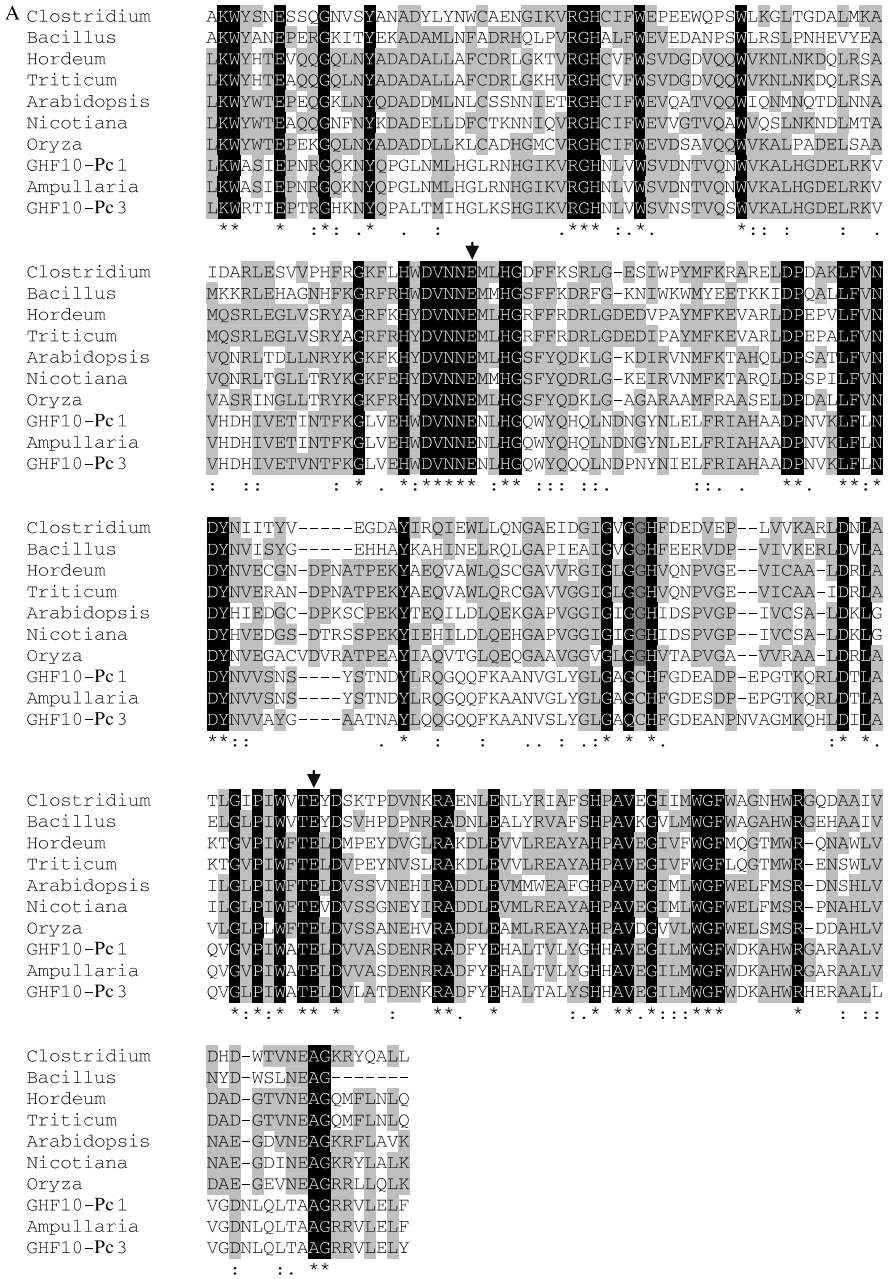
donor or nucleophile in GHF10 enzymes were conserved in GHF10-*Pc*1 and GHF10-*Pc*3 as E167 (putative catalytic acid/base) and E268 (putative catalytic nucleophile). Thus, we concluded that GHF10-*Pc*1 and GHF10-*Pc*3 is an enzyme belonging to GHF10. Conservation of the nucleophilic glutamate and the considerable sequence similarities between the other enzymes in glycosyl hydrolases family 10 suggested that the GHF10-*Pc*1 and GHF10-*Pc*3 should also be retaining enzymes.

