WHOLE GENOME SEQUENCING OF *Bacillus altitudinis* AGE-B8, A POTENTIAL CELLULOLYTIC BACTERIUM ISOLATED FROM ROTTED SUGAR CANE BAGASSE

A. AGEEZ¹,²,*

1. Agricultural Genetic Engineering Research Institute (AGERI), Agricultural Research Center (ARC), 9 Gamaa St, Giza, 12619, Egypt
2. Faculty of Biotechnology, MSA University, 6 October City, Egypt

* Corresponding author: aageez@msa.eun.eg. Tel: +20 12 2766 0233 Fax: +20 23 568 9519

**Keywords:** lignocellulosic degrading, *B. altitudinis*, Phylogenetic analysis, CMCase, FPase, Whole-genome sequencing, genome annotation,

List of **Abbreviations**

*B. altitudinis*: *Bacillus altitudinis*; CMC carboxymethyl cellulose; CMCase: carboxymethyl cellulase; FPase: filter-paper enzyme; *gyrB*: gyrase B subunit; *rpoB*: RNA polymerase β subunit; SCB: Sugarcane Bagasse; *pycA*: pyruvate carboxylase.

Biofuel produced from agricultural wastes can be an important renewable and sustainable source to replace the depleting fossil fuels (Luo et al., 2019). Successful conversion of lignocellulose waste to fermentable sugars involves three successive steps: (1) size reduction, (2) pretreatment/fractionation, and (3) enzymatic hydrolysis (Kucharska et al., 2018). Enzymatic hydrolysis of natural lignocellulosic materials is considered the most difficult technological challenges that needed to be overcome (Mosier et al., 2005). Due to the limited success of protein engineering to enhance the performance of existing lignocellulose degrading enzymes, retrieving enzymes from naturally evolved biomass-degrading microbial communities offers a promising strategy for the identification of new lignocellulolytic enzymes with potentially improved activities (Rubin, 2008).

Cellulase enzyme cocktails from fungus species are favored over bacteria due to the ability of the former to produce abundant amounts of cellulolytic enzymes and often less complex than bacterial cellulase (Gusakov, 2011). Compared to its counterparts, bacteria have many favored characteristics such as i) it has a higher growth rate than fungi that led to higher
recombinant production of enzymes, ii) bacterial cellulases are more complex than provide increased function and synergy; iii) bacteria inhabit a wide variety of environments, which results in cellulolytic strains that are extremely resistant to environmental stresses, and hence able to produce cellulolytic enzymes stable under extreme conditions (Maki et al., 2009; Pham et al., 2019).

Bagasse has various higher-value uses as a second-generation biofuel, fibers for paper, xylan-based products such as xylooligosaccharides, substrate for single-cell protein, and enzymes, or other high-value microbial products (Gebbie et al., 2020; Pandey et al., 2000). Culturing, and metagenomics analysis of bagasse samples has discovered a unique microbial community compared to other lignocellulosic environments, which lead to the identification of new biomass-degrading microbes and enzymes (García-Huante et al., 2017; Mhuantong et al., 2015; Rattanachomsri et al., 2011).

Whole-genome sequencing is an innovative technology to analyze all DNA sequence of an organism efficiently. Whole-genome sequencing technology was used to investigate several potential lignocellulose degrading bacteria during recent years. Full genome sequencing was performed for various bacterial strains such as Amycolatopsis sp. strain ATCC 39116 (Davis et al., 2012), Novosphingobium sp. strain MBES04 isolated from sunken wood (Ohta et al., 2015), Pseudomonas sp. strain YS-1p (Prabhakaran et al., 2016), the β-proteobacterium Cupriavidus basilensis B-8 (Shi et al., 2013), Burkholderia sp. Strain LIG30 isolated from wet tropical forest soil (Woo et al., 2014b), and Klebsiella sp. strain BRL6-2 (Woo et al., 2014a). These studies reveal a tremendous amount of high-quality information about the agricultural waste degrading mechanisms adopted by several microorganisms.

The current report describes the isolation of cellulose-degrading bacterial isolate, AGE-B8, isolated from rotted sugarcane bagasse. Various parameters such as pH, incubation temperature, incubation time, and different carbon sources were tested for identifying the best cellulases induction conditions for the isolate. A draft assembly of the bacterial isolate genome was performed. The isolate was molecularly identified using multilocus sequence analysis as B. altitudinis. A total of 3,540 genes were identified in the produced bacterial genome assembly.

