Screening, cloning, enzymatic properties of a novel thermostable cellulase enzyme, and its potential application on water hyacinth utilization

Xiaoshen Zhao¹,² · Liyang Liu³ · Zujun Deng¹,² · Shan Liu⁴ · Jeonyun Yun⁴ · Xiong Xiao⁴ · He Li¹,²

Received: 7 December 2020 / Revised: 2 February 2021 / Accepted: 2 March 2021 / Published online: 8 March 2021
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2021

Abstract
Cellulose is the cheapest, natural, renewable organic substance that is used as a carbon source in various fields. Water hyacinth, an aquatic plant rich in cellulose, is often used as a raw material in fuel production. However, natural cellulase can be hardly used in industrial production on account of its low thermal stability and activity. In this study, a metagenomic library was constructed. Then, a new cellulase gene, cel1029, was screened by Congo red staining and expressed in the prokaryotic system. Enzymatic properties of Cel1029 were explored, including optimum temperature and pH, thermal and pH stability, and tolerance against organic solvents, metal ions, and salt solutions. Finally, its ability of degrading water hyacinth was identified and evaluated. Cel1029 displayed high homology with endoglucanase in the glycoside hydrolase family 5 (GH5) and had high stability across a broad temperature range. More than 86% of its enzymatic activities were retained between 4 and 60 °C after 24 h of incubation. Single-factor analysis and orthogonal design were further conducted to determine the optimal conditions for the highest reducing sugar yield of water hyacinth. Interestingly, Cel1029 efficiently transformed water hyacinth with a reducing sugar yield of 430.39 mg/g in 22 h. These findings may open the door for significant industrial applications of a novel GH5 cellulase (NCBI Reference Sequence: MK051001, Cel1029) and help identify more efficient methods to degrade cellulose-rich plants.

Keywords Biotransformation · Cellulase · Metagenomic library · Thermostability · Water hyacinth

Introduction
Cellulase, including endoglucanase (endo-1,4-D-glucanase, EC 3.2.1.4), exoglucanase (exo-1,4-D-glucanase, EC 3.2.1.91), and glucosidase (1,4-D-glucosidase, EC 3.2.1.21), is a type of enzymes that can hydrolyze the β-1,4-glycosidic bond between cellulose molecules (Lin et al. 2016). Endocellulase mainly catalyzes degradation of the amorphous region in cellulose, providing the starting site for further catalysis of exocellulase; so, it plays an important role in efficient degradation of cellulose. Cellulase is distributed in many families, and the endoglucanases belonging to the glycoside hydrolase family 5 (GH5) have high similarity in both amino acid sequence and two strict conserved catalytic residues of glutamic acid (Yuan et al. 2019; Matsuyama et al. 1999). The two residues are considered catalytic proton donors and active site nucleophiles, respectively. They can maintain the structure of the substrate isocarbon of the enzyme (Valérie et al. 1995). Cellulase is widely used in many fields, such as biofuel production, papemaking, textiles, food processing, brewing, and extraction of active ingredients of traditional Chinese medicine (Garg et al. 2016; Chia et al. 2016; Samkelo et al. 2020). Endocellulase has attracted attention because of its catalytic efficiency. So far, although various cellulase has been isolated from bacteria, fungi (Okada 1976), plants, and higher animals, thermostable cellulase with excellent properties still needs to be discovered (Morteza et al. 2020a, b; Zheng et al. 2018).

Ferula asafoetida is a rare medicinal plant resource distributed in Xinjiang in China. The distribution area of Ferula...
asafoetida in Shihezi on the southern edge of Junggar Basin experiences much dramatic temperature changes (e.g., long time sunshine, large difference of daily temperature) and high content of salt-alkali. Therefore, microbes and enzymes in the soils of this area are characteristic of salt-alkali resistance and high thermal stability. Hence, the soils here are a good sample for screening new enzymes with industrial application potential. However, most of the microorganisms in the environment cannot be cultured with the existing techniques or be used for biotechnology or basic research (Amann et al. 1995; Schmitz 2004). For example, only 1% or even less of prokaryotes in the soils are readily cultivatable (Griffiths et al. 1996). Metagenomic libraries containing DNA extracted directly from environmental samples provide genomic sequences, and phylogenetic and functional information (Jo et al. 2002). Thus, metagenomics, the genomic analysis of collective genomes in an assemblage of organisms, is suitable for screening new functional biocatalysts and molecules with industrial application potential from soils (Handsman 2004). Many new enzymes have been discovered by using metagenomics, such as lipases (Tirawongsaroj et al. 2008; Bayer et al. 2009), amylases (Yun et al. 2004), and cellulase (Nimchua et al. 2012; Ko et al. 2013; Amitha et al. 2013). To our knowledge, no study has reported any highly thermostable cellulase that is found through metagenomic screening from the soils of Ferula asafoetida distribution area.

Burning fossil fuels can cause environmental pollution (Bao et al. 2011). Cellulose is the most abundant component of lignocellulose in the biosphere, and the cheapest renewable and natural organic substance (Liu et al. 2011). Cellulosic ethanol is the best alternative to fossil fuels (Raj and Krishnan 2018; Costa et al. 2018). However, many intermolecular, intramolecular hydrogen bonds and other intermolecular interactions result in the long-lasting chain conformation and tight chain filling of cellulose, which cannot be dissolved by ordinary solvents (Medronho et al. 2012). Moreover, the hydrolysis of cellulose by chemical methods to produce cellulosic ethanol is complicated and expensive (Salinas et al. 2011). The utilization of lignocellulosic biomass lacks effective low-cost means (Lynd et al. 2002).