Cellulolytic enzymes, on the basis of sequence similarities, can be classified into families. One of these families is known as the GHF10 (Henrissat 1991) or as the cellulase family F. This enzyme family is a large and diverse group of hydrolases that also comprises enzymes with a number of known activities; xylanase (EC 3.2.1.8), endo-l,3-β-xylanase (EC 3.2.1.32) and cellobiohydrolase (EC 3.2.1.91). It has been shown that known animal cellulases mostly belong to three glycosyl hydrolase families: GHF5 (plant parasitic nematodes), GHF9 (termites, cockroaches and crayfish) and GHF45 (mussels and beetles) but *A. crossean* EGX and *P. canaliculata* GHF10-*Pc*1 and GHF10-*Pc*3 possessed the family 10



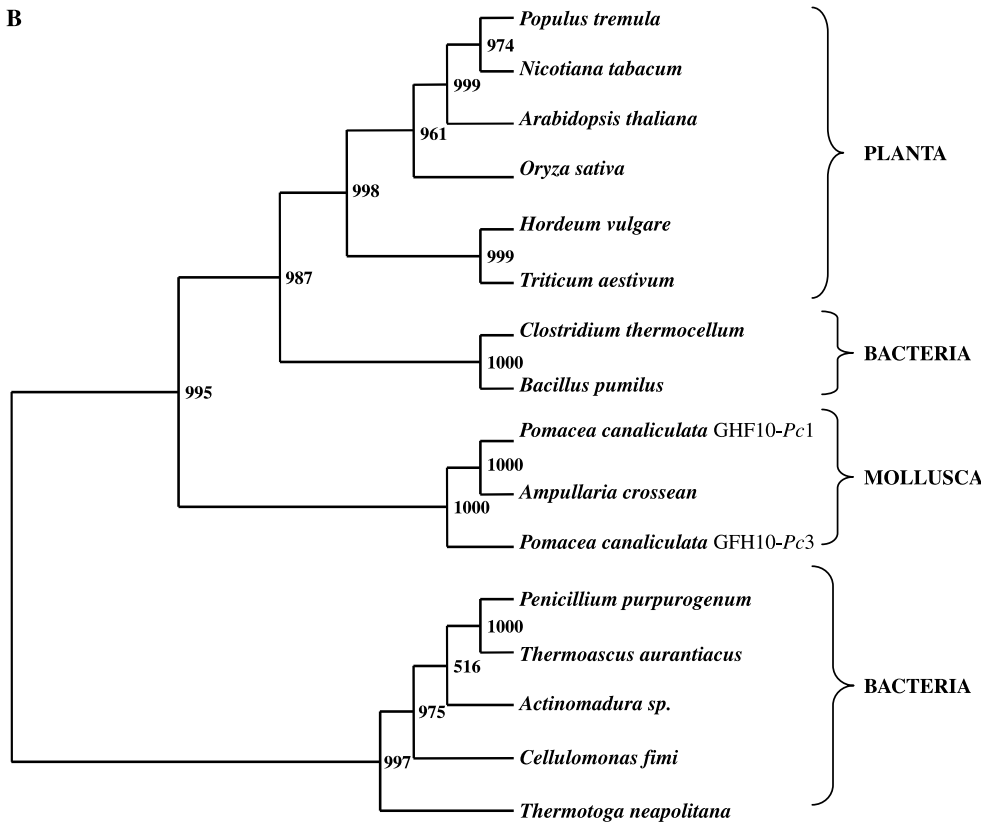


Figure 2. Multiple alignments of deduced amino acid sequence of catalytic GHF1O domain from various species (A) and a bootstrapped neighbour-joining tree (B). *P. canaliculata* cellulase gene (GHF10-*Pc*1 and GHF10-*Pc*3, this study), *A. crossean* (AAP31839), *Oryza sativa* (BAD52808), *Arabidopsis thaliana* (BAB83869), *Populus tremula* × *Populus tremuloides* (AAX33301), *Nicotiana tabacum* (AAZ79232), *Triticum aestivum* (AF156977), *Hordeum vulgare* (AAB51668), *Bacillus pumilus* (AF466829), *Clostridium thermocellum* (EAM47104), *Thermotoga neapolitana* (Q60041), *Actinomadura sp.* (AAA17888), *Penicillium purpurogenum* (AAF71268), *Cellulomanas fimi* (AAZ76373) and *Thermoascus aurantiacus* (AF 127529). Black indicates complete conservation. Grey indicates conservation in three or more species. The conserved catalytic nucleophile and proton donor of Glu268 and the Glu167 are arrowed. Values at the node of a bootstrapped neighbour-joining tree indicate the percentage of times that the particular node occurred in 1000 trees generated by bootstrapping the original deduced protein sequences.

GHF. Accordingly, *P. canaliculata* GHF10-*Pc*1 and GHF10-*Pc*3 *is* considered to be the second animal cellulase that belongs to family 10 GHF.