MATERIALS AND METHODS

Bacterial isolation and qualitative screening of their cellulolytic activity.

In this study, cellulose-degrading bacteria were collected from rotted bagasse. Twenty grams of the rotted bagasse were added to 180 ml sterilized distilled water in a 250 ml Erlenmeyer flask and shooked for one hour at 30°C. One ml of sample suspension was transferred into 9 ml of sterilized distilled water; serial di-
whole genome sequencing of *B. altitudinis* age-b8, a potential cellulolytic bacterium isolated from rotted sugar cane bagasse

luted to $10^{-7}$, and one hundred µl were plated on nutrient agar plates then incubated at 30°C for four days. Bacterial colonies were picked and sub-cultured to obtain a pure culture. Cellulase producing bacteria were screened on selective carboxymethyl cellulose agar containing (g/L): NaNO$_3$ 2.0, KH$_2$PO$_4$ 1.0, MgSO$_4$ .7H$_2$O 0.5, KCl 0.5, carboxymethyl cellulose sodium salt 10.0, peptone 0.2, agar 17.0. Plates were spot inoculated with a suspension of pure culture and incubated at 30°C. The endoglucanase activity of the isolated bacterial isolates was tested through the detection of hydrolyzing clear zones on agar plates containing CMC-Na and Iodine staining (Kasana et al., 2008). After 4 days, plates were flooded with gram’s iodine stain (2.0 g KI and 1.0 g iodine in 300 ml distilled water) for 3 to 5 minutes. The diameter of the zone of decolorization around each well was recorded.

**Quantitative screening of cellulolytic activity of isolated bacteria**

The pure cultures of bacterial isolates were inoculated with 3 ml (1 x $10^7$ CFU/ml) in modified basal mineral medium (g/L): KH$_2$PO$_4$, 2g; (NH$_4$)$_2$SO$_4$, 1.4g; Urea, 0.3g; CaCl$_2$·2H$_2$O, 0.3 g; MgSO$_4$·7H$_2$O, 0.3 g; peptone , 1.0 g; Tween 80, 0.2%; CoCl$_2$·6H$_2$O, 2.0 mg; FeSO$_4$·7H$_2$O, 5.0 mg; MnSO$_4$·2H$_2$O, 1.6 mg; ZnSO$_4$·7H$_2$O, 1.4 mg; Distilled water , 1000 ml. One gram of treated sugar cane bagasse (Gutierrez-Correa and Tengerdy, 1997) was added into 100 ml of the production medium (pH 5.5) in 250 ml Erlenmeyer flask then incubated at 30°C with an agitation speed of 180 rpm in a rotary shaking incubator. The cultures filtrate were collected after four days by centrifugation at 6000 rpm for 15 min and the supernatant was used as a source of crude cellulase enzyme (Vega et al., 2012). filter-paper enzyme (FPase) and carboxymethyl cellulase (CMCase) activities were determined using the method of Ghose, (1987). The activity of the enzymes was expressed in U/ml. One unit (IU) of enzyme activity is defined as the amount of enzyme required to release 1µmol of glucose in the reaction mixture per minute under specified conditions.

**Optimization of parameters for enzymatic hydrolysis**

The cellulolytic activities of the bacterial isolates were tested at varying pH ranged from 4-10, varying temperatures ranged from 30°C - 37°C, and varying incubation periods ranged from 4-10 days, while all other parameters were kept constant. For carbon source induction experiments; sugar cane bagasse (SCB), as natural cellulose and carbon methylcellulose (CMC), as artificial cellulose, were tested individually as a sole carbon source in a basal mineral media at 30°C and 170 rpm. The production media was inoculated by 3ml (1 x $10^7$ CFU/ml) and all the flasks were incubated in a shaking incubator at 180 rpm (New Brunswick, USA). Enzyme assay of both FPase and CMCase was determined for each experiment as described before.
Identification of bacterial isolate using 16S rRNA sequencing

Genomic DNA of AGE-B8 isolate was purified using GeneJET Genomic DNA Purification Kit (ThermoFisher®, USA). The Primers 27F (AGAGTTGTGCTCAG) and 1492R (TACGTTACCCTGTTACGACTT) were used to amplify nearly full-length 16S rRNA gene (Youseif et al., 2021). The PCR product was purified, and the sequencing reaction of the PCR product was carried out using an ABI BigDye Terminator V3.1 cycle sequencing kit and ABI PRISM 3500 genetic analyzer (Applied Biosystem, USA).