Water hyacinth (Eichhornia crassipes) has abundant utilisable lignocellulosic biomass (Ismail et al. 1995). It can tolerate seasonal changes in flow rate, water level, pH, nutrient availability, temperature, and toxic substances (Asrofi et al. 2018; Sumrith and Dangtungee 2019). Water hyacinth can multiply quickly at suitable temperature (between 10 and 35 °C) in a nutrition balanced environment (Mayo and Hanai 2017). The infestation of water hyacinth affects water transport and ecological balance and causes secondary pollution of water bodies. These limitations are great challenges to the ecosystem. To solve these problems and make reasonable use of water hyacinth resources, some researchers have studied enzymatic degradation of water hyacinth. The reactions of 170 U/g cellulase from Trichoderma reesei with alkali-treated water hyacinth for 36 h (Ganguly et al. 2013), 30 FPU/g cellulase from Aspergillus fumigatus with acid-treated water hyacinth for 22 h (Das et al. 2013, 2016), and 10 FPU/g cellulase from Trichoderma atroviride with acid-treated water hyacinth for 70 h (Rajesh et al. 2014) lead to the maximum sugar yields of 299.13, 425.6, and 380.97 mg/g, respectively. In this work, after 121.25 U/g Cell1029 (the enzyme loading of 1 g of water hyacinth is 121.25 U) reacts with acid-treated water hyacinth for 22 h, the reducing sugar yield is 430.39 mg/g (sugar yield per gram of water hyacinth).

We metagenomically screened a novel type of cellulase through from the soils of Ferula asafoetida distribution area. The cellulase shows excellent thermal stability, tolerance against metal ions, organic solvents, and salt solutions, and high ability of degrading water hyacinth, which together prove its great potential for industrial application.

Materials and methods

Strains, materials, and chemicals

Escherichia coli BL21 (DE3) and E. coli DH5α (both from TSINGKE Biological Technology, Guangzhou, China) were used as the expression host and cloning host, respectively. The pET-32a (+) (Novagen) was used for protein expression. T4 DNA ligases, restriction endonucleases, and pUC118/BamHI (BAP) were purchased from TaKaRa (Dalian, China). Ampicillin (Amp), 5-bromo-4-chloro-3-indolyl-β-d-galactopyranoside (X-Gal), and isopropyl-β-d-thiogalactopyranoside (IPTG) were bought from Sigma. All other chemicals and reagents were produced by RUISHU Biological Technology (Guangzhou, China) unless otherwise stated. Water hyacinth was gathered locally. Molecular biology reagents from TaKaRa were used according to the manufacturer’s instructions.

Preparation of WHB

Water hyacinth was collected from the lake of Guangdong Pharmaceutical University. Then, the water hyacinth was chopped, dried at 105 °C, and finally ground into 0.1–1 mm powder (Amrani et al. 2016). A certain amount of WHB was added with dilute H2SO4 (2%) at the ratio of 1:10 (g/mL) under soaking for 1 h at 40 °C. Then, the dilute H2SO4 was drained, and WHB was washed to neutralize the pH with distilled water.
DNA extraction, metagenomic library construction, and cellulase gene screening

A soil sample was collected from the wild *Ferula asafoetida* distribution area in Shihezi on the southern edge of Junggar Basin in Xinjiang. The sample was taken from 5 to 10 cm under the surface, sealed in a sterile bag, and preserved at ~20 °C until DNA extraction. Total genomic DNA from the soil sample was extracted according to a reported method (Zhou et al. 1996). The amount of the DNA extraction buffer was increased from 2.7 to 3.5 mL/g, and the time of vortex shocking was prolonged from 30 to 45 min. Genomic DNA was partially digested with *EcoR* V/*Pst* I/*BamH* I/*Sau3A* I. After determination of the optimal endonuclease, the enzyme digestion time of the soil DNA was optimized.

The DNA fragments obtained by the above method were purified using the OMEGA gel extraction kit. Then, the purified DNA fragments were ligated with a pUC118/*BamH* I vector overnight at 16 °C for 12 h with T4 DNA ligase. This recombinant plasmid was transformed into *E. coli* DH5α and cultivated in a solid medium containing 50 μg/mL Amp, 20 μg/mL IPTG, and X-Gal (50 μg/mL) at 37 °C overnight. The white colonies were expanded and cultured at 37 °C and 220 rpm for 14 h. White colonies were randomly picked for the cellulase screening culture (5 mg/mL CMC, 50 μg/mL Amp) at 37 °C for 24 h. Then, the hydrolysis circle of CMC was observed by Congo red staining. Firstly, 15 mL of a Congo red dyeing solution (1 mg/mL) was added for 1 h of dyeing, and after the waste liquid was discarded, and 15 mL of NaCl (1 mol/L) was added for 1 h of decolorization. An obvious hydrolysis circle demonstrating a potential for cellulase production was identified, as shown on the plates of solid CMC stained with Congo red (1 mg/mL). Colonies were purified by repeated streaking. Finally, the DNA of positive colonies was sequenced by TSINGKE Biological Technology (Guangzhou, China) and analyzed using bioinformatics programs, such as ORFfinder, BLAST, and MEGA-X.