To determine the relatedness of the GHF10-*Pc*1 and GHF10-*Pc*3 with other sequences, a phylogenetic tree was constructed based on the amino acid sequence of catalytic GHF 10 domain of GHF10-*Pc*1 and GHF10-*Pc*3 and other GHF 10 members. A bootstrapped NJ tree constructed from a sequence divergence of deduced amino acids (Figure 2B) clearly indicates that GHF10-*Pc*1 and GHF10-*Pc*3 is more closely linked to mollusca (*A. crossean*) cellulase than with plant cellulases and bacterial cellulases.

Genomic organization of cellulase GHF10-Pc1 and GHF10-Pc3 genes

To identify the genomic structure of GHF10-*Pc*1 and GHF10-*Pc*3 genes and to confirm the existence of the GHF10-*Pc*1 and GHF10-*Pc*3 gene in the *P. canaliculata* chromosomal DNA, a genomic polymerase chain reaction (PCR) approach was performed. The genomic PCR amplification generated two bands of approximately 4.6 and 5-kb fragment. PCR products were cloned and sequenced. The GHF10-*Pc*1 and GHF10-*Pc*3 gene sequences have been deposited in the GenBank nucleotide sequence database under the accession no. DQ848669 and DQ848670.

The GHF10-*Pc*1 and GHF10-*Pc*3 genomic sequences were 4937 and 4512 bp in length, respectively. The intron–exon boundaries of the GHF10-*Pc*1 and GHF10-*Pc*3 genes were determined by aligning the genomic DNA with cDNA of each corresponding gene (Figure 3A, B). Both genes were composed of 9 exons interrupted by 8 introns. The sequence of GHF10-*Pc*1 and GHF10-*Pc*3 gene was in good agreement with the cDNA sequence. The sixth intron was the shortest (234 bp in GHF10-*Pc*1 and 232 bp in GHF10-*Pc*3), the last eighth intron was the longest (1071 bp in GHF10-*Pc*1 and 671 bp in GHF10-*Pc*3) (Table II). The size of the introns between *P. canaliculata* GHF10 and *A. crossean* EGX was quite

different. The length of intron 2 in *P. canaliculata* GHF10-*Pc*1 and GHF10-*Pc*3 (283 and 418 bp, respectively) was approximately 3- or 2-fold smaller than that of intron 2 (1032 bp) in *A. crossean* EGX.

The sequences of all the exon–intron boundaries conformed to the typical eukaryotic splice sites, including an invariant GT at the intron 5′ boundary and an invariant AG at its 3′ boundary. In contrast, the exon–intron boundaries of corresponding introns in *A. crossean* EGX conformed to GT–AG rule only in intron number 1, 5, 7 and 8 (Wang et al. 2003). Therefore, in spite of the high similarity of *P. canaliculata* GHF10-*Pc*1 and GHF10-*Pc*3 cDNA to *A. crossean* EGX cDNA, the difference in genomic organization of these genes was observed.

