumina, USA) and later sequenced via the Illumina NovaSeq 6000 platform (2 × 150 bp) (Illumina, USA).

The sequencing reads were processed according to Del Angel et al. (2018). First, adapter trimming and filtering of low-quality reads were performed using Trimmomatic version 0.39 to improve the quality of the raw reads with the following parameters: sliding window: 5 bp, minimal length: 50 bp and average quality: 25 (Bolger et al. 2014). The processed reads were then subjected to de novo assembly using SPAdes version 3.9.0 to reconstruct the genome based on the overlapping reads (Prjibelski et al. 2020). Subsequently, functional gene prediction was performed by annotating the assembled genome using the Prokaryotic Genome Annotation Pipeline (PGAP) version 6.4 (Tatusova et al. 2016). Comparative analysis was then performed by comparing the strain CR1 genome and the retrieved *Hahella* genomes in the NCBI database, in terms of their genome features and sequence identities. Different computational tools were employed for assessing nucleotide and protein sequence identity; Average Nucleotide Identity (ANI) was conducted using OrthoANI version 1.33 (Lee et al. 2016), Average Amino Acid Identity (AAI) was determined using EzAAI version 1.2.2 (Kim et al. 2021), and Digital DNA–DNA Hybridization (dDDH) between strain CR1 and other members was performed using GGDC 3.0 (Meier-Kolthoff et al. 2022). The percentage of conserved proteins (POCP) of strain CR1 and other genomes of the same genus was determined according to Qin et al. (2014).

Pangenome analysis was carried out using Pipeline for Bacterial Pan Genome Analysis (BPGA) version 1.3 in mapping core, accessory and unique genes of the *Hahella* members and strain CR1 (Chaudhari et al. 2016). A phylogenomic tree was constructed using the RealPhy web server (https://realphy.unibas.ch/realphy/) and refined using MEGA11 by the maximum likelihood method with 1000 bootstrap iterations. A cluster of Orthologous Groups of protein (COG) analysis was then performed to compare strain CR1 with other *Hahella* members using EggNOG-mapper version 2.1.9 (Cantalapiedra et al. 2021). The presence of CAZymes in the studied genomes was predicted using dbCAN2 version 3.0.7 against HMMER, DIAMOND and eCAMI (Zhang et al. 2018). Proteins belonging to GH families and associated with cellulases were identified from the annotated CAZymes in strain CR1. The selected proteins were screened using InterPro (https://www.ebi.ac.uk/interpro/search/sequence/) and BLAST against the Swiss-Prot database using DIAMOND version 2.1.0 (Buchfink et al. 2021). Default parameters were used for each software in the analysis unless specified.

Production of cellulase

A loop of strain CR1 was inoculated into 50 mL MB in a 250 mL conical flask and incubated in an incubator shaker overnight at 30 ℃ and 150 rpm. Subsequently, the cell turbidity of the overnight culture was measured in a spectrophotometer and adjusted to OD_{600} of 1.0. A 5.0% (v/v) inoculum [5.5×10^8 (CFU) of bacteria cells] was pipetted into 50 mL fresh MB supplemented with 0.2% (w/v) CMC to initiate the cellulase production. The fresh culture was incubated for another 18 h at 30 ℃ and 150 rpm. The bacterial culture was centrifuged at 10,000 g for 20 min at 4 ℃, and the supernatant was harvested. The supernatant containing cellulases was dialyzed in a dialysis tubing with 14.3/26 mm diameter and a pore size of 25 Å (BT Lab, China). The dialysis was performed in Tris–Hydrochloric acid (Tris–HCl) buffer (pH 7.2) at 4 ℃ for 18–20 h.

Enzyme activity for cellulases

The cellulase activity was determined qualitatively by measuring the reducing sugar released via the 3,5-dinitrosalicylic acid (DNS) assay. First, harvested cellulases (0.25 ml) were added into an equal volume of 1.0% (w/v) CMC and dissolved in 0.5 M Tris–HCl buffer solution (pH 7.2). The reaction mixture was then incubated at optimal temperature for 30 min. The reaction was stopped by adding an equal volume of DNS reagent and boiled for 5 min. The absorbance was determined using a spectrophotometer (DLAB Scientific, China) at 540 nm. The experiment was repeated with boiled cellulases as a negative control. One unit (U) of enzyme activity was expressed as the amount of enzyme that liberated reducing sugar equivalent to 1 ug of glucose per mL per minute under the assay conditions. All experiments were performed in triplicates. The results were analyzed using

one-way analysis of variance (ANOVA) and expressed as mean ± standard error (Dadheech et al. 2018).

Characterization of cellulases

Effects of temperature, pH and salinity

The cellulase activity was tested at different temperatures (5–80 ℃). The enzymatic assays were performed in different buffers (0.1 M) to determine the effect of pH on the harvested cellulase, including citrate phosphate buffer (pH 3.0–6.0), tris–HCl buffer (pH 7.0–8.0) and glycine–NaOH (pH 10.0–12.0). Meanwhile, the effect of NaCl concentration was determined by incubating the enzyme in different salinity [0.0% to 18.0% (w/v)]. The enzyme was characterized by an adaptive one-factor-at-a-time method (Seddouk et al. 2022).

Effects of metal ions and other chemicals

The effect of chemical additives on the harvested cellulases was determined by incubating the enzyme and substrate in 5.0 mM potassium chloride, iron (II) sulfate, copper (II) sulfate, magnesium chloride, cobalt (II) chloride, and 1.0% (w/v) Tween 40, Triton X-100, sodium dodecyl sulfate (SDS), and ethylene-diamine-tetra-acetic acid (EDTA). The assays were performed under optimized conditions.

Saccharification of cellulosic agricultural biomass

The enzymatic saccharification of agricultural biomass using cellulase from strain CR1 and cellulase enzyme blend, Cellic CTec2 (Sigma, USA) was performed according to Dadheech et al. (2018) with modifications. The crude enzyme and cellulase blend were diluted 2X and 10000X, respectively, to standardize the enzyme concentration to 0.1 U/mL. Oil palm empty fruit bunch (EFB), coconut husk (CH), and sugarcane bagasse (SB) were used as substrates for saccharification, which were ground and autoclaved before the experiment. The saccharification was performed in a 10 mL reaction system containing 1.0% (w/v) of biomass with three different enzymes setups: A. 2.5 mL cellulases from strain CR1 + 7.5 mL buffer, B. 2.5 mL commercial cellulase + 7.5 mL buffer, and C. 2.5 mL cellulases from strain CR1 + 2.5 mL commercial cellulase + 5.0 mL buffer. All mixtures were incubated for 24 h at 60 ℃, pH 7.0, with and without 15.0% (w/v) salinity. Finally, the mixtures were centrifuged at 4 ℃ and 10,000 g before the supernatant was harvested, and the sugar content was determined using DNS assay.

Results and discussion

Strain CR1 as member of genus *Hahella*

Based on the amplified 16S rRNA gene partial sequence, strain CR1 demonstrated the highest similarity to *Hahella chejuensis* KCTC 2396 ^{T} with an identity percentage of 99.78%. The genus *Hahella* was proposed in 2001 under the family *Hahellacaea* and order Oceanospirillales (Lee et al. 2001). To date, 15 strains have been deposited in the NCBI taxonomy database, comprising only two classified species and six genome data. Studies have been done on *Hahella chejuensis* to investigate its potential in secondary metabolite production. *Hahella chejuensis* NBU794 has reportedly produced new compounds of prodiginine derivatives, exhibiting anti-cancer and anti-microbial bioactivities (He et al. 2022; Li et al. 2022). Additionally, this species produces abundant exopolysaccharides and biosurfactants in bioremediation (Lee et al. 2001; Poli et al. 2010).