Library amplification and genome sequencing of AGE-B8 isolate.

Genomic DNA extraction was performed using the GeneJET Genomic DNA Purification Kit (ThermoFisher®, USA). DNA extraction quality was performed using DeNovix DS-11 (DeNovix, USA). Sequencing library was prepared by random fragmentation of the DNA, followed by 5’ and 3’ adapter ligation using trueq nano DNA library kit (Illumina, USA), and the size of PCR enriched fragments was confirmed by electrophoresis on Agilent Technologies 2100 Bioanalyzer (Agilent, USA). For cluster generation, the library was loaded into a flow cell where fragments are captured on a lawn of surface-bound oligos complementary to the library adapters. Each fragment is then amplified into distinct, clonal clusters through bridge amplification. The sequencing was performed using NovaSeq 6000 (Illumina, USA). Library preparation and sequencing were performed at macrogen (Seoul, Korea).

Genome assembly and genome annotation of AGE-8B

FastQC (V0.11.8) was used to check the high throughput sequencing data. Adapter sequences and low quality reads quality filtering of the fastaq raw files was performed using Trimmomatic (Bolger et al., 2014). Map with BWA-MEM (Galaxy Version 0.7.17.1) was used to map the filtered reads to the Bacillus altitudinis reference genome (Li and Durbin, 2009). SAMTools was used to manipulate the SAM/BAM files produced from the BWA-MEM software (Li et al., 2009). Variants (SNPs and short indels) were identified by SnpEff (v4.3t) (Cingolani et al., 2012). Genome annotation was performed using Prokka Prokaryotic genome annotation (Seemann, 2014), and RAST (Rapid Annotation using Subsystem Technology (Aziz et al., 2008).

Phylogenetic analysis

The 16S rRNA gene sequencing reads were edited and assembled using Sequencher (Gene Code Corporation, MI, USA). The relevant sequences were obtained using the nucleotide BLAST program (https://blast.ncbi.nlm.nih.gov/Blast.cgi) and aligned using ClustalW version1.8 (Altschul et al., 1997). The gryB, rpoB, pycA housekeeping genes were identified from the draft genome assembly, BLASTed in GenBank
whole genome sequencing of B. altitudinis age-b8, a potential cellulolytic bacterium isolated from rotted sugar cane bagasse

(www.ncbi.nlm.nih.gov) using the nucleotide BLAST program, and were aligned using ClustalW version1.8 (Altschul et al., 1997). The maximum likelihood (ML) algorithm (Saitou and Nei, 1987) was selected for performing the phylogenetic analyses. Phylogenetic analyses were performed using MEGA 6.06 (Kumar et al., 2016) using the Tamura-Nei model. Bootstrap support (BT) for each node was evaluated with 1000 replicates.

RESULTS AND DISCUSSION

Isolation, screening, and identification of most efficient cellulase producing bacteria

Twenty bacterial isolates were acquired based on their ability to grow on basal mineral medium containing sugarcane bagasse pulp as the sole carbon source on solid media. The hydrolyzing zone diameter and colony diameter were recorded for all tested isolates, and only isolates hydrolyzing zone diameter larger than 10 mm, ten isolates, were selected for quantitative analysis. The ten bacterial isolates showing the highest clear zone were further quantitatively screened using DNS (3,5-dinitrosalicylic acid) method (Adney and Nrel, 1996). Bacterial isolates were inoculated in production media containing Sugarcane Bagasse (SCB) as the sole carbon source for seven days at 30ºC, and the induced exoglucanase (FPase) and endoglucanase (CMCase) were recorded (Table 1). The bacterial isolates showed FPase activity in a range from 0.022 to 0.092 U/ml, while the CMCase activity was ranged from 0.31 to 0.527 U/ml. The bacterial isolate AGE-B8 showed the highest CMCase activity with 0.527 U/ml in addition to 0.092 U/ml for FPase, and was selected for further investigations. The results show a direct relationship between the enzyme concentration and the diameter of the cleared zone in the Congo-red staining method. The results are in accordance with other reports that confirmed this pattern (Bradner et al., 1999; Florencio et al., 2012; Teather and Wood, 1982).

Optimization of cellulase production

Cellulase production from AGE-B8 isolate was assessed after fermentation at the different parameters. The effect of pH on the induction of exoglucanase (FPase) and endoglucanase (CMCase) was assessed for the AGE-B8 bacterial isolate at 30ºC (Fig. 1 A). The results illustrated by Fig. (1. A) show that both Fbase and CMCase gradually increases as the pH values increased and reached maximum values at pH from 5.5 to 6.0 . At a pH of 5.5, the values were 0.532 and 0.094 U/ml for CMCase and FBase, respectively.

The results illustrated by Fig. (1) clearly show that CMCase production, expressed as enzyme activity, gradually increased as the incubation period increased reaching the maximum (0.72 U/ml ±0.004) at the tenth day, when incubated at 30ºC. The maximum activity for CMCase was 0.713 U/ml ± 0.02 on day 7 when the bacterial isolate incubated at 34ºC, and 0.80 U/ml ± 0.05 when incubated at 37ºC. The FPase production showed a similar pattern. It gradually increased as
the incubation period increased reaching the maximum (0.082 U/ml ±0.004) on the tenth day, when incubated at 30°C. The maximum activity for FPase was 0.274 U/ml ± 0.04 on day 7 when the isolate incubated at 34°C, and 0.191 U/ml ± 0.03 when incubated at 37°C.

Sugarcane bagasse and carboxymethyl cellulose were tested for their effects on cellulase production. Data presented in Fig (2) showed that Bagasse is a better sole source of carbon for cellulase production as cellulase activity was 0.072U/ml for FPase and 0.54 U/ml for CMCase (Fig. 2).

**Identification of AGE-B8 using 16S rRNA**

Nearly full-length 16S rRNA gene (1500 bp) was successfully amplified and sequenced. The 16S rRNA sequence of AGE-B8 was submitted to the GenBank under accession numbers LC593153. Nucleotide BLAST of the 16S rRNA showed the highest similarity to *B. altitudinis* strain GQYP101 (Accession number CP040514.1), *B. altitudinis* strain BIM (Accession number CP063360.1), *B. altitudinis* strain CHB19 (Accession number CP043559.1), *B. pumilus* strain TUAT1 (Accession number AP044928.1), *B. pumilus* strain SH B11 (Accession number C010997.1P), *B. cellulasensis* strain NJ-M2 (Accession number CP012329.1), *B. cellulasensis* strain GQYP101 (Accession number CP0405141.1), *B. cellulasensis* strain NJ-V2 (Accession number CP012482.1). Maximum likelihood (ML) phylogenetic tree based on 16S rRNA was generated to determine the evolutionary relatedness of the bacterial isolates AGE-B8 (Fig. 3). According to the sequence similarity, the AGE-B8 isolate was grouped and could not be differentiated from the aforementioned strains (Fig. 3). *B. pumilus*, *B. cellulasensis* and *B. altitudinis*, form one group that can be differentiated from other closely related group. Some Bacillus isolates closely related to the *B. pumilus*, are not easily distinguished from *B. cellulasensis* and *B. altitudinis* by 16S rRNA gene sequence alone, which are nearly identical in 16S rRNA gene sequence, sharing similarity over 99.5% (Liu et al., 2013).

**Genome assembly and genome annotation of AGE-B8 bacterial isolate**

For sequencing, genomic DNA was extracted, and one paired-ends library was performed. The genome was sequenced using the Novaseq 6000 (Illumina). In total, 15,882,070 reads with 2,398,192,570 bases were generated, achieving about 600X-fold coverage of the *B. pumilus, B. cellulasensis* and *B. altitudinis* reference genomes. Trimmomatic was used to remove adapter sequences and low quality reads to reduce biases in further analysis. After quality trimming at Q30, (~93.5%) of total length passed the filtering. Table (2) summarizes the resulting sequencing data and related quality metrics, before and after the quali-
whole genome sequencing of B. altitudinis age-b8, a potential cellulolytic bacterium isolated from rotted sugar cane bagasse

ty trimming. The GC% of the reads was 41.57%.

The reads were submitted to NCBI under project number PRJNA675729, BioSample number SAMN16712403, and SRA accession number SRR13084424. To my knowledge, this is the first draft genome for Bacillus altitudinis with ability for lignocellulosic degrading. The reference genome strategy was applied for building the draft genome assembly. The NCBI reference genome of the B. altitudinis (accession number GCF_004563755.2) was selected as reference for aligning the reads sequences. More than 94% of the filtered reads could be aligned to it, compared to 85% match in case of B. pumilus, B. cellulasensis. The BWA-MEM (Galaxy Version 0.7.17.1) was used to map the filtered reads to the Bacillus altitudinis reference genome. The reads bases can be aligned to ninety-two percent of the 3,867,833 bases of the B. altitudinis reference genome (Table 3). The average alignment depth was 504.14.

Genome annotation was performed using Prokka Prokaryotic genome annotation (Seemann, 2014). The resulting assembly generated 208 contigs, with 3,575,132 bases, with N50 of 51,914. The assembly identified twenty-two rRNA, 74 tRNA, one tmRNA, and 3,540 coding regions. The functional annotation was carried out by RAST (Rapid Annotation using Subsystem Technology (Aziz et al., 2008). The annotated genome revealed 232 genes potentially involved in carbohydrates including 76 Central carbohydrate metabolism genes (Fig. 4). These genes, with the mutations identified within it, can be characterized for the identification of new biomass-degrading enzymes, and testing its efficient conversion processes on bagasse that is characterized by the recalcitrant nature of its polymers.

Identification of AGE-B8 using Multilocus Sequence Analysis

To confirm the taxonomy of the AGE-B8 isolate, phylogenetic analysis based upon the Multilocus Sequence Analysis (MLSA) of gyrB, rpoB, pycA housekeeping genes was constructed. The sequence of gyrB, rpoB, pycA housekeeping genes were extracted from the AGE-B8 assembly and aligned to those of bacterial strains that showed the highest similarity of 16S rRNA to AGE-B8 strain. The maximum likelihood (ML) phylogenetic tree based on gyrB, rpoB, pycA housekeeping genes were generated and confirmed the identification of the AGE-B8 isolated as B. altitudinis isolate. The AGE-B8 can be grouped in a separate clade and could be differentiated from the B. pumilus, B. cellulasensis strains (Fig. 5). The bacteria of the B. pumilus, groups are repeatedly misnamed at the species level due to the high similarity in their 16S rRNA gene sequence. The use of housekeeping genes can be used as a molecular marker to distinguish these closely related strains (Liu et al., 2013). Inside the B. subtilis group, the gyrB gene has a better discrimination power than 16S rRNA (Wang et al., 2007). Also three concatenated housekeeping genes was used to
distinguish closely related rhizobial isolates (Youseif et al., 2014). In this study, the gyrB, rpoB, pycA housekeeping gene could successfully identify the *B. altitudinis* isolates from the *B. pumilus, B. cellulasesensis* strains.

**Genetic variation analysis of AGE-8B bacterial isolate**

Comparing with the reference genome, *B. altitudinis* (accession number GCF_004563755.2), the produced genome assembly was tested for genetic variation using the SNPeff software. After removing duplicates and identifying variants, information on each variant was gathered and classified. Compared with the reference genome, 375 deletions, 398 insertions, and 60,009 SNP were identified. Supplementary file 1 lists all deletions, insertions, and SNPs along with their annotations (Supp. File 1).

Base changes (DNA substitution) are of two types. Interchanges of purines (A <-> G), or of pyrimidines (C <-> T) are transitions, while interchanges of purine for pyrimidine bases, and vice versa, are transversions. Generally, transversions are more likely to cause amino acid sequence changes. The number of transition (Ts) and transversion (Tv), and the Ts/Tv ratio were calculated using the base change count. Base changes (DNA substitution) are of two types. Interchanges of purines (A <-> G), or of pyrimidines (C <-> T) are transitions, while interchanges of purine for pyrimidine bases, and vice versa, are transversions. A total of 40,521 transitions and 19,488 transversions were identified (Table 4). Generally, transversions are more likely to cause amino acid sequence changes.

To find out the annotation information such as amino acid changes by variants, SNPeff was used. Table (5) shows the number of variants per type (based on the representative transcript). Seventy-eight percent of the identified 53,790 variations were synonymous. Twenty-one percent were nonsense_variant (Table 5). A list of 60,765 identified variants is listed in supp. file (2).

**SUMMARY**

Bagasse is a major source of lignocellulose that is produced in a huge amount in Egypt. It can be used as a second-generation biofuel starting material, fibers for paper, and substrate for high-value microbial products. The present study describes the isolation of cellulose-degrading bacterial strain, AGE-B8, isolated from rotted sugar cane bagasse. AGE-B8 bacterium was tested for the cellulases induction under various parameters such as pH, incubation temperature, incubation time, carbon sources. The best filter-paper enzyme (FPase) activity was obtained after incubation for 7 days at 34°C, while best carboxymethyl cellulase (CMCase) activity was obtained after incubation for 7 days at 37°C. The strain was identified using multilocus sequence analysis as *Bacillus altitudinis*. A draft assembly of the AGE-B8 genome was performed. The assembly was submitted to NCBI under project number
whole genome sequencing of *B. altitudinis* age-b8, a potential cellulolytic bacterium isolated from rotted sugar cane bagasse

PRJNA675729, BioSample number SAMN16712403, and SRA accession number SRR13084424. A total of 3,540 genes were identified in the produced bacterial genome assembly. Gene annotation and genetic variation analysis was performed for the identified isolate. A detailed list of all mutations, in reference to the NCBI *Bacillus altitudinis* reference genome, was identified. The results will significantly contribute to a better understanding of the genetics of the lignocellulosic degrading bacteria.

**REFERENCES**


Florencio C., Couri S., and Farinas C. S., (2012). Correlation between agar plate screening and solid-state fermentation for the prediction of


whole genome sequencing of B. altitudinis age-b8, a potential cellulolytic bacterium isolated from rotted sugar cane bagasse


whole genome sequencing of *B. altitudinis* age-b8, a potential cellulolytic bacterium isolated from rotted sugar cane bagasse


Supporting Files

- **S1 File**: list of all deletions, insertions, and SNPs along with their annotations in the *Bacillus altitudinis* AGE-B8 genome. https://drive.google.com/file/d/1CfiYq1rrvax0AErYguNcy3MWqbqoJUjd/view?usp=sharing

- **S2 File**: A list of 60,765 identified variants in the *B. altitudinis* AGE-B8 genome. https://drive.google.com/file/d/1BfZIrlZ5g7kBule5Dg044NgtAn-5aG/view?usp=sharing
Table (1): Evaluation of Exoglucanase (FPase) activity and Endoglucanase (CMCase) activity of the cellulose-degrading bacterial isolates.

<table>
<thead>
<tr>
<th>Isolate code</th>
<th>D/d</th>
<th>FPase (IU/ml)</th>
<th>CMCase (IU/ml)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGE-B1</td>
<td>27/2.1</td>
<td>0.037 ± 0.002</td>
<td>0.352 ± 0.002</td>
</tr>
<tr>
<td>AGE-B2</td>
<td>28/2.2</td>
<td>0.032 ± 0.001</td>
<td>0.425 ± 0.004</td>
</tr>
<tr>
<td>AGE-B3</td>
<td>18/2.0</td>
<td>0.022 ± 0.001</td>
<td>0.310 ± 0.003</td>
</tr>
<tr>
<td>AGE-B4</td>
<td>29/2.1</td>
<td>0.075 ± 0.01</td>
<td>0.499 ± 0.002</td>
</tr>
<tr>
<td>AGE-B5</td>
<td>24/2.3</td>
<td>0.042 ± 0.004</td>
<td>0.400 ± 0.003</td>
</tr>
<tr>
<td>AGE-B6</td>
<td>23/1.8</td>
<td>0.056 ± 0.002</td>
<td>0.325 ± 0.001</td>
</tr>
<tr>
<td>AGE-B7</td>
<td>24/2.2</td>
<td>0.053 ± 0.003</td>
<td>0.325 ± 0.004</td>
</tr>
<tr>
<td>AGE-B8</td>
<td>31/1.9</td>
<td>0.092 ± 0.001</td>
<td>0.527 ± 0.008</td>
</tr>
<tr>
<td>AGE-B9</td>
<td>17/2.0</td>
<td>0.025 ± 0.002</td>
<td>0.410 ± 0.005</td>
</tr>
<tr>
<td>AGE-B10</td>
<td>29/2.1</td>
<td>0.086 ± 0.001</td>
<td>0.458 ± 0.002</td>
</tr>
</tbody>
</table>

“D/d”: hydrolyzed zone diameter/colony diameter, in millimeter, on agar media containing CMC as a sole carbon source.

Table (2): Raw data Stats before and after trimming using the trimmomatic software.

