Annals of Microbiology (2018) 68:163–174
https://doi.org/10.1007/s13213-018-1327-1

ORIGINAL ARTICLE

Genetic diversity detection and gene discovery of novel glycoside hydrolase family 48 from soil environmental genomic DNA

Mei Hua[1,2] · Shanshan Yu[3] · Yu Ma[2] · Shan Chen[2,4] · Fan Li[2,4]

Received: 18 September 2017 / Accepted: 13 February 2018 / Published online: 17 March 2018
© Springer-Verlag GmbH Germany, part of Springer Nature and the University of Milan 2018

Abstract

Sequence diversity within a family of functional enzymes provides a platform for novel gene development and protein engineering to improve the properties of these enzymes for further applications. Glycoside hydrolase family 48 (GH48) is an important group of microbial cellulases. However, the genetic diversity and gene discovery of GH48 enzyme in natural environments are rarely reported. In this study, the genetic diversity of GH48 from Changbai Mountain soil was evaluated by building a clone library via a culture-independent molecular method for the first time. Results showed that the genetic diversity of GH48 in Changbai Mountain soil was different from that in thermophilic compost and marine sediment libraries, and more than 80% of the sequences exhibited the highest identity with cellulase genes from Chloroflexi. Novel GH48 genes were also cloned, and the recombinants Cel48_hm01 and Cel48_hm02 were prokaryotically expressed, purified, and characterized. Characterization results suggested that they were probably endocellulases that adopted a catalytic mechanism similar to the GH48 cellulase from *Clostridium*. This study revealed the genetic distribution of glycoside hydrolases in soil environment, described Changbai Mountain soil as a valuable source for glycoside hydrolase gene screening, and presented supplementary property data on novel GH48 from natural soil environments.

Keywords Culture-independent approach · Enzymatic characterization · Genetic diversity · GH48 · Glycoside hydrolase

Introduction

As an important component of plant cell walls, cellulose is the most abundant source of biocarbon and remarkable renewable energy (Ragauskas et al. 2006). Glycoside hydrolase (GH) is a class of enzymes in different families, which play a key role in the hydrolysis of cellulose and hemicelluloses (Segato et al. 2014; Berlemont and Martiny 2015; Talamantes et al. 2016). Glycoside hydrolase family 48 (GH48) is an important part of microbial cellulases and is present in various cellulolytic systems, including free-enzyme systems, multifunctional enzymes, and cellulosomes of cellulose-degrading microorganisms, such as *Thermobifida fusca*, *Caldicellulosiruptor saccharolyticus*, and *Piromyces equi* (Irwin et al. 2000; Guimarães et al. 2002; Steenbakkers et al. 2002). In thermophilic anaerobic bacteria, *Clostridium thermocellum* is a widely studied species that exhibits the highest known rate of cellulose degradation, and GH48 cellulases are upregulated during their growth on crystalline cellulose. Knockout of *cel48S* and *cel48Y* of *C. thermocellum* significantly inhibits cellulolytic activity (Olson et al. 2010).

It is worth noting that only one (rarely two or three) gene(s) encoding GH48 enzymes in the genomes of cellulolytic bacteria generally, which make the gene sequence a representative of excellent cellulolytic capacity to screen new genes and cellulolytic bacteria through degenerate primer amplification (Sukharnikov et al. 2012). Izquierdo et al. (2010) described a

Mei Hua and Shanshan Yu contributed equally to this work and should be considered co-first authors.
Fan Li and Shan Chen contributed equally to this work and should be considered co-corresponding authors.

✉ Shan Chen
chens093@nenu.edu.cn

✉ Fan Li
lif885@nenu.edu.cn

[1] Institute of Special Animal and Plant Sciences, Chinese Academy of Agricultural Sciences, Changchun, Jilin, China

[2] Department of Life Science, Northeast Normal University, Changchun, Jilin, China

[3] Jilin Ginseng Academy, Changchun University of Chinese Medicine, Changchun, Jilin, China

[4] Changchun, People's Republic of China

genetic diversity assessment of enriched microbial communities from a thermophilic compost by using 16S rRNA and GH48 gene. A pair of novel degenerate primers has also been designed according clostridial GH48 gene sequences to assess functional diversity in cellulolytic bacterial communities, and this pair is a desirable tool for the molecular detection and characterization of GH48 gene (Ji et al. 2012).

With advanced technologies, novel enzymes and modified enzymes have been explored for microorganism resource development in natural environments and specific industrial applications (Borrelli and Trono 2015; Santiago et al. 2016). Conventional polymerase chain reaction (PCR), combined with the degenerate primer and thermal asymmetric interlaced (TAIL)-PCR of genomic walking technique, has been extensively developed to clone library construction and gene screening in environmental microbial genomic DNA (Huang et al. 2010; Ahmad et al. 2015). However, current knowledge on GH families and different roles of their specific enzymes in natural ecosystems remains insufficient (Zhang et al. 2017; de Mendez et al. 2015; Book et al. 2014; Sizova et al., 2011). Molecular research on the genetic diversity and gene discovery of GH48 in natural environments, especially soil environments, has also been rarely performed. The well-preserved Changbai Mountain is one of the largest pristine temperate forest ecosystem in China (Zou et al. 2015). This mountain is an ideal soil sample collection zone with humus-rich and carbon circulation-complete system that provides optimal environments necessary to investigate the abundant biology diversity resources (Li et al. 2017). Meanwhile, studies related to microbial glycoside hydrolase detection of the mountain has yet to be reported. Using the PCR-based genomic walking technique, the genetic diversity of specialized cellulase family in soil environments can be revealed and target genes can be discovered via a culture-independent approach (Barbi et al. 2014). Majority of studies on cellulase production have been studied on fungi, while the degradation of bacteria cellulase is also attracting more and more attention (Payne et al. 2015; Saraihom et al. 2016; Jain and Agrawal 2018). For characterizing the bacterial cellulase gene distribution in Changbai Mountain soil, we have chosen the GH48 family gene to research, which is mainly from bacteria in CAZy database and considered to be "true cellulase" (Sukharnikov et al. 2012).

In this study, the genetic diversity of conservative regions in GH48 was detected from the microbial genomic DNA of Changbai Mountain soil. Novel GH48 genes were also cloned through the genomic walking, and the enzymatic property of recombinants was detected. Genetic diversity results confirmed that Changbai Mountain soil is a distinctive and essential complementary source of GH48 gene distributed in the environment. This study provided supplementary property data on novel GH48 from natural soil environments.

Materials and methods

Soil region

The soil region is located in the Changbai Mountain (41° 23′–42° 36′ N, 126° 55′–129° 00′ E) in Jilin Province of China. Reserve belongs to typical temperate continental climate affected by the monsoon climate. The mean annual temperature maintains from 2.9 to 7.3 °C, and the mean annual precipitation increases from 750 to 1340 mm (Li et al. 2017). Different forest soils can be distinguished according to the altitude and vegetation in Changbai Mountain, which can generally be divided into dark brown forest soil, brown coniferous forest soil, mountain tundra soil and swamp soil, and meadow soil (Zou et al. 2015).

Sample collection and genomic DNA extraction

In July 2012 (short rainy season), we sampled along the northern slope of Changbai Mountain (520–2100 m) according to different soil types. Soils from two sampling sites (> 50 m distance between) with five independent replicates were selected randomly for four soil types (10 cm in depth and 5 cm in diameter, ~ 300 g in weight each). A total of 40 samples were stored in a sterile ziplock bag at − 80 °C until further use.

A MoBio PowerSoil kit (MoBio, Carlsbad, USA) was used to extract the genomic DNA from 0.25 g of soil samples according to the manufacturer's instructions as the template for GH48 gene amplification. The approximate size of soil DNA was determined through 0.8% agarose gel electrophoresis.