Biochemical characteristics of strain CR1

Strain CR1 was characterized using different biochemical tests and compared with the reference strains of genus *Hahella* (see Table 1). Strain CR1 exhibited circular, volcanic, orange–red pigment colonies resembling *Hahella chejuensis* KCTC 2396 ^{T}. Furthermore, strain CR1 is a Gram-negative, rod-shaped bacterium measuring 0.2–0.3 µm wide and 1.0–1.8 µm long, and demonstrated the ability to hydrolyze various complex molecules, including casein, starch, CMC, xylan, LBG, azure-B, and Tween(s) surfactants. The API® 20 E demonstrated that strain CR1 produced gelatinase but showed negative results in other tests, with notable distinctions from two reference strains. Moreover, strain CR1 was resistant to ampicillin, bacitracin, carbenicillin, lincomycin, minocycline, novobiocin, oleandomycin, oxacillin, penicillin G, and piperacillin and susceptible to gentamycin, neomycin, polymyxin B, rifampicin, and streptomycin.

General genome features

Table 2 details the genomic analysis of strain CR1. The assembled genome had a total length of 7,106,771 bp with 212.59 × coverage, 62 contigs, and a GC content of 53.50%. The largest contig produced was 1,241,827 bp with N_{50} of 449,172 bp. Precisely, the genome size of strain CR1 is similar to *Hahella* sp. HN01 (7,128,576 bp) compared to other existing genomes. The GC content of this bacterium also falls within the range of the other *Hahella* strains (49.00–54.00%). A total of 6,397 genes were annotated by PGAP for strain CR1, including 6293 protein-coding genes,

Table 1 Morphological and biochemical characteristics of strain CR1 and its reference strains of the genus *Hahella*

Description	CR1	KCTC 2396 ^{T}	KCTC 12277 ^{T}
Colony morphology			
Shape	Circular	ND	Circular
Elevation	Volcanic	Volcanic	Convex
Margin	Entire	ND	Entire
Surface/texture	Hard	ND	Smooth
Pigmentation	Orange-red	Pinkish red	Creamish color
Gram stain	Negative	Negative	Negative
Size (μm)	0.2–0.3 × 1.0–1.8	0.7–0.8 × 1.4–1.7	0.4–0.5 × 1.0–1.5
Production of			
Catalase	+	+	+
Oxidase	+	+	+
Hydrolysis of			
Bile esculin	–	+	+
Casein	+	ND	ND
Starch	+	ND	ND
Xylan	+	ND	ND
LBG	+	ND	ND
Azure B	–	ND	ND
Tween(s)	+	ND	ND
API® 20 E Test[a]			
Lysine decarboxylase	–	–	–
Citrate utilization	–	–	–
H_2S production	–	–	–
Tryptophan deaminase	–	–	–
Acetoin production	–	–	–
Gelatinase	+	+	+
Fermentation/oxidation:			
D-Glucose	–	+	+
D-Mannitol	–	+	–
Inositol	–	+	–
D-Sorbitol	–	+	–
D-Sucrose	–	+	–
D-Melibiose	–	–	–
Amygdalin	–	NR	–

CR1 *Hahella* sp. CR1, KCTC 2396 ^{T} *Hahella chejuensis* KCTC 2396 ^{T} (Lee et al. 2001), *KCTC 12277* ^{T} *Hahella ganghwensis* KCTC 12277 ^{T} (Baik et al. 2005). + positive result. – Negative result. *ND* not determined. [a]API® 20 E consists of 20 different biochemical tests, results that were not shown in the table were reported negative in all strains

42 pseudogenes, three rRNA, 55 tRNA, and four ncRNAs. In addition, the total number of genes from this bacterium was among the highest compared to the other four genomes, except for *Hahella* sp. HN01 (6528 genes) and *Hahella chejuensis* KCTC 2396 ^{T} (6402 genes).

The phylogenomic tree (see Fig. 1) indicated *Hahella chejuensis* KCTC 2396 ^{T}, *Hahella* sp. KA22 and *Hahella* sp. HN01 are grouped in a similar clade with strain CR1. Meanwhile, the pan-genome analysis (see Fig. 2) suggested that strain CR1 shared 1,751 core genes and 6,212 representative accessory genes with other genus members. Table 2 exhibited the ANI value between strain CR1 with *Hahella* sp. HN01 and strain KA22 are 97.04% and 92.83%, respectively. Other genomes showed < 90% identity values, including the two reference strains: *Hahella chejuensis* KCTC 2396 ^{T} (88.80%) and *Hahella ganghwensis* DSM 17046 ^{T} (69.30%). Furthermore, *Hahella* sp. HN01 (75.2%) exhibited the highest dDDH (identities/HSP length) values with strain CR1. The amino acid profile of strain CR1 is highly similar to *Hahella* sp. HN01; AAI = 97.95% and conserved protein percentage (POCP) = 91.0%. Collectively, strain CR1 and *Hahella* sp. HN01 were probably a novel species in genus *Hahella*.

Table 2 Comparison of basic genome features of strain CR1 with available genome data in genus *Hahella*

Strains	CR1	KCTC 2396 ^{T}	DSM 17046 ^{T}	KA22	CCB-MM4	HN01	PL25b_bin.18*
Sequencing platform	Illumina NovaSeq 6000	NA	Illumina HiSeq 2000	Illumina HiSeq; PacBio RSII	Illumina MiSeq	Illumina MiSeq	DNBSEQ
Assembly	SPAdes v. 3.9.0	NA	Velvet v. 1.1.04	MECAT v. V1.3	SPAdes v. 3.9.0	SPAdes v. 2.0	MegaHIT v. v1.2.9
Finishing strategy	High-quality draft	Complete Genome	High-Quality Draft	Complete Genome	High-Quality Draft	High-Quality Draft	High-Quality Draft
Genome coverage	212.59×	NA	NA	47.00×	132.00×	30.00×	8.34×
Genome quality	No contamination	No contamination	No contamination	No contamination	No contamination	No contamination	No contamination
Relevance	Environmental	Environmental	Environmental	Environmental	Environmental	Environmental	Environmental
Source	Mangrove sediment	Marine sediment	Tidal flat	Coastal water	Mangrove sediment	Mangrove sediment	Corals
Genome Features							
Genome size (bp)	7,106,771	7,215,267	6,564,965	6,960,771	6,663,740	7,128,576	6,676,736
GC content (%)	53.70	53.50	49.16	53.90	49.80	53.90	49.00
Number of contigs	62	1	144	1	161	134	151
Total genes	6397	6402	6232	6203	5936	6545	5876
Total CDS	6335	6316	–	6117	5866	6479	5825
Protein coding genes	6293	6259	6155	6068	5803	6298	–
RNA genes	62	86	77	86	70	66	47
rRNA genes	3	15	7	15	8	4	2
tRNA genes	55	67	58	67	58	58	44
ncRNA genes	4	4	12	4	4	4	1
Pseudogenes	42	57	–	49	63	181	–
Identity against CR1							
OrthoANI (%)	–	88.80	69.36	92.83	69.39	97.04	67.80
EzAAI (%)	–	93.46	67.31	95.97	67.57	97.95	64.49
dDDH (%)	–	36.50	18.80	49.80	18.90	75.20	17.10
POCP (%)	–	87.03	58.30	89.19	59.16	91.00	43.96
Accession number							
Genbank	JAPELK000000000	GCA_000012985.1	AQXX00000000	GCA_004135205.1	MRYI00000000	JAHMIN000000000	JANQMN000000000
BioProject	PRJNA896735	PRJNA16064	PRJNA182405	PRJNA512371	PRJNA356491	PRJNA736840	PRJNA857095
Biosample	SAMN31564167	SAMN02603483	SAMN02440417	SAMN10790512	SAMN06114548	SAMN19666839	SAMN29486655
Locus Tag	ONV78	HCH	F566	EUZ85	BTA51	KP814	NA
Reference	This study	Lee et al. (2001)	Baik et al. (2005)	Feng et al. (2019)	Sam et al. (2017)	Cri et al. (2019)	Tandon et al. (2022)