<table>
<thead>
<tr>
<th></th>
<th>Total read bases (bp)</th>
<th>Total number of reads</th>
<th>Total reads GC(%)</th>
<th>Q20(%)</th>
<th>Q30(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before trimming</td>
<td>2,398,192,570</td>
<td>15,882,070</td>
<td>41.58</td>
<td>96.20</td>
<td>90.73</td>
</tr>
<tr>
<td>After trimming</td>
<td>2,170,161,529</td>
<td>14,675,394</td>
<td>41.57</td>
<td>98.29</td>
<td>93.54</td>
</tr>
</tbody>
</table>
whole genome sequencing of *B. altitudinis* age-b8, a potential cellulolytic bacterium isolated from rotted sugar cane bagasse

Table (3): Mapped data Stats.

<table>
<thead>
<tr>
<th>Ref. Length (GCF_004563755.2)</th>
<th>Mapped Sites (&gt;=1x)</th>
<th>Total Reads</th>
<th>Mapped Reads</th>
<th>Mapped Bases Mean Depth</th>
</tr>
</thead>
<tbody>
<tr>
<td>3,867,833</td>
<td>3,586,315 (92.73%)</td>
<td>14,675,394</td>
<td>13,813,626 (94.13%)</td>
<td>504.14</td>
</tr>
</tbody>
</table>

Table (4): Transition, Transversion information table.

<table>
<thead>
<tr>
<th>Ref. Length</th>
<th>Number of SNPs</th>
<th>Transition</th>
<th>Transversion</th>
<th>Ts/Tv</th>
</tr>
</thead>
<tbody>
<tr>
<td>3,867,833</td>
<td>60,009</td>
<td>40,521</td>
<td>19,488</td>
<td>2.08</td>
</tr>
</tbody>
</table>

Table (5): Annotation type count.

<table>
<thead>
<tr>
<th>Type of annotation</th>
<th>Count</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>synonymous_variant</td>
<td>42,117</td>
<td>78.3%</td>
</tr>
<tr>
<td>missense_variant</td>
<td>11,398</td>
<td>21.19%</td>
</tr>
<tr>
<td>stop_gained</td>
<td>53</td>
<td>0.1%</td>
</tr>
<tr>
<td>frameshift_variant</td>
<td>50</td>
<td>0.09%</td>
</tr>
<tr>
<td>downstream_gene_variant</td>
<td>40</td>
<td>0.07%</td>
</tr>
<tr>
<td>splice_region_variant</td>
<td></td>
<td></td>
</tr>
<tr>
<td>and stop_retained_variant</td>
<td>31</td>
<td>0.06%</td>
</tr>
<tr>
<td>conservative_inframe_insertion</td>
<td>18</td>
<td>0.03%</td>
</tr>
<tr>
<td>start_lost</td>
<td>14</td>
<td>0.03%</td>
</tr>
<tr>
<td>conservative_inframe_deletion</td>
<td>14</td>
<td>0.03%</td>
</tr>
<tr>
<td>upstream_gene_variant</td>
<td>10</td>
<td>0.02%</td>
</tr>
</tbody>
</table>
Fig. (1): Effect of pH and incubation time at different incubation temperatures on the activity of FPase and CMCase for *B. altitudinis* AGE-B8. (A): Effect of different pH, at 30°C, on the production of cellulases by *B. altitudinis* AGE-B8. Effect of incubation period on the production of cellulases at incubation temperature 30°C (B), 34°C (C) and 37°C (D).

Fig. (2): Effect of SCB and CMC as sole carbon sources for the induction of FPase and CMCase.
whole genome sequencing of *B. altitudinis* age-b8, a potential cellulolytic bacterium isolated from rotted sugar cane bagasse

Fig. (3): Phylogenetic tree based upon the 16S rRNA sequences obtained by maximum likelihood (ML) method. GenBank accession numbers are indicated next to the strain name. Bootstrap values were calculated from 1000 trees compared with the sequence of standard strains. Bootstraps are shown at the branching points.
Fig. (4): Subsystem distribution of AGE-8B genes based on RAST annotation server.
whole genome sequencing of *B. altitudinis* age-b8, a potential cellulolytic bacterium isolated from rotted sugar cane bagasse

Fig. (5): ML tree based on three concatenated *gyrB*, *rpoB*, *pycA* housekeeping genes. GenBank accession numbers are indicated next to the strain name. Bootstrap values were calculated from 1000 trees compared with the sequence of standard strains. Bootstraps are shown at the branching points.