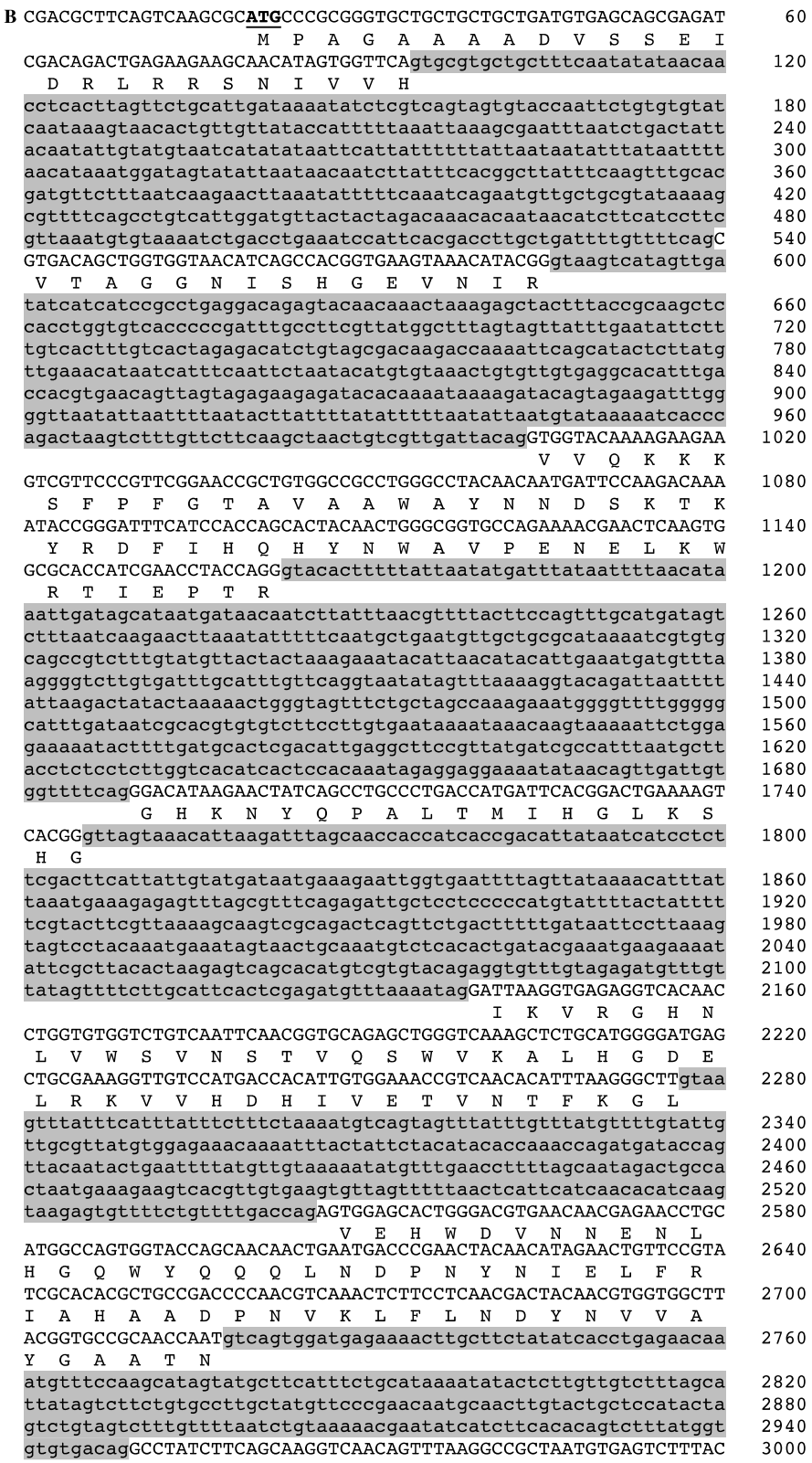
Itakura et al. (2006) reported successful identification of termite species by using cDNA and genomic DNA sequences of termite origin. A pair of common primers that amplified 381-bp fragments from cDNAs encoding the endo-β-1,4-glucanases (EGases) of *Coptotermes formosanus* Shiraki and *Reticulitermes speratus* (Kolbe). The cDNAs from *C. formosanus* and *R. speratus* and genomic DNA from *R. speratus*, were amplified by using this primer pair and then cloned and sequenced. Sequences amplified from *C. formosanus* cDNA displayed 97–99% identity to cDNA encoding the EGase of *C. formosanus* (*CfEG*), and 92–94% identity to cDNA encoding the EGase of *R. speratus* (*RsEG*). By contrast, cDNA from *R. speratus* displayed 99–100% identity to *RsEG* cDNA and 93–94% identity to *CfEG* cDNA. *CfEG* and *RsEG* cDNAs can therefore be used as markers for the identification of these termite species. Accordingly, the different in intron size of the cellulase genes in two snail species may be used as markers for species identification.

The GC content analysis showed a slightly greater thermal stability in exons (31–58%) than in introns (26–35%) (Table II). Intron 1, 4 and 5 interrupt ORFs after the 2nd of the codons (type 2 intron) whereas the remaining introns interrupt ORFs between two codons (type 0 intron).