CR1 *Hahella* sp. CR1, *KCTC 2396 ^{T} Hahella chejuensis* KCTC 2396 ^{T}, *DSM 17046 ^{T} Hahella ganghwensis* DSM 17046 ^{T}, *KA22 Hahella* sp. KA22, *CCB-MM4 Hahella* sp. CCB-MM4, *HN01 Hahella* sp. HN01, *PL25b_bin.18 Hahella* sp. PL25b_bin.18 (MAG). *Annotation was done using PROKKA version 1.14.5. *NA* Not Available

COGs and CAZymes analysis

A total of 4,974 protein-coding genes in strain CR1 were successfully annotated in different COGs (see Fig. 3A). Majority of the annotated proteins (1108, 22.28%) were classified as "Function Unknown" due to the lack of similarities with the COG database, indicating that part of the strain CR1 proteins remains unexplored. Meanwhile, 206 genes (4.14%) were annotated with functions related to carbohydrate transport and metabolism, highlighting its potential for carbohydrate utilization. Genes encoding for carbohydrate-related proteins were further analyzed via annotation against the CAZymes database (see Fig. 3B), revealing a higher number of CAZymes in strain CR1 compared to other genomes. There are 88 glycosyl transferases (GT), 54 glycosyl hydrolase (GH), 11 carbohydrate esterases (CE), seven auxiliary activity enzymes (AA), two polysaccharide lyases (PL), and 48 carbohydrate-binding modules (CBM) annotated in strain CR1. GHs, responsible for degrading different carbohydrates, were found in all *Hahella* strains, with at least 25% of the total CAZymes count. AAs proteins, such as lytic polysaccharide monooxygenases (LPMOs), were also annotated

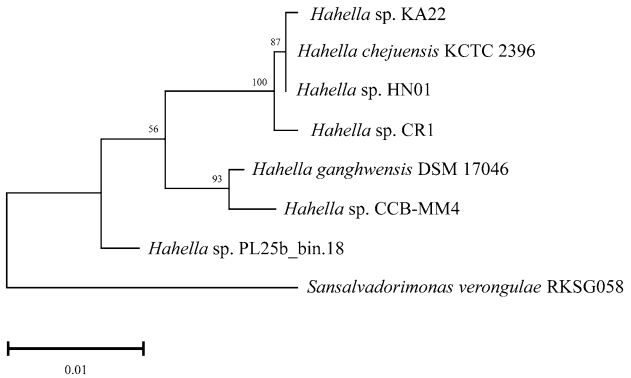

Fig. 1 Phylogenomic tree of *Hahella* members with *Sansalvadorimonas verongulae* RKSG058 as the outgroup. The tree was undergone bootstrap test (1000 replicates) with the percentage of replicate (bootstrap value) showed at each node. The tree scale (0.01) represents evolutionary distances in units of base substitutions per site as computed by Kimura-2 parameter method

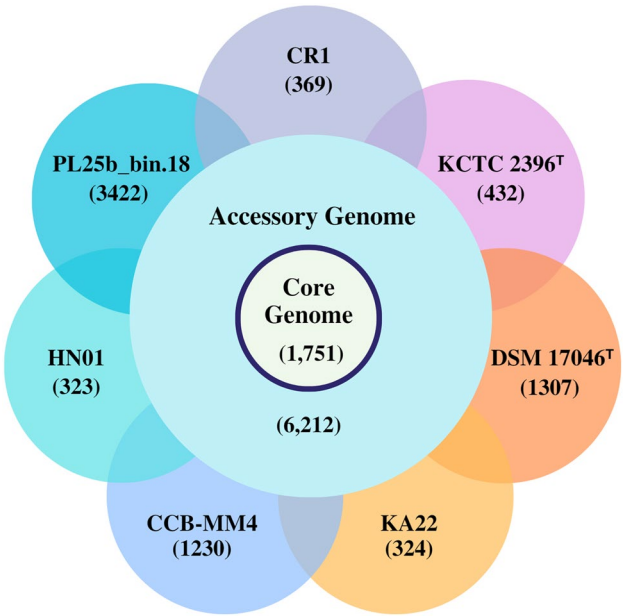

Fig. 2 Pangenome analysis of strain CR1 with other *Hahella* genomes, including *Hahella chejuensis* KCTC 2396 [T], *Hahella ganghwensis* DSM 17046 [T], *Hahella* sp. KA22, *Hahella* sp. CCB-MM4, *Hahella* sp. HN01, *Hahella* sp. PL25b_bin.18

in strain CR1 and the other *Hahella* genomes. In addition, AA proteins, such as lytic polysaccharide monooxygenases (LPMOs), were also annotated in strain CR1 and the other *Hahella* genomes. These enzymes are gaining attention from researchers due to the ability to cleave glycosidic bonds in cellulose and xylan through the oxidation reaction (Sun et al. 2022). Notably, strain CR1 and the genus *Hahella* are enzyme producers worth exploring due to their ability to produce various enzyme groups.

Mining of potential cellulase proteins

Annotated GHs were analyzed to identify potential cellulase-related proteins, and 11 GHs were related to cellulose degradation (see Table 3). The selected proteins were > 94.0% similar to the annotated proteins from *Hahella* sp. HN01, which remains understudied. Furthermore, the identity percentage toward selected genes ranged from 51.15% to 67.58% for other genera. Three main types of cellulases associated with endoglucanase function were found to be the most abundant, followed by exoglucanase and β-glucosidase.

Seven endoglucanase-encoding proteins associated with GH5, 8, and 9 (ONV78_14920, ONV78_25055, ONV78_18065, ONV78_01600, ONV78_03115, ONV78_01610, ONV78_05915) were identified in strain CR1. Exoglucanase-encoding proteins GH48 and GH6 (ONV78_05910, ONV78_01605) and two different families of β-glucosidases, GH1 and 3 (ONV78_02680, ONV78_02960) were also found in this bacterium. Moreover, several CAZymes contain CBMs (CBM2, 3, 6) responsible for cellulose binding (Sidar et al. 2020). The HcCel5 protein characterized from *Hahella chejuensis* KCTC 2396 [T] demonstrated a high association (95.15%) with protein ONV78_25055, which was also identified in strain CR1 as GH5 endoglucanase linked to two CBM6 modules (Ghatge et al. 2014). Genes from other GH families in the genus *Hahella* are still understudied. Eight of the selected genes possess signal peptides, indicating that these proteins are possibly expressed extracellularly in strain CR1. This information provides the fundamentals of extracellular cellulases produced by strain CR1 for further applications.

Characterization of cellulases produced from *Hahella* sp. CR1

The cellulases produced by strain CR1 were characterized quantitatively by elucidating the enzymatic activities at different temperatures, pH, and salinity. The optimum temperature for strain CR1 cellulase activity was 60 ℃ (see Table 4 and Fig. S1), similar to cellulase produced from the marine bacterium *Microbulbifer* sp. (Tanaka et al. 2021). Cellulase with optimal activity higher than 50 ℃ is favorable because temperatures lower than that potentially lead to slow reaction, lower yield of simple sugars and high susceptibility to microbial contamination (Patel et al. 2019). Nonetheless, there was a dramatic decline in the relative activity of strain CR1 cellulase to 22% at 80 ℃, but the cellulase remained active from 5 ℃ to 70 ℃ with a relative activity of > 50%. Likewise, purified HcCel5 endoglucanase from *H. chejuensis* exhibited optimal activity at 55 ℃ (Ghatge et al. 2014).

The cellulase activity of strain CR1 was optimal at pH 7, similar to cellulases from *Bacillus vallismortis* RG-07 (Gaur et al. 2015), *Bacillus amyoliquefaciens* (Lee et al. 2008), and

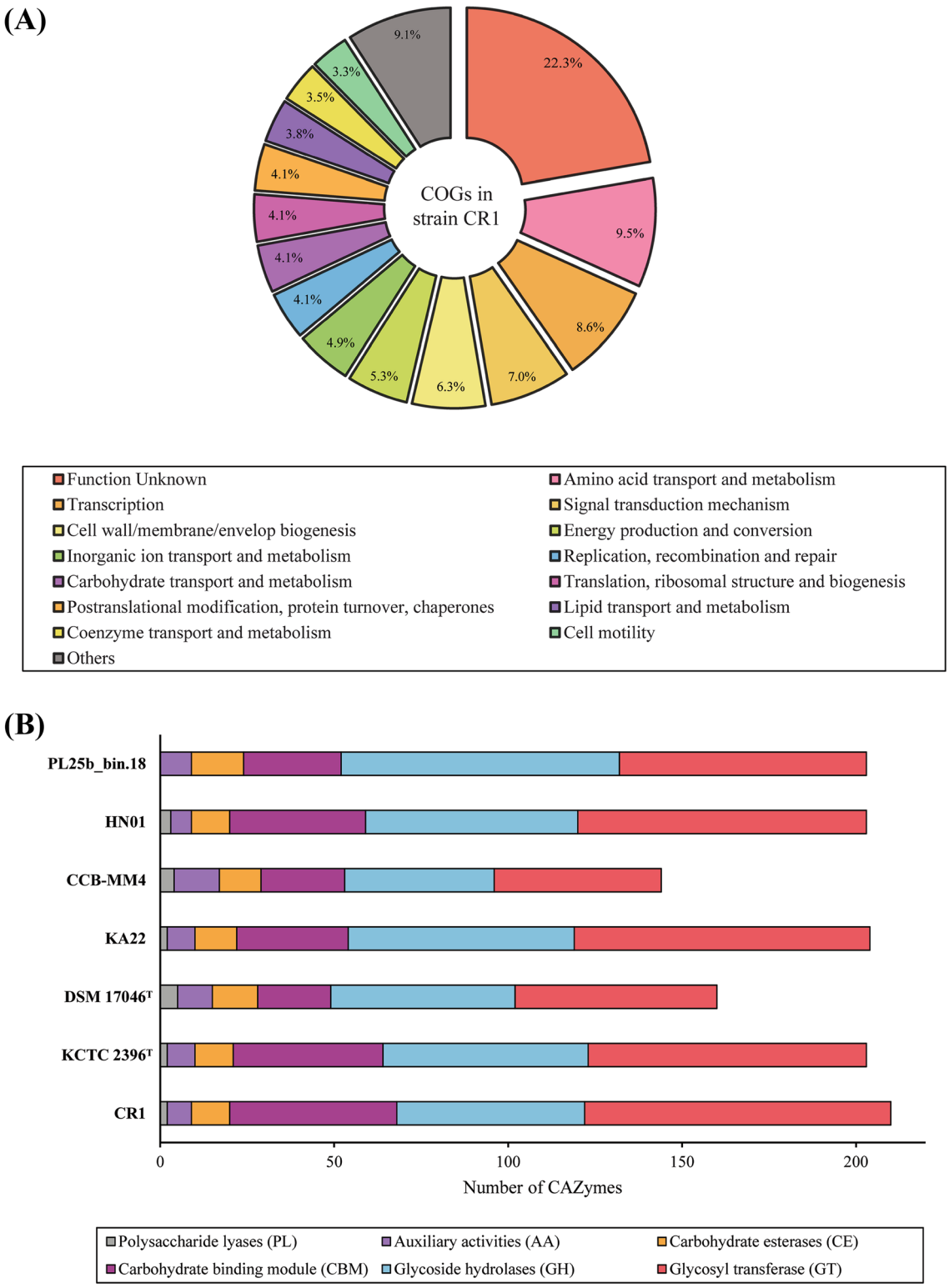

Fig. 3 Functional genes annotation in strain CR1 according to COG categories (**A**) and CAZymes families (**B**), including genomes of *Hahella chejuensis* KCTC 2396 ^{T}, *Hahella ganghwensis* DSM 17046 ^{T}, *Hahella* sp. KA22, *Hahella* sp. CCB-MM4, *Hahella* sp. HN01, *Hahella* sp. PL25b_bin.18

Table 3 List of potential cellulases from *Hahella* sp. CR1

Annotation	Locus tag accession number	Cazyme family	Carbohydrate binding module family	SignalP	Closest sequence to *Hahella* sp. HN01 (% identity)	Closest sequence to other species (% identity)
Endoglucanase	ONV78_14920	GH5		No	MBU6951461.1 99.35%	WP_181919329.1 [*Alkalilimnocola ehrlichii*] 51.23%
	ONV78_25055	GH5 subfamily 2	CBM 6 + CBM 6	Yes	MBU6953193.1 98.38%	WP_207862596.1 [*Acanthopleuribacter pedis*] 64.7%
	ONV78_18065	GH5 subfamily 48		Yes	MBU6955081.1 99.12%	WP_199775613.1 [*Microbulbifer pacificus*] 60.78%
	ONV78_01600	GH5 subfamily 8	CBM 2	No	MBU6950256.1 94.81%	TBR43115.1 [*Marinomonas agarivorans*] 53.01%
	ONV78_03115	GH8	CBM 2	Yes	MBU6950982.1 95.67%	MBI3898014.1 [*Gammaproteobacteria bacterium*] 65.63%
	ONV78_01610	GH9	CBM 2	Yes	MBU6950254.1 98.74%	WP_052830268.1 [G*yunella sunshinyii*] 59.74%
	ONV78_05915	GH9	CBM 3 + CBM 2	Yes	MBU6950838.1 98.40%	WP_230438080.1 [*Microbulbifer celer*] 60.04%
Exoglucanase	ONV78_05910	GH48	CBM 2	Yes	MBU6950839.1 99.05%	WP_251262880.1 [*Echinomonas agarilytica*] 51.15%
	ONV78_01605	GH6	CBM 2	Yes	MBU6950255.1 98.52%	WP_265689411.1 [*Vibrio* sp.] 54.93%
β-glucosidase	ONV78_02680	GH1		No	MBU6951664.1 99.77%	WP_183631697.1 [*Niveibacterium umoris*] 67.58%
	ONV78_02960	GH3		Yes	MBU6950951.1 99.53%	WP_189612581.1 [*Saccharospirillum salsuginis*] 60.33%

Pseudomonas fluorescens (Bakare et al. 2005). In addition, cellulases from strain CR1 adapted well in alkaline conditions with a relative activity of 45% at pH 12. In contrast, the cellulase activity decreased significantly at lower pH, with only 26% relative activity at pH 6. This finding contradicted the action of HcCel5 endoglucanase from *H. chejuensis* that functions well in acidic and < pH 7 environments at > 60% relative activity (Ghatge et al. 2014).

Cellulases from strain CR1 also demonstrated halotolerant characteristics. The cellulase relative activity was 60% at 0% (w/v) salt concentration, increased gradually at 2% to 15% (w/v) salinity and achieved maximum activity at 15% (w/v) salinity. The relative activity of cellulase also increased by 44% in 15% (w/v) NaCl, approximately two-fold than in 0% (w/v) NaCl. In a different study, the endoglucanase (BG-CS10) produced by halophilic *Bacillus* sp. isolated from Salt Lake exhibited optimal activity at 15% (w/v) NaCl (Zhang et al. 2012). Besides, this finding is comparable to the endoglucanase HcCel5 derived from *H. chejuensis*, with enzymatic inhibition up to 5 M NaCl and equivalent to 30% (w/v) salinity.

The effects of metal ions and other chemical additives on the cellulase of strain CR1 were also investigated in this study. The presence of metal ions, such as K^+, Fe^{2+}, Mg^{2+}, and Co^{2+}, significantly increased the cellulase activity of strain CR1, consistent with the performance of endoglucanase HcCel5 from *H. chejuensis* in the presence of 5 mM of Mg^{2+} and K^+ (Ghatge et al. 2014). However, earlier studies reported the inhibitory effect of Fe^{2+} on cellulase (Nazir et al. 2009; Bai et al. 2013; Bagewadi et al. 2015), contradicting the current study where metal ions significantly activated the cellulase activity of strain CR1 by 66%. Surfactant, Tween 40 was also observed in enhancing the cellulase activity of strain CR1. Previously, surfactants have been reported to enhance enzymatic activity by promoting substrate-binding on the active sites

Table 4 Physicochemical characteristics of crude cellulase produced from *Hahella* sp. CR1

Characteristics	Description
Enzyme incubation conditions*	
Temperature	5–70℃ (Optimal: 60 ℃)
pH	6.0–12.0 (Optimal: 7.0)
Salinity (%) (w/v)	0–15.0 (Optimal: 15.0)
Effect of metal ions (5 mM)	Relative Activity (%)
Control	100.0 ± 0
K^+	172.1 ± 15.6
Fe^{2+}	183.4 ± 16.5
Cu^{2+}	92.5 ± 6.6
Mg^{2+}	155.7 ± 7.8
Co^{2+}	177.7 ± 3.8
Effect of chemical additives	
EDTA	70.8 ± 6.2
Tween 40	165.0 ± 14.5
Triton X-100	119.9 ± 28.9
SDS	71.7 ± 17.2

(*) Refer to Fig. S1 showing the optimal cellulase activity against different conditions

(Hwang et al. 2008). In contrast, adding a chelating agent, such as EDTA, reduced the cellulase activity of strain CR1 by ~ 30%, consistent with the literature (Nazir et al. 2009; Vasconcellos et al. 2016; Okonkwo 2019). This observation is possibly caused by the chelation of metal ions and other enzyme cofactors, which are crucial for maintaining enzyme stability and catalytic activity (Nazir et al. 2009; Okonkwo 2019).

Saccharification of agricultural biomass

Three different agricultural biomasses (EFB, CH, and SB) were used as substrates for the saccharification by cellulases from strain CR1 and commercial cellulase blend (Cellic CTec2). Figure 4 illustrated that saccharification efficiency was the highest when strain CR1 and Cellic CTec2 cellulases worked synergistically. The reducing sugar increased more than two-fold in CH (0.11–0.35 mg/mL), SB (0.70–1.49 mg/mL), and EFB (0.23–0.62 mg/mL) in the mixture compared to the commercial cellulase blend setup. Additionally, the saccharification was significantly enhanced with salt, producing more reducing sugars from the biomass than in 0% (w/v) salt condition. Previously, crude cellulolytic enzymes from *Bacillus* sp. and *Exiguobacterium* sp. reportedly increased the reduced sugar content up to 94.7% when added to commercial cellulase in saccharifying corn stover (Wu et al. 2022a, b). Another study also showed that the addition of 5.0% (w/v) NaCl had increased the sugar released from sugarcane bagasse by 1.76 times (Gundupalli et al. 2021).

Significance and limitations of strain CR1 as a cellulase producer

This study has demonstrated that strain CR1 is a promising source of cellulase, with 11 identified proteins

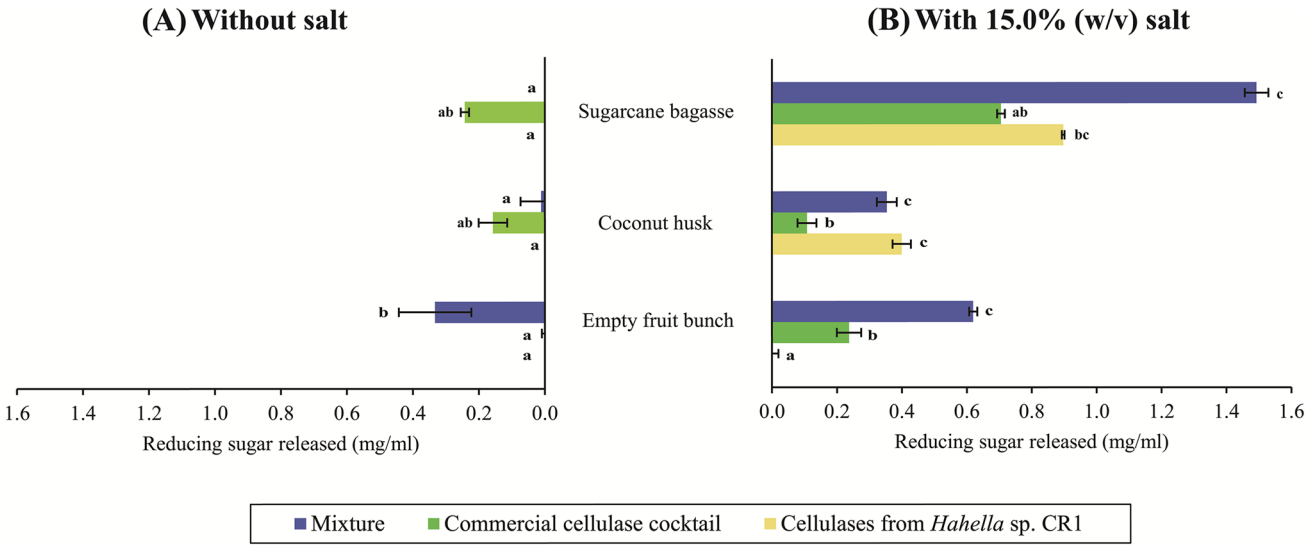

Fig. 4 Reducing sugars released in saccharification of various cellulosic agricultural substrate under similar optimal conditions with and without 15.0% (w/v) NaCl. Mean values ($n = 3$) are expressed and standard deviations is indicated as error bars. Superscript letters (**a**, **b**, **c**) above bars indicate significant difference of mean ($p < 0.05$), taken among the six enzyme set-ups in each type of substrate

associated with cellulose degradation through functional annotation of the sequenced genome. The enzyme mixture derived from strain CR1 exhibited remarkable cellulolytic activity, particularly at elevated temperatures (50 to 70 ℃) and salinity of up to 15.0% (w/v). These findings suggest the potential of this bacterium as an efficient cellulase system, particularly for combined pre-treatment of bioconversion involving ionic liquids or other physicochemical approaches at elevated temperatures (Wu et al. 2022a, b). Additionally, enzymatic saccharification performed at higher temperatures can enhance the yield of fermentable sugars (Akram et al. 2021). However, there are limitations to consider regarding the enzyme mixture investigated as the specific role of each predicted cellulolytic proteins from genome annotation could not be identified without further experimental studies. Therefore, future studies should focus on synthesizing the specific protein in the purified form via molecular cloning to characterize its physicochemical properties and applications.

Conclusion

This study provided insights into the halophilic bacterium *Hahella* sp. CR1 as a potential novel species for cellulase production that are stable in extreme conditions. A total of 11 cellulolytic proteins from different GH families were identified in the sequenced genome, eight of which were possibly extracellularly produced in the enzyme mixture of strain CR1. The enzyme mixture exhibited a high level of cellulolytic activity at 60 ℃ and 15% (w/v) salinity. Moreover, the enzyme mixture significantly improved the saccharification of agricultural biomass when working synergistically with a commercial cellulase blend. Future studies should focus on optimizing the working parameters and exploring the synergy between strain CR1 and the commercial cellulase blend for the saccharification of different agricultural wastes. The predicted cellulolytic proteins in strain CR1 should be further purified and characterized for potential scientific and industrial applications.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00203-023-03617-6.

Acknowledgements M.C.Y.T. would like to thank UTMNexus of Universiti Teknologi Malaysia for scholarship support. K.J.L. is a researcher under the Post-Doctoral Fellowship Scheme provided by Universiti Teknologi Malaysia.

Author contributions MCYT performed the genome analysis, carried out the enzyme experiments and wrote the manuscript. MRZ collected the sequencing data. KJL reviewed and edited the manuscript. CSC conceived the project. All authors reviewed and approved the final manuscript.

Funding This work was supported by the Ministry of Higher Education Malaysia under Fundamental Research Grant Scheme (FRGS/1/2019/WAB13/UTM/02/1) as main sponsor. This work was also supported by Professional Development Research University Grant (Q.J130000.21A2.06E00).

Data availability The 16S rRNA partial gene sequence obtained from Sanger sequencing can be retrieved in GenBank (accession number OQ607710). The raw reads from the Illumina sequencing were deposited in NCBI Sequence Read Archive (SRA) database with accession number SRR22726834. The genome assembly can be retrieved from the NCBI database under the accession number JAPELK000000000 (BioProject ID: PRJNA896735; BioSample ID: SAMN31564167).

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

Akram F, Ul-Haq I, Aqeel A, Ahmed Z, Shah FI (2021) Thermostable cellulases: Structure, catalytic mechanisms, directed evolution and industrial implementations. Renew Sustain Energy Rev 151:111597. https://doi.org/10.1016/j.rser.2021.111597

Andlar M, Rezić T, Marđetko N, Kracher D, Ludwig R, Šantek B (2018) Lignocellulose degradation: An overview of fungi and fungal enzymes involved in lignocellulose degradation. Eng Life Sci 18(11):768–778. https://doi.org/10.1002/elsc.201800039

Astuti T, Akbar SA, Rofiq MN, Jamarun N, Huda N, Fudholi A (2022) Activity of cellulase and ligninase enzymes in a local bioactivator from cattle and buffalo rumen contents. Biocat Agricult Biotechnol 45:102497. https://doi.org/10.1016/j.bcab.2022.102497

Aytaş ZG, Tunçer M, Kul ÇS, Cilmeli S, Aydın N, Doruk T, Adıgüzel AO (2023) Partial characterization of β-glucosidase, β-xylosidase, and α-l-arabinofuranosidase from *Jiangella alba* DSM 45237 and their potential in lignocellulose-based biorefining. Sustain Chem Pharmacy 31:100900. https://doi.org/10.1016/j.scp.2022.100900

Bagewadi Z, Ninnekar H (2015) Production, purification and characterization of endoglucanase from *Aspergillus fumigatus* and enzymatic hydrolysis of lignocellulosic waste. Int J Biotechnol Biomed Sci 1:25–32. https://doi.org/10.1007/s12010-013-0227-x

Bai H, Wang H, Sun J, Irfan M, Han M, Huang Y, Han X, Yang Q (2013) Purification and characterization of beta 1, 4-glucanases from *Penicillium simplicissimum* H-11. BioResources 8(3):3657–3671. https://doi.org/10.15376/biores.8.3.3657-3671

Baik KS, Seong CN, Kim EM, Yi H, Bae KS, Chun J (2005) Hahella ganghwensis sp. nov., isolated from tidal flat sediment. Int J Syst Evol Microbiol 55(2):681–684. https://doi.org/10.1099/ijs.0.63411-0

Bakare M, Adewale I, Ajayi A and Shonukan O (2005) Purification and characterization of cellulase from the wild-type and two improved mutants of *Pseudomonas fluorescens*. Afr J Biotechnol 4(9):898–904. https://doi.org/10.5897/AJB2005.000-3178

Behera BC, Sethi B, Mishra R, Dutta S, Thatoi H (2017) Microbial cellulases–diversity & biotechnology with reference to mangrove environment: a review. J Genetic Eng Biotechnol 15(1):197–210. https://doi.org/10.1016/j.jgeb.2016.12.001

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30(15):2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Buchfink B, Reuter K, Drost H-G (2021) Sensitive protein alignments at tree-of-life scale using DIAMOND. Nat Methods 18(4):366–368. https://doi.org/10.1038/s41592-021-01101-x

Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J (2021) eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. Mol Biol Evol 38(12):5825–5829. https://doi.org/10.1093/molbev/msab293

Carbonaro M, Aulitto M, Gallo G, Contursi P, Limauro D, Fiorentino G (2023) Insight into CAZymes of *Alicyclobacillus mali* FL18: characterization of a New Multifunctional GH9 Enzyme. Int J Mol Sci 24(1):243. https://doi.org/10.3390/ijms24010243

Chaudhari NM, Gupta VK, Dutta C (2016) BPGA-an ultra-fast pan-genome analysis pipeline. Sci Rep 6(1):24373. https://doi.org/10.1038/srep24373

Cri P, Gallantiswara D, Dong H, Kayansamruaj P, Senapin S, Rodkhum C (2019) Genome characterization on the virulence determinants of *Hahella chejuensis* HN01 causing red eggs disease in Tilapia hatcheries.

Dadheech T, Shah R, Pandit R, Hinsu A, Chauhan PS, Jakhesara S, Kunjadiya A, Rank D, Joshi C (2018) Cloning, molecular modeling and characterization of acidic cellulase from buffalo rumen and its applicability in saccharification of lignocellulosic biomass. Int J Biol Macromol 113:73–81. https://doi.org/10.1016/j.ijbiomac.2018.02.100

Dar MA, Pawar KD, Rajput BP, Rahi P, Pandit RS (2019) Purification of a cellulase from cellulolytic gut bacterium, Bacillus tequilensis G9 and its evaluation for valorization of agro-wastes into added value byproducts. Biocatal Agric Biotechnol 20:101219

Del Angel VD, Hjerde E, Sterck L, Capella-Gutierrez S, Notredame C, Pettersson OV, Amselem J, Bouri L, Bocs S, Klopp C, Gibrat JF (2018) Ten steps to get started in genome assembly and annotation. F1000Research 2018, 7(ELIXIR):148. https://doi.org/10.12688/f1000research.13598.1

Feng J, Hu Z, Wang H (2019) Complete genome sequence of Hahella sp KA22, a prodigiosin-producing algicidal bacterium. Mar Genomics 47:100678. https://doi.org/10.1016/j.margen.2019.04.003

Gaur R, Tiwari S (2015) Isolation, production, purification and characterization of an organic-solvent-thermostable alkalophilic cellulase from *Bacillus vallismortis* RG-07. BMC Biotechnol 15(1):1–12. https://doi.org/10.1186/s12896-015-0129-9

Ghatge SS, Telke AA, Kang S-H, Arulalapperumal V, Lee K-W, Govindwar SP, Um Y, Oh D-B, Shin H-D, Kim S-W (2014) Characterization of modular bifunctional processive endoglucanase Cel5 from *Hahella chejuensis* KCTC 2396. Appl Microbiol Biotechnol 98(10):4421–4435. https://doi.org/10.1007/s00253-013-5446-0

Gundupalli, M., P. Tantayotai, K. Rattanaporn, W. Pongprayoon, T. Phusantisampan and M. Sriariyanun (2021). Effects of inorganic salts on enzymatic saccharification kinetics of lignocellulosic biomass for biofuel Production. In: 2021 The 10th International Conference on Informatics, Environment, Energy and Applications.

He S, Li P, Wang J, Zhang Y, Lu H, Shi L, Huang T, Zhang W, Ding L, He S (2022) Discovery of new secondary metabolites from marine bacteria *Hahella* based on an omics strategy. Mar Drugs 20(4):269. https://doi.org/10.3390/md20040269

Henciya S, Vengateshwaran TD, Gokul MS, Dahms H-U, James RA (2020) Antibacterial activity of halophilic bacteria against drug-resistant microbes associated with diabetic foot infections. Curr Microbiol 77(11):3711–3723

Holmes B, Willcox W, Lapage S (1978) Identification of Enterobacteriaceae by the API 20E system. J Clin Pathol 31(1):22–30

Holt J, Krieg N, Snealth P, Staley J (1994). Bergey's Manual of Determinative Bacteriology. Nocardioform Actinomycetes I

Houfani AA, Anders N, Spiess AC, Baldrian P, Benallaoua S (2020) Insights from enzymatic degradation of cellulose and hemicellulose to fermentable sugars–a review. Biomass Bioenergy 134:105481. https://doi.org/10.1016/j.biombioe.2020.105481

Hwang IH, Lee CH, Kim SW, Sung HG, Lee SY, Lee SS, Hong HO, Kwak Y-C, Ha JK (2008) Effects of mixtures of Tween80 and cellulolytic enzymes on nutrient digestion and cellulolytic bacterial adhesion. Asian Australas J Anim Sci 21(11):1604–1609. https://doi.org/10.5713/ajas.2008.80333

Kim D, Park S, Chun J (2021) Introducing EzAAI: a pipeline for high throughput calculations of prokaryotic average amino acid identity. J Microbiol 59:476–480. https://doi.org/10.1007/s12275-021-1154-0

Lee HK, Chun J, Moon EY, Ko S-H, Lee D-S, Lee HS, Bae KS (2001) Hahella chejuensis gen nov., sp. nov., an extracellular-polysaccharide-producing marine bacterium. Int J Syst Evolut Microbiol 51(2):661–666. https://doi.org/10.1099/00207713-51-2-661

Lee Y-J, Kim B-K, Lee B-H, Jo K-I, Lee N-K, Chung C-H, Lee Y-C, Lee J-W (2008) Purification and characterization of cellulase produced by *Bacillus amyoliquefaciens* DL-3 utilizing rice hull. Biores Technol 99(2):378–386. https://doi.org/10.1016/j.biortech.2006.12.013

Lee IC, Kim OY, Park SC, Chun JS (2016) OrthoANI: an improved algorithm and software for calculating average nucleotide identity. Int J Syst Evol Microbiol 66(2):1100–1103. https://doi.org/10.1099/ijsem.0.000760

Li P, He S, Zhang X, Gao Q, Liu Y, Liu L (2022) Structures, biosynthesis, and bioactivities of prodiginine natural products. Appl Microbiol Biotechnol. https://doi.org/10.1007/s00253-022-12245-x

Li X, Li K, Wang Y, Huang Y, Yang H, Zhu P, Li Q (2023) Diversity of lignocellulolytic functional genes and heterogeneity of thermophilic microbes during different wastes composting. Biores Technol 372:128697

Liu M, Huang H, Bao S, Tong Y (2019) Microbial community structure of soils in Bamenwan mangrove wetland. Sci Rep 9(1):1–11. https://doi.org/10.1038/s41598-019-44788-x

Mamangkey J, Suryanto D, Munir E, Mustopa AZ, Sibero MT, Mendes LW, Hartanto A, Taniwan S, Ek-Ramos MJ, Harahap A (2021) Isolation and enzyme bioprospection of bacteria associated to *Bruguiera cylindrica*, a mangrove plant of North Sumatra. Indonesia. Biotechnology Reports 30:e00617. https://doi.org/10.1016/j.btre.2021.e00617

Meier-Kolthoff JP, Carbasse JS, Peinado-Olarte RL, Göker M (2022) TYGS and LPSN: a database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. Nucleic Acids Res 50(D1):D801–D807. https://doi.org/10.1093/nar/gkab902

Mujtaba, M., L. Fraceto, M. Fazeli, S. Mukherjee, S. M. Savassa, G. A. de Medeiros, A. d. E. Santo Pereira, S. D. Mancini, J. Lipponen and F. Vilaplana (2023). Lignocellulosic biomass from agricultural waste to the circular economy: A review with focus on biofuels, biocomposites and bioplastics. Journal of Cleaner Production: 136815.