Gene cloning and sequencing

The degenerate primers (GARGCNCCNGAYYAYGGICA and CCNCGYTGRWAIGTRTTDAT) revealed a conserved region between EAPD[Y/H]GH and INTFQRG (Sheng et al. 2015; Wang et al., 2010). These primers were used to detect and amplify the conserved regions of the GH48 gene by touchdown PCR program. The optimized PCR procedure was performed as follows: 5 min at 94 °C, followed by 15 cycles of 94 °C for 30 s, 65 to 50 °C (decreasing by 1 °C after each cycle) for 30 s, and 72 °C for 30 s, followed by 30 cycles of 94 °C for 30 s, 50 °C for 30 s, and 72 °C for 30 s. Final extension was at 72 °C for 10 min.

The PCR products were visualized on agarose gel, and the target bands were purified with a E.Z.N.A.® gel extraction kit (Omega, Norcross, USA). To sequence the GH48 gene fragments, we ligated the purified PCR products into the pMD18-T vector (TaKaRa, Dalian, China) and transformed them into competent *Escherichia coli* DH5α cells (TaKaRa, Dalian, China). The cells were grown on Luria–Bertani (LB) agar plates containing 100 μg ml^{-1} ampicillin, 80 μg ml^{-1} X-Gal, and 0.5 μM isopropyl-β-D-1-thiogalactopyranoside (IPTG,

Sigma-Aldrich, USA) at 37 °C for 12 h. A total of 220 positive transformants (white clones) were selected randomly, amplified by bacterial PCR with primers M13–47 (CCAGGGTT TTCCCAGTCACGAC) and RV-M (GAGCGGATAACAAT TTCACACAGG) and sequenced by Sangon Biotech Co., Ltd. (Shanghai, China) for confirmation.

Phylogenetic analysis

Based on the BLAST results (http://www.ncbi.nlm.nih.gov/Blast), the correct open reading frames were identified for each GH48 gene fragment after translation into amino acid sequence by the online ExPasy-Translate tool (http://web.expasy.org/translate/) and then aligned with the ClustalW software (Wang et al. 2014). The PHYLIP package was used to classify all the sequences into different operational taxonomic units (OTUs) and the rarefaction curves defined by 95% similarity cutoff was calculated (Izquierdo et al. 2010; Wang et al. 2014). For each OTU, one representative sequence was selected to build the phylogenetic tree by MEGA 5.0 using the neighbor-joining method (Kumar et al. 2004). Confidence for tree topologies was estimated by bootstrap values based on 1000 replicates.

Cloning and identification of novel GH48 genes

After analyzing the genetic diversity of the GH48 sequences, we confirmed the target gene for full-length sequence amplification and obtained the 5′- and 3′- flanking regions through TAIL-PCR with six cycles of nested insertion-specific primers (Huang et al. 2010) from the Genome Walking Kit according to the manufacturer's instructions (TaKaRa, Dalian, China, D316). In brief, we used the genome walking approach with the designed efficient-binding specific upstream and downstream primers (GH48Tail_Fsp1-4/GH48Tail_Rsp1-2) to anchor on the known conservative sequence. Then we selected five random degenerate primers (AD1, AD2, AD3, AD4, and AD5, Zhou et al. 2010; Liu and Whittier 1995) from the Genome Walking Kit for unknown region amplification (Table 1). The PCR products were sequenced and assembled with the conserved region to produce the full-length gene sequences. The primers GH48FL_01*Nde*IFsp and GH48FL_01*Not*IRsp with restriction enzyme sites for *Nde*I and *Not*I were designed on the basis of the sequencing results. We subsequently cloned the full-length *cel48_hm01* from the Changbai Mountain soil genomic DNA performed homology analysis by BLAST in NCBI (http://www.ncbi.nlm.nih.gov/Blast) and predicted the signal peptide by using SignalP 4.0 (http://www.cbs.dtu.dk/services/SignalP/). *cel48_hm02*, another functional gene, with different structures was cloned in the same manner with the primers GH48FL_02*Nde*IFsp and GH48FL_01*Not*IRsp.

The recombinant protein sequences were translated by online software ExPASy-Translate tool (http://web.expasy.org/translate/) and submitted to NCBI for BLASTp program in nonredundant (nr) protein sequence database. The glycoside hydrolase family, protein structure, and identity with known sequences or strains in NCBI were listed in BLAST results. Various properties and parameters, including molecular weight, theoretical molecular weight, and theoretical pI, were analyzed using ProtParam (http://web.expasy.org/protparam/).

Expression and purification of recombinant proteins

The full-length gene sequence without the signal peptide was cloned into plasmid pET-22b (+) and transformed into competent cells *E. coli* BL21 (DE3) for prokaryotic expression. The recombinant was grown in LB medium containing 100 μg ml^{-1} ampicillin at 37 °C to an A_{600} of 0.6. The expression was induced by 0.8 mM IPTG at 37 °C for 7 h. The cells were harvested and resuspended in lysis buffer (50 mM Tris–HCl buffer, 500 mM NaCl, pH 7.6) at 4 °C and lysed by sonication on ice. Centrifugation was performed at 14,000×*g* for 20 min to separate the inclusion bodies from soluble proteins. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was used to identify the recombinant protein expression location. The soluble proteins were purified by Ni-NTA affinity chromatography (GenScript, NJ, USA). The inclusion bodies were washed (50 mM Tris–HCl, 500 mM NaCl, 1 M urea, 1 mM EDTA, 0.5% TritonX-100, pH 7.6), dissolved in urea-dissolved buffer (6 M urea in wash buffer), purified by Ni-NTA with elution buffer (50 mM Tris–HCl, 500 mM NaCl, 6 M urea, 10% glycerol, 20–200 mM imidazole, and pH 7.6), gradient dialysis, and refolded (6, 4, 2, 1, and 0 M urea in elution buffer for 4 h each) at 4 °C. The refolded proteins were concentrated by a 3-kDa ultrafiltration membrane (PALL, NY, USA), and the activity was tested. SDS-PAGE was used to determine the purity and apparent molecular mass of recombinant proteins. The purified enzyme was quantified by the Bradford method (1976) with bovine serum albumin as the standard protein.

Activity assay and enzymatic property characterization

Recombinant activity was measured as D-glucose equivalents by the DNS method (Miller 1959). Substrate sodium carboxymethylcellulose (CMC-Na) (Sigma-Aldrich, USA) was dissolved in 50 mM citrate-phosphate buffer (pH 7.0). A volume of 1.9 ml of 1% (*w*/*v*) CMC-Na and 0.1 ml of purified recombinant solution were pooled together at 50 °C for 30 min. Absorbance of the reaction system was read at 540 nm. The concentration of the reducing sugar was determined against a glucose standard. One enzymatic activity (IU) was measured

Table 1 Primers used in cloning and expression of GH48 genes

Primers	Sequences (5′ → 3′)	Description
AD1	NTCGASTWTSGWGTT	For TAIL-PCR amplification of GH48 gene
AD2	WGTGNAGWANCANAGA	
AD3	AGWGNAGWANCAWAGG	
AD4	CAWCGNCNGANASGAA	
AD5	TTGNAGNACNANAGG	
GH48Tail_Fsp1	CAACCTACGGCACCTCGAAC	
GH48Tail_Fsp2	CCTGCGCTACTCGATGCTCG	
GH48Tail_Fsp3	CGGCATGTTTTATGACG	
GH48Tail_Fsp4	GGAGTGACGGCAGCCTACG	
GH48Tail_Rsp1	CCAGCGGTTGCCAGTTACC	
GH48Tail_Rsp2	GCACACGACACCATTCAACC	
GH48FL_01*Nde*IFsp[a]	GGAATTCCATATGTGTT GTCTGGTCACCTATCG	For full-length gene clone with restriction enzyme sites
GH48FL_01*Not*IRsp[a]	ATAAGAATGCGGCCGCGGCC ATTGCGATGTCG	
GH48FL_02*Nde*IFsp[a]	GGAATTCCATATGTACG CGCAACGTTTTC	
GH48FL_02*Not*IRsp	Same as GH48FL_01*Not*IRsp	

[a] Restriction enzyme cutting sites in primers were marked with underline

as 1 μmol of glucose released per minute. The specific activity was measured as the enzyme activity per milligram protein.