All of intron sequences from GHF10-*Pc*1 and GHF10-*Pc*3 genes were examined to identify potential homology between introns. A strong sequence homology existed between the corrresponding intron of GHF10-*Pc*1 and GHF10-*Pc*3. Intron numbers 1, 4–7 of GHF10-*Pc*1 and GHF10-*Pc*3 shared 86–97% identical internal sequence as well as similarity in size. In contrast, the size of the last GHF10-*Pc*1 intron

```
A CGACGCTTCAGTCAAGCGCATGCCCGCTGGTGCTGCTGGTGCTGGGGTGACCAGCGACAT      60
                      M  P  A  G  A  A  G  A  G  V  T  S  D  I
  CGACAGACTGAGAAGAAGCGACATAACGGTTCAgtgcgtgctgctttcaatatataacaa     120
    D  R  L  R  R  S  D  I  T  V  H
  cctcacttagttctgcattgataaaatatctcgtcagtagtgtaccaattctgtgtgtat     180
  caagaaagtaacaccgttgttataccatttttaaagtaaagcgaatttaatctggaatag     240
  tacaatacttgtatgtaatgcatatataattcatgattttggtagtaataataagtataa     300
  ttgtaacataaacggatagaatattaataacaatcttatgtcacgggtcatttcaagttt     360
  gcacgatatcctggaatcaagaaccgaaatagtaattaaggccgaacgttgcggcgtaga     420
  aaatcgtgtgcagcccgtcatgggatgttaatactaaagaaacacaataacatccggcag     480
  ccgtcgtcaagtgtgtaaaatctgacctgaaatccattcacgaagttgatgatttcttat     540
  atcagCGTGAATGTTGGTGGTAACATCAACCACGGTCAAGTGAGCATTCGTgtaagtcat     600
         V  N  V  G  G  N  I  N  H  G  Q  V  S  I  R
  agatgagagcatcatccgcctgaggacaggggcaataaaaagagcaagctcatcaaagg     660
  tgtcaacctcgagaatgtgccttaagtaaacttgattgcgaggtacagttgaccacgtga     720
  acagataggagagaagagagacaaaaaatgaaggatacaatagaagatgtaggaataata     780
  ttaattttaatacttatttttatatttaatgttaatgtataaaagccacccagactaaat     840
  ctttgttcttcaaggtaactgtcgttgattacagGTGTTACAAAAGAAAAAGGCATTCCC     900
                                    V  L  Q  K  K  K  A  F  P
  GTTCGGGACATGTGTGGCCGCCTGGGCCTACAACGATGGGTCCAAAGGAGCATACCGGGA     960
    F  G  T  C  V  A  A  W  A  Y  N  D  G  S  K  G  A  Y  R  D
  TTTCATCCACCAGCACTACAACTGGGCCGTGCCAGAAAACTCACTCAAGTGGGCTAGCAT    1020
    F  I  H  Q  H  Y  N  W  A  V  P  E  N  S  L  K  W  A  S  I
  CGAACCTAACAGGgtacatttttttattaatgagatttataaatttaacataaattgatag    1080
    E  P  N  R
  tatattaataaaaatcttctttaacgttttacttccagtttgcacgatagtctttaatca    1140
  agaactgaaatatttttcaatactgaatgttgctgcgcataaaatcgtttgcagcctgtc    1200
  atttgatgttactactaaagaaatacaataacatatattgaaatgatgtttaaggggtct    1260
  tctgatttatatctgttcagggagagaatagtttaaaaggtactccattcacagcaagga    1320
  aagacaaagttacccgttgtcatattcatagttttactttctaaatcgtgtgtgtgtgtg    1380
  tatatatatttatgtctgagcgcgcttgaatgcttaagtcgacaacttaatgtaacagat    1440
  tagttttattaagagtatactaaaaactgggttgttcctgctagccaaagaaatggggtt    1500
  ttgggggtatttgataatcgcacgtgtgtcttccttgtgaataaaataaacaagtaaaaa    1560
  ttctggagggaacatacttttgaggcactcgacattgaggcttccgttatgatcgccatt    1620
  taatgcttgcctctcctcttggacatatagaaggggaaaatataatgactgttgattgtg    1680
  ttttcagGGACAAAAGAATTATCAGCCTGGCCTAAACATGCTTCACGGACTGAGAAATC    1740
           G  Q  K  N  Y  Q  P  G  L  N  M  L  H  G  L  R  N
  ACGGgttagtaaacattaggatactgcaacccccatcaccgacattataatcatcatcgt    1800
  H  G
```

```
cgacttcattatcgtataataatgaaagaattggcgaattttaattatttaagaattatt   1860
aacataaatttaattaacatttaataaatcaaagaaagtttagcgtttcagagattgctc   1920
ctcccccatgtattttactatttctatcacttcgttaaaagcaagtctcagactcagttc   1980
tgtctttttgataattccttaaaatagtcctacaaatgaaatagtaactgaaatgtctca   2040
cattgatacgaaattaagaaaatattcacttacactaagagttaacacatgtcgtgtaca   2100
gaggtgtttgtagaaatgtttgttatagttttcttgcatccactcgagatgtttaaaaca   2160
gGATTAAGGTGAGAGGTCACAACCTGGTGTGGTCTGTCGACAATACGGTGCAGAACTGGG   2220
   I  K  V  R  G  H  N  L  V  W  S  V  D  N  T  V  Q  N  W
TGAAGGCTCTGCATGGGGATGAGCTGCGAAAGGTTGTCCATGACCACATCGTGGAAACCA   2280
V  K  A  L  H  G  D  E  L  R  K  V  V  H  D  H  I  V  E  T
TCAACACATTTAAGGGCCTgtaagtttatatcatttgttactttctttctaaