Antonie van Leeuwenhoek (2008) 94:563–571
DOI 10.1007/s10482-008-9274-7

ORIGINAL PAPER

A flax-retting endopolygalacturonase-encoding gene from *Rhizopus oryzae*

Zhizhuang Xiao · Shaozhao Wang · Hélène Bergeron · Jianchun Zhang · Peter C. K. Lau

Received: 12 May 2008 / Accepted: 31 July 2008 / Published online: 14 August 2008

Abstract A polygalacturonase from the filamentous fungus *Rhizopus oryzae* strain sb (NRRL 29086), previously shown to be effective in the retting of flax fibers, was shown by the analysis of its reaction products on polygalacturonic acid to be an endo-type. By zymogram analysis, the enzyme in the crude culture filtrate appeared as two active species of 37 and 40 kD. The endopolygalacturonase-encoding gene was cloned in *Escherichia coli* and its translated 383-amino acid sequence found to be identical to that of a presumed exopolygalacturonase found in *R. oryzae* strain YM9901 and 96% identical to a hypothetical protein (RO3G_04731.1) in the sequenced genome of *R. oryzae* strain 99–880. Phylogenetic analysis revealed the presence of an unique cluster of *Rhizopus* polygalacturonase sequences that are separate from other fungal polygalacturonases. Conservation of 12 cysteines appears to be a special feature of this family of *Rhizopus* polygalacturonase sequences.

Keywords Natural bast fibers · Glycosyl hydrolase · Enzyme-retting · *Rhizopus* genomics · Fungal pectinases

Electronic supplementary material The online version of this article (doi:10.1007/s10482-008-9274-7) contains supplementary material, which is available to authorized users.

Z. Xiao · S. Wang · H. Bergeron · J. Zhang · P. C. K. Lau (✉)
Biotechnology Research Institute, National Research Council Canada, 6100 Royalmount Avenue, Montreal, QC, Canada H4P 2R2
e-mail: peter.lau@cnrc-nrc.gc.ca

Introduction

Rhizopus oryzae (synonym: *R. arrhizus*) is one of the 13 species of the genus *Rhizopus* (Ribes et al. 2000; Abe et al. 2006). Strains of *R. oryzae* are important in organic acid and lipase production (Abe et al. 2003; Saito et al. 2004; Di Lorenzo et al. 2005; Roa Engel et al. 2008), and *R. oryzae* glucoamylase is especially known for its saccharification of starch to produce alcohol (Ashikari et al. 1986; Merten and Skory 2007a, b). In a clinical setting, *R. oryzae* is the most common causative agent of zygomycosis, a rhinocerebral infection that usually occurs in patients with diabetic ketoacidosis (Ribes et al. 2000; Ibrahim et al. 2003). Until the recent *R. oryzae* genome sequencing project (http://www.broad.mit.edu.), there has been a general lack of DNA sequence and protein information on *Rhizopus* species or other zygomycetes (Ashikari et al. 1986; Friedberg et al. 1995; Farley and Sullivan 1998; Saito et al. 2004; Yoshida et al. 2004; Skory and Ibrahim 2007; Mertens and Skory 2007a, b).

R. oryzae strain sb (initially known as *Rhizomucor pusillus*) was one of the seven strains of filamentous fungi isolated from dew-retted flax in South Carolina,

USA, and found to contain an extracellular polygalacturonase (PG), whose activity was capable of producing good quality flax fibers (Henriksson et al. 1997,1999; Akin et al. 2002). This enzyme (herein referred to as RoPG) was recently purified and characterized to have a molecular mass of 37,436 Da, and an apparent substrate preference for unesterified polygalacturonates providing supportive evidence for its pectinolytic activity (Zhang et al. 2005). The enzyme was also classified among family 28 of glycoside hydrolases based on some sequence motifs (Markovic and Janecek 2001).

Akin et al. (2002) suggested the cloning of the RoPG-encoding gene from *R. oryzae* strain sb for preparation of a simplified enzyme formulation for bast fiber retting, vis-à-vis a mixture or cocktail of hydrolytic enzymes, such as some of those available commercially, eg., Viscozyme, Flaxzyme or Lyvelin (Akin et al. 2001, 2004; Kozlowski et al. 2006; Antonov et al. 2007). Mixed-component formulations contain cellulases or xylanases among other hydrolytic enzymes that can weaken flax fibers by cellulose digestion (Akin et al. 2004). Herein, the cloning and sequencing of the RoPG-encoding gene from strain sb is reported, and a detailed analysis of the predicted RoPG amino acid sequence as a prototype *Rhizopus* PG is provided. This is new information that was missed in the list of fungal PG sequences recently reviewed by Niture (2008). The RoPG enzyme is established to be the endo-type.

Materials and methods

Fungal growth

R. oryzae strain sb was obtained from ARS Culture Collection United States Department of Agriculture NRRL 29086, and provided by Kerry O'Donnell. It was grown in Vogel medium (2.5 g sodium citrate, 5 g KH_2PO_4, 2 g NH_4NO_3, 0.4 g $MgSO_4 \cdot 7H_2O$, 0.2 g $CaCl_2 \cdot H_2O$, 2.5 mg $FeSO_4$, 0.98 mg $MnSO_4 \cdot H_2O$, 0.83 mg $ZnCl_2$, and 1.0 mg $CoCl_2$ in 1,000 ml) supplemented with 10 g/l citrus pectin (Sigma) as carbon and energy source (Zhang et al. 2005). The pH was adjusted to 5 using concentrated HCl. Flasks (500 ml) containing 100 ml of fungal culture were shaken on a rotary shaker at 30°C, 200 rpm. Two-day-old culture was harvested and cells washed with 1% NaCl prior to extraction of total RNA.

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and zymogram

Deglycosylation of protein mixture secreted from *R. oryzae* was performed by PNGase F (BioLabs) according to the manufacturer's instruction. Proteins were heated for 10 min at 45°C in sample loading buffer before separation on 10% SDS–PAGE containing 0.1% polygalacturonic acid (PGA). After electrophoresis, protein bands were stained with Coomassie blue. For zymogram staining, the gel was soaked in 2.5% Triton X-100 for 30 min, followed by washing with 100 mM sodium acetate buffer, pH 5.2 for 30 min. After 20 h of incubation at 40°C in acetate buffer containing 1 g/l PGA, the gel was stained with 0.05% (w/v) ruthenium red for 20 min and washed with water for 60 min.

Preparation of RNA and first-strand cDNA synthesis

First-strand cDNA was synthesized from total RNA prepared with EZNA fungal RNA miniprep kit (Omega Bio-Tek) using oligo dT anchor primer (Supplemental material online, Table 1) provided in a 5′/3′ RACE kit (Roche Applied Science) according to the manufacturer instructions.

Amplification and cloning of ROPG cDNA

Standard molecular techniques were used for isolation of plasmid DNA, restriction enzyme digestion, ligation, cloning and transformation (Sambrook et al. 1989). For gene amplification, degenerate oligos, 20 and 17DGF (Supplementary material online, Table 1) were designed from two peptide sequences (P1/NTDAIDV; P2/NNADTP) as reported in Zhang et al. (2005). Failing this, two gene-specific primers 23SPF and 23SPR designed based on the sequence of the *rpg1* gene (Yoshida et al. 2004) were used to amplify the full-length RoPG cDNA. PCR conditions were 30 cycles of 1 min at 94°C, 1 min at 50°C, and 1.5 min at 72°C. All primers used in this work are listed (Supplementary material online, Table 1).

After digestion with *Eco*RI and *Hin*dIII, the PCR products were ligated with digested plasmid pGEM-

3Zf (+) (Promega), generating the recombinant plasmid R-1. The recombinant plasmids were transformed into competent *E. coli* Top10 cells (Invitrogen). Two plasmid universal primers (pUC/M13 forward and reverse primers) and a gene-specific primer Sqn-20INR (Supplementary material online Table 1) were used to sequence the clones harboring the desired RoPG-encoding gene. DNA sequencing was performed using a Big Dye DNA sequencing kit (Applied Biosystems) with an automated DNA sequencer (Model 377, ABI Prism).

Bioinformatic programs

CLUSTALW2 program (http://www.ebi.ac.uk/Tools/clustalw2/index.html) was used to perform sequence alignment. *N*-Glycosylation sites were predicted with the NetNGlyc 1.0 Server (http://www.cbs.dtu.dk/services/NetNGlyc/). A search for conserved domains and related sequences was performed at the National Center for Biotechnology Information BLAST website (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). SignalP 3.0 server (http://www.cbs.dtu.dk/services/SignalP/) was used to analyze signal peptide sequences. Phylogenetic analysis was carried out using PHYLIP (version 3.67; http://evolution.genetics.washington.edu.edu/phylip.html) programs. Bootstrapping was done with 100 repetitions and the tree displayed by TreeView for Windows.

Determination of PG activity

The activity of PG was determined by measuring the amount of the reducing sugars released from PGA using the modified 3,5-dinitrosalicylic acid method (Xiao et al. 2005; Miller 1959). One unit (IU) of PG activity was defined as the amount of enzyme that produced 1 μmol of galacturonic acid or equivalent in 1 min under the assay conditions. All data were the average of triplicate measurements.

Thin-layer chromatography (TLC)

Galacturonic acid (G1), digalacturonic acid (G2), and trigalacturonic acid (G3) were purchased from Sigma. Hydrolysis products and standard G1, G2 and G3 were separated by TLC on a silica gel 60-plastic sheet (Merck) with a solvent of 1-butanol/water/acetic acid (5:3:2), and the oligosaccharides were visualized by spraying the plate with 10% (v/v) sulfuric acid in ethanol (Tamaru and Doi 2001).

Nucleotide sequence of RoPG

The 1,199 bp genomic DNA sequence including a 47 bp intron has been deposited in the GenBank with Accession Number EU432156.

Results

Analysis of PG-active components in the culture filtrate of *R. oryzae* strain sb

Zymogram analysis (Fig. 1) indicated that two protein bands with PG activity and estimated M_r of 37 and 40 kD exist in the crude culture filtrate. After 4 h of deglycosylation with PNGaseF, the 40 kD band disappeared. Although the 37 kD band persisted at the 4th h, its PG activity was completely lost after 16 h of deglycosylation (not shown).

Cloning of RoPG-encoding gene

Initially, two degenerate oligonucleotides (20 and 17DGF; Supplementary material online, Table 1) were

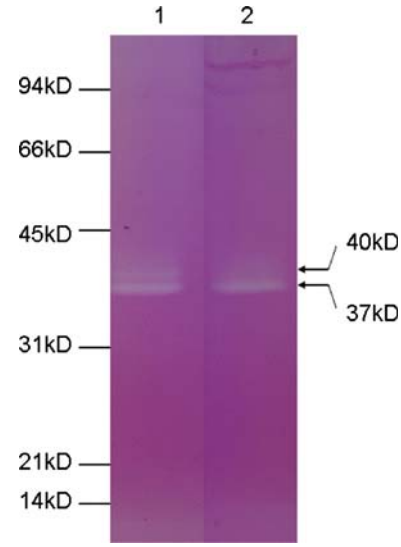


Fig. 1 Zymogram analysis of PG-active components in the culture of *Rhizopus oryzae*. Lane 1: native RoPG; lane 2: RoPG after 4 h deglycosylation with PNGaseF. Molecular weight standard is indicated alongside

designed as possible primers for amplification of the desired cDNA from strain sb based on the available peptide sequence information (Zhang et al. 2005). These experiments were unsuccessful. Subsequently, two DNA fragments of 1,152 and 1,199 bp were amplified using specific primers 23SPF and 23SPR (Supplementary material online, Table 1) derived from a putative exo-PG gene sequence (AB127052) from *R. oryzae* strain YM9901, an isolate from a mulberry rot (Yoshida et al. 2001, 2004).

Sequencing and analysis of the 1,152 bp cDNA indicated that, except for two nucleotide transitions (G to A at nucleotide position 891; and C to T at 1,047), this sequence is identical to the PG-encoding cDNA from *R. oryzae* YM9901 (AB127052). The 1,199 bp fragment was the result of amplification of genomic DNA from strain sb containing the RoPG gene and one intron of 47 bp.

Sequence characteristics of RoPG

The predicted amino acid sequence of RoPG (EU432156) consists of 383 amino acids residues (M_r 39,705.4) and is identical to that of *R. oryzae* YM9901 (BAD67423) since the two base changes are silent mutations, both involving Thr codons at amino acid positions 297 and 349. In the recently sequenced genome of *R. oryzae* strain RA 99–880, a putative PG gene (RO3G 04731.1) that contains a 49 bp intron was found to share 94% nucleotide identity and 96% amino acid sequence identity with the RoPG sequence. Besides the two base insertions in the RA 99-880 intron, there are eight other base changes. In all three sequences, the splice junction is localized at the same amino acid codon, TGG encoding Trp124. In the latter genome, there are 17 other putative PG sequences that displayed 44–88% amino acid sequence identity with respect to RoPG (Supplementary material online Fig. 1). Other than *Rhizopus* sequences, a hypothetical protein AN8761.2 (XP_682030.1) from the *Aspergillus nidulans* FGSC A4 genome provided 34% sequence identity. In the structure database (Protein Data Base), the highest score of 31% identity was that of an endo-PG of a fungal *Colletotrichum lupini* (EF094978; Bonivento et al. 2008).

A putative signal peptide of 27 amino acids was predicted for RoPG (Supplementary material online, Fig. 1). The mature protein therefore contains 356 amino acids with a calculated M_r of 36,928 and a theoretical pI of 7.65. This molecular mass is very close to that determined by mass spectrometry, a value of 37,436 with unknown glycosylation moiety (Zhang et al. 2005). The mature RoPG was found to contain the four conserved active site segments of numerous PGs of the glycosyl hydrolases family 28 (Markovic and Janecek 2001; Niture 2008). These sequences are 185NTD, 208DD, 230GHGXXXGS and 265RIK. Residues D187, D208, D209, H231, R265, and K267 are equivalent to D180, D201, D202, H223, R256, and K258 of *Aspergillus niger* endo-PG II where the three aspartates were involved in the catalytic mechanism of the enzyme (van Santen et al. 1999; Armand et al. 2000). Y299 in RoPG is equivalent to the invariant Y (Y291 in the *A. niger* PG sequence) when 115 sequences of plant, fungal, insect or bacterial PGs were aligned (Markovic and Janecek 2001).

Analysis of the position of the potential *N*-linked glycosylation sites in RoPG showed two likely possibilities—NITL at N170 and N219. A third possibility at position N254 with sequence NSSC, although likely, scored just marginally above the program threshold value.

A notable feature of RoPG is the presence of 12 cysteines that is concentrated mostly at the C-terminus of the protein. These cysteines are at positions 30, 50, 210, 227, 257, 260, 300, 308, 342, 348, 369, and 381 (Supplementary material on line Fig. 1). Eleven of the 12 cysteines are fully conserved in all the 17 PGs in *R. oryzae* strain 99-880. Seven of these cysteines (at positions 210, 227, 300, 308, 342, 348, and 369) are conserved in the plants (*Medicago sativa*; GenBank Accession Number Y11118) and tomato (GenBank Accession Number AF118567) PGs (Bergey et al. 1999; Munoz et al. 1998). However, only four (at positions 210, 227, 342 and 348) are found to be conserved in fungal PGs such as *A. niger* and *A. tubigensis* PGII (Bussink et al. 1991).

None of the following sequence segments typically found in fungal exo-PGs were found in the RoPG: 121_SFKxxFQN, 166_LRPiL, 225_WDTYR, 248_SFKPN, and 319_GGGG, sequence and numbering according to *Aspergillus tubigensis* exo-PG (Markovic and Janecek 2001; Abbott and Boraston 2007).

Phylogenetic analysis of RoPG

To gain an understanding of the phylogenetic positioning of RoPG and *R. oryzae* PGs at large, we included the recently available PGs from the genome of another filamentous fungus, *Aspergillus oryzae* (16 sequences, labelled as BAE as in Park et al. 2008) and various other PGs of plant origin or other fungal sequences in the NCBI database (Fig. 2). Included in this dataset are PGs that have been characterized biochemically: 14 endo-PGs and 8 exo-PGs. Two features are evident: (i) all the *Rhizopus* PGs appear as a cluster even though in a few sub-clades and there are distinct from those of *A. oryzae* PGs; (ii) there appears to be no distinct separation of endo- and exo-type PGs from the various phyla. The well-characterized *Erwinia carotovora* endo-PG (Pickersgill et al. 1998; P26509 = CAA35998 *Pectobacterium carotovorum*) is clustered with two presumed endo-PGs of *Xanthomonas* origin. (Xiao et al. 2008).

Enzyme mode of action for RoPG

In order to determine whether RoPG is of the endo- or exo-type, we carried out RoPG digestion on substrates that include digalacturonic acid (G2), trigalacturonic acid (G3) and PGA (Fig. 3). Analysis by TLC indicated the formation of trimers, dimers and small amount of monomers with preference for trimers. RoPG did not degrade dimer (G2) or trimer (G3), consistent with this enzyme having an endo-mode of action.

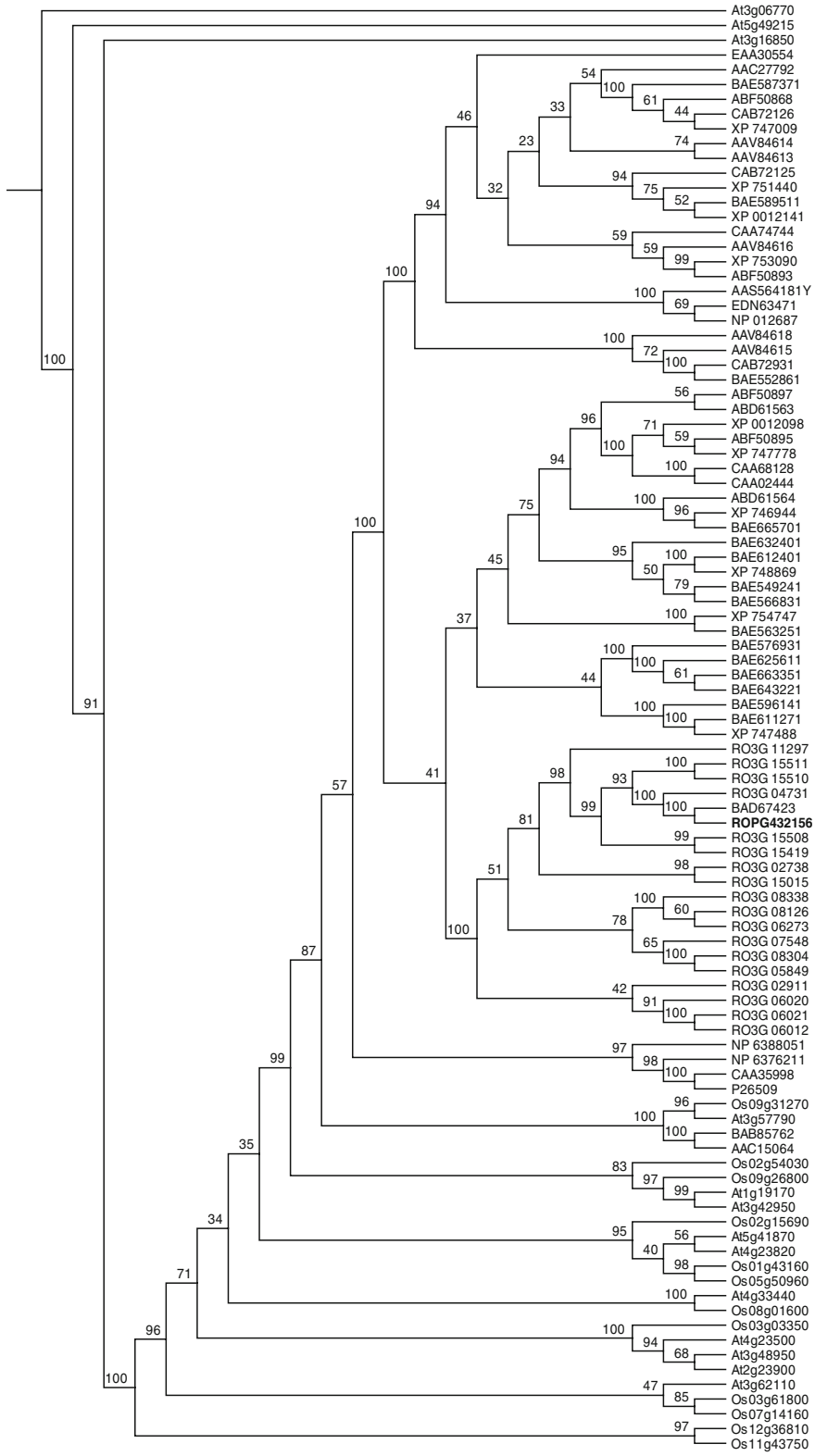
Discussion

In this study we identified the gene and predicted amino acid sequence of a flax-retting endo-PG derived from *R. oryzae* strain sb and provided a comparative view of related sequences in two other *Rhizopus* strains of distinct geographical origin. Strain sb, appearing as airy white mycelium and black fruit bodies, came from dew-retted flax in South Carolina, USA (Henriksson et al. 1997); YM strain was isolated from rotted cortical tissues of the stocks of grafted saplings of mulberry in Tsukuba, Japan (Yoshida et al. 2001); whereas strain 99–880 was a clinical isolate from a brain abscess of a diabetic patient with rhinocerebral zygomycosis (San Antonio, Texas, USA; Ibrahim et al. 2003). It is evident that the three *R. oryzae* strains share a near identical PG gene. In strain YM9901, a single copy of the putative exo-PG encoding gene (*rpg1*) was established by Southern blot analysis (Yoshida et al. 2004), whereas in strain sb it was only established at the biochemical level that a single PG was present in the culture filtrate that released reducing sugar from citrus pectin (Zhang et al. 2005). In the case of strain 99-880, multiple PG-encoding genes exist; the number of introns range from 1 to 3. In the phylogenetic analysis, it is interesting that the 17 other sequences, despite having a sequence identity of 44–88% with respect to RoPG, are not dispersed but clustered in the "*Rhizopus*" clade. We have also analyzed the PG genes in *A. oryzae* genome (16 sequences; Park et al. 2008) and found them to be in separate sub-clads than those of *Rhizopus* PGs.

The complete amino acid sequence identity of the RoPG of *R. oryzae* strain sb to the presumed exo-PG of strain YM 9901 was unexpected since Zhang and coworkers (2005) had reported 10 peptide sequences of the former protein that were different. Our analysis showed that only one of the 10 peptide sequences (containing the family 28 signature glycosyl hydrolase sequence), matched with 81% identity and no gap to the RoPG sequence (not shown). One likely cause of discrepancy is associated with *de novo* sequencing to the effect that at least 50% of product ion spectra produced by collision-induced dissociation in mass spectrometry are only partially interpretable and that they do not result in confident peptide identification (Kinter and Sherman 2000; Aebersold and Mann 2003; Frank et al. 2007).

The presence of two protein bands with PG activity in the crude culture filtrate of *R. oryzae* strain sb is consistent with the observation of PG activity in the crude filtrate of *R. oryzae* YM9901 culture (Yoshida et al. 2004). In the latter case, the two proteins were found to share identical N-terminal 22 amino acid residues starting at amino acid 27 of the proprotein. It was proposed that both proteins originated from a single PG gene. Two forms of active PG encoded by the same single-copy gene might result from the different extents of glycosylation at the two of the predicted N-glycosylation sites at positions N170 and N219.

Apparently, it is not possible to correlate the position of PG in the phylogenetic tree with enzyme


At3g06770
At5g49215
At3g16850
EAA30554
AAC27792
BAE587371
ABF50868
CAB72126
XP 747009
AAV84614
AAV84613
CAB72125
XP 751440
BAE589511
XP 0012141
CAA74744
AAV84616
XP 753090
ABF50893
AAS564181Y
EDN63471
NP 012687
AAV84618
AAV84615
CAB72931
BAE552861
ABF50897
ABD61563
XP 0012098
ABF50895
XP 747778
CAA68128
CAA02444
ABD61564
XP 746944
BAE665701
BAE632401
BAE612401
XP 748869
BAE549241
BAE566831
XP 754747
BAE563251
BAE576931
BAE625611
BAE663351
BAE643221
BAE596141
BAE611271
XP 747488
RO3G 11297
RO3G 15511
RO3G 15510
RO3G 04731
BAD67423
ROPG432156
RO3G 15508
RO3G 15419
RO3G 02738
RO3G 15015
RO3G 08338
RO3G 08126
RO3G 06273
RO3G 07548
RO3G 08304
RO3G 05849
RO3G 02911
RO3G 06020
RO3G 06021
RO3G 06012
NP 6388051
NP 6376211
CAA35998
P26509
Os09g31270
At3g57790
BAB85762
AAC15064
Os02g54030
Os09g26800
At1g19170
At3g42950
Os02g15690
At5g41870
At4g23820
Os01g43160
Os05g50960
At4g33440
Os08g01600
Os03g03350
At4g23500
At3g48950
At2g23900
At3g62110
Os03g61800
Os07g14160
Os12g36810
Os11g43750

◀ **Fig. 2** Phylogenetic analysis of *Rhizopus oryzae* polygalacturonases among other PGs. Bootstrap values are indicated at each node. ROPG432156 (in bold) is from this study; BAD67423 is from *R. oryzae* strain YM9901, and RO3G (18 candidates) are from the sequenced genome of *R. oryzae* strain 99-880. The 16 BAE sequences are from the genome of *A. oryzae* (Park et al. 2008); Os and At sequences are from *Oryza sativa* and *Arabidopsis thaliana*, respectively, as referenced in Park et al. 2008. Biochemically characterized endo- and exo-PGs are as indicated together with the accession number of the source organism: *Aspergillus niger* (exo-ABD61563; exo-ABD61564; endo-CAA74744; endo-CAB72125; endo-CAB72126; endo-CAB72931); *Aspergillus tubingensis* (exo-CAA68128; exo-CAA02444); *Bacillus* sp. P-358 (exo-BAB85762); *Botryotinia fuckeliana* (endo-AAV84613; endo-AAV84614; endo-AAV84615; endo-AAV84616; endo-AAV84618); *Emericella nidulans* (exo-ABF50895; exo-ABF50897; endo-ABF50868; endo-ABF50893); *Erwinia carotovora* (endo-P26509); *Fusarium oxysporum* (endo-AAC27792); *Pectobacterium carotovorum* (endo-CAA35998); *Yersinia enterocolitica* (exo-AAC15064); Uncharacterized PGs are from: *Aspergillus fumigatus* (XP_747778; XP_747488; XP_746944; XP_748869; XP_754747; XP_751440; XP_753090; XP_747009); *Aspergillus terreus* (XP_001209838; XP_001214169); *Neurospora crassa* (EAA30554); *Saccharomyces cerevisiae* (NP_012687; EDN63471); and NP637621 and NP638805 from *Xanthomonas campestris* (Xiao et al. 2008)

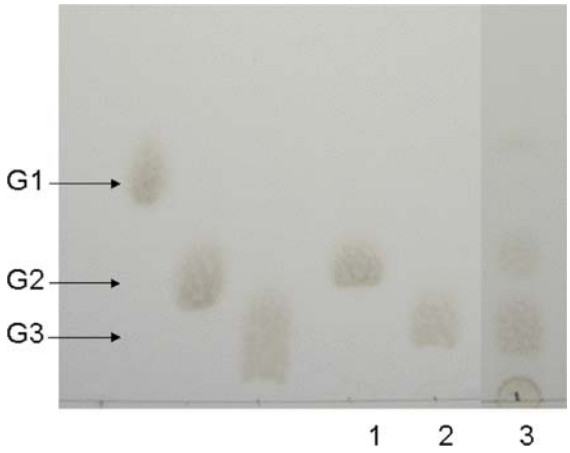


Fig. 3 TLC analysis of the reaction products of *Rhizopus oryzae* PG on the substrates. 1: G2 + RoPG; 2: G3 + RoPG; 3: PGA + RoPG. In lane 3, a very faint band appears near the gel front that corresponds to position of G1. G1, G2 and G3 are galacturonic acid, digalacturonic acid and trigalacturonic acid, respectively

specificity. We confirmed the positioning of RoPG, an endo-PG, amidst the exo-PG clade as previously observed by Yoshida et al. (2004) who used a smaller sequence dataset. Moreover, a distinct cluster of *Rhizopus* PG sequences, although organized in subclades, emerges for the first time among various PGs of other microbial origin. This is not the case of *A. oryzae* PGs where three of the BAE sequences were found in this analysis to be loosely associated with the main *A. oryzae* clade. Park et al. (2008) had previously reported a loose association of five *A. oryzae* PGs with some plant PGs which is no the case with *R. oryzae* PGs. The presence of 12 cysteines in RoPG and among the *Rhizopus* clade PGs appears to be a new feature among fungal PGs although plant PGs contain 10–17 cysteines (Tebbutt et al. 1994; Bergey et al. 1999). It has been pointed out that the conservation of cysteines among PG sequences is a reflection of the taxonomy position (Niture 2008). This appears to hold true for the *Rhizopus* clade of PGs that has a dozen conserved cysteines among the presently available 20 sequences of *Rhizopus* origin. Saito et al. (2004) had purified an endo-PG from *R. oryzae* NBRC 4707 with a M_r of 29.7–31 kDa, but unfortunately this sequence is not known.

Preliminary heterologous expression of RoPG-encoding gene in either *E. coli* or *Pichia pastoris* has met with limited success (results not shown). The dozen cysteines in RoPG and codon usage are among factors of consideration for a dedicated gene expression study to effect a possible over expression of RoPG for the purpose of enzyme-retting, bioscouring of natural fibres in straw management, and other industrial applications (Zhang et al. 2000; Akin et al. 2007; Foulk et al. 2008; Xiao et al. 2008).

Acknowledgment Funding and support to ZX and JZ by the Climate Change Technology and Innovation Biotechnology Program of the Canadian Biomass Innovation Network of Natural Resources Canada is gratefully acknowledged. We thank K. Abokitse for help in the phylogenetic analysis and A. Matte for suggestions and careful reading of the manuscript.

References

Abbott WD, Boraston AB (2007) The structural basis for exo-polygalacturonase activity in a family 28 glycoside hydrolase. J Mol Biol 368:1215–1222. doi:10.1016/j.jmb.2007.02.083

Abe A, Oda Y, Asano K, Sone T (2006) The molecular phylogeny of the genus *Rhizopus* based on rDNA sequences. Biosci Biotechnol Biochem 70:2387–2393. doi:10.1271/bbb.60101

Abe A, Sone T, Sujaya N, Saito K, Oda Y, Asano K et al (2003) rDNA ITS sequence of *Rhizopus oryzae*: its

application to classification and identification of lactic acid producers. Biosci Biotechnol Biochem 67:1725–1731. doi:10.1271/bbb.67.1725

Aebersold R, Mann M (2003) Mass spectrometry-based proteomics. Nature 422:198–207. doi:10.1038/nature01511

Akin DE, Condon B, Sohn M, Foulk JA, Dodd RB, Rigsby LL (2007) Optimization for enzyme-retting of flax with pectate lyase. Ind Crops Prod 25:136–146. doi:10.1016/j.indcrop.2006.08.003

Akin DE, Foulk JA, Dodd RB, McAlister DD (2001) Enzyme-retting of flax and characterization of processed fibers. J Biotechnol 89:193–203. doi:10.1016/S0168-1656(01)00298-X

Akin DE, Slomczynski D, Rigsby LL, Eriksson KL (2002) Retting flax with endopolygalacturonase from *Rhizopus oryzae*. Text Res J 72:27–34. doi:10.1177/004051750207200105

Akin DE, Henriksson G, Evans JD, Adamsen APS, Foulk JA, Dodd RB (2004) Progress in enzyme-retting of flax. J Nat Fibers 1:21–47. doi:10.1300/J395v01n01_03

Antonov V, Marek J, Bjelkova M, Smirous P, Fischer H (2007) Easily available enzymes as natural retting agents. Biotechnol J 2:342–346. doi:10.1002/biot.200600110

Armand S, Wagemaker MJM, Sanchez-Torres P, Kester HCM, van Santen Y, Dijkstra BW et al (2000) The active site topology of *Aspergillus niger* endopolygalacturonase II as studied by site-directed mutagenesis. J Biol Chem 275:691–696. doi:10.1074/jbc.275.1.691

Ashikari T, Nakamura N, Tanaka Y, Kiuchi N, Shibano Y, Tanaka T et al (1986) *Rhizopus* raw-starch degrading glucoamylase: its cloning and expression in yeast. Agric Biol Chem 50:957–964

Bergey DR, Orozco-Cardenas M, de Moura DS, Ryan CA (1999) A wound- and system in-inducible polygalacturonase in tomato leaves. Proc Natl Acad Sci USA 96:1756–1760. doi:10.1073/pnas.96.4.1756

Bonivento D, Pontiggia D, Di Matteo A, Fernandez-Recio J, Salvi G, Tsernoglou D et al (2008) Crystal structure of the endopolygalacturonase from the phytopathogenic fungus *Colletotrichum lupini* and its interaction with polygalacturonase–inhibiting proteins. Proteins 70:294–299. doi:10.1002/prot.21610

Bussink HJD, Buxton FP, Visser J (1991) Expression and sequence comparison of the *Aspergillus niger* and *Aspergillus tubigensis* genes encoding polygalacturonase II. Curr Genet 19:467–474. doi:10.1007/BF00312738

Di Lorenzo M, Hidalgo A, Haas M, Bornscheuer UT (2005) Heterologous production of functional forms of *Rhizopus oryzae* lipase in *Escherichia coli*. Appl Environ Microbiol 71:8974–8977. doi:10.1128/AEM.71.12.8974-8977.2005

Farley PC, Sullivan PA (1998) The *Rhizopus oryzae* secreted aspartic proteinase gene family: an analysis of gene expression. Microbiol 144:2355–2366

Foulk JA, Akin DE, Dodd RB (2008) Influence of pectinolytic enzymes on retting effectiveness and resultant fiber properties. BioResources 3:155–169

Frank AM, Savitski MM, Nielsen ML, Zuberev RA, Pevzner PA (2007) *De novo* peptide sequencing and identification with precise mass spectrometry. J Proteome Res 6:114–123. doi:10.1021/pr060271u

Friedberg D, Peleg Y, Monsonego A, Maissi S, Battat E, Rokem JS et al (1995) The *fumR* gene encoding fumarase in the filamentous fungus *Rhizopus oryzae*: cloning, structure and expression. Gene 163:139–144. doi:10.1016/0378-1119(95)00367-F

Henriksson G, Akin DE, Hanlin RT, Rodriguez C, Archibald DD, Rigsby LL et al (1997) Identification and retting efficiencies of fungi isolated from dew-retted flax in the United States and Europe. Appl Environ Microbiol 63:3950–3956

Henriksson G, Akin DE, Slomczynski D, Eriksson KL (1999) Production of highly efficient enzymes for flax retting by *Rhizomucor pusillus*. J Biotechnol 68:115–123. doi:10.1016/S0168-1656(98)00192-8

Ibrahim AS, Avanessian V, Spellberg B, Edwards JE Jr (2003) Liposomal amphoteric in B, and not amphotericin B deoxycholate, improves survival of diabetic mice infected with *Rhizopus oryzae*. Antimicrob Agents Chemother 47:3343–3344. doi:10.1128/AAC.47.10.3343-3344.2003

Kinter M, Sherman NE (2000) Protein sequencing and identification using tandem mass spectrometry. Wiley Interscience, New York