Optimal pH for the recombinant was tested at 37 °C in buffers of pH 3.0 to 12.0 (0.2 M Na_2HPO_4 and 0.1 M citric acid for pH 3.0–8.0, 0.2 M Tris–HCl for pH 8.0–9.0, and 0.2 M glycine–NaOH for pH 9.0–12.0). The optimal temperature was tested at 20 to 80 °C in the optimal pH buffer. After incubating the recombinant at pH 3.0–12.0 at 4 °C for 24 h, the pH stability was estimated by measuring the residual activity. Thermostability was estimated at 4 to 70 °C in the optimal pH for 3 h. The highest activity was calculated as 100%. Other relative activities were converted to percentage in each test, except thermostability test, which treated the enzyme activity at 4 °C as 100%. Error bars represent the mean ± standard deviation (SD) ($n = 3$).

Substrate specificity and kinetic parameters

The substrate specificity was determined with 1.0% CMC-Na, laminarin (Sigma-Aldrich, USA), beechwood xylan (Sigma-Aldrich, USA), filter paper (Whatman No. 1, GE Healthcare, USA), *p*NPG (Sigma-Aldrich, USA), and Avicel (Sigma-Aldrich, USA) as the substrates. Kinetic parameters (K_m and V_{max}) were determined at 0.5–10 mg/ml CMC-Na and at the optimal pH and temperature for 10 min. K_m and V_{max} were plotted according to the Lineweaver–Burk method.

Homology modeling and key catalytic residue analysis

To speculate the possible catalytic mechanism, DS Sequence, DS Modeler, and DS Validate Protein Structure modules of Discovery Studio 3.5 software (DS3.5, Accelrys, Inc., USA) were used for construction and analysis of the model of single-catalytic domain recombinant Cel48_hm02. The sequences and key catalytic residues of Cel48_hm01, Cel48_hm02, and other homologous GH48 proteins, which determined the tertiary structure and activity, were compared using DNAMAN.

Nucleotide sequence accession numbers

Accession numbers KX164806–KX164852 in GenBank were provided for the nucleotide sequences of the GH48-conserved gene sequences in this article. Accession numbers ALF08529.1 (Cel48_hm01) and ALF08530.1 (Cel48_hm02) were also assigned for the nucleotide sequences of the novel full-length genes of GH48.

Results

Conservative fragments amplification and sequences analysis

Using TAIL-PCR, the template DNA of Changbai Mountain soil obtained specific amplification products of ~420 bp

(Fig. 1, line 1). The fragments were cloned, and all the GH48 gene fragments were constructed on LB^{Amp}/X-Gal/IPTG agar plates, screened by bacterial PCR with primers M13–47 and RV-M (Fig. 1, lines 2–21).

A total of 220 clones with the correct-size insert were selected for sequencing. After removal of the mismatch and redundant sequences, 214 clones with an average size of 423 bp and 59–78% amino acid identity were detected with known GH48 cellulases. The representative sequence was ZJ_GH1, which showed the highest identity with the cellulase sequence (ABX02947.1) from *Herpetosiphon aurantiacus*.

The genetic distance was set as 0.05 (species level) for all 214 clone abundance analysis, and the OTU number was 30. Rarefaction analysis showed that 11 new OTUs were generated after 184 clones and formed a curve plateau with the number of retrievable OTUs. However, additional sampling of 30 more clonal sequences contributed no new sequences as singletons (indicating 1 clone per out) Izquierdo et al. 2010). The clone library coverage was calculated as follows (Izquierdo et al. 2010): $C\% = [1 - (n/N)] \times 100\%$, where n is the number of singletons, 11 in our results, and N is the total clone number, 214 in our results. This finding indicated that 95% sequences were characterized. In addition, the sequencing data can fully represent the genetic diversity of GH48 clone library. All the protein sequences showed identity lower than 80% with GH48 cellulase in NCBI database, suggesting the conservation and novelty of these sequences. The OTU groups with less than 75% identity were nearly 50% in all groups (14/30). The OTU groups with less than 65% identity were nearly 7% in all groups (2/30). In general, 87% of the OTU groups (26/30) showed the highest identity with endoglucanase (ABX02947.1, 78%) from *H. aurantiacus*.

An unrooted phylogenetic tree of the GH48 protein sequences was constructed from the representative sequences of the OTU groups of the genomic DNA of Changbai Mountain soil and the known reference sequences of NCBI database (Fig. 2). As shown in the tree, all the protein sequences, including the references of GH48, were divided into four groups, namely, group I, Chloroflexi; group II, Actinobacteria; group III, Proteobacteria; and group IV, Firmicutes. Chloroflexi is the major genetic groups in this tree, and 80% sequences showed the highest identity with the reference sequence ABX02947.1 (*H. aurantiacus*). All the representative sequences in this group clustered together and made a distinction between the three reference sequences. Therefore, these new protein sequences of Chloroflexi may possess close genetic relationships with the GH48 sequences of unculturable bacterium. Furthermore, the close genetic relationship became directly the basis of novel GH48 genes cloning from the genomic DNA. In group II, Actinobacteria, the representative sequences HN_GH47 and ZJ_GH24 were clustered with the reference sequences WP_014689550.1 (*Actinoplanes* sp. SE50/110), EWM68846.1 (*Micromonospora* sp. M42), KDS84721.1 (*Streptomyces fradiae*), WP_013227017.1 (*Amycolatopsis mediterranei*), WP_020543322.1 (*Nonomuraea coxensis*), and WP_029069559.1 (*Jonesia quinghaiensis*). These reference sequences were mainly from the typical marine actinomycetes (*Micromonospora* sp. and *A. mediterranei* M42) (Zhao et al. 2010; Still et al. 2014), rare actinomycetes in mangroves (*N. coxensis*) (Wang et al. 2011), and coryneform strain isolated from soda lake (*J. quinghaiensis*) (Schumann et al. 2004). As the smallest group in the phylogenetic tree, group III, Proteobacteria, only included one representative sequence and two reference sequences. Sequence ZJ_GH97 showed the closest genetic relationship with the reference sequences AEM44250.1 from uncultured bacterium and WP_018608878.1 from *Uliginosibacterium gangwon*, suggesting that this protein may have originated from β-Proteobacteria. Group IV, Firmicutes, included two representative sequences and five reference sequences. The representative sequences HN_GH76 and ZJ_GH58 displayed the closest genetic relationship with WP_027621856.1 from *Clostridium clariflavum* and ACT46163.1 from *Ruminiclostridium thermocellum*.

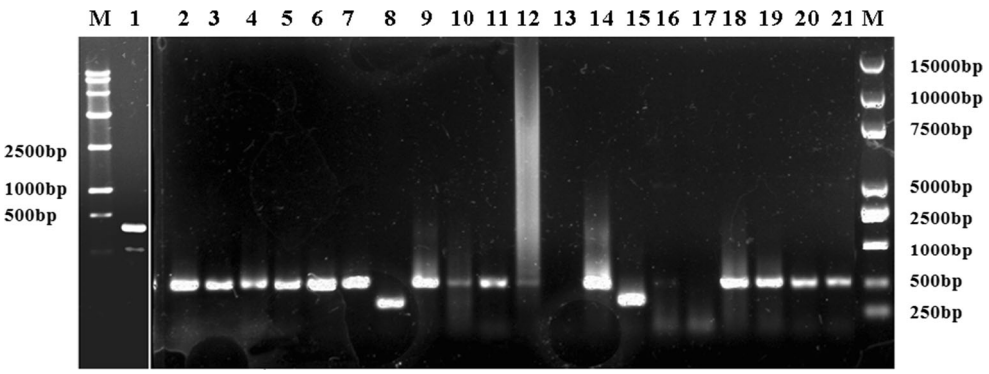

Fig. 1 Amplification and bacterial PCR of GH48 gene fragments from the Changbai Mountain soil genomic DNA. 1: GH48 gene fragments by touchdown PCR from the template DNA; 2–21: GH48 gene fragments by bacterial PCR of the template DNA; M: DL 15 kb DNA marker (TaKaRa, Dalian, China)

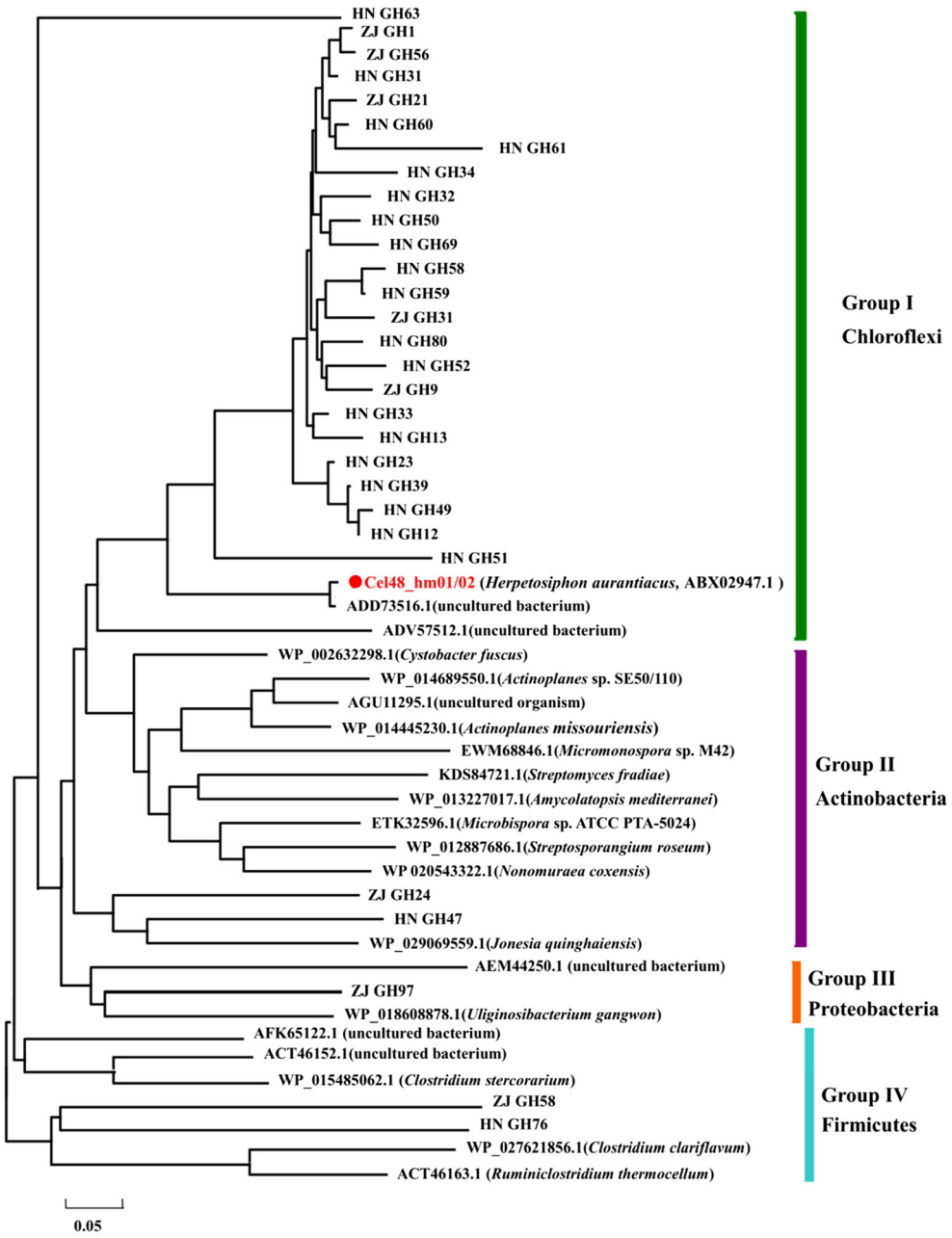

Fig. 2 Phylogenetic analysis based on the partial amino acid sequences of GH48 genes detected in the genomic DNA of Changbai Mountain soil. The full-length GH48 gene (Cel48_hm01/02) cloned from the partial amino acid sequences was marked with a closed circle in red. The reference sequences BLASTed in NCBI for Cel48_hm01/02 were marked with source strains and GenBank accession numbers in parentheses. Each sequence was analyzed by NCBI BLASTp (version 2.6.0) against the GenBank nr database. An *E* score (expect value) cutoff of 10^{-10} (default) was applied. The bootstrap values based on 1000 replications. The *scale bar* represents 0.05 amino acid substitutions per position

Cloning and sequence analysis of novel GH48 functional genes

A novel GH48 functional gene, *cel48_hm01*, was directly amplified from the soil genomic DNA by genomic walking technology with a series of primers (Table 1). Sequence analysis showed that the complete gene contained an open reading frame of 2406 bp, which encoded a putative cellulase composed of 802 amino acid residues with the N-terminal binding domain CBM_2 and the C-terminal typical catalytic domain of GH48. The molecular formula of mature protein is $C_{3962}H_{5893}N_{1065}O_{1180}S_{19}$, with a theoretical molecular weight of 87.9 kDa and theoretical pI of 6.83. Homology alignment showed that the proteins share the highest identity (62%, 100% query cover, and 0.0 *e* value with 10^{-10} cutoff) with an endoglucanase (ABX02947.1) of *H. aurantiacus* without a signal peptide.

We also amplified *cel48_hm02*, another novel functional gene, from the genomic DNA in the same way as *cel48_hm01*. The complete gene contained an open reading frame of 1917 bp, which encoded a single structural domain, or GH48 catalytic domain, in the C-terminal. Homology alignment showed that the proteins shared the highest identity (67%, 100% query cover, and 0.0 *e* value with 10^{-10} cutoff) with an endoglucanase (ABX02947.1) of *H. aurantiacus* without a signal peptide.

Expression and purification of Cel48_hm01 and Cel48_hm02

The plasmids pET-22b_*cel48_hm01* and pET-22b_*cel48_hm02* were constructed by restriction enzymes *Nde*I/*Not*I and produced in *E. coli* BL21 (DE3). Cel48_hm01 and Cel48_hm02 expressed significant inducing belts in the 90 and 70 kDa sites of SDS-PAGE test (Fig. 3). The purified Cel48_hm01 and Cel48_hm02 were recovered after renaturation by urea gradient dialysis at 4 °C and concentrated by a 3-kDa ultrafiltration membrane.

Characterization of Cel48_hm01 and Cel48_hm02

Effects of pH and temperature on the Cel48_hm01 and Cel48_hm02

The optimal pH of Cel48_hm01 was 6.0 at 50 °C, suggesting that it was a leaning acidophilic cellulase (Fig. 4a). The optimal temperature of Cel48_hm01 was 50 °C (Fig. 4b). More than 50% activity can be retained at pH 5.0–11.0 (Fig. 4c). At 4–70 °C, Cel48_hm01 was stable with activity over 70%. More than 80% activity can be retained at 60 °C for 3 h (Fig. 4d). The optimal pH and temperature level of Cel48_hm01 were very close to that of the processive endocellulase Cel48F, a major component of the *Clostridium cellulolyticum* ATCC 35319 cellulosome (Reverbel-Leroy et al. 1997). In addition, Cel48_hm01 thermostability was better than the second cellobiohydrolase of *C. thermocellum*, Cel48Y, which displayed 50% or more activities between 50 and 71 °C but no activity above 75 °C (Berger et al. 2007). Cel48_hm02 showed the same optimal pH and temperature as those of Cel48_hm01 but exhibited pH stability and thermostability slightly lower than those of Cel48_hm01.

Substrate specificity and kinetic parameters of Cel48_hm01 and Cel48_hm02

The substrate specificity of Cel48_hm01 and Cel48_hm02 was determined under optimal pH and temperature as shown in Table 2. Recombinants showed the highest activity on CMC-Na and certain activity on beechwood xylan and laminarin. Cel48_hm02 also showed weak activity on insoluble substrates (filter paper and Avicel). This phenomenon may be attributed to the homology of Cel48_hm02 protein structure with some cellobiohydrolase genes.

The kinetic parameters were determined as K_m of 0.78 mg/ml and V_{max} of 173.01 μmol/(min mg) for Cel48_hm01 and K_m of 0.66 mg/ml and V_{max} of 155.79 μmol/(min mg) for Cel48_hm02. The absence of CBM domain in Cel48_hm02 may affect the V_{max} value, which is lower than that of Cel48_hm01. K_m of Cel48_hm01 and Cel48_hm02 was better than the truncated CelF (PDB: 1fbw, 2.4 g/l toward PAS cellulose and 35 g/l toward CMC), which is a nonthermostable form of endocellulase CelF that lacks a part of or the entire duplicated C-terminal segment (Reverbel-Leroy et al. 1997).

Homology modeling of Cel48_hm02

The endoglucanase Cel48F (PDB: 1FBO) of *C. cellulolyticum* H10 was selected as the template to build the homologous model of Cel48_hm02 (Fig. 5a). A Ramachandran plot obtained using the model indicated that approximately 90% of the enzyme's residues were in the core and allowed regions. The overall verity score for the protein was 293.93. This value is higher than the expected high score of 290.689, indicating that the model is reliable.

Cel48_hm02 modeling consisted of the monomer catalytic structure $(\alpha/\alpha)_6$ barrel with a Ca^{2+} ligand, and an inverting mechanism may be adopted (Hamid et al. 2015). Key catalytic

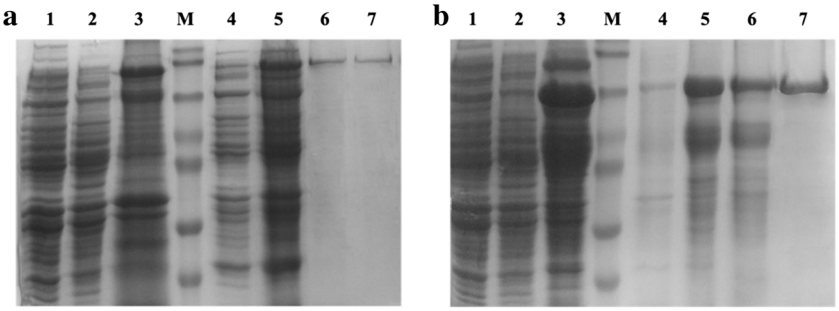

Fig. 3 SDS-PAGE of Cel48_hm01 (**a**) and Cel48_hm02 (**b**). **a** Purification of Cel48_hm01. 1: expression of the empty plasmid pET-22b(+) induced by IPTG; 2: supernatant of Cel48_hm01 without induced by IPTG; 3: total protein of Cel48_hm01 induced by IPTG; M: Blue Plus I Protein Marker (Transgen, Beijing, China; from top to bottom, 100, 70, 50, 40, 30, and 25 kDa); 4: supernatant of Cel48_hm01 induced by IPTG; 5: insoluble precipitation of Cel48_hm01 induced by IPTG; 6: purified Cel48_hm01 after Ni-NTA affinity chromatography with 20 mM imidazole; 7: purified Cel48_hm01 after Ni-NTA affinity chromatography with 200 mM imidazole. **b** Purification of Cel48_hm02 with the same order as Cel48_hm01

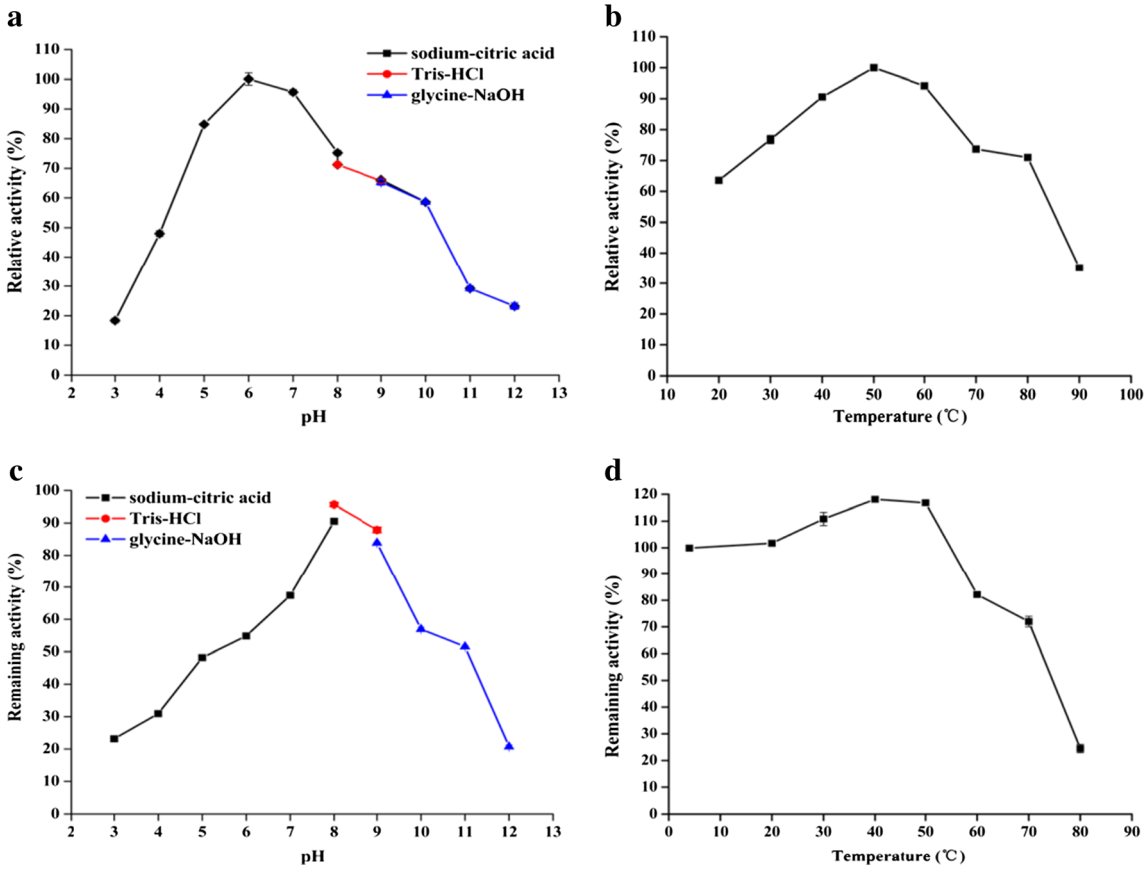

Fig. 4 Effects of pH and temperature on Cel48_hm01 activity. Activity dependence on pH and temperature (**a**, **b**), pH stability at 24 h, and thermostability at 3 h (**c**, **d**). Residual activities were assayed under optimum conditions. The buffers were marked in corresponding colors. The optimal pH was tested at 37 °C in buffers of pH 3.0 to 12.0. The optimal temperature was tested at 20 to 80 °C in the optimal pH buffer. After incubating the recombinant at pH 3.0–12.0 at 4 °C for 24 h, the pH stability was estimated by measuring the residual activity. Thermostability was estimated at 4 to 70 °C in the optimal pH for 3 h. The highest activity was calculated as 100%. Other relative activity was converted to percentage in each test, except for the thermostability test, which treated the enzyme activity at 4 °C as 100%. Error bars represent the mean ± SD ($n = 3$)

residues Glu39, Glu50, Asp238, and Tyr340 were also conserved in Cel48_hm01 and Cel48_hm02 (Glu44, Glu55, Asp230, and Tyr323 in 1FBO) according to the homologous BLAST result with known GH48 cellulases (Fig. 5b).

Discussion

In recent years, studies have used molecular biology techniques to exploit soil microbial resources (Takasaki et al. 2013; Hua et al. 2015). GH48 gene, which is not as conservative as 16S rRNA genes, often exists in genomes as a horizontally transferred gene in different species or even phyla. Cellulases from GH48 exhibit a distinct evolutionarily conserved sequence and structural features, which can be used as effective tools to differentiate cellulases from noncellulases in genomic data sets (Sukharnikov et al. 2012).

This study selected GH48 gene as a representative of cellulose-degrading functional proteins to detect the genetic diversity of GHs in Changbai Mountain soil. The phylogenetic tree constructed in this article covered Actinobacteria, Chloroflexi, and Firmicutes, the phyla where the GH48 gene all distributed (Izquierdo et al. 2010). As one of the largest bacterial phyla in Changbai Mountain soil (Shen

Table 2 Substrate specificity of Cel48_hm01 and Cel48_hm02

Substrate (1%)	Specific activity[a, b] (U/mg)	
	Cel48_hm01	Cel48_hm02
CMC-Na	243.73 ± 0.59	106.04 ± 0.43
Filter paper	31.07 ± 0.07	6.53 ± 0.04
Laminarin	104.55 ± 0.59	64.60 ± 0.26
Beechwood xylan	98.84 ± 0.78	44.21 ± 0.33
Avicel	7.56 ± 0.15	2.10 ± 0.02
*p*NPG	0	0

[a] The activity was measured under optimum pH and temperature conditions, except for the filter paper and Avicel incubation at 50 °C for 120 min

[b] Values represent the mean ± SD ($n = 3$) relative to control samples

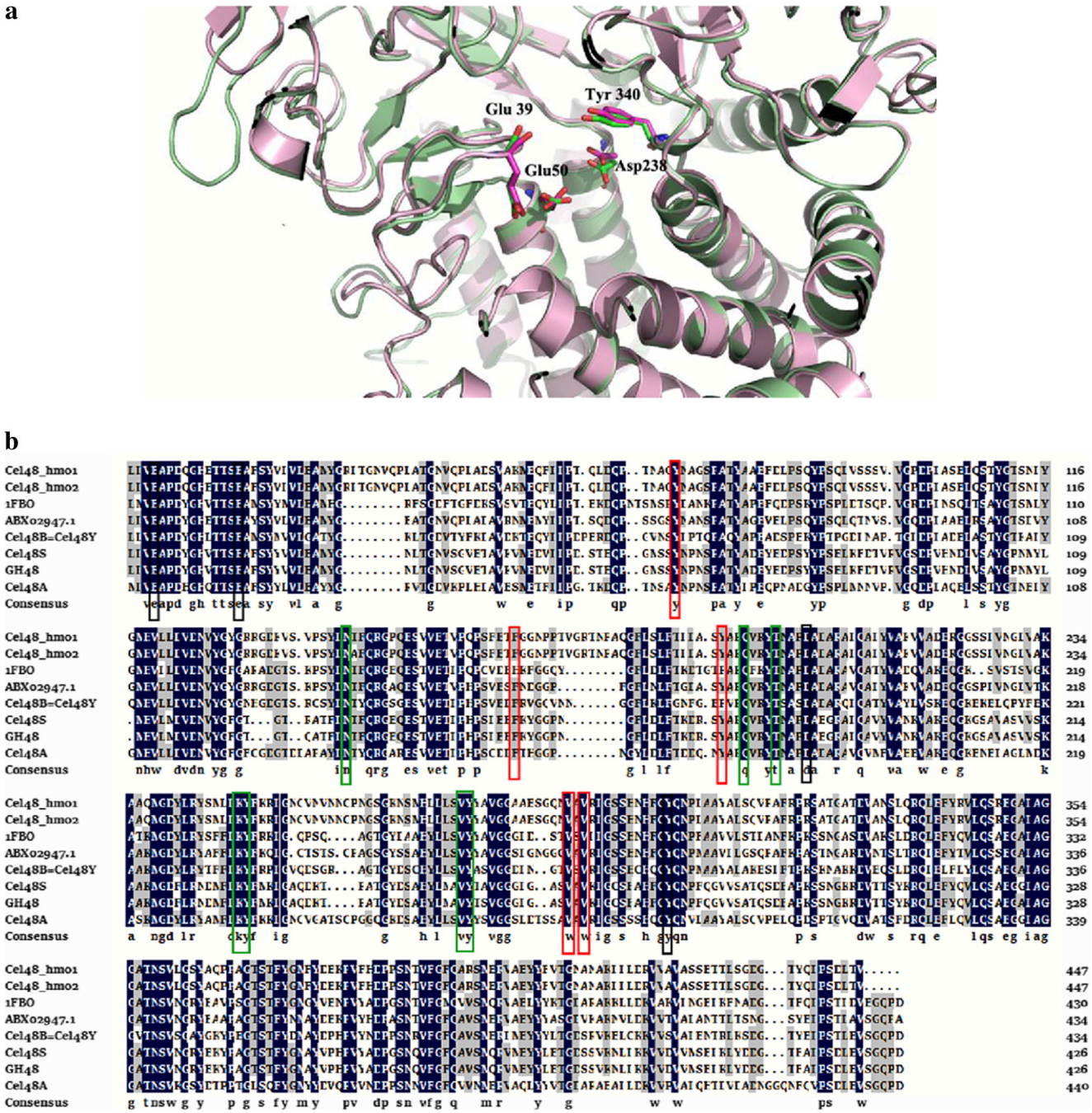

Fig. 5 Structure analysis and sequence BLAST of Cel48_hm01 and Cel48_hm02. **a** Partial superimposed polypeptide backbones of Cel48_hm02 (green) and Cel48F (PDB: 1FBO, pink). Possible key catalytic residue sites of Cel48_hm02 (Glu39, Glu50, Asp238, and Tyr340) conserved with Cel48F were marked. **b** Sequence BLAST of Cel48_hm01 and Cel48_hm02 with known GH48 cellulases. Key catalytic residues were marked with different color boxes. 1FBO: modeling template. The conservative catalytic residues (Glu44, Glu55, Asp230, and Tyr323) were marked with black boxes; ABX02947.1: endoglucanase from *H. aurantiacus* DSM 785; Cel48B = Cel48Y: endo-β-1,4-glucanase from *C. thermocellum* ATCC 27405; Cel48S: cellulase from *C. thermocellum* ATCC 27405 with the conservative catalytic residues (N204, Q247, T251, K301, Y302, W326, and Y327) marked with green boxes; GH48: cellulase from *R. thermocellum* DSM 1313; Cel48A: cellulase from *T. fusca* YX with the conservative catalytic residues (Tyr97, Phe195, Tyr213, Trp313, and Trp315) marked with red boxes

et al. 2013), Proteobacteria only accounts for 3.5% sequences in the phylogenetic tree. Consistent with the conclusions of Berlemont and Martiny (2013), this result suggested that the microbial diversity and GH48 genetic diversity of Changbai Mountain soil were different and not clearly correlated. More than 80% sequences in the phylogenetic tree showed the highest identity with cellulase belonging to *H. aurantiacus* of Chloroflexi, which match the result of Pereyra et al.

(2010). This finding suggested that *H. aurantiacus* and related species belonging to Chloroflexi may be the dominant strains as sources of GH48 genes in Changbai Mountain soil. To our knowledge, no specialized identity research on *H. aurantiacus* cellulase is available. However, the NCBI database includes the whole-genome sequence and annotation information for *H. aurantiacus* strain DSM 785, which contains 11 genes related to cellulose decomposition (Kiss et al. 2011). In our study, sequence information may be an important supplementary source of cellulase genes from *Herpetosiphon* strains in soil.

At present, few studies on GH48 cellulase gene analysis offered applicable primers, such as the report of Izquierdo et al. and de Menezes et al. In the enrichment thermophilic compost library of Izquierdo et al., all the GH48 sequences come from the clostridia organisms (Izquierdo et al. 2010). In the enriched marine sediment library of Ji et al., beyond the clostridium sequences, clones accounting for 28% showed 71% amino acid sequence similarity to the GHF48 identified in *Herpetosiphon aurantiacus* of Chloroflexi (Ji et al. 2012). In our research, we also detected a large number of sequences having close phylogenetic relationship with *H. aurantiacus* of Chloroflexi. The results suggested that the sequences from *H. aurantiacus* of Chloroflexi may be an important component of GH48 gene. It also suggested that the genetic diversity of GH48 gene of Changbai Mountain soil is different from that of unnatural environments, such as the biocompost (Izquierdo et al. 2010). However, it is worth noting that, if more diverse or universal primers can be chosen, we may have higher abundance sequences beyond Chloroflexi. So the sequences detected in our library could be a useful complement for the GH48 gene and be used for more universal GH48 primers design.