Nanda S, Mohammad J, Reddy SN, Kozinski JA, Dalai AK (2014) Pathways of lignocellulosic biomass conversion to renewable fuels. Biomass Convers Biorefinery 4:157–191

Naresh S, Kunasundari B, Gunny AAN, Teoh YP, Shuit SH, Ng QH, Hoo PY (2019) Isolation and partial characterisation of thermophilic cellulolytic bacteria from north Malaysian tropical mangrove soil. Tropical Life Sci Res 30(1):123

Nazir A, Soni R, Saini H, Manhas R, Chadha B (2009) Purification and characterization of an endoglucanase from *Aspergillus terreus* highly active against barley β-glucan and xyloglucan. World J Microbiol Biotechnol 25(7):1189–1197. https://doi.org/10.1007/s11274-009-0001-y

Okolie JA, Nanda S, Dalai AK, Kozinski JA (2021) Chemistry and specialty industrial applications of lignocellulosic biomass. Waste Biomass Valorizat 12:2145–2169

Okonkwo I (2019) Effect of metal ions and enzyme inhibitor on the activity of cellulase enzyme of *Aspergillus flavus*. Int J Environ, Agricult Biotechnol. https://doi.org/10.22161/ijeab/4.3.20

Palit K, Rath S, Chatterjee S, Das S (2022) Microbial diversity and ecological interactions of microorganisms in the mangrove ecosystem: Threats, vulnerability, and adaptations. Environ Sci Pollut Res. https://doi.org/10.1007/s11356-022-19048-7

Patel AK, Singhania RR, Sim SJ, Pandey A (2019) Thermostable cellulases: current status and perspectives. Biores Technol 279:385–392. https://doi.org/10.1016/j.biortech.2019.01.049

Paul M, Mohapatra S, Mohapatra PKD, Thatoi H (2021) Microbial cellulases–An update towards its surface chemistry, genetic engineering and recovery for its biotechnological potential. Bioresource Technol 340:125710. https://doi.org/10.1016/j.biortech.2021.125710

Poli A, Anzelmo G, Nicolaus B (2010) Bacterial exopolysaccharides from extreme marine habitats: production, characterization and biological activities. Mar Drugs 8(6):1779–1802

Prawisut A, Choknud S, Cairns JRK (2020) Expression of rice β-exoglucanase II (OsExoII) in Escherichia coli, purification, and characterization. Protein Exp Purif 175:105708. https://doi.org/10.1016/j.pep.2020.105708

Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A (2020) Using SPAdes de novo assembler. Current Protocols Bioinformat 70(1):e102. https://doi.org/10.1002/cpbi.102

Qin Q-L, Xie B-B, Zhang X-Y, Chen X-L, Zhou B-C, Zhou J, Oren A, Zhang Y-Z (2014) A proposed genus boundary for the prokaryotes based on genomic insights. J Bacteriol 196(12):2210–2215. https://doi.org/10.1128/jb.01688-14

Reiner K (2010) Catalase test protocol. Am Soc Microbiol: 1–6

Sam K-K, Lau N-S, Furusawa G, Amirul A-AA (2017) Draft genome sequence of halophilic *Hahella* sp. strain CCB-MM4, isolated from Matang Mangrove Forest in Perak Malaysia. Genome Announc 5(42):e01147-e11117. https://doi.org/10.1128/genomea.01147-17

Seddouk L, Jamai L, Tazi K, Ettayebi M, Alaoui-Mhamdi M, Aleya L, Janati-Idrissi A (2022) Isolation and characterization of a mesophilic cellulolytic endophyte *Preussia africana* from *Juniperus oxycedrus*. Environ Sci Pollut Res 29(30):45589–45600. https://doi.org/10.1007/s11356-022-19151-9

Shields P, Cathcart L (2010) Oxidase test protocol. Am Soc Microbiol: 1–9

Sidar A, Albuquerque ED, Voshol GP, Ram AF, Vijgenboom E, Punt PJ (2020) Carbohydrate binding modules: diversity of domain architecture in amylases and cellulases from filamentous microorganisms. Front Bioeng Biotechnol 8:871. https://doi.org/10.3389/fbioe.2020.00871

Sun C, Meng X, Sun F, Zhang J, Tu M, Chang J-S, Reungsang A, Xia A, Ragauskas AJ (2022) Advances and perspectives on mass transfer and enzymatic hydrolysis in the enzyme-mediated lignocellulosic biorefinery: a review. Biotechnol Adv. https://doi.org/10.1016/j.biotechadv.2022.108059

Tanaka D, Ohnishi K-I, Watanabe S, Suzuki S (2021) Isolation of cellulase-producing Microbulbifer sp. from marine teleost blackfish (*Girella melanichthys*) intestine and the enzyme characterization. J General Appl Microbiol 67(2):47–53. https://doi.org/10.2323/jgam.2020.05.001

Tandon K, Ricci F, Costa J, Medina M, Kühl M, Blackall LL, Verbruggen H (2022) Genomic view of the diversity and functional role of archaea and bacteria in the skeleton of the reef-building corals *Porites lutea* and *Isopora palifera*. GigaSci. https://doi.org/10.1093/gigascience/giac127

Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J (2016) NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res 44(14):6614–6624. https://doi.org/10.1093/nar/gkw569

Thaochan N, Drew R, Hughes J, Vijaysegaran S, Chinajariyawong A (2010) Alimentary tract bacteria isolated and identified with API-20E and molecular cloning techniques from Australian tropical fruit flies, *Bactrocera cacuminata* and *B. tryoni*. J Insect Sci 10(1):131. https://doi.org/10.1673/031.010.13101

Tripathi N, Sapra A (2021) Gram Staining. StatPearls, Treasure Island (FL)

Vasconcellos V, Tardioli P, Giordano R, Farinas C (2016) Addition of metal ions to a (hemi) cellulolytic enzymatic cocktail produced in-house improves its activity, thermostability, and efficiency in the saccharification of pretreated sugarcane bagasse. New Biotechnol 33(3):331–337. https://doi.org/10.1016/j.nbt.2015.12.002

Wayne, P. (2015). CLSI. Performance standards for antimicrobial susceptibility testing; twenty-fifth informational supplement: CLSI Document M100-S25, Clinical and Laboratory Standards Institute; 2015, Clinical and Laboratory Standards Institute Wayne, PA.

Weisburg WG, Barns SM, Pelletier DA, Lane DJ (1991) 16S ribosomal DNA amplification for phylogenetic study. J Bacteriol 173(2):697–703

Wu Y, Shresth S, Guo H, Zhang J, Wang H, Qin W (2022a) Enhancement of saccharification of corn stover by cellulolytic enzyme produced from biomass-degrading bacteria. BioResources. https://doi.org/10.15376/biores.17.1.1301-1318

Wu Z, Peng K, Zhang Y, Wang M, Yong C, Chen L, Qu P, Huang H, Sun E, Pan M (2022b) Lignocellulose dissociation with biological pretreatment towards the biochemical platform: a review. Materials Today Bio 28:100445

Xie R, Dong C, Wang S, Danso B, Dar MA, Pandit RS, Pawar KD, Geng A, Zhu D, Li X (2023) Host-specific diversity of culturable bacteria in the gut systems of fungus-growing termites and their potential functions towards lignocellulose bioconversion. InSects 14(4):403. https://doi.org/10.3390/insects14040403

Zhang G, Li S, Xue Y, Mao L, Ma Y (2012) Effects of salts on activity of halophilic cellulase with glucomannanase activity isolated from alkaliphilic and halophilic *Bacillus* sp. BG-CS10. Extremophiles 16(1):35–43. https://doi.org/10.1007/s00792-011-0403-2

Zhang H, Yohe T, Huang L, Entwistle S, Wu P, Yang Z, Busk PK, Xu Y, Yin Y (2018) dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. Nucleic Acids Res 46(W1):W95–W101. https://doi.org/10.1093/nar/gky418