Appl Microbiol Biotechnol (2007) 75:319–328
DOI 10.1007/s00253-006-0820-9

Cloning and identification of novel cellulase genes from uncultured microorganisms in rabbit cecum and characterization of the expressed cellulases

Yi Feng · Cheng-Jie Duan · Hao Pang · Xin-Chun Mo · Chun-Feng Wu · Yuan Yu · Ya-Lin Hu · Jie Wei · Ji-Liang Tang · Jia-Xun Feng

Received: 28 October 2006 / Revised: 19 December 2006 / Accepted: 20 December 2006 / Published online: 11 January 2007

Abstract A metagenomic cosmid library was prepared in *Escherichia coli* from DNA extracted from the contents of rabbit cecum and screened for cellulase activities. Eleven independent clones expressing cellulase activities (four endo-β-1,4-glucanases and seven β-glucosidases) were isolated. Subcloning and sequencing analysis of these clones identified 11 cellulase genes; the encoded products of which shared less than 50% identities and 70% similarities to cellulases in the databases. All four endo-β-1,4-glucanases and all seven β-glucosidases, respectively, belonged to glycosyl hydrolase family 5 (GHF 5) and family 3 (GHF 3) and formed two separate branches in the phylogenetic tree. Ten of the 11 cloned cellulases exhibited highest activities at pH 5.5~7.0 and 40~55°C, a condition similar to that in the rabbit cecum. All the four endo-β-1,4-glucanases could hydrolyze a wide range of β-1,4-, β-1,4/β-1,3- or β-1,3/β-1,6-linked polysaccharides. One endo-β-1, 4-glucanase gene, *umcel5G*, was overexpressed in *E. coli*, and the purified recombinant enzyme was characterized in detail. The enzymes cloned in this work represented at least some of the cellulases operating efficiently in the rabbit cecum. This work provides the first snapshot on the cellulases produced by bacteria in rabbit cecum.

Y. Feng · C.-J. Duan · H. Pang · X.-C. Mo · C.-F. Wu · Y. Yu · Y.-L. Hu · J. Wei · J.-L. Tang · J.-X. Feng
Guangxi Key Laboratory of Subtropical Bioresources Conservation and Utilization, The Key Laboratory of Ministry of Education for Microbial and Plant Genetic Engineering, and College of Life Science and Technology, Guangxi University, Nanning, Guangxi 530005, People's Republic of China

J.-X. Feng (✉)
College of Life Science and Technology, Guangxi University, The Eastern Campus, 75 Xiuling Road, Nanning, Guangxi 530005, People's Republic of China
e-mail: feng@public.nn.gx.cn

Keywords Cellulase · Cloning · Metagenome · Rabbit cecum · Expression · Enzyme · Properties

Introduction

Cellulose is the most abundant biomass in nature, composed of repeating cellobiose units linked by β-1,4-glucosidic bonds. Cellulose can be degraded to glucose through the synergistical hydrolysis of three classes of cellulase, including endo-β-1,4-glucanase (EC3.2.1.4), cellobiohydrolase (EC3.2.1.91) and β-glucosidase (EC3.2.1.21; Lynd et al. 2002). Glucose can be easily fermented into useful chemicals, such as ethanol, which can be used as an environmentally friendly biofuel.

Cellulase can be produced by a broad range of organisms, including microbes, plants and animals. The digestive tracts of herbivores harbor symbiotic microorganisms that help the hosts to digest the cellulosic feed (Mackie 1997). Large numbers of microbes found in rumen are responsible for the degradation of cellulose (Krause et al. 2003). Cecum is the most important digestive organ for some nonruminants, such as horse and rabbit. Symbiotic microorganisms in rabbit cecum produce fibrolytic enzymes to break down the plant cell walls (McLaughlin and Chiasson 1990). Cellulolytic microorganisms in rumen and their cellulases have been intensively studied (Krause et al. 2003; Kamra 2005), whereas those in cecum have been less often investigated. Only studies concerning the distribution of cellulase activity in the rabbit digestive tracts, the changes of the enzyme activity in rabbit cecum during growth and the impacts of diet on the enzyme activity and the microflora have been conducted (Gidenne 1997). No study on cloning cellulase gene from the symbionts in rabbit cecum has been reported.

It is generally considered that more than 99% of the microorganisms in nature are unculturable (Amann et al. 1995). Even for the intensively studied rumen, more than 85% of its microorganisms have not yet been cultivated in laboratory (Krause et al. 2003). The study on the diversity of the bacteria presenting in rabbit cecum also indicated that the rabbit cecal flora contained lots of microorganisms not described previously (Abecia et al. 2005). Culture-independent approaches are developed to access the genetic resources of microbial species that have escaped scientific investigation so far. These approaches normally involve the extraction of metagenomic DNA from environmental samples to obviate the cultivation of each organism. The environmental DNA is subsequently cloned into certain vectors to construct metagenomic libraries that are screened for genes of interest by DNA–DNA hybridization, polymerase chain reaction (PCR) or activity screening (Cowan et al. 2005). Cellulase genes cloned from metagenomes have been described in several reports. One endo-β-1,4-glucanase gene was identified from gene libraries constructed with DNA isolated from the microbial consortia in a thermophilic, anaerobic digester maintained on lignocellulosic feedstocks (Healy et al. 1995). Rees et al. (2003) constructed DNA libraries of lake water samples and their enrichment cultures and sequenced two endo-β-1,4-glucanase genes screened from the libraries. One endo-β-1,4-glucanase gene was cloned from the microbial consortia in the enrichment cultures of soil (Voget et al. 2003). Nine endo-β-1,4-glucanase genes and one β-glucosidase gene were cloned from rumen metagenome (Ferrer et al. 2005). Walter et al. (2005) cloned one β-glucosidase gene from the large-bowel microbiota of mouse. Very recently, a soil metagenome-derived halotolerant endoglucanase was carefully characterized (Voget et al. 2006).

The goal of this study was to clone novel cellulase genes from the microbes in rabbit cecum by an activity-based metagenomic approach, characterize the diversity of the cellulases in cecal system and try to search for cellulases with potential application in biocatalysis.

Materials and methods

Construction and screening of the metagenomic library of the contents of rabbit cecum

Before the sample was collected, adult rabbits were fed with only grass for 2 weeks to enrich the fibrolytic microorganisms in their ceca. Metagenomic DNA was prepared according to Zhou et al. (1996) with minor modifications. DNA was isolated from cecum contents that had been washed with potassium–phosphate buffer (0.18 M, pH 7.2) to remove the microbial cells that did not bind to the fibrous pellet. Cells in the pellet were lysed with 5 mg/ml lysozyme and 1.0% sodium dodecyl sulfate (SDS). Crude DNA was purified with Sephadex G200 and polyvinylpolypyrrolidone (PVPP) as described by Kuske et al. (1998). DNA fragments of 30~50 kb were recovered by electroelution for the construction of the library by using pWEB::TNC Cosmid Cloning Kit (Epicentre) according to the manufacturer's instructions.

Colonies of the library were replica plated onto Luria-Bertani (LB) agar plates containing different substrates for screening for cellulase activity. Carboxymethyl cellulose (CMC, low viscosity, Sigma) was applied for screening for endo-β-1,4-glucanase activity as described by Teather and Wood (1982). Esculin hydrate and ferric ammonium citrate (Sigma) were used for detecting β-glucosidase activity according to the method described by Eberhart et al. (1964). To detect cellobiohydrolase activity, 4-methylumbelliferyl-β-D-cellobioside (4-MUC, Sigma) was employed as described by Reinhold-Hurek et al. (1993). The clones showing activity toward 4-MUC, but not esculin hydrate and ferric ammonium citrate, were assumed to express cellobiohydrolase activity. The plasmids of the positive clones were retransformed into host *Escherichia coli* EPI100 to confirm that the activities were due to the cloned DNA.

Molecular analyses of the cellulase genes

Subcloning was performed to localize the cellulase genes and to shorten the inserts for effective sequencing and biochemical characterization. The smallest fragments expressing cellulase activities were sequenced. Possible open reading frames (ORFs) were identified with the ORF finder at National Center for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov). Comparisons of sequences with those in the databases were made with BlastN, BlastX and BlastP (BLAST, basic local alignment search tool) at NCBI. Module structures of the enzymes were predicted by simple modular architecture research tool (SMART; http://smart.embl-heidelberg.de). Amino acid sequences of some selected glycosyl hydrolase family (GHF) 3 and GHF 5 members, the cloned cellulases and their highest matches from the Blast analyses were used for the phylogenetic analyses. Multialignments were preformed online at ClustalW (http://www.ebi.ac.uk/clustalw), and phylogenetic tree was generated with ClustalX1.83 and MEGA2.1 using neighbor-joining method.

Protein and enzyme assays

Cytoplasmatic extracts were prepared with the clear sonicated cell lysates of the subclones harboring the shortest inserts expressing cellulase activity. Cells of 150-ml overnight cultures of the subclones were harvested, resuspended in 2.5-ml citrate–phosphate buffer (McIlvaine buffer) of

pH 7.0 and sonicated. The sonicates were centrifuged, and the supernatants were collected as cytoplasmatic extracts for the subsequent experiments.

Protein was quantified by the method of Bradford (1976) with bovine serum albumin as a standard. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was performed according to Laemmli (1970). To measure endoglucanase activity, 10-μl cytoplasmatic extract that contained 20~30 mg/ml proteins was added to 200-μl-indicated buffer containing 1% (*w*/*v*) CMC. Reactions were carried out for 15 min at temperatures according to each experiment. The released reducing sugars were measured as D-glucose equivalents as described by Miller (1959). One unit (U) of the endoglucanase activity was defined as the amount of enzyme releasing 1 μmol of reducing sugar per minute. Specific activity was defined as the number of activity units per gram or milligram of protein. *p*-Nitrophenyl-β-D-glucopyranoside (*p*-NPG, Sigma) was used as a substrate for the characterizations of the β-glucosidases. β-Glucosidase activity was determined by measuring the amount of *p*-nitrophenol (*p*-NP) generated from *p*-NPG according to Odoux et al. (2003). One unit (U) of the β-glucosidase activity was defined as the amount of enzyme releasing 1 μmol of *p*-NP per minute.

Zymogram analyses of the active subclones

SDS-PAGE was performed with the cytoplasmatic extracts of the active subclones harboring the shortest inserts to fractionate the proteins. The gels were then soaked in a refolding buffer to allow the renaturations of the denatured proteins as described by Feng et al. (2000). For the zymogram analysis of the endoglucanases, the treated gel was placed on a fresh LB agar plate containing 0.5% (*w*/*v*) CMC, and the gel covered plate was incubated at 37°C for 1 h followed by staining and destaining as described above in the library screening. For the zymogram analysis of the β-glucosidases, 1 ml of a solution containing 0.1% (*w*/*v*) esculin hydrate and 0.2% (*w*/*v*) ferric ammonium citrate was evenly spread on the treated gel. The black bands appearing on the gel showed the activity of β-glucosidase after 1 h incubation at 37°C.

Characterization of the cloned endoglucanases and β-glucosidases

The cytoplasmatic extracts of the active subclones were employed for the characterization of the cloned cecal cellulases. Optimal pH were tested by the enzyme assays at 37°C in buffers of different pH (citrate–phosphate buffer, pH 3.0~7.0; 0.1 M sodium–phosphate buffer, pH 6.0~8.0; 0.1 M Tris–HCl buffer, pH 7.0~9.0; 0.1 M glycine–NaOH buffer, pH 8.6~10.0). Optimal temperatures were determined by the enzyme assays at the optimal pH and temperatures from 25 to 65°C.

To investigate the substrate specificity of the endoglucanases, CMC was replaced in enzyme assays by other polysaccharides under optimal conditions with different reaction time, 15 min for lichenan (Sigma) and barley glucan (Sigma) and 120 min for 2-hydroxyethyl cellulose (Sigma), methyl cellulose (Sigma), oat spelt xylan (Sigma), birch wood xylan (Sigma), laminarin (Sigma) and Avicel (Fluka). One percent (*w*/*v*) *p*-NPG and *p*-nitrophenyl-β-D-cellobioside (*p*-NPC, Sigma) were also used in the substrate tests of the endoglucanases. The enzyme assays were run for 120 min under optimal conditions, and the enzyme activities were determined by the measurement of the generated *p*-NP as described above.

Expression of *umcel5G* and purification of the recombinant Umcel5G

The endoglucanase gene harbored in clone RC1, designated *umcel5G*, was heterologously expressed in *E. coli* as a fusion protein with a 6×His tag on the N-terminal to facilitate the purification. The complete ORF of *umcel5G* was amplified by PCR. The expression plasmid was constructed by cloning the PCR product into vector pET-30a(+) (Novagen) and transformed into host *E. coli* BL21 (DE3) pLysS (Novagen). The expression cells were cultivated in LB medium containing 25 μg/ml kanamycin and 34 μg/ml chloramphenicol at 37°C and induced with 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG) when the OD_{600nm} of 0.5 was reached. The recombinant protein in the cytoplasmatic extract of the expression cell was purified primarily by affinity chromatography with nickel–nitrilotriacetic acid agarose resin (Ni–NTA, Qiagen), then by size-exclusion chromatography on an AKTA Explorer 100 system (Amersham Biosciences) with column Superdex 75 (Amersham Biosciences).

Characterization of the recombinant Umcel5G

Ten microliters of the diluted purified enzyme (protein concentration=0.016 mg/ml) was applied in the enzyme assay described above to characterize the biochemical properties of the recombinant Umcel5G. pH stability of the recombinant Umcel5G was studied as described by Eckert and Schneider (2003). Thermal stability was evaluated according to the method of Inoue et al. (2005). For the determination of Km and Vmax values, reactions were carried out under optimal condition with CMC of different concentrations, ranging from 3.33 to 33.33 mg/ml.

Hydrolysis products of cello-oligosaccharides by the recombinant Umcel5G were detected by thin layer chromatography (TLC) as described by Sugimura et al. (2003)

after 1-μl diluted purified enzyme was incubated, respectively, with 10-μl sugar solutions containing 10 mg/ml of cellobiose (G2), cellotriose (G3), cellotetraose (G4) or cellopentaose (G5; Sigma) at 55°C for 2 h. Metal ions, chelating agent (ethylenediamine tetraacetic acid [EDTA]) and surfactants (SDS and Triton X-100) were added to optimal reaction systems to investigate their effects on the enzyme activity.

Nucleotide sequence accession number

The sequences of the cloned cellulases were deposited in the GenBank database, and their accession numbers were listed in Table 1.

Results

Construction and screening of the metagenomic library of the contents of rabbit cecum

A cosmid library containing about 32,500 clones was constructed with the metagenomic DNA isolated from the contents of the rabbit cecum. *Bam*HI restriction analyses of the plasmids of 14 randomly chosen clones showed that the inserts ranged from 22 to 47 kb with different restriction patterns (data not shown), which indicated that the library presented a good randomness of the cloned DNA. The average size of inserted DNA was 35.1 kb, and the total capacity of the library was estimated to be 1.14×10^9 bp.

Four independent clones expressing endoglucanase activity and seven independent clones expressing β-glucosidase activity were isolated from the library. All these positive clones were confirmed to harbor cellulase genes by retransformation. No clone expressing cellobiohydrolase activity was obtained.

Sequence analyses of the cloned cellulase genes

Inserts in the initial positive clones were subcloned to 2.1~7.0 kb for sequencing. Sequence analyses of the cloned cellulase genes were showed in Table 1. Each subclone harbored an insert that contained a complete ORF coding for either an endo-β-1,4-glucanase or a β-glucosidase. No nucleotide homology for the overall sequence of each ORF to known cellulase genes was found by BlastN. The deduced products of the four cloned endo-β-1,4-glucanase genes consisted of 395~526 amino acids (aa) with the predicted molecular masses of 43,697~58,785 Da and were most similar to endoglucanases from *Bacillus* and *Pecto-*

Table 1 Sequence analyses of the cellulase genes cloned from the rabbit cecum

Subclone name (insert size of initial clone, kb)	Accession numbers of cellulase genes	Most homologous proteins	aa Identity/ similarity (%)	Conserved domains and catalytic residues
RC1 (41.5)	DQ182491	*Bacillus cellulosilyticus* endoglucanase B (P06565)	46/64	Cellulase (GHF5): aa94~360 E226 (P), E316 (N)
RC2 (39.2)	DQ182492	*Bacillus subtilis* DLG endo-β-1,4-glucanase (M16185)	46/66	Cellulase (GHF5): aa116~376 E242 (P), E332 (N)
RC3 (35.5)	DQ916112	*Pectobacterium atrosepticum* endoglucanase N precursor (Q59394)	46/61	Cellulase (GHF5): aa216~482 E348 (P), E438 (N)
RC5 (38.0)	DQ916113	*Pectobacterium atrosepticum* endoglucanase N precursor (Q59394)	48/64	Cellulase (GHF5): aa225~491 E357 (P), E447 (N)
RG2 (27.0)	DQ182493	*Clostridium beijerincki* NCIMB 8052 β-glucosidase-related glycosidase (EAP61663)	45/61	GHF 3 C domain: aa31~305 GHF 3 domain: aa545~807 D772 (N)
RG3 (37.0)	DQ182494	*Clostridium beijerincki* NCIMB 8052 β-glucosidase-related glycosidase (EAP61663)	48/64	GHF 3 C domain: aa31~294 GHF 3 domain: aa545~807 D772 (N)
RG11 (29.0)	DQ916114	*Butyrivibrio fibrisolvens* H17c β-glucosidase A (P16084)	32/48	GHF 3 C domain: aa35~309 GHF 3 domain: aa551~811 D776 (N)
RG12 (28.5)	DQ916115	*Butyrivibrio fibrisolvens* H17c β-glucosidase A (P16084)	33/46	GHF 3 C domain: aa35~307 GHF 3 domain: aa547~809 D774 (N)
RG14 (31.8)	DQ916116	*Butyrivibrio fibrisolvens* H17c β-glucosidase A (P16084)	33/48	GHF 3 C domain: aa45~320 GHF 3 domain: aa563~822 D777 (N)
RG20 (22.0)	DQ916117	*Reinekea* sp. MED297 β-glucosidase-related glycosidase (EAR10786)	33/48	GHF 3 C domain: aa49~321 GHF 3 domain: aa561~823 D788 (N)
RG25 (26.3)	DQ916118	*Reinekea* sp. MED297 β-glucosidase-related glycosidase (EAR10786)	32/49	GHF 3 C domain: aa38~312 GHF 3 domain: aa555~814 D779 (N)

The functions of the catalytic residues are indicated in the brackets after each residue
P represents the proton donor and N represents the nucleophile

bacterium (46~48% aa identity and 61~66% aa similarity; Table 1). All four endo-β-1,4-glucanases belonged to GHF 5 based on their module analyses. Each of them contained only one catalytic domain (cellulase domain). The putative products of all cloned β-glucosidase genes had similar protein lengths (854~865 aa) and theoretical molecular masses (91,296~93,653 Da), and their amino acid sequences shared 32~48% identity and 46~64% similarity with their closest known β-glucosidases (Table 1). All of the seven β-glucosidases belonged to GHF 3, possessing a GHF 3 C-terminal domain and a GHF 3 domain. The multialignment using ClustalW revealed the catalytic

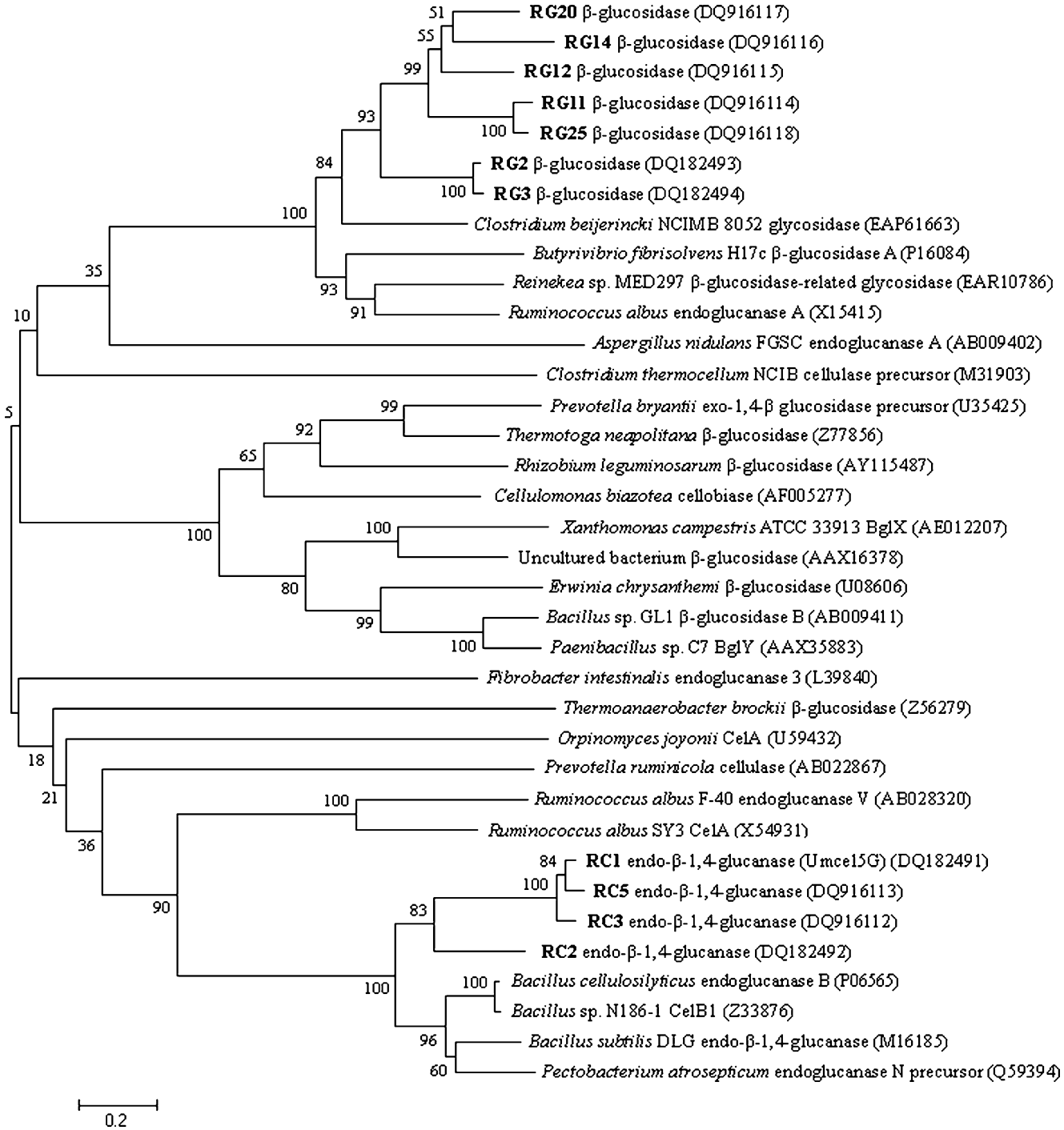

Fig. 1 Phylogenetic analyses of GHF 5 endo-glucanases and GHF 3 β-glucosidases. The multialignments were performed in ClustalX1.83. The resulted alignments were then used in MEGA2.1 to establish the phylogenetic tree by the method of neighbor-joining. One thousand boostrap replications and Poission correction were carried out. This is an unrooted phylogenetic tree of GHF 5 and GHF 3 enzymes. Database accession numbers are shown in brackets after each enzyme

Table 2 Optimal pH and temperature of the cytoplasmatic extracts of the active subclones

Subclone name and encoded enzyme	Endo-β-1,4-glucanase				β-Glucosidase						
	RC1	RC2	RC3	RC5	RG2	RG3	RG11	RG12	RG14	RG20	RG25
Optimal pH	6.5~7.0	6.0~7.0	6.0~7.0	6.5~7.0	6.5~7.0	5.5~6.0	5.0~6.0	5.5~6.0	5.5~7.0	7.5~8.5	5.5
Optimal temperature (°C)	45~55	45~55	50	50	40~45	50~55	45~50	45	50~55	35~45	45~50

The cytoplasmatic extracts that contained 20~30 mg/ml proteins were used to determine the optimal pH and temperatures of the cloned endo-β-1,4-glucanases and the β-glucosidases with 1% (*w*/*v*) CMC and *p*-NPG as substrates, respectively. pH optima were tested by the enzyme assays at 37°C in buffers of different pH (citrate–phosphate buffer, pH 3.0~7.0; 0.1 M sodium–phosphate buffer, pH 6.0~8.0; 0.1 M Tris–HCl buffer, pH 7.0~9.0; 0.1 M glycine–NaOH buffer, pH 8.6~10.0). Temperature optima were determined by the enzyme assays at the optimal pH and temperatures from 25 to 65°C. No activity was detected for the *E. coli* EPI100 harboring the empty cosmid vector.

residues of these cellulases. The putative proton donor (Glu) and nucleophile (Glu) were predicted for each cloned endoglucanase. As the catalytic proton donor of the GHF 3 β-glucosidase (Glu or His) is very uncertain and needs to be investigated and confirmed (Bhatia et al. 2005), only the nucleophiles (Asp) were predicted for the cloned β-glucosidases in this paper (Table 1).

Phylogenetic analyses of GHF5 endoglucanases and GHF 3 β-glucosidases were illustrated in Fig. 1. The four cecal endo-β-1,4-glucanases fell into one branch of the phylogenetic tree mainly consisting of endoglucanases from *Bacillus* species. Seven cloned β-glucosidases fell into another branch shared by the β-glucosidases from *Clostridium*, *Butyrivibrio* and *Ruminococcus* species.

Characterization of the cloned cellulases with the cytoplasmatic extracts of the active subclones

As the subclones for sequencing harbored 2.1- to 7.0-kb inserts and might specify more than one cellulase activity, zymograms were performed for each of the 11 active subclones. The cytoplasmatic extract of each subclone displayed only one apparent active band with the expected molecular size in the respective zymograms (data not shown), suggesting that the enzyme produced by each cloned cellulase gene made a dominant contribution to the cellulase activity of the corresponding cytoplasmatic extract. The optimal pH and temperatures of all the cytoplasmatic extracts, respectively, fell into pH 5.5~7.0 and 40~55°C, except that of RG20 that showed highest activity at pH 7.5~8.5 (Table 2).

In the substrate specificity tests for the endo-β-1,4-glucanases, all the employed polysaccharides could be hydrolyzed to different extents (Table 3). The four endo-β-

Table 3 Substrate specificity of the cytoplasmatic extracts of the subclones expressing the cloned rabbit cecal endo-β-1,4-glucanases

Substrate	Specific activity (U/g)			
	RC1	RC2	RC3	RC5
Lichenan (β-1,3/4-glucan)	184.4(3.36)	98.3 (3.21)	126.8 (2.47)	83.2 (1.96)
Barley glucan (β-1,3/4-glucan)	87.2 (0.75)	21.7 (2.26)	35.1 (1.86)	28.8 (1.47)
Carboxymethylcellulose (β-1,4-glucan)	53.4 (2.20)	74.8 (2.81)	53.4 (1.90)	55.5 (2.45)
2-Hydroxyethyl cellulose (β-1,4-glucan)	2.8 (0.16)	3.6 (0.09)	2.0 (0.14)	3.0 (0.09)
Methyl cellulose (β-1,4-glucan)	1.8 (0.15)	3.6 (0.23)	1.4 (0.13)	2.7 (0.12)
Avicel (β-1,4-glucan)	1.8 (0.13)	3.3 (0.20)	2.1 (0.18)	3.1 (0.10)
Xylan from birchwood (β-1,4-xylan)	2.7 (0.15)	3.5 (0.19)	2.2 (0.13)	3.7 (0.25)
Xylan from oat spelt (β-1,4-xylan)	2.1 (0.09)	2.2 (0.15)	1.4 (0.08)	2.4 (0.10)
Laminarin (β-1,3/6-glucan)	0.3 (0.02)	0.9 (0.12)	0.4 (0.05)	0.8 (0.06)
p-Nitrophenyl-D-cellobioside	2.3 (0.11)	1.6 (0.15)	–	–
p-Nitrophenyl-D-glucopyranoside	0.3 (0.03)	0.2 (0.03)	–	–

Ten microliters cytoplasmatic extracts of the active subclones with the protein concentration of 20~30 mg/ml was added to 200-μl citrate–phosphate buffer containing 1% (*w*/*v*) substrate, and enzyme assays were carried out under each optimal condition (Table 2). Reaction time were 15 min for lichenan, barley glucan and CMC, and 120 min for the other substrates. No activity was detected for the *E. coli* EPI100 harboring the empty cosmid vector. The specific activity is given in units per gram of proteins of the crude enzymes. Standard deviations were shown in the brackets behind the specific activities.

1,4-glucanases exhibited the highest activity toward lichenan (83.2~184.4 units per gram of protein, U/g) followed by CMC (53.4~74.8 U/g) and barley glucan (21.7~87.2 U/g). Activities toward 2-hydroxyethyl cellulose, methyl cellulose, Avicel, oat spelt xylan and birch wood xylan were low (1.4~3.7 U/g). All four endoglucanases could also hydrolyze laminarin, a β-1,3/1,6-glucan, but to a lesser extent (0.3~0.9 U/g). The endoglucanases expressed by RC1 and RC2 could hydrolyze *p*-NPG and *p*-NPC with low activities, whereas those expressed by RC3 and RC5 could not (Table 3).

Expression of *umcel5G* and purification of the recombinant Umcel5G

Among the four endoglucanases, the one encoded by subclone RC1 was selected for further study, as it presented the highest activity toward lichenan and the widest substrate spectrum (Table 3). The expression of the endoglucanase gene *umcel5G* harbored in RC1, and the purification of the recombinant enzyme Umcel5G were analyzed with SDS-PAGE (Fig. 2). The extract of the induced expression cells (Fig. 2, lane 4) was distinct from that of the uninduced cells (Fig. 2, lane 3) with a specific band of about 45 kDa as expected. During the purification, the specific activity of the enzyme was raised from 4.22 U/mg for the clear cell lysate (Fig. 2, lane 5) to 49.51 U/mg after Ni–NTA affinity chromatography (Fig. 2, lane 6) and 56.56 U/mg after size-exclusion chromatography (Fig. 2, lane 7). Thus, the enzyme had been purified by 13-fold. This purification procedure was repeated five times, and similar results were obtained.

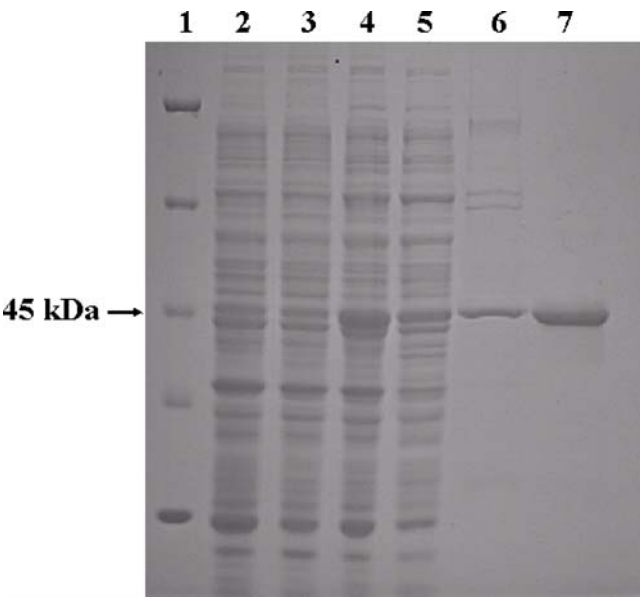

Fig. 2 Analyses of the expression of *umcel5G* and the purification of the recombinant protein on SDS-PAGE (8% polyacrylamide gel) stained with Coomasie brilliant blue G250. *Lane 1*, protein molecular weight marker (116.0, 66.2, 45.0, 35.0, 25.0 kDa); *lane 2*, total protein of BL21(DE3) pLysS harboring empty pET-30a(+) as control; *lane 3*, total protein of uninduced expression cells, BL21(DE3) pLysS harboring the recombinant *umcel5G* in pET-30a(+); *lane 4*, total protein of expression cells induced by 0.5 mM IPTG; *lane 5*, cleared lysate of induced expression cells; *lane 6*, sample purified by Ni–NTA column; *lane 7*, sample purified by size-exclusion chromatography

Characterization of the purified recombinant Umcel5G

The optimal pH of the recombinant Umcel5G toward CMC was approximately pH 6.0~6.5 (Fig. 3a), similar as tested with the cytoplasmatic extract of the subclone of RC1 (Table 2). Little activity was detected at pH 10.0 and below pH 3.5. The enzyme retained more than 90% activity after storage for 24 h at broad pH ranging from 5.0 to 9.0 (Fig. 3b). The recombinant Umcel5G acted efficiently at 35~55°C, and the optimal temperature was 55°C (Fig. 3c). The enzyme was stable for the 30-min incubation at temperatures below 50°C with more than 90% of the activity remained and was unstable at temperatures above 55°C (Fig. 3d). Km and Vmax of the recombinant Umcel5G toward CMC were 16.07 mg/ml and 417.5 U/mg, respectively.

The hydrolysis of cello-oligosaccharides was illustrated on a TLC plate (Fig. 4). G5 was completely degraded by the recombinant Umcel5G to produce predominantly G2 and G3 with a tiny amount of G1 (glucose; Fig. 4, lane 5). G4 was partially hydrolyzed to release mainly G2 with G3 as a minor product and a trace amount of G1 (Fig. 4, lane 4). G3 was slightly degraded to G1 and G2 (Fig. 4, lane3). No hydrolysis was observed for G2 (Fig. 4, lane 2).

The addition of $CoCl_2$, $CaCl_2$ and $CrCl_2$ resulted in a strong stimulation of activity to 137, 122 and 119%, respectively, whereas $CuCl_2$ and $ZnCl_2$ dramatically reduced the enzyme activity to 18 and 53%, respectively. Chelating agent EDTA and anionic surfactant SDS were also effective inhibitors, reducing the enzyme activity to 47 and 34%, respectively. KCl, LiCl, $MgCl_2$, $FeCl_2$, $FeCl_3$ and nonionic surfactant Triton X-100 showed no significant effect on the enzyme activity.

Discussion

In this study, metagenomic DNA was isolated from the contents of the rabbit cecum by direct lysis method without separating the microbial cells from the sample because cellulolytic microorganisms often bind tightly on the surface of the fibrous pellet (Michalet-Doreau et al. 2001), purified by size exclusion chromatography using a column containing Sephadex G200 and PVPP and cloned into a cosmid vector. Eleven cellulase genes encoding endo-β-1,4-glucanase and β-glucosidase were identified from the metagenomic library. The predicted products of these genes share less than 50% aa identities and 70% aa similarities

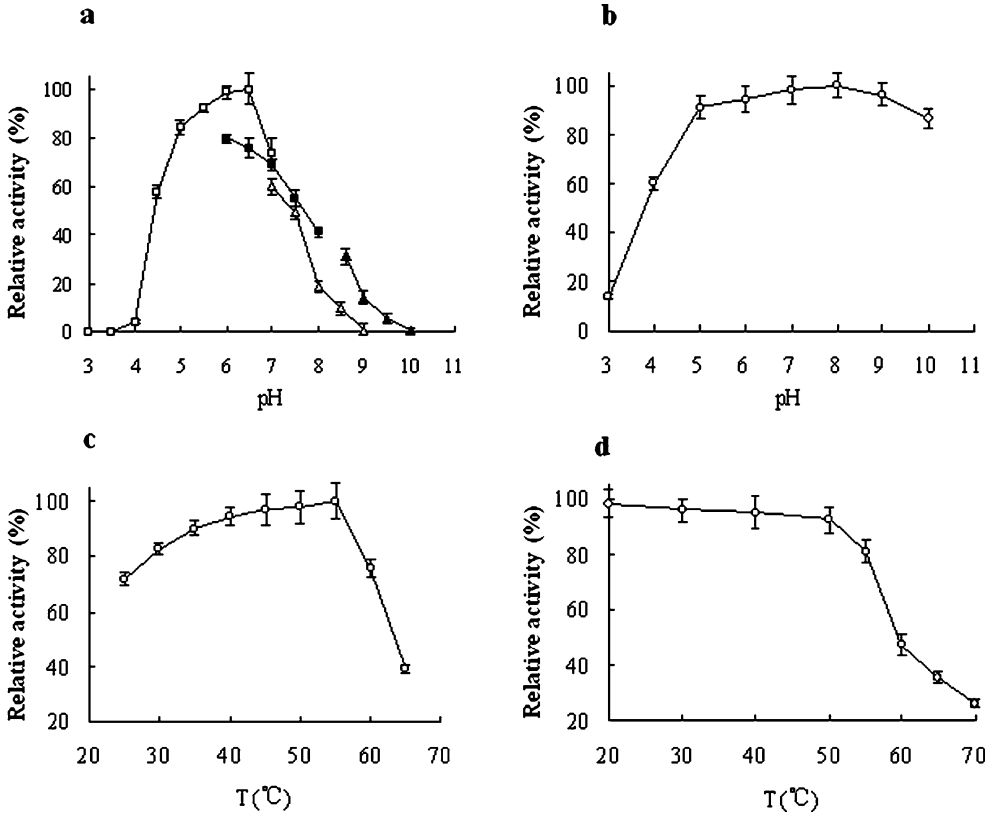

Fig. 3 Effects of pH and temperature on enzyme activity and stability of the recombinant Umcel5G. **a** Determination of the optimal pH of the recombinant Umcel5G. Enzyme assays were performed at indicated pH at 37°C for 15 min. *Open squares*, citrate–phosphate buffer, pH 3.0~7.0; *closed squares*, 0.1 M sodium–phosphate buffer, pH 6.0~8.0; *open triangles*, 0.1 M Tris–HCl buffer, pH 7.0~9.0; *closed triangles*, 0.1 M glycine–NaOH buffer, pH 8.6~10.0. **b** pH stability of the recombinant Umcel5G. Activity was measured under optimal condition (citrate–phosphate buffer of pH 6.5, 55°C, 15 min) after the purified enzyme was incubated in buffers (citrate–phosphate buffer, pH 3.0~7.0; 0.1 M Tris–HCl buffer, pH 8.0~9.0; 0.1 M glycine–NaOH buffer, pH 10.0) at 4°C for 24 h. **c** Determination of the optimal temperature of the recombinant Umcel5G. Activity was measured at pH 6.5 (citrate–phosphate buffer) at the indicated temperatures for 15 min. **d** Thermal stability of the recombinant Umcel5G. Activity was measured under optimal condition (citrate–phosphate buffer of pH 6.5, 55°C, 15 min) after the incubation of the enzyme at indicated temperatures for 30 min

with their closest homologues in the database, indicating their novelty. The source bacteria of these genes were unknown because the full-length original clones were not sequenced, and not any 16S rRNA genes linked with these identified cellulase genes were revealed by sequencing. Phylogenetic analysis revealed that these cloned cellulases cluster into two independent groups separated from the other known cellulases, but they are quite similar within the groups. Ferrer et al. (2005) found that the cellulases cloned from the rumen of a dairy cow fall into different families. Among the nine ruminal endo-β-1,4-glucanases they identified, six belong to GHF 5, two belong to GHF 26 and one may belong to an unidentified family. One ruminal glycosyl hydrolase is a member of GHF 3. In this study, however, the diversity of the cloned rabbit cecal cellulases is limited because all four cloned endoglucanases belong to GHF 5 and all seven β-glucosidases are members of GHF 3.

Most of the cloned cecal cellulases prefer to work in a condition (pH 5.5~7.0 and temperature 40~55°C) similar to that in the rabbit cecum (the pH of the fluid of the rabbit cecum was 6.5, and the body temperature of the rabbit was 39°C, data not shown). For example, the β-glucosidase expressed by subclone RG2 showed highest activity at pH 6.5~7.0 and temperature 40~45°C. The four cecal endo-β-1,4-glucanases can degrade a wide range of soluble and insoluble polysaccharides linked by β-1,4, β-1,4/β-1,3, or β-1,3/β-1,6 glucosidic bonds. Therefore, the cellulases cloned in this work represented at least some of the cellulases operating efficiently in the rabbit cecum.

All the cloned endo-β-1,4-glucanases of this work can hydrolyze laminarin, which is a glucan primarily linked by β-1,3 glucosidic bonds with some β-1,6 bonds at the branch points. Only one GHF 5 endo-β-1,4-glucanase cloned from *Bacillus* sp. BP-23 was reported to have this ability (Blanco et al. 1998). All the four cloned endo-β-1,4-glucanases can also hydrolyze Avicel and xylan, whereas many of cloned GHF 5 cellulases, including the metagenome-derived endoglucanase Cel5A, cannot (Voget et al. 2006). The purified recombinant Umcel5G cuts randomly on the chain of cello-oligosaccharides, revealing its endo-mode action. G5 is

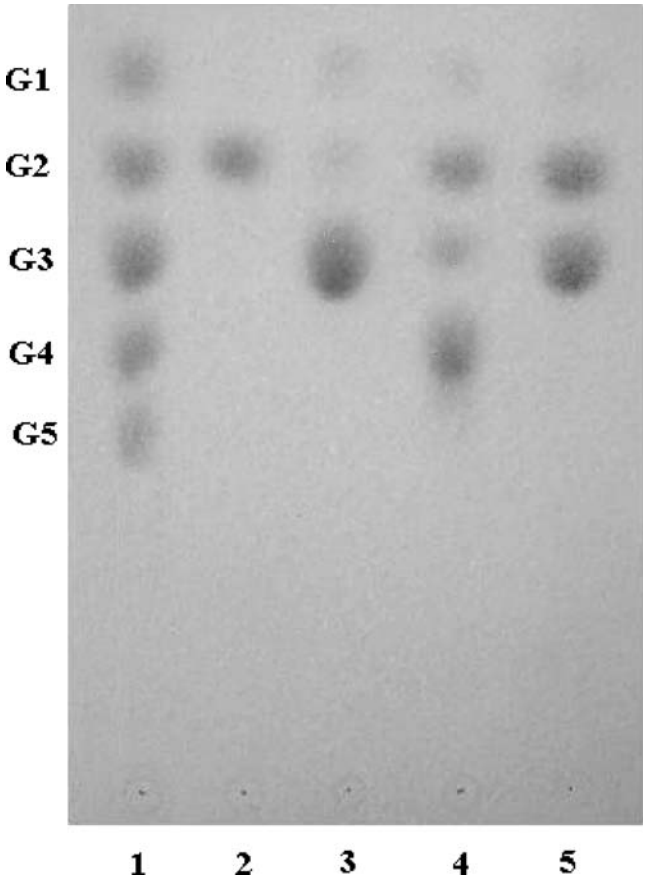

Fig. 4 Hydrolysis products of cello-oligosaccharides by the recombinant Umcel5G. *Lane 1*, mixed standard sugars: glucose (*G1*), cellobiose (*G2*), cellotriose (*G3*), cellotetraose (*G4*) and cellopentaose (*G5*); *lane 2~5*, G2~G5 treated with the recombinant Umcel5G. 1-μl purified enzyme (protein concentration=0.016 mg/ml) was mixed, respectively, with 10-μl sugar solutions that contained 10 mg/ml of G2, G3, G4 and G5 in citrate–phosphate buffer of pH 6.5. The mixtures were loaded onto the TLC plate after reactions at 55°C for 2 h. TLC was performed with the solvent system of butanol/acetic acid/water (2:1:1, *v*/*v*/*v*). Sugars were visualized by incubating the TLC plate at 100°C for 10 min after spraying with 50% (*v*/*v*) H_2SO_4 in methanol

completely degraded, but G4 and G3 are only partially hydrolyzed by Umcel5G, which suggests that the enzyme may acts less effectively on shorter cello-oligosaccharide.

Genes encoding cellobiohydrolase and multidomain endoglucanase were not cloned in this study. The main possible reasons may be as follows: (1) The chance to clone a cellobiohydrolase gene is much less than to get an endoglucanase gene from a bacterium by shotgun cloning. The whole genome sequencing analysis provides evidence that the number of cellobiohydrolase genes is much less than that of endoglucanase genes in the genomes of some cellulolytic bacteria. For example, 12 endoglucanase genes and only one cellobiohydrolase gene were predicted in the genome of marine cellulolytic bacterium *Saccharophagus degradans* (Taylor II et al. 2006). Twenty genes were predicted to encode endoglucanase, and only four ORFs were proven or predicted to encode cellobiohydrolase in the genome of soil bacterium *Clostridium thermocellum* (Taylor II et al. 2006; Zverlov et al. 2005). (2) The cellobiohydrolase genes or multidomain endoglucanase genes might not be expressed as active enzymes because of the problems of heterologous gene expression in *E. coli* (Sørensen and Mortensen 2005). (3) The relatively small capacity of the library we constructed may have restricted the full representation. Three hundred and seventy five phylotypes of bacteria were identified in the pig gastrointestinal tract by culture-independent comparative 16S ribosomal DNA sequence analysis (Leser et al. 2002). Assuming similar number of bacterial phylotypes presented in the rabbit cecum and average genome sizes of 4.0×10^6 bp, the metagenome would be 1.5×10^9 bp. Thus, the number of clones in the constructed library with average insert size of 35.1 kb would have to be 197,000 to represent the entire metagenome based on Seed et al. (1982), assuming that all bacterial species are equally represented in the isolated metagenomic DNA and the constructed library.

In summary, this is so far the first study on the cloning and the characterization of cellulase genes from the microbes in rabbit cecum, which helps to understand the properties, the diversity and the potential biotechnological applications of the cellulases in the cecal system of rabbit.

Acknowledgment This work was supported by the National Natural Science Foundation of China (30560003), International Science and Technology Collaborating Key Project Program of China (2002AA217121) and Program for New Century Excellent Talents in University of China (NCET-05-0752). The authors are grateful to Kunio Ohmiya and Shuichi Karita for their helpful discussions.

References

Abecia L, Fondevila M, Balcells J, Edwards JE, Newbold CJ, McEwan NR (2005) Molecular profiling of bacterial species in the rabbit cecum. FEMS Microbiol Lett 244:111–115

Amann RI, Ludwig W, Schleifer KH (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiol Rev 59:143–169

Bhatia Y, Mishra S, Bisaria V (2005) Purification and characterization of recombinant *Escherichia coli*-expressed *Pichia etchellsii* β-glucosidase II with high hydrolytic activity on sophorose. Appl Microbiol Biotechnol 66:527–535

Blanco A, Diaz P, Martinez J, Vidal T, Torres AL, Pastor FIJ (1998) Cloning of a new endoglucanase gene from *Bacillus* sp. BP-23 and characterisation of the enzyme. Performance in paper manufacture from cereal straw. Appl Microbiol Biotechnol 50:48–54

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

Cowan D, Meyer Q, Stafford W, Muyanga S, Cameron R, Wittwer P (2005) Metagenomic gene discovery: past, present and future. Trends Biotechnol 23:321–329

Eberhart B, Cross DF, Chase LR (1964) β-Glucosidase system of *Neurospora crassa* I. β-Glucosidase and cellulase activities of mutant and wild-type strains. J Bacteriol 87:761–770

Eckert K, Schneider E (2003) A thermoacidophilic endoglucanase (CelB) from *Alicyclobacillus acidocaldarius* displays high sequence similarity to arabinofuranosidases belonging to family 51 of glycoside hydrolases. Eur J Biochem 270:3593–3602

Feng JX, Karita S, Fujino E, Fujino T, Kimura T, Sakka K, Ohmiya K (2000) Cloning, sequencing, and expression of the gene encoding a cell-bound multi-domain xylanase from *Clostridium josui*, and characterization of the translated product. Biosci Biotechnol Biochem 64:2614–2624

Ferrer M, Golyshina OV, Chernikova TN, Khachane AN, Reyes-Duarte D, Martins Dos Santos VAP, Strompl C, Elborough K, Jarvis G, Neef A, Yakimov MM, Timmis KN, Golyshin PN (2005) Novel hydrolase diversity retrieved from a metagenomic library of bovine rumen microflora. Environ Microbiol 7:1996–2010

Gidenne T (1997) Caeco-colic digestion in the growing rabbit: impact of nutritional factors and related disturbances. Livest Prod Sci 51:73–88

Healy FG, Ray RM, Aldrich HC, Wilkie AC, Ingram LO, Shanmugam KT (1995) Direct isolation of functional genes encoding cellulases from the microbial consortia in a thermophilic, anaerobic digester maintained on lignocellulose. Appl Microbiol Biotechnol 43: 667–674

Inoue T, Moriya S, Ohkuma M, Kudo T (2005) Molecular cloning and characterization of a cellulase gene from a symbiotic protist of the lower termite, *Coptotermes formosanus*. Gene 349:67–75

Kamra DN (2005) Rumen microbial ecosystem. Curr Sci 89:124–135

Krause DO, Denman SE, Mackie RI, Morrison M, Rae AL, Attwood GT, McSweeney CS (2003) Opportunities to improve fiber degradation in the rumen: microbiology, ecology, and genomics. FEMS Microbiol Rev 27:663–693

Kuske CR, Banton KL, Adorada DL, Stark PC (1998) Small-scale DNA sample preparation method for field PCR detection of microbial cells and spores in soil. Appl Environ Microbiol 64:2463–2472

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680–685

Leser TD, Amenuvor JZ, Jensen TK, Lindecrona RH, Boye M, Moller K (2002) Culture-independent analysis of gut bacteria: the pig gastrointestinal tract microbiota revisited. Appl Environ Microbiol 68:673–690

Lynd LR, Weimer PJ, Willem HZ, Pretorius IS (2002) Microbial cellulose utilization: fundamentals and biotechnology. Microbiol Mol Biol Rev 66:506–577

Mackie RI (1997) Gut environment and evolution of mutualistic fermentative digestion. In: Mackie RI, White BA (eds) Gastrointestinal Microbiology, vol 1. Chapman and Hall, New York, pp 13–35

McLaughlin CA, Chiasson RB (1990) Laboratory Anatomy of the Rabbit, 3rd edn. William C. Brown, Dubuque, IA, pp 59–64

Michalet-Doreau B, Fernandez I, Peyron C, Millet L, Fonty G (2001) Fibrolytic activities and cellulolytic bacterial community structure in the solid and liquid phases of rumen contents. Reprod Nutr Dev 41:187–194

Miller GL (1959) Use of dinitrosalicyclic acid reagent for determination of reducing sugar. Biotechnol Bioeng Symp 5:193–219

Odoux E, Escoute J, Verdeil L, Brillouet M (2003) Localization of β-glucosidase activity and glucovanillin in vanilla bean (*Vanilla planifolia* Andrews). Ann Bot 92:437–444

Rees HC, Grant S, Jones B, Grant WD, Heaphy S (2003) Detecting cellulase and esterase enzyme activities encoded by novel genes present in environmental DNA libraries. Extremophiles 7:415–421

Reinhold-Hurek B, Hurek T, Claeyssens M, Montagu M (1993) Cloning, expression in *Escherichia coli*, and characterization of cellulolytic enzymes of *Azoarcus* sp., a root-invading diazotroph. J Bacteriol 175:7056–7065

Seed B, Parker RC, Davidson N (1982) Representation of DNA sequences in recombinant DNA libraries prepared by restriction enzyme partial digestion. Gene 19:201–209

Sørensen HP, Mortensen KK (2005) Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. J Biotechnol 115:113–128

Sugimura M, Watanabe H, Lo N, Saito H (2003) Purification, characterization, cDNA cloning and nucleotide sequencing of a cellulase from the yellow-spotted longicorn beetle, *Psacothea hilaris*. Eur J Biochem 270:3455–3460

Taylor LE II, Henrissat B, Coutinho PM, Ekborg NA, Hutcheson SW, Weiner RM (2006) Complete cellulase system in the marine bacterium *Saccharophagus degradans* strain 2-40T. J Bacteriol 188:3849–3861

Teather RM, Wood PJ (1982) Use of Congo red-polysaccharide interactions in enumeration and characterization of cellulolytic bacteria from the bovine rumen. Appl Environ Microbiol 43:777–780

Voget S, Leggewie C, Uesbeck A, Raasch C, Jaeger KE, Streit WR (2003) Prospecting for novel biocatalysts in a soil metagenome. Appl Environ Microbiol 69:6235–6242

Voget S, Steele HL, Streit WR (2006) Characterization of a metagenome-derived halotolerant cellulase. J Biotechnol 126:26–36

Walter J, Mangold M, Tannock GW (2005) Construction, analysis, and beta-glucanase screening of a bacterial artificial chromosome library from the large-bowel microbiota of mice. Appl Environ Microbiol 71:2347–2354

Zhou JZ, Bruns MA, Tiedje JM (1996) DNA recovery from soils of diverse composition. Appl Environ Microbiol 62:316–322

Zverlov VV, Kellermann J, Schwarz WH (2005) Functional subgenomics of *Clostridium thermocellum* cellulosomal genes: identification of the major catalytic components in the extracellular complex and detection of three new enzymes. Proteomics 5:3646–3653