Thorough structural studies on GH48 enzyme have focused on type strains (Reverbel-Leroy et al. 1997; Guimarães et al. 2002; Berger et al. 2007), but obtaining a novel GH48 gene independent of microbial culture has yet to be reported. This study used molecular cloning technology to identify novel GH48 genes directly from the soil sample genomic DNA, thereby offering a feasible approach for future enzyme exploitation. The cloned recombinants conservatively shared the same polysaccharide protein catalytic residues with cellobiohydrolase Cel48S from *C. thermocellum* ATCC 27405 (Asn204, Gln247, Thr251, Lys301, Tyr302, Trp326, and Tyr327) (Guimarães et al. 2002) and exocellulase Cel48A from *T. fusca* YX (Tyr97, Phe195, Tyr213, Trp313, and Trp315) (Kostylev et al. 2014). According to characterization and homologous model results, Cel48_hm01 and Cel48_hm02 exhibited some similar properties and the senior structural homology with Cel48F, endocellulase from *C. cellulolyticum* with exo- and endo-activities (Reverbel-Leroy et al. 1997). This phenomenon indicated that the recombinants probably adopt the same catalytic mechanism as that of Cel48F. GH48 genes in these important cellulolytic thermophiles mainly exist in the form of partial cellulosome or independent soluble cellulases and play an essential role in these bacterial cellulolytic systems (Devillard et al. 2004; Kyeremeh et al. 2016). As the only characterized cellulases, which are not from thermophilic anaerobic microbes, Cel48_hm01/Cel48_hm02 were suitable for the exploration of evolutionary relationships and the enzymatic engineering of superior industrial cellulase preparation.

Our results showed that Changbai Mountain soil is rich in glycoside hydrolase genetic resources. The synergistic biodegradation of cellulose involves various enzymes. Thus, we speculated that Changbai Mountain soil must contain several other GHs family gene, such as GH9 cellulase genes, which are often paired with the GH48 gene, GH10 and GH11 hemicellulase genes, which are exclusive for xylan degradation, and GH5, which includes the most widespread and multifunctional cellulase genes. These enzymes act coordinately to achieve highly efficient cellulose degradation and recycling of organic carbon sources in the Changbai Mountain soil environment. We amplified several GH5 cellulase genes directly from the genomic DNA of Changbai Mountain soil and successfully achieved their prokaryotic expression in *E. coli* (Hua et al. 2015). A follow-up research on the structure–function relationship of these cellulases and the exploitation of more cellulase genes from Changbai Mountain soil is underway.

Conclusions

The genetic diversity of GH48 gene from the microbial genomic DNA of Changbai Mountain soil in China was investigated, and our results revealed that the genetic diversity of Changbai Mountain soil is different from that of thermophilic compost and marine sediment libraries. Changbai Mountain soil is a valuable GH48 gene source for screening. Novel GH48 genes from the soil environment were also cloned, and the enzymatic properties of Cel48_hm01 and Cel48_hm02 were examined. They exhibited the substrate specificity of endocellulases and possessed conservative key amino acids of GH48 cellulases. This study provided available resource for novel GH48 gene discovery and supplementary property data on novel GH48 in soil environments.

Acknowledgments This work was supported by the National Natural Science Foundation of Jilin, China (No. 20140101137JC).

References

Ahmad A, Niwa Y, Goto S, Kobayashi K, Shimizu M, Ito S, Usui Y, Nakayama T, Kobayashi H (2015) Genome-wide screening of salt tolerant genes by activation-tagging using dedifferentiated calli of *Arabidopsis* and its application to finding gene for *Myo*-inositol-1-P-synthase. PLoS One 10:e0115502

Barbi F, Bragalini C, Vallon L, Prudent E, Dubost A, Fraissinet-Tachet L, Marmeisse R, Luis P (2014) PCR primers to study the diversity of expressed fungal genes encoding lignocellulolytic enzymes in soils using high-throughput sequencing. PLoS One 9:e116264

Berger E, Zhang D, Zverlov VV, Schwarz WH (2007) Two noncellulosomal cellulases of *Clostridium thermocellum*, Cel9I and Cel48Y, hydrolyse crystalline cellulose synergistically. FEMS Microbiol Lett 268:194–201

Berlemont R, Martiny AC (2013) Phylogenetic distribution of potential cellulases in bacteria. Appl Environ Microbiol 79:1545–1554

Berlemont R, Martiny AC (2015) Genomic potential for polysaccharide deconstruction in bacteria. Appl Environ Microbiol 81:1513–1519

Book AJ, Lewin GR, McDonald BR, Takasuka TE, Doering DT, Adams AS, Blodgett JAV, Clardy J, Raffa KF, Fox BG, Currie CR (2014) Cellulolytic *Streptomyces* strains associated with herbivorous insects share a phylogenetically linked capacity to degrade lignocellulose. Appl Environ Microbiol 80:4692–4701

Borrelli GM, Trono D (2015) Recombinant lipases and phospholipases and their use as biocatalysts for industrial applications. Int J Mol Sci 16:20774–20840

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248–254

De Menezes AB, Prendergast-Miller MT, Poonpatana P, Farrell M, Bissett A, Macdonald LM, Toscas P, Richardson AE, Thrall PH (2015) C/N ratio drives soil actinobacterial cellobiohydrolase gene diversity. Appl Environ Microbiol 81:3016–3028

Devillard E, Goodheart DB, Karnati SKR, Bayer EA, Lamed R, Miron J, Nelson KE, Morrison M (2004) *Ruminococcus albus* 8 mutants defective in cellulose degradation are deficient in two processive endocellulases, Cel48A and Cel9B, both of which possess a novel modular architecture. J Bacteriol 186:136–145

Guimarães BG, Souchon H, Lytle BL, David Wu JH, Alzari PM (2002) The crystal structure and catalytic mechanism of cellobiohydrolase CelS, the major enzymatic component of the *Clostridium thermocellum* cellulosome. J Mol Biol 320:587–596

Hamid SBA, Islam MM, Das R (2015) Cellulase biocatalysis: key influencing factors and mode of action. Cellulose 22:2157–2182

Hua M, Zhao SB, Zhang LL, Liu DB, Xia HM, Li F, Chen S (2015) Direct detection, cloning and characterization of a glucoside hydrolase from forest soil. Biotechnol Lett 37:1227–1232

Huang H, Wang G, Zhao Y, Shi P, Luo H, Yao B (2010) Direct and efficient cloning of full-length genes from environmental DNA by RT-qPCR and modified TAIL-PCR. Appl Microbiol Biotechnol 87: 1141–1149

Irwin DC, Zhang S, Wilson DB (2000) Cloning, expression and characterization of a family 48 exocellulase, Cel48A, from *Thermobifida fusca*. Eur J Biochem 267:4988–4997

Izquierdo JA, Sizova MV, Lynd LR (2010) Diversity of bacteria and glycosyl hydrolase family 48 genes in cellulolytic consortia enriched from thermophilic biocompost. Appl Environ Microbiol 76:3545–3553

Jain L, Agrawal D (2018) Performance evaluation of fungal cellulases with dilute acid pretreated sugarcane bagasse: a robust bioprospecting strategy for biofuel enzymes. Renew Energy 115: 978–988

Ji S, Wang S, Tan Y, Chen X, Schwarz W, Li F (2012) An untapped bacterial cellulolytic community enriched from coastal marine sediment under anaerobic and thermophilic conditions. FEMS Microbiol Lett 335:39–46