Kozlowski R, Batog J, Konczewicz W, Mackiewicz-Talarczyk M, Muzyczek M, Sedelnik N et al (2006) Enzymes in bast fibrous plant processing. Biotechnol Lett 28:761–765. doi:10.1007/s10529-006-9044-4

Markovic O, Janecek S (2001) Pectin degrading glycoside hydrolases of family 28: sequence-structural features, specificities and evolution. Protein Eng 14:615–631. doi:10.1093/protein/14.9.615

Mertens JA, Skory CD (2007a) Isolation and characterization of a second glucoamylase gene without a starch binding domain from *Rhizopus oryzae*. Enzyme Microb Technol 40:874–880. doi:10.1016/j.enzmictec.2006.07.003

Mertens JA, Skory CD (2007b) Isolation and characterization of two genes that encode active glucoamylase without a starch binding domain from *Rhizopus oryzae*. Curr Microbiol 54:462–466. doi:10.1007/s00284-006-0655-8

Miller GL (1959) Use of dinitrosalicylic acid reagent for determination of reducing sugar. Anal Chem 31:426–428. doi:10.1021/ac60147a030

Munoz JA, Coronado C, Perez-Hormaeche J, Kondorosi A, Ratel R, Palomares AJ (1998) *MsPG3*, a *Medicago sativa* polygalacturonase gene expressed during the alfalfa-*Rhizobium meliloti* interaction. Proc Natl Acad Sci USA 95:9687–9692. doi:10.1073/pnas.95.16.9687

Niture SK (2008) Comparative biochemical and structural characterizations of fungal polygalacturonases. Biologia 63:1–19. doi:10.2478/s11756-008-0018-y

Park K, Kwon S, Kim P, Bureau T, Kim N (2008) Gene structure dynamics and divergence of the polygalacturonase gene family of plants and fungus. Genome 51:30–40. doi:10.1139/G07-093

Pickersgill R, Smith D, Worboys K, Jenkins J (1998) Crystal structure of polygalacturonase from *Erwinia carotovora* ssp. *carotovora*. J Biol Chem 273:24660–24664. doi:10.1074/jbc.273.38.24660

Ribes JA, Vanover-Sams CL, Baker DJ (2000) Zygomycetes in human disease. Clin Microbiol Rev 13:236–301. doi:10.1128/CMR.13.2.236-301.2000

Roa Engel CA, Straathof AJJ, Zijlmans TW, van Gulik WM, van der Wielen LAM (2008) Fumaric acid production by fermentation. Appl Microbiol Biotechnol 78:379–389. doi:10.1007/s00253-007-1341-x

Saito K, Takakuwa N, Oda Y (2004) Purification of the extracellular pectinolytic enzyme from the fungus *Rhizopus oryzae* NBRC 4707. Microbiol Res 159:83–86. doi:10.1016/j.micres.2004.01.001

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Skory CD, Ibrahim AS (2007) Native and modified lactate dehydrogenase expression in a fumaric acid producing isolate *Rhizopus oryzae* 99-880. Curr Genet 52:23–33. doi:10.1007/s00294-007-0135-0

Tamaru Y, Doi RH (2001) Pectate lyase A, an enzymatic subunit of the *Clostridium cellulovorans* cellulosome. Proc Natl Acad Sci USA 98:4125–4129. doi:10.1073/pnas.071045598

Tebbutt SJ, Rogers HJ, Lonsdale DM (1994) Characterization of a tobacco gene encoding a pollen-specific polygalacturonase. Plant Mol Biol 25:283–297. doi:10.1007/BF00023244

van Santen Y, Benen JAE, Schroter KH, Kalk KH, Armand S, Visser J et al (1999) 1.68-A crystal structure of endopolygalacturonase II from *Aspergillus niger* and identification of active site residues by site-directed mutagenesis. J Biol Chem 274:30474–30480. doi:10.1074/jbc.274.43.30474

Xiao Z, Storms R, Tsang A (2005) Microplate-based carboxymethyl-cellulose assay for endoglucanase activity. Anal Biochem 342:176–178. doi:10.1016/j.ab.2005.01.052

Xiao Z, Boyd J, Grosse S, Beauchemin M, Coupe E, Lau PCK (2008) Mining *Xanthomonas* and *Streptomyces* genomes for new pectinase-encoding sequences and their heterologous expression in *Escherichia coli*. Appl Microbiol Biotechnol 78:973–981. doi:10.1007/s00253-008-1389-2

Yoshida S, Murakami R, Watanabe T, Koyama A (2001) *Rhizopus* rot of mulberry-grafted saplings caused by *Rhizopus oryzae*. J Gen Plant Pathol 67:291–293. doi:10.1007/PL00013033

Yoshida S, Suzuki F, Tsukiboshi T, Shinohara H (2004) Cloning and characterization of a gene rpg1 encoding polygalacturonase of *Rhizopus oryzae*. Mycol Res 108:1407–1414

Zhang J, Henriksson G, Johansson G (2000) Polygalacturonase is the key component in enzyme-retting of flax. J Biotechnol 81:85–89

Zhang J, Henriksson H, Szabo IJ, Henriksson G, Johansson G (2005) The active component in the flax-retting system of the zygomycete *Rhizopus oryzae* sb is a family 28 polygalacturonase. J Ind Microbiol Biotechnol 32:431–438. doi:10.1007/s10295-005-0014-y