Construction of recombinant plasmid and transformation

During PCR amplification, pUC118-cell1029 plasmid as the template and cell1029-F and cell1029-R as primers were used. The amplified product and vector, pET-32a (+), was digested with *BamH* I and *Hind* III and then linked by T4 DNA Ligase. Finally, the constructed vector, pET32a-cell1029, was transferred into *E. coli* BL21 (DE3). PCR primers for cell1029 amplification were as follows: cell1029-F, 5′-CGGGATCCATGGGATCCGAAATCATCATCATCATACCACT-3′; cell1029-R, 5′-CCGGAACCAGCTTTCGCCCCTTTCCCTTTGCGTTCA-3′ (the *BamH* I and *Hind* III restriction sites were italicized). The PCR process was as follows: denaturation 2 min at 98 °C; 30 cycles of 10 s at 98 °C, 5 s at 64 °C, and 5 s at 72 °C; and elongation for 8 min at 72 °C. The expression vector pET-32a (+) was digested using *BamH* I and *Hind* III at 37 °C for 20 min and then ligated to the PCR products, which were treated with the same restriction endonuclease. The double-digested pET-32a (+) and the PCR products were ligated by TaKaRa T4 DNA ligase at 16 °C for 14 h. This recombinant plasmid was transformed into *E. coli* BL21 (DE3) by a heat shock method.

Expression, protein purification, and electrophoresis

The protein was expressed by growing the *E. coli* BL21(DE3) cells in an LB medium (50 μg/mL, Amp). For purification of *Cell1029* expressed by *E. coli*, the sediment was collected by centrifugation at 8000 rpm and 4 °C for 10 min. The sediment was washed with sterile water twice and broken using an ultrasonicator with an amplitude set at 30 W for 15 min. Recombinant cellulase *Cell1029* was purified using the His-tag protein purification kit. The molecular weight of denatured protein was determined by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) (Cheeseman et al. 2001). The molecular weight of the expressed protein was determined based on the protein markers (TaKaRa) as the standards. According to the predicted molecular weight of protein and its relationship with different concentrations of polyacrylamide gel, separation and concentration gels with the contents of 12% and 5%, respectively, were selected. The protein was stained with Coomassie brilliant blue G-250 for 10–12 h and then decolorized with eluent. For the CMC zymogram analysis, 2.75 mL of a 1% CMC solution was added into 12% separation gel. Then, 5% concentrated gel was prepared, and the purified *Cell1029* was loaded. After electrophoresis, the *Cell1029* was fixed with glacial acetic acid for 1 min, then reacted for 3 h, dyed with a 1 mg/mL Congo red solution for 1 h, and finally decolorized with 1 mol/L NaCl for 1 h.

Cellulase activity assay

Cellulase activity was measured using CMC as the substrate. Cellulase hydrolyzed CMC into reducing sugar. The absorbance peak of the reducing sugar was 540 nm after oxidation with dinitrosalicylic acid (DNS) (Miller 1959) in a boiling water bath. The absorption was linearly related with the content of reducing sugar. A standard curve was drawn using glucose concentration as the abscissa (*X*) axis, and *A* 540 as the ordinate (*Y*) axis. Reducing sugar content can be detected from the standard curve after the *A* 540 of the solution was measured. One unit was defined as the amount of enzyme that released 1 μmol of reducing sugar per minute under the optimal conditions (Kang et al. 2007). All assays were performed in triplicate.
Evaluation of enzymatic properties

Cellulase activity was determined using 1% (w/v) CMC as the substrate, as previously mentioned. At 50 °C, the enzyme activity was measured within pH 2–10. Into 90 μL of 1% (w/v) CMC prepared with the 40 mM BR buffer of different pHs, 10 μL of purified enzyme solution was added for 30 min of reaction. Then, 150 μL of the DNS solution was added to react in boiling water for 10 min, and the supernatant was collected to determine the A540 and thereby the optimal pH. The optimal temperature was investigated from 4 to 80 °C at the optimal pH. In brief, 2 mL of a crude enzyme solution was prepared with BR buffer solution of different pHs and stored at 4 °C for 24 h. Then, 10 μL of the crude enzyme solution was taken every 2 h, and the residual enzyme activity was determined under the optimum temperature and pH. The maximum enzyme activity was set as 100%. The purified enzyme solution was placed at different temperatures (4–80 °C) for 24 h and sampled every 2 h for measurement of residual enzyme activity at the optimum temperature and pH according to the method above.

The effects of various metal ions (K+, Na+, Fe2+, Mg2+, Mn2+, Ca2+, Cu2+, Zn2+, Ni2+, and Co2+) and chemicals (SDS, EDTA, urea, imidazole, methanol, ethanol, isopropanol, DMSO, and Triton X-100) on Cel1029 activity were investigated by preincubating the enzyme with the reagents for 20 min at the optimum temperature and pH. The enzyme activity without any reagent was set as 100%. NaCl was used to study the salt tolerance of the recombinant enzyme. The activity of the enzyme was determined in the optimal pH buffer with a 0 to 3 mM salt solution at the optimum temperature and pH. The enzyme without adding agent was used as a control.

CMC solutions in concentrations from 0.0 to 5.0 mg/mL were prepared with the buffer solution of the optimal pH. Then, 90 μL of the CMC solution at each concentration was added with 10 μL of the 0.649 mg/mL enzyme solution and reacted for 30 min under the optimum temperature. After that, the reaction solution was taken out and added with 150 μL of the DNS solution for 10 min of reaction in boiling water, followed by measurement of A540. The absorbance at 0.0 mM was taken as a negative control, and the concentration of glucose was determined. According to the substrate concentration [S] and the measured initial reaction rate V, the Lineweaver Burk double reciprocal plot (V = Vmax [S]/(Km + [S]) was plotted for 1/V, and the enzymatic kinetic parameters Km and Vmax of cellulase were calculated.

Enzymatic hydrolysis of WHB

The optimal conditions were tested with WHB. The water hyacinth contained high amounts of cellulose that was converted to a reducing sugar by cellulase. First, single-factor analysis was used to study the WHB degradation by changing the pH, temperature, enzyme dosage, and enzymatic hydrolysis time. Then, an L9(34) orthogonal design (Table 1) was used to determine an optimal degradation combination. The reaction mixtures contained the WHB (0.2 g), Britton-Robinson buffer (40 mM), and cellulase in a 20-mL system at 200 rpm. Enzymatic extraction of the WHB was determined using spectrophotometry. The supernatant was centrifuged to measure the absorbance at 540 nm. The same amount of an inactivated enzyme solution was used as a negative control. All reactions were performed in triplicate.

Nucleotide sequence accession number

This nucleotide sequence was declared to the GenBank at the registration number of MK051001.

Results

Screening for cellulase using the soil metagenomic library

Metagenomic strategies are applied to isolate and identify novel enzymes with new catalytic or secondary metabolites from different environmental samples (Tang et al. 2018; Fan et al. 2017). The method of carboxymethyl cellulose sodium (CMC)-Congo red staining (Teather and Wood 1982) is often used to screen active clones of endoglucanase. In this study, we screened 7200 positive clones from a metagenomic library, which was built by extracting the total DNA from soil microbes. Results showed that the optimum endonuclease was BamH I and the optimum time was 80 min. The average fragment size of the total DNA was about 2–2.5 kb. Cellulases were screened from the metagenomic library by Congo red staining, and one putative cellulase clone was visually identified by a transparent hydrolytic circle with a diameter of 1.25 cm that appeared on the Congo red plate (Fig. 1).

The clone was sequenced and analyzed with ORFfinder of the National Center for Biotechnology Information (NCBI), which showed the full-length 1029-bp fragment contained a cellulase gene named cel1029. The cel1029 had an open

| Factor/level | 1 | 2 | 3 |
|-------------|---|---|---|
| A—temperature/°C | 45 | 50 | 55 |
| B—enzyme amount/mL | 2 | 4 | 6 |
| C—pH | 4.0 | 5.0 | 6.0 |
| D—time/h | 20 | 22 | 24 |
reading frame of 927 bp and encoded a 34.21 kDa protein consisting of 309 amino acids. Conserved domain analysis revealed that *Cel1029* had a domain belonging to GH5 (312–1013 residues). Through multiple sequence alignments and protein BLAST search by BioEdit, ESPript 3.0, and NCBI, *Cel1029* was found to be similar with endo-1,4-D-glucanase (62.81%, NCBI accession P07103.2), endo-1,4-D-glucanase (48.48%, P15704.1), endo-1,4-D-glucanase (41.06%, P10475.1), endo-1,4-D-glucanase (41.06%, P07983.2), endo-1,4-D-glucanase (40.07%, P23549.1), endo-1,4-D-glucanase (42.57%, P06565.1), endo-1,4-D-glucanase (48.07%, O85465.1), and endo-1,4-D-glucanase (48.07%, P06566.1) (Fig. 2). These sequences were strictly conserved at two catalytic residue sites—Glu176 and Glu264. The two residues served as catalytic proton donors and active site nucleophiles, respectively, and can maintain the structure of the substrate isocarbon for the enzyme. A phylogenetic tree was constructed by neighbor-joining to verify the evolutionary relationship of *Cel1029* with 14 known endoglucanases. Results show that *Cel1029* is closely related to three endocellulases...
from Bacillus sp. (Fig. 3). Multiple sequence alignments and phylogenetic tree analysis indicate that Cel1029 is an endo-1,4-\(\beta\)-glucanase (EC3.2.1.4) belonging to GH5.

**Cloning and overexpression of cellulase gene in E. coli and purification of recombinant protein**

PCR was conducted by using plasmids pUC118-cel1029 as the template and cel1029-F and cel1029-R as primers. The amplified fragment length was consistent with the predicted target DNA fragment length of 1029 bp. pET32a-cel1029 was identified by the verified enzyme digestion (Fig. 4) and the confirmed DNA sequencing. Cel1029 was expressed in E. coli BL21 (DE3). According to the optimal induction temperature (30 °C) and time (14 h), the optimal induction concentration of IPTG was 0.9 mM. Expression strain was induced by IPTG and then broken by ultrasound. The crude enzyme was purified and then analyzed by SDS-PAGE. The enzyme was purified by about 9-fold, with a specific activity of 883 ± 25 U/mg. Figure 5 shows the approximate target band (37.8 kDa), including the theoretically calculated cellulose enzyme molecular weight (34.21 kDa) and protein molecular tag weight (3.63 kDa). The recombinant protein was highly soluble expressed in E. coli, which is also the premise and basis for industrial application. The results of CMC zymogram analysis show a translucent zone (Fig. 6) that indicates CMCase activity.

**Characterization of recombinant Cel1029**

According to the relationship between the concentration of the standard glucose solution and the absorbance at 540 nm, we obtained the regression equation of the standard curve \(Y = 0.7609x + 0.00916, R^2 = 0.998\). Cel1029 was active over a wide pH range, maintaining 40% of its optimum activity at pH 2.0 and stable between pH 5.0 and 7.0 (Fig. 7a). The optimal reaction temperature of Cel1029 at the optimal reaction pH was measured with CMC as the substrate (Fig. 7b). The maximum enzyme activity was defined as 100%. Cel1029 retained 85.3% of the maximum activity at 4 °C and 68.5% at 80 °C. Maximum activity occurred at 55 °C, and more than 50% activity was retained at 4–80 °C, indicating that Cel1029 has high adaptability at a wide temperature range. More than 86% of maximum activity was retained at 4–60 °C after 24 h (Fig. 7d), indicating that Cel1029 has excellent thermal stability. Cellulase activity was obviously enhanced by the addition of Co\(^{2+}\) and slightly inhibited by Ni\(^{2+}\) (Fig. 8). The influence of Co\(^{2+}\) concentration on Cel1029 was further studied, and the CAZy. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches.

![Fig. 3 Phylogenetic tree of Cel1029 and 14 known endoglucanases built by neighbor-joining. Evolutionary history was inferred using neighbor-joining by MEGA-X. Sequences of all endoglucanases were cited from the CAZy. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches.](image-url)
results showed 10 mM Co$_{2+}$ promoted enzyme activity to 156.18%.

Cel1029 was stable in some common organic and nonionic solvents, as more than 50% activity was retained in all tested solvents at different concentrations (Fig. 9). Especially, the residual enzyme activity of Cel1029 was 95.3% when the SDS concentration was 1 mM, and still reached about 70% and 50% when the SDS concentration increased to 10 and 100 mM, respectively. Cel1029 was highly resistant against methanol and isopropanol, and its activity increased in the 1% (v/v) concentration.

NaCl (0 to 3 M) was used to study the salt tolerance of the recombinant enzyme under the optimal temperature and pH. The enzyme activity was maintained above 80% when the salt concentration was 3.0 M (Fig. 10).

Determining $K_m$, the most important characteristic constant of enzymes, is an important method to study enzyme kinetics and reflects the magnitude of affinity between an enzyme and the substrate. In the past, Michaelis-Menten parameters were used to test the $K_m$ and $V_{max}$ of cellulase (King et al. 2009). Then, Lineweaver and Burk used the equation derived from the Michaelis-Menten equation to determine $K_m$ and $V_{max}$ to be 0.535 mM and 1.588 mM/min, respectively.

Enzymatic hydrolysis of water hyacinth

The content of reducing sugar was 17.3 ± 0.7 mg/g in the negative control and was 59 ± 3.8 mg/g in WHB before elution, indicating a part of reducing sugar was lost during the pretreatment of water hyacinth with H$_2$SO$_4$. The optimal conditions of enzymatic hydrolysis to WHB were studied by spectrophotometry (Fig. 11). The sugar yield (sugar production/WHB quality, mg/g) reached 331.99 mg/g at pH 5 (Fig. 11a) and decreased almost linearly with the increasing pH. The sugar yield was only 213.65 mg/g at pH 8 and minimized to 230.92 mg/g at 35 °C. The sugar yield also increased with temperature rise and maximized to 383.57 mg/g at 50 °C (Fig. 11b). However, the sugar yield decreased when the temperature exceeded 50 °C. Therefore, the sugar yield of WHB was the highest at 50 °C. The maximum sugar yield reached 356.30 mg/g with 2 mL of enzyme (24.25 U, Fig. 11c).

Figure 11D shows the effect of enzymatic hydrolysis time on the sugar yield. The sugar yield of WHB maximized to 395.19 mg/g when the enzymatic hydrolysis time was 24 h. Table 2 shows the results of the orthogonal test. The primary factor affecting the sugar yield of WHB is enzymatic hydrolysis time, followed by the secondary factors, including pH,
additional amount, and temperature. The optimal conditions for a peak sugar yield (430.39 mg/g) of WHB are as follows: 50 °C, 4 mL of enzyme, pH 6.0, and 22 h of enzymatic hydrolysis.

**Discussion**

Metagenomic strategies have been applied to isolate and identify novel enzymes with new catalytic or secondary metabolites from different environmental samples. A metagenomic library was built successfully using DNA extracted from crop residues of red rice with castor bean cake and a new phytase was screened (Farias et al. 2018). A novel β-glucosidase gene bgl2238 was screened from the macrogenomic library by functional screening (Tang et al. 2018). A novel gene (aai810) encoding an N-acylhomoserine lactonase was isolated from the Mao-tofu metagenome for the first time (Fan et al. 2017). CMC-Congo red staining (Teather and Wood 1982) is often used to screen active clones of endoglucanase. The main limitation of metagenomic libraries is the need of suitable hosts to promote heterologous expression (Meneses et al. 2016). This limitation is mainly attributed to the difficulties in identifying regulatory elements and the existence of different codons (Wang et al. 2016). In an arid agricultural system, the dominant flora include spore-forming strains of *Bacillus* and *Paenibacillus*, while *Streptomyces* is more abundant in natural desert systems (Schrempf 2013). The codons of the actinomycete shine Dalgarno promoter region are different from those of *E. coli*; so, heterologous expression is difficult.
for the cellulase mainly from *Actinomycetes* screened in natural desert systems. In this study, the cellulase gene *cel1029* was successfully expressed in *E. coli*. Direct extraction of soil samples was adopted, and the amount of DNA extraction buffer and the time of vortex shock were increased to reduce microbial adsorption by soils and enhance the degree of DNA dissolution. In addition, a Congo red screening medium was used here, which can show a transparent hydrolysis circle, increase the recognition, and reduce the probability of missing positive clones from the metagenomic library.

*Cel1029* has high adaptability at wide ranges of temperature and pH, and excellent thermal and pH stability. Several reports on thermostable cellulase show similar catalytic activity at high temperature. The optimal temperature of *Cel1029* is slightly higher than that of *Lp-egl-1* from *Lyrodus pedicellatus* (Hiroshi et al. 2008) and *PersiCell* from camel rumen metagenome (Morteza et al. 2020a, b). On the basis of hot activity, *Cel1029* outperforms *PersiCel2* with about 10% activity at 80 °C (Morteza et al. 2020a, b) but is worse than *R63* from *Thermoanaerobacterium* sp. with about 80% activity at 75 °C, (Harnvoravongchai et al. 2020) and *PW2* from *Bacillus* sp. with 100% activity at 80 °C (Divyanshi et al. 2019). The residual activity at 70 °C is 48.1% after incubation for 24 h, which is better than that of *MaCel* (Tong et al. 2020). In terms of pH adaptability, the optimum pH of *PersiCell* with good pH adaptability is 5 (Morteza et al. 2020a, b), and the relative enzyme activity is about 40% at pH 4 and pH 10. The optimum pH of *ZFYN184* from soil metagenomic library is 4 (Chai et al. 2020), and the relative enzyme activity is lower than 10% at pH 3 and pH 6. The optimum pH of *Cel1029* is 6, and the relative enzyme activity is slightly higher than 40% at pH 2, and above 45% at pH 10. In conclusion, *Cel1029* has obvious advantages in temperature and pH adaptability.

Activity of *Cel1029* is obviously enhanced by the addition of Co²⁺ and slightly inhibited by Ni²⁺. Similar phenomena have been reported, in which 0.5 mM Co²⁺ double-increased the enzyme activity of *MaCel* (Tong et al. 2020) and 1 mM Co²⁺ promoted the activity of AgCMCase (Kumar and Parikh 2015). The results suggest that *Cel1029* is resistant against most metal ions. The enzyme activity can be increased by adding appropriate metal ions in industrial production. To determine the usefulness of *Cel1029* in industrial applications, it is necessary to study the effects of organic and nonionic solvents on its activity and stability. Interestingly, when the SDS concentration increases to 10 and 100 mM, the residual enzyme activity is still about 70% and 50%, respectively. *BC1*, an alkalophilic cellulase from the symbiotic *Bacillus subtilis*, retains about 30% at the presence of 1% SDS (Dehghanikhah et al. 2020). *GAC 16.2* from *Acinetobacter junii* reserves 66.1% activity with the presence of 5 mM SDS (Sandipan et al. 2020). *Cel1029* has obvious advantages in tolerance against SDS. On the one hand, SDS increases the surface negative charge of cellulase and the substrate, increasing the electrostatic repulsion between cellulase and the substrate. Finally, the binding ability between cellulase and the substrate and the
hydrolysis ability of cellulase is decreased. On the other hand, the ineffective adsorption between cellulase and the substrate makes the enzyme irreversibly adsorbed on the substrate surface (Palonen et al. 2004), which results in the change of protein conformation, protein denaturation, and decrease of cellulase activity (Norde and Favier 1992). SDS is adsorbed onto the surface of the substrate through hydrogen bonding (Borjesson et al. 2007) and hydrophobic interaction, which significantly reduce the ineffective adsorption of the substrate and enzyme and enhances the cellulase activity in the liquid phase (Bálint et al. 2011). The relative activity of Cel1029 is still about 50% with the presence of 100 mM SDS, indicating that the negative charge of SDS has a weak effect on the binding degree of Cel1029 and the substrate, which may be affected by the charge properties of Cel1029.

Table 2 Orthogonal test results and analysis. $K_1$, $K_2$, and $K_3$ represent the average value of the corresponding factor in each column in three experiments at levels 1, 2, and 3, respectively. $R$ represents the difference between the maximum and minimum values of $K$ in each column.

| Test No. | A—temperature/°C | B—enzyme amount/mL | C—pH | D—time/h | Sugar yield/mg/g |
|----------|------------------|--------------------|-------|----------|-----------------|
| 1        | 1                | 1                  | 1     | 1        | 278.02          |
| 2        | 1                | 2                  | 2     | 2        | 430.39          |
| 3        | 1                | 3                  | 3     | 3        | 212.67          |
| 4        | 2                | 1                  | 2     | 3        | 335.01          |
| 5        | 2                | 2                  | 3     | 1        | 327.28          |
| 6        | 2                | 3                  | 1     | 2        | 312.72          |
| 7        | 3                | 1                  | 3     | 2        | 365.52          |
| 8        | 3                | 2                  | 1     | 3        | 285.90          |
| 9        | 3                | 3                  | 2     | 1        | 362.00          |
| $K_1$    | 308.36           | 326.19             | 266.70| 321.10   |                 |
| $K_2$    | 337.81           | 346.52             | 375.80| 369.54   |                 |
| $K_3$    | 323.67           | 297.13             | 301.82| 279.19   |                 |
| $R$      | 29.45            | 49.39              | 109.10| 89.75    |                 |

Fig. 11 Effects of pH, temperature, enzyme amount, and time on enzymatic saccharification of WHB. Data points and error bars represent mean and standard deviation respectively. a Effect of pH tested in 40 mM B-R buffer (pH 4.0–8.0) at 55 °C. b Effect of temperature tested in 40 mM B-R buffer (pH 5.0) from 35 to 60 °C. c Effect of enzyme amount tested in 40 mM B-R buffer (pH 5.0) at 50 °C. d Effect of time tested in 40 mM B-R buffer (pH 5.0) at 50 °C. Values are shown as reducing sugar yield/WHB quality in reaction systems. Each reaction involved 0.2 g of WHB and 24.25 U cellulase, and the same amount of an inactivated enzyme solution was used as a negative control.
largely reduce the ineffective adsorption of Cel1029 onto the substrate. Cel1029 can well resist methanol and isopropanol, and its activity is increased at the concentration of 1% (v/v). Similar performances have been reported. The activity of HSC7 from Bacillus sonorensis increases to 166% in the presence of methanol (Azadian et al. 2017). The activity of Est906 (Xiaolin et al. 2020) from paper mill wastewater sediments is enhanced to 119.8% in the presence of isopropanol. These results suggest that adding appropriate organic or nonionic solvents into industrial production may strengthen the enzyme activity. Recombinant Cel1029 reserves above 80% activity when the salt concentration is 3.0 M and thus is more stable than SR22 from Bacillus sp. strain (Santos et al. 2018), indicating that Cel1029 has high tolerance against salt solutions. The $K_m$ and $V_{max}$ of cellulase from Aspergillus fumigatus are 10.052 mM and 1.38 mM/min, respectively (Prabhpreet et al. 2020).

Water hyacinth is rich in cellulose, which is hydrolyzed into cellobiose or glucose by enzymes such as cellulase. Cellulase can degrade cellulose into cellulosic ethanol and then produce biofuels by microbial fermentation, which can optimize the use of WHB as a resource (Ma et al. 2010). In this study, the degradation of water hyacinth by cellulase Cel1029 was studied preliminarily. The reducing sugar yield of WHB through biotransformation of Cel1029 under optimal conditions is 84.36% of the theoretical maximum reducing sugar yield, which is about 510.2 mg/g (Klass and Ghosh 1981; Nigam 2002; Kumar et al. 2009). So far, it has not been reported that the cellulase degrades water hyacinth with such high efficiency.

Conclusions

A soil metagenomic library was constructed, and a new cellulase gene, cel1029, was cloned by using an activity-based functional method. Cel1029 was successfully expressed in the prokaryotic system to explore its stability, enzymatic properties, and ability to degrade water hyacinth. It showed high activity in a wide pH range, good thermostability, and excellent tolerance against organic solvents and salts. The recombinant enzyme was purified and characterized. The selected new cellulase was used to degrade and transform water hyacinth and directly damaged the plant cell walls. Cellulase directly converted water hyacinth into glucose during the short test period, indicating a potential application of water hyacinth for use in energy production. We present new cellulase from the soils of Ferula asafoetida distribution area, enrich the source of cellulase, and make full use of cellulase. Meanwhile, this study provides a research basis for improving the utilization rate of cellulose-rich plants.

Acknowledgements This work was financially supported by the Guangzhou Basic Clean Cosmetics Manufacturing Co. Ltd.

Availability of data and material Not applicable

Code availability Not applicable

Author contribution Conceptualization, X.S.Z.; methodology, J.Y.; software, Z.J.D.; validation, X.S.Z., S.L., and J.Y.; formal analysis, X.S.Z.; investigation, Z.J.D.; resources, Z.J.D.; data curation, X.X.; writing—original draft preparation, X.S.Z.; writing—review and editing, L.Y.L.; visualization, X.X.; supervision, H.L.; project administration, H.L.; funding acquisition, S.L. All authors have read and agreed to the published version of the manuscript.

Funding This research has been funded by Natural Science Foundation of China (31400680), Science and Technology Plan Project of Guangzhou (201802030009), the Innovation and Strengthening School Project from Guangdong Pharmaceutical University (2016KTSX067 and 2016SFKC_28), Science and Technology Plan Project of Guangdong Province (2017A010105011, 2014A020208134, 2014A020212602), and Education Project of Guangdong Province (2013KJCX0107).

Declarations

Consent to participate Not applicable

Consent for publication Not applicable

Conflict of interest The authors declare that they have no conflict of interest.

References

Amann RI, Ludwig W, Schleifer KH (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiol Rev 59:143–169
Amitha PR, Christopher WS, Patrik D, Jane K, Helcio B, Masood H, Blake AS, Steven WS, Michael PT, Jean SVG (2013) Discovery of microorganisms and enzymes involved in high-solids decomposition of rice straw using metagenomic analyses. Plots One 8:210
Amrani F, Salim FA, Iskandaria I (2016) Physical and biophysical pretreatment of water hyacinth biomass for cellulase enzyme production. Chem Biochem Eng Q 30:237–244
Asrofi M, Abral H, Putra YK, Kim HJ (2018) Effect of duration of sonication during gelatinization on properties of tapioca starch water hyacinth fiber biocomposite. Int J Biol Macromol 108:167–176
Azadian F, Badoei-Dalfard A, Namaki-Shoushtari A, Karami Z, Hassanshahian M (2017) Production and characterization of an acido-thermophilic, organic solvent stable cellulase from Bacillus sonorensis HSC7 by conversion of lignocellulosic wastes. J Genet Eng Biotechnol 15(1):187–196
Meneses C, Silva B, Medeiros B, Serrato R, Johnston-Monje D (2016) A Metagenomic Advance for the Cloning and Characterization of a Cellulase from Red Rice Crop Residues. Molecules 21:831
Miller GL (1959) Use of dinitrosalicylic acid reagent for determination of reducing sugar. Anal Chem 31:426–428
Morteza M, Mehdi FS, Kaveh K, Shohreh A, Ghasem HSAC (2020a) A novel thermostable cellulase cocktail enhances lignocellulosic bioconversion and biorefining in a broad range of pH. Int J Biol Macromol 154:349–360
Morteza MA, Mehdi FS, Kaveh K, Shohreh A, Ghasem HS (2020b) A novel thermostable cellulase cocktail enhances lignocellulosic bioconversion and biorefining in a broad range of pH. Int J Biol Macromol 154:349–360
Nigmat JN (2002) Bioconversion of water-hyacinth (Eichhornia crassipes) hemicellulose acid hydrolysate to motor fuel ethanol by xylose-fermenting yeast. J Biotechnol 97:107–116
Nimchua T, Thongaram T, Uengwetwanit T, Somchai P, Lily E (2012) Metagenomic Advance for the Cloning and Characterization of a novel thermostable cellulase cocktail for the production of fermentable sugars. Catalysts 2:450–459
Okada G (1976) Enzymatic studies on a cellulase system of Trichoderma viride IV Purification and properties of a less-random type cellulase. J Biochem 80:913–922
Palonen H, Tjerneld F, Zacchi G, Tenkanen M (2004) Adsorption of cellulase from Red Rice Crop Residues. Molecules 21:831–849
Kaur P, Taggar MS, Kalia A (2020) Characterization of magnetic nanoparticle–immobilized cellulases for enzymatic saccharification of rice straw. Biomass Conv Bioref. https://doi.org/10.1007/s13399-020-00628-x
Raj K, Krishnan C (2018) High sugar yields from sugarcane (Saccharum officinarum) bagasse using low-temperature aqueous ammonia pretreatment and laccase-mediated assisted enzymatic hydrolysis. Ind Crop Prod 111:673–683
Rajesh D, Anil D, Sakshi (2014) Bioconversion of water hyacinth to ethanol by using cellulase from Trichoderma atroviride AD-130. Adv Mater Res 3135:145–148
Salinas A, Vega M, Maria EL, Alejandro G, Rene C, Oriana S (2011) Cloning of novel cellulases from cellulolytic fungi: heterologous expression of a family 5 glycoside hydrolase from Trametes versicolor in Pichia pastoris. Enzym Microb Technol 49:485–491
Samkelo M, Shauñita HR, Willem HZ, Brett IP (2020) Enzymatic hydrolysis of softwood derived paper sludge by an in vitro recombinant cellulase cocktail for the production of fermentable sugars. Catalysts 10:775–792
Sandipan B, Tushar KM, Raj NR (2020) Production, purification, and characterization of cellulase from Acinetobacter junii GAC 16.2, a novel cellulolytic gut isolate of Gryllotalpa africana, and its effects on cotton fiber and sawdust. Ann Microbiol 70(1):853–866
Santos YQD, Veras BOD, Jácome FAF, Gorlach-Lira K, Elizeu ADS (2018) A new salt-tolerant thermostable cellulase from a marine Bacillus sp strain. J Microbiol Biootechnol 28(7):1078–1085
Schmitz A (2004) Metagenomics—the key to the uncultured microbes. Curr Opin Microbiol 7:492–498
Schrempf H (2013) Actinobacteria within soils: capacities for mutualism, symbiosis and pathogenesis. FEMS Microbiol Lett 342:77–78
Sumrith N, Dangtungee R (2019) Mechanical properties of water hyacinth fiber reinforced bio-based epoxy composite. Fibers 48297–11
Tang LL, Wang XM, Wu XL, Huang Q, Li H (2018) Screening, cloning and characterization of a novel β-Glucosidase from soil metagenome library. FoodSci 39:118–124
Teather RM, Wood PJ (1982) Use of Congo red-polysaccharide interactions in enumeration and characterization of cellulolytic bacteria from the bovine rumen. Appl Environ Microbiol 43(4):777–780
Tirawongsaroj P, Sriprang R, Harnpicharnchai P, Taksawarn T, Verawat C, Sutipa T, Kusol P, Lily E (2008) Novel thermophilic and thermostable lipolytic enzymes from a Thailand hot spring metagenomic library. J Biotechnol 133:42–49
Tong Z, Ruiyfeng L, Jinyuan S, Yinglu C, Bian W (2020) Characterization and efficient production of a thermostable, halostable and organic solvent-stable cellulase from an oil reservoir. Int J Biol Macromol 159:622–629
Valère D, Mirjam C, Anne B, Christian G, Henri-Pierre F, Jean-Pierre B, Gideon JD, Richard HC (1995) Crystal structure of the catalytic domain of a bacterial cellulase belonging to family 5. Structure 3:939–949
Wang C, Dong D, Wang H, Müller K, Qin Y, Wang H, Wu W (2016) Metagenomic analysis of microbial consortia enriched from compost: new insights into the role of Actinobacteria in lignocellulose decomposition. Biotechnol Biofuels 9(1):22
Xiaolin Z, Yongzhen T, Meilu J, Yide L, Gang L (2020) Characterization and purification via nucleic acid aptamers of a novel esterase from the metagenome of paper mill wastewater sediments. Int J Biol Macromol 153:441–450
Yuan YL, Zhi YW, Ya TZ, Guo PZ, Ling L (2019) Enzymatic identification and functional sites study of a novel cold-active cellulase (McCel5) from Microbacterium kitamiense. Biotechnol Biotec Equip 33:739–747
Yun J, Kang S, Park S, Yoon H, Kim MJ, Heu S, Ryu S (2004) Characterization of a novel amylolytic enzyme encoded by a gene from a soil-derived metagenomic library. Appl Environ Microbiol 70:7229–7235
Zheng F, Tu T, Wang XYY, Ma R, Su XXM, Yao B, Luo HY (2018) Enhancing the catalytic activity of a novel GH5 cellulase GtCel5 from Gloeophyllum trabeum CBS 90073 by site-directed mutagenesis on loop 6. Biotechnol Biofuels 11:76
Zhou JZ, Bruns MA, Tiejie JM (1996) DNA recovery from soils of diverse composition. Appl Environ Microbial 62:316–322

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