MOLECULAR ANALYSIS AND CLONING OF A CELLULASE GENE FROM *Bacillus velezensis* RB.IBE29

Tran Minh Dinh[1], Pham Thi Thu Thuy[2], Nguyen Thi Huyen[1]

Received Date: 25/6/2024; Revised Date: 02/8/2024; Accepted for Publication: 03/8/2024

ABSTRACT

Bacillus velezensis RB.IBE29 was originally isolated from the rhizospheric soil of black pepper cultivated in the Central Highlands. Previous studies showed that this bacterium was a good chitinase producer, biocontrol agent, and biofertilizer. The complete genome sequence of strain RB.IBE29 was reported and revealed that it harbors the gene coding for the family 5 cellulase of glycoside hydrolases, however, this gene has not been experimentally characterized. This small work aimed to analyze and clone the gene encoding the family 5 cellulase of strain RB.IBE29 for subsequent studies. Results showed that strain RB.IBE29 possesses the *celA* encoding the family 5 cellulase. The ORF of the *celA* contains 1500 bp and encodes cellulase with 499 aa and 55.0 kDa. Primary structure analysis found that the enzyme contains a signal peptide sequence at the N-terminus, a GH5 catalytic domain, and a CBM3 domain at C-terminus. A fragment (1410 bp) without signal peptide sequence was successfully amplified and cloned in *Escherichia coli* DH5α. Sequencing analysis confirmed that no mutations or frameshift mutations were found in the insert of the recombinant vector from the positive clone. This work provided valuable material for our next expression, purification, and characterization of the cellulase from strain RB.IBE29.

Keywords: *Cellulase; the glycoside hydrolase family 5; gene cloning; Bacillus velezensis RB.IBE29*

1. INTRODUCTION

Cellulose is the most abundant biopolymer on Earth and the primary structural component of plant and algal cell walls. Cellulases are enzymes that hydrolyze β-1,4-glycosidic linkage in cellulose. Various microorganisms produce cellulases, such as fungi and bacteria. In agriculture, it was reported that cellulase-producing microorganisms can treat agricultural wastes and convert them to biofertilizers for crop production (Juturu et al., 2014).

Bacillus velezensis is a strong biocontrol and plant-growth promoting agent, and has been applied for agricultural cultivation (Cai et al. 2024; Fan et al., 2018). *B. velezensis* RB.IBE29 was isolated from the rhizospheric soil of black pepper cultivated in the Central Highlands. Cells of strain RB.IBE29 exhibited high activities of chitinase, β-glucanase, and protease; produced indole-3-acetic acid; solubilized insoluble phosphate; and suppressed the growth of *Phytophthora*, a fungal pathogen of black pepper (Trinh et al., 2019). According to field experiments, strain RB.IBE29 was a good biological agent for producing black pepper sustainably (Nguyen et al., 2021). Strain RB.IBE29 possessed two novel family 18 chitinases. These enzymes have been expressed in *E. coli* cells, and purified recombinant chitinases strongly inhibited the germination of fungal spores and the hatching of nematode eggs (Tran et al., 2022a,b). A xylanase gene of strain RB.IBE29 has been expressed, and purified recombinant xylanase was active at a wide range of temperatures and pHs (Tran et al., 2024). The complete genome of strain RB.IBE29 was sequenced and provide valuable genomic information for biotechnological application (Tran et al., 2023). Among them, one gene encoding cellulase belonging to family 5 was found in the genome of strain RB.IBE29; however, this enzyme has not been experimentally characterized. Hence, this small work reported the molecular analysis and cloning of the gene encoding the family 5 cellulase from *B. velezensis* RB.IBE29 in *E. coli* DH5α. The cloned gene was valuable for further expressing, purifying, and characterizing this enzyme.

2. MATERIALS AND METHODS

2.1. Materials

The genome sequence of *B. velezensis* RB.IBE29 (DDBJ/EMBL/Genbank accession AP028932) was used to search for and analyze the cellulase gene. *B. velezensis* RB.IBE29 (Trinh et al., 2019) was used to isolate the cellulase gene. *E. coli* DH5α (New England Biolabs, USA) was the competent cell. pUC19 (Thermo Fisher Scientific, USA) was used as the cloning vector. Luria-Bertani (LB) medium was used for bacterial growth.

[1]*Institute of Biotechnology and Environment, Tay Nguyen University;*
[2]*University of Science, Viet Nam National University - Ho Chi Minh City;*
Corresponding author: Tran Minh Dinh; Tel: 0943095079; Email: tmdinh@ttn.edu.vn

2.2. Methods

2.2.1. Searching and analyzing the cellulase gene

The open reading frame (ORF) was examined using ORF finder (www.ncbi.nlm.nih.gov/orffinder/). The signal peptide sequence was identified using SignalP 5.0 (Almagro Armenteros et al., 2019). The primary structure was analyzed using the SMART 9.0 (Letunic et al., 2020). Molecular weight was calculated using the Compute pI/Mw tool (Gasteiger et al., 2005). The BLASTp program (Johnson et al., 2008) was used to examine the homology of nucleotide and amino acid sequences. Phylogenetic analysis was conducted using MEGA 6.0 (Tamura et al., 2013).

2.2.2. Extraction of the genomic DNA

The genomic DNA of *B. velezensis* RB.IBE29 (Trinh et al., 2019) was obtained from an overnight culture (16 h) by boiling for 5 min followed by centrifuge (13,000 rpm, 1 min, 4°C). The supernatant was used as a template for polymerase chain reaction (PCR) (Tran et al., 2018; Pentekhina et al. 2020).

2.2.3. Gene cloning and sequencing analysis

A fragment without signal peptide sequence (1410 bp) of the *celA* was amplified from the genomic DNA of *B. velezensis* RB.IBE29 by PCR. A 50-µL reaction mixture contained Cel-F 5'-GGTGGTGGATCCGCAGGGACAAAAACCCCAGTAG-3'(underlines indicate cleavage sites for BamHI and HindIII, respectively), and Phusion High-Fidelity DNA polymerase (Thermo Fisher Scientific, USA), per the supplier's instructions. Primers were designed based on the gene ending the family 5 cellulase (accession RBIBE_18160) of *B. velezensis* RB.IBE29 using the online OligoAnalyzer tool (www.idtdna.com/pages/tools/oligoanalyzer). The reaction mixture was incubated under a schedule consisting of predenaturation at 98°C for 3 min; followed by 40 cycles of 98°C for 10 sec, 55°C for 20 sec, and 72°C for 45 sec; and 72°C for 5 min. Electrophoresis was used to separate the amplified products on a 1.2% agarose gel. The target band was then cut out and purified using the Wizard SV Gel and PCR Clean-Up System (Promega, USA), following the manufacturer's protocol.

The vector pUC19 and insert were incubated individually with the FastDigest BamHI and HindIII (Thermo Fisher Scientific, USA), following the manufacturer's instructions. The product was then analyzed using electrophoresis on 1.5% agarose gel. The treated vector and insert were ligated using the Mighty Mix (Takara, Japan) to generate the recombinant vector utilizing the manufacturer's instruction. The recombinant vector was introduced into *E. coli* DH5α by heat shock. Transformants were spread on LB agar plates containing X-gal (0.04 mg/mL), ampicillin (100 µg/mL), and Isopropyl ***β***-D-thiogalactopyranoside (0.1 mM) (IPTG) and incubated at 37°C for 24 h.

The recombinant vector from white colonies was examined by colony-PCR. A 10-µL reaction mixture contained bacterial cells of the white colony, primers Cel-F and Cel-R, and *Taq* DNA polymerase (Bioline, USA), per the supplier's instructions. The reaction mixture was incubated under a schedule consisting of predenaturation at 95°C for 3 min; followed by 40 cycles of 95°C for 10 sec, 55°C for 15 sec, and 72°C for 15 sec; and 72°C for 5 min. Electrophoresis was used to separate the amplified products on a 1.0% agarose gel. After that, recombinant vectors from positive colonies were isolated and purified using the GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific, USA), per the supplier's instruction. The insert from purified recombinant vectors was sequenced at the 1st BASE Company (Selangor, Malaysia) using primers M13-F 5'-CCCAGTCACGACGTTGTAAAACG-3' and M13-R 5'-AGCGGATAACAATTTCACACAGG-3'.

3. RESULTS AND DISCUSSIONS

3.1. Sequence analysis of deduced cellulase gene

Analysis based on the complete genome sequence of strain RB.IBE29 showed that the genome of strain RB.IBE29 possesses the gene (*celA*) coding for the family 5 cellulase.

Fig. 1A shows that the ORF of the *celA* contains 1500 bp and encodes the cellulase with 499 aa. The molecular weight of the deduced cellulase was calculated to be 55.0 kDa. Sequence analysis exhibited that amino acids of the *B. velezensis* RB.IBE29 cellulase shared the maximum identity (99.6%) to those of a cellulase (WP_302242883.1) of *B. velezensis*, followed by 99.19% to cellulase (WP_256890651.1) of *B. amyloliquefaciens*, 98.79% to a cellulase (ABS70712.1) of *B. subtilis*, 97.98% to a cellulase (WP_133489245.1) of *B. inaquosorum*, and 97.98% to a cellulase (WP_047474757.1) of *B. siamensis*. The primary structure of the enzyme contains a signal peptide at the N-terminus, a GH5 catalytic domain, and a CBM3 domain at the C-terminus (Fig. 1B).

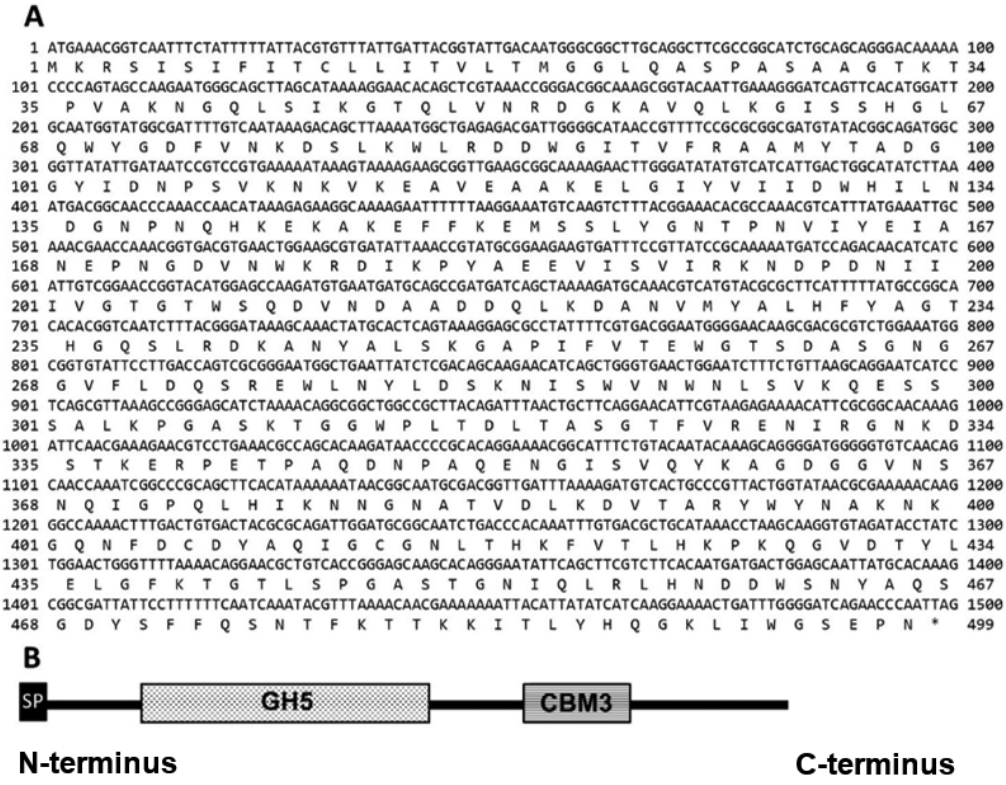

Fig. 1. Nucleotide of cellulase gene and its amino acid sequence, and primary structure of cellulase from strain RB.IBE29.

Note: A, the nucleotide sequence of cellulase gene and its amino acid sequence; B, the primary structure of deduced cellulase; SP: signal peptide sequence; GH5, GH5 catalytic domain; CBM3, carbohydrate-binding module family 3.

To visualize the relationship between cellulase of strain RB.IBE29 and reported bacterial enzymes, a phylogenetic tree was analyzed using MEGA 6.0. The result showed that cellulase of strain RB.IBE29 was grouped with bacterial cellulases (Fig. 2). These analyses confirmed the enzyme is cellulase.

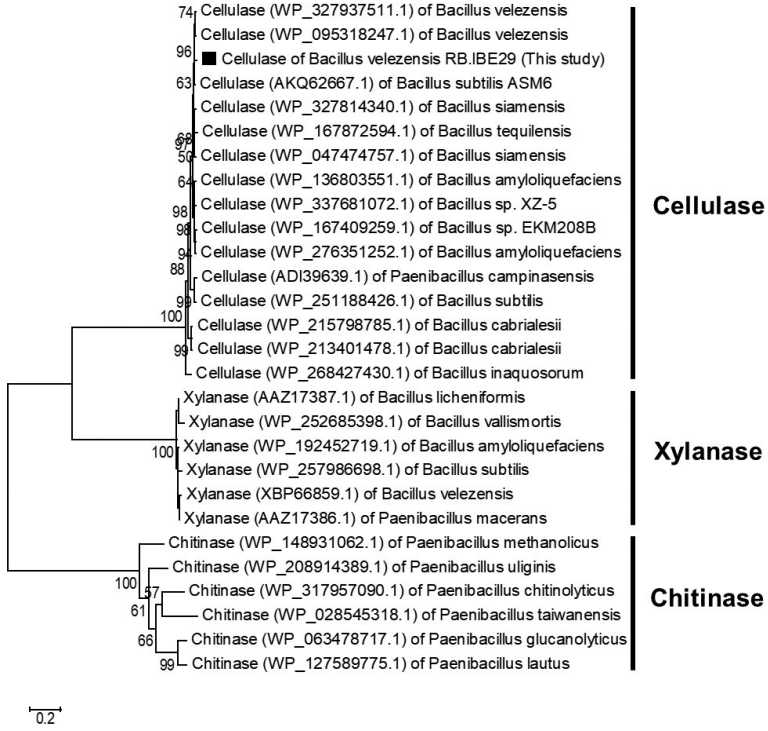

Fig. 2. Phylogenetic analysis of the cellulase from strain RB.IBE29 and reported enzymes

3.2. Cloning of the cellulase gene

Fig. 3A shows that several white colonies grew on the LB agar plate containing X-gal, ampicillin, and IPTG. Using primers M13-F and M13-R, the insert was identified from the white colonies using colony-PCR (Fig. 3B). These results indicated the *celA* (without signal peptide sequence) of *B. velezensis* RB.IBE29 was successfully cloned in *E. coli* DH5α.

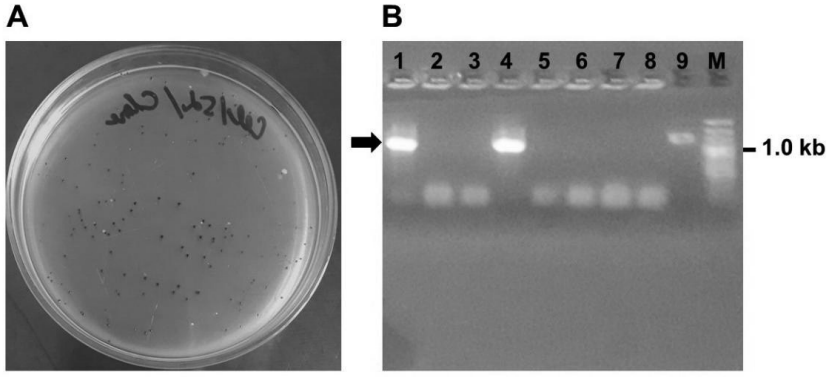

Fig. 3. Transformed cells grown on the LB agar plate containing X-gal, ampicillin, and IPTG, and examination of positive clones by colony-PCR.

Note: A, the transformed cells grown on the LB agar plate containing X-gal, ampicillin, and IPTG; B, the examination of positive clones by colony-PCR. The arrow indicates the target PCR product; Lane M indicates in marker in kilobase; Lanes 1-9 indicate samples.

3.3. Examination of the insert from the clone harboring the recombinant vector

To examine any mutations caused by amplification of the insert, the insert from purified recombinant vectors was sequenced using primers M13-F and M13-R. The result showed that no mutations or frameshift mutations were found in the insert of the recombinant vector from the positive clone (Fig. 4). This result suggested that the recombinant vector is valuable for subsequent studies concerning the expression and purification of the cellulase.

```
Insert          GCAGGGACAAAAACCCCAGTAGCCAAGAATGGGCAGCTTAGCATAAAAGGAACACAGCTCGTAAACCGGGACGGCAAAGCGGTACAATTGAAAGGGATCAGTTCACATGGATTGCAATGG
Intact celA     GCAGGGACAAAAACCCCAGTAGCCAAGAATGGGCAGCTTAGCATAAAAGGAACACAGCTCGTAAACCGGGACGGCAAAGCGGTACAATTGAAAGGGATCAGTTCACATGGATTGCAATGG
                ************************************************************************************************************************

Insert          TATGGCGATTTTGTCAATAAAGACAGCTTAAAATGGCTGAGAGACGATTGGGGCATAACCGTTTTCCGCGCGGCGATGTATACGGCAGATGGCGGTTATATTGATAATCCGTCCGTGAAA
Intact celA     TATGGCGATTTTGTCAATAAAGACAGCTTAAAATGGCTGAGAGACGATTGGGGCATAACCGTTTTCCGCGCGGCGATGTATACGGCAGATGGCGGTTATATTGATAATCCGTCCGTGAAA
                ************************************************************************************************************************

Insert          AATAAAGTAAAAGAAGCGGTTGAAGCGGCAAAAGAACTTGGGATATATGTCATCATTGACTGGCATATCTTAAATGACGGCAACCCAAACCAACATAAAGAGAAGGCAAAAGAATTTTTT
Intact celA     AATAAAGTAAAAGAAGCGGTTGAAGCGGCAAAAGAACTTGGGATATATGTCATCATTGACTGGCATATCTTAAATGACGGCAACCCAAACCAACATAAAGAGAAGGCAAAAGAATTTTTT
                ************************************************************************************************************************

Insert          AAGGAAATGTCAAGTCTTTACGGAAACACGCCAAACGTCATTTATGAAATTGCAAACGAACCAAACGGTGACGTGAACTGGAAGCGTGATATTAAACCGTATGCGGAAGAAGTGATTTCC
Intact celA     AAGGAAATGTCAAGTCTTTACGGAAACACGCCAAACGTCATTTATGAAATTGCAAACGAACCAAACGGTGACGTGAACTGGAAGCGTGATATTAAACCGTATGCGGAAGAAGTGATTTCC
                ************************************************************************************************************************

Insert          GTTATCCGCAAAAATGATCCAGACAACATCATCATTGTCGGAACCGGTACATGGAGCCAAGATGTGAATGATGCAGCCGATGATCAGCTAAAAGATGCAAACGTCATGTACGCGCTTCAT
Intact celA     GTTATCCGCAAAAATGATCCAGACAACATCATCATTGTCGGAACCGGTACATGGAGCCAAGATGTGAATGATGCAGCCGATGATCAGCTAAAAGATGCAAACGTCATGTACGCGCTTCAT
                ************************************************************************************************************************

Insert          TTTTATGCCGGCACACACGGTCAATCTTTACGGGATAAAGCAAACTATGCACTCAGTAAAGGAGCGCCTATTTTCGTGACGGAATGGGGAACAAGCGACGCGTCTGGAAATGGCGGTGTA
Intact celA     TTTTATGCCGGCACACACGGTCAATCTTTACGGGATAAAGCAAACTATGCACTCAGTAAAGGAGCGCCTATTTTCGTGACGGAATGGGGAACAAGCGACGCGTCTGGAAATGGCGGTGTA
                ************************************************************************************************************************

Insert          TTCCTTGACCAGTCGCGGGAATGGCTGAATTATCTCGACAGCAAGAACATCAGCTGGGTGAACTGGAATCTTTCTGTTAAGCAGGAATCATCCTCAGCGTTAAAGCCGGGAGCATCTAAA
Intact celA     TTCCTTGACCAGTCGCGGGAATGGCTGAATTATCTCGACAGCAAGAACATCAGCTGGGTGAACTGGAATCTTTCTGTTAAGCAGGAATCATCCTCAGCGTTAAAGCCGGGAGCATCTAAA
                ************************************************************************************************************************

Insert          ACAGGCGGCTGGCCGCTTACAGATTTAACTGCTTCAGGAACATTCGTAAGAGAAAACATTCGCGGCAACAAAGATTCAACGAAAGAACGTCCTGAAACGCCAGCACAAGATAACCCCGCA
Intact celA     ACAGGCGGCTGGCCGCTTACAGATTTAACTGCTTCAGGAACATTCGTAAGAGAAAACATTCGCGGCAACAAAGATTCAACGAAAGAACGTCCTGAAACGCCAGCACAAGATAACCCCGCA
                ************************************************************************************************************************

Insert          CAGGAAAACGGCATTTCTGTACAATACAAAGCAGGGGATGGGGGTGTCAACAGCAACCAAATCGGCCCGCAGCTTCACATAAAAAATAACGGCAATGCGACGGTTGATTTAAAAGATGTC
Intact celA     CAGGAAAACGGCATTTCTGTACAATACAAAGCAGGGGATGGGGGTGTCAACAGCAACCAAATCGGCCCGCAGCTTCACATAAAAAATAACGGCAATGCGACGGTTGATTTAAAAGATGTC
                ************************************************************************************************************************

Insert          ACTGCCCGTTACTGGTATAACGCGAAAAACAAGGGCCAAAACTTTGACTGTGACTACGCGCAGATTGGATGCGGCAATCTGACCCACAAATTTGTGACGCTGCATAAACCTAAGCAAGGT
Intact celA     ACTGCCCGTTACTGGTATAACGCGAAAAACAAGGGCCAAAACTTTGACTGTGACTACGCGCAGATTGGATGCGGCAATCTGACCCACAAATTTGTGACGCTGCATAAACCTAAGCAAGGT
                ************************************************************************************************************************

Insert          GTAGATACCTATCTGGAACTGGGTTTTAAAACAGGAACGCTGTCACCGGGAGCAAGCACAGGGAATATTCAGCTTCGTCTTCACAATGATGACTGGAGCAATTATGCACAAAGCGGCGAT
Intact celA     GTAGATACCTATCTGGAACTGGGTTTTAAAACAGGAACGCTGTCACCGGGAGCAAGCACAGGGAATATTCAGCTTCGTCTTCACAATGATGACTGGAGCAATTATGCACAAAGCGGCGAT
                ************************************************************************************************************************

Insert          TATTCCTTTTTTCAATCAAATACGTTTAAAACAACGAAAAAAATTACATTATATCATCAAGGAAAACTGATTTGGGGATCAGAACCCAAT
Intact celA     TATTCCTTTTTTCAATCAAATACGTTTAAAACAACGAAAAAAATTACATTATATCATCAAGGAAAACTGATTTGGGGATCAGAACCCAAT
                ***************************************************************************************
```

Fig. 4. Comparison of nucleotides of the insert from the positive clone and those of the intact *celA*

4. CONCLUSIONS

B. velezensis RB.IBE29 possesses a gene (*celA*) encoding cellulase belonging to family 5. The ORF of the *celA* contains 1500 bp in length and encodes cellulase with 499 aa and 55.0 kDa. The primary structure of the enzyme contains the signal peptide at the N-terminus, the GH5 catalytic domain, and the CBM3 domain at the C-terminus. A fragment without signal peptide (1410 bp) was successfully cloned in *E. coli* DH5α. No mutations or frameshift mutations were found in the insert of the recombinant vector from the positive clone. This work provided valuable material for the next expression, purification, and characterization of the cellulase from strain RB.IBE29.

Acknowledgments

This work was supported by the Ministry of Education and Training under grant number B2023-TTN-05.

PHÂN TÍCH PHÂN TỬ VÀ TẠO DÒNG GENE MÃ HÓA CELLULASE CỦA *Bacillus velezensis* RB.IBE29

Trần Minh Định[1], Phạm Thị Thu Thủy[2], Nguyễn Thị Huyền[1]

Ngày nhận bài: 25/6/2024; Ngày phản biện thông qua: 02/8/2024; Ngày duyệt đăng: 03/8/2024

TÓM TẮT

Bacillus velezensis RB.IBE29 được phân lập từ đất vùng rễ cây hồ tiêu trồng tại Tây Nguyên. Các nghiên cứu trước đây cho thấy, đây là tác nhân sản xuất chitinase, kiểm soát sinh học và phân bón sinh học. Trình tự bộ gen của chủng RB.IBE29 đã được công bố và sở hữu gen mã hóa cellulase họ 5, tuy nhiên, gen này chưa được nghiên cứu bằng thực nghiệm. Mục tiêu của nghiên cứu nhằm phân tích phân tử và tạo dòng gen mã hóa cellulase họ 5 của chủng RB.IBE29 để tạo vật liệu cho các nghiên cứu tiếp theo. Kết quả nghiên cứu cho thấy, chủng RB.IBE29 sở hữu *celA* mã hóa cellulase họ 5. ORF của *celA* gồm 1500 bp và mã hóa cellulase với 499 aa và 55,0 kDa. Cấu trúc bậc một của cellulase gồm chuỗi peptide tín hiệu ở đầu N, domain xúc tác họ 5 và domain CBM3 ở đầu C của chuỗi polypeptide. Trình tự (1410 bp) nhưng không bao gồm đoạn peptide tín hiệu được khuếch đại và tạo dòng thành công trong *Escherichia coli* DH5α. Phân tích giải trình tự cho thấy không có đột biến cũng như đột biến dịch khung xảy ra ở trình tự gen mục tiêu trong vectơ tái tổ hợp của khuẩn lạc dương. Nghiên cứu này cung cấp nguồn vật liệu có giá trị cho nghiên cứu tiếp theo về biểu hiện, tinh sạch và xác định đặc tính sinh học của cellulase.

Từ khóa: *Cellulase; enzyme thủy phân đường họ 5; tạo dòng gene; Bacillus velezensis RB.IBE29.*

REFERENCES

Almagro Armenteros, J.J. et al. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nature Biotechnology*, *37*(4), 420–423.

Cai, Z., et al. (2024). Introduction of Cellulolytic Bacterium *Bacillus velezensis* Z2.6 and Its Cellulase Production Optimization. *Microorganisms*, *12*, 979.

Fan, B. et al. (2018). *Bacillus velezensis* FZB42 in 2018: the gram-positive model strain for plant growth promotion and biocontrol. *Front Microbiol*, *9*, 2491.

[1]Viện Công nghệ Sinh học và Môi trường, Trường Đại học Tây Nguyên;
[2]Trường Đại học Khoa học Tự nhiên, Đại học Quốc gia TP. Hồ Chí Minh;
Tác giả liên hệ: Trần Minh Định; ĐT: 0943095079; Email: tmdinh@ttn.edu.vn.

Gasteiger, E. et al. (2005). *Protein Identification and Analysis Tools on the Expasy Server;* (In) John M. Walker (ed): The Proteomics Protocols Handbook. Humana Press.

Johnson, M. et al. (2008). NCBI BLAST: a better web interface. *Nucleic Acids Research, 36*, W5–W9.

Juturu, V. et al. (2014). *Microbial cellulases engineering, production and applications*. *Renewable and Sustainable Energy Reviews, 33*, 188–203.

Letunic, I. et al. (2021). SMART: recent updates, new developments and status in 2020. *Nucleic Acids Research, 49*(D1), D458–D460.

Nguyen, S.D. et al (2021). Combined application of rhizosphere bacteria with endophytic bacteria suppresses root diseases and increases productivity of black pepper (*Piper nigrum* L.). *Agriculture*, *11*(1),15.

Pentekhina, I. et al. (2020). Chitinase system of *Aeromonas salmonicida*, and characterization of enzymes involved in chitin degradation. *Bioscience, Biotechnology, and Biochemistry*, *84*, 1347–6947.

Tamura, K. et al. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution, 30*, 2725–2729.

Tran, D.M. et al. (2018). Identification and characterization of chitinolytic bacteria isolated from a freshwater lake. *Bioscience, Biotechnology, and Biochemistry*, *82*, 343–355.

Tran, D.M. et al. (2022a). Molecular analysis of genes involved in chitin degradation from the chitinolytic bacterium *Bacillus velezensis*. *Antonie van Leeuwenhoek, 115*, 215–231.

Tran, D.M. et al. (2022b). Expression, purification, and basic properties of a novel domain structure possessing chitinase from *Escherichia coli* carrying the family 18 chitinase gene of *Bacillus velezensis* strain RB.IBE29. *Mol Biol Rep*, *49*, 4141–4148.

Tran, D.M. et al. (2023). Complete genome sequence data of chitin-degrading *Bacillus velezensis* RB. IBE29. *Data Brief, 51*, 109815.

Tran, D.M. et al. (2024). Data on gene cloning, expression, purification, and characterization of the glycoside hydrolase family 11 from *Bacillus velezensis*. *Data Brief, 52*, 109834.

Trinh, T.H.T. et al. (2019). A potent antifungal rhizobacteria *Bacillus velezensis* isolated from black pepper (*Piper nigrum* L.). *Research on Chemical Intermediates*, *45*(11), 5309–5323.