Kiss H, Nett M, Domin N, Martin K, Maresca JA, Copeland A, Lapidus A, Lucas S, Berry KW, Glavina Del Rio T, Dalin E, Tice H, Pitluck S, Richardson P, Bruce D, Goodwin L, Han C, Detter JC, Schmutz J, Brettin T, Land M, Hauser L, Kyrpides NC, Ivanova N, Göker M, Woyke T, Klenk HP, Bryant DA (2011) Complete genome sequence of the filamentous gliding predatory bacterium *Herpetosiphon aurantiacus* type strain (114-95^T). Stand Genomic Sci 5:356–370

Kostylev M, Alahuhta M, Chen M, Brunecky R, Himmel ME, Lunin VV, Brady J, Wilson DB (2014) Cel48A from *Thermobifida fusca*: structure and site directed mutagenesis of key residues. Biotechnol Bioeng 111:664–673

Kumar S, Tamura K, Nei M (2004) MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform 5:150–163

Kyeremeh IA, Charles CJ, Rout SP, Laws AP, Humphreys PN (2016) Microbial community evolution is significantly impacted by the use of calcium isosaccharinic acid as an analogue for the products of alkaline cellulose degradation. PLoS One 11:e0165832

Li L, Xing M, Lv J, Wang X, Chen X (2017) Response of rhizosphere soil microbial to *Deyeuxia angustifolia* encroaching in two different vegetation communities in alpine tundra. Sci Rep 7:43150

Liu YG, Whittier RF (1995) Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. Genomics 25:674–681

Miller GL (1959) Use of dinitrosalicylic acid reagent for determination of reducing sugar. Anal Chem 31:426–428

Olson DG, Tripathi SA, Giannone RJ, Lo J, Caiazza NC, Hogsett DA, Hettich RL, Guss AM, Dubrovsky G, Lynd LR (2010) Deletion of the Cel48S cellulase from *Clostridium thermocellum*. Proc Natl Acad Sci 107:17727–17732

Payne CM, Knott BC, Mayes HB, Hansson H, Himmel ME, Sandgren M, Ståhlberg J, Beckham GT (2015) Fungal cellulases. Chem Rev 115: 1308–1448

Pereyra LP, Hiibel SR, Prieto Riquelme MV, Reardon KF, Pruden A (2010) Detection and quantification of functional genes of cellulose-degrading, fermentative, and sulfate-reducing bacteria and methanogenic archaea. Appl Environ Microbiol 76:2192–2102

Ragauskas AJ, Williams CK, Davison BH, Britovsek G, Cairney J, Eckert CA, Frederick WJ Jr, Hallett JP, Leak DJ, Liotta CL, Mielenz JR, Murphy R, Templer R, Tschaplinski T (2006) The path forward for biofuels and biomaterials. Science 311:484–489

Reverbel-Leroy C, Pages S, Belaich A, Belaich JP, Tardif C (1997) The processive endocellulase CelF, a major component of the *Clostridium cellulolyticum* cellulosome: purification and characterization of the recombinant form. J Bacteriol 179:46–52

Santiago M, Ramírez-Sarmiento CA, Zamora RA, Parra LP (2016) Discovery, molecular mechanisms, and industrial applications of cold-active enzymes. Front Microbiol 7:140

Saraihom S, Kobayashi DY, Lotrakul P, Prasongsuk S, Eveleigh DE, Punnapayak H (2016) First report of a tropical *Lysobacter enzymogenes* producing bifunctional endoglucanase activity towards carboxymethylcellulose and chitosan. Ann Microbiol 66: 907–919

Schumann P, Cui X, Stackebrandt E, Kroppenstedt RM, Xu L, Jiang C (2004) *Jonesia quinghaiensis* sp. nov., a new member of the suborder Micrococcineae. Int J Syst Evol Microbiol 54:2181–2184

Segato F, Damásio ARL, De Lucas RC, Squina FM, Prade RA (2014) Genomics review of holocellulose deconstruction by aspergilli. Microbio Mol Biol Rev 78:588–613

Shen C, Xiong J, Zhang H, Feng Y, Lin X et al (2013) Soil pH drives the spatial distribution of bacterial communities along elevation on Changbai Mountain. Soil Biol Biochem 57:204–211

Sheng P, Li Y, Marshall SDG, Zhang H (2015) High genetic diversity of microbial cellulase and hemicellulase genes in the hindgut of *Holotrichia parallela* larvae. Int J Mol Sci 16:16545–16559

Sizova MV, Izquierdo JA, Panikov NS, Lynd LR (2011) Cellulose-and xylan-degrading thermophilic anaerobic bacteria from biocompost. Appl Environ Microbiol 77:2282–2291

Steenbakkers PJ, Freelove A, Van Cranenbroek B, Sweegers BM, Harhangi HR, Vogels GD, Hazlewood GP, Gilbert HJ, Op Den

Camp HJM (2002) The major component of the cellulosomes of anaerobic fungi from the genus *Piromyces* is a family 48 glycoside hydrolase. DNA Seq 13:313–320

Still PC, Johnson TA, Theodore CM, Loveridge ST, Crews P (2014) Scrutinizing the scaffolds of marine biosynthetics from different source organisms: Gram-negative cultured bacterial products enter center stage. J Nat Prod 77:690–602

Sukharnikov LO, Alahuhta M, Brunecky R, Upadhyay A, Himmel ME, Lunin VV, Zhulin IB (2012) Sequence, structure, and evolution of cellulases in glycoside hydrolase family 48. J Biol Chem 287: 41068–41077

Takasaki K, Miura T, Kanno M, Tamaki H, Hanada S, Kamagata Y, Kimura N (2013) Discovery of glycoside hydrolase enzymes in an Avicel adapted forest soil fungal community by a metatranscriptomic approach. PLoS One 8:e55485

Talamantes D, Biabini N, Dang H, Abdoun K, Berlemont R (2016) Natural diversity of cellulases, xylanases, and chitinases in bacteria. Biotechnol Biofuels 9:133

Wang F, Xu XX, Qu Z, Wang C, Lin HP, Xie QY, Ruan JS, Sun M, Hong K (2011) Nonomuraea wenchangensis sp. nov., isolated from mangrove rhizosphere soil. Int J Syst Evol Micrbiol 61:1304–1308

Wang G, Huang X, Ng TB, Lin J, Ye XY (2014) High phylogenetic diversity of glycosyl hydrolase family 10 and 11 xylanases in the sediment of Lake Dabusu in China. PLoS One 9:e112798

Wang GZ, Wang YR, Yang PL, Luo HY, Huang HQ, Shi PJ, Meng K, Yao B (2010) Molecular detection and diversity of xylanase genes in alpine tundra soil. Appl Microbiol Biotechnol 87:1383–1393

Zhang Q, Liang G, Guo T, He P, Wang X, Zhou W (2017) Evident variations of fungal and actinobacterial cellulolytic communities associated with different humified particle-size fractions in a long-term fertilizer experiment. Soil Biol Biochem 113:1–13

Zhao W, Zhong Y, Yuan H, Wang J, Zheng H, Wang Y, Cen X, Xu F, Bai J, Han X, Lu G, Zhu Y, Shao Z, Yan H, Li C, Peng N, Zhang Z, Zhang Y, Lin W, Fan Y, Qin Z, Hu Y, Zhu B, Wang S, Ding X, Zhao GP (2010) Complete genome sequence of the rifamycin SV-producing *Amycolatopsis mediterranei* U32 revealed its genetic characteristics in phylogeny and metabolism. Cell Res 20:1096–1008

Zhou JP, Huang HQ, Meng K, Shi PJ, Wang Y, Luo HY, Yang PL, Bai YG, Yao B (2010) Cloning of a new xylanase gene from *Streptomyces* sp. TN119 using a modified thermal asymmetric interlaced-PCR specific for GC-rich genes and biochemical characterization. Appl Biochem Biotechnol 160:1277–1292

Zou Y, Sang W, Wang S, Thomas EW, Liu Y, Yu Z, Wang C, Axmacher JC (2015) Diversity patterns of ground beetles and understory vegetation in mature, secondary, and plantation forest regions of temperate Northern China. Ecol Evol 5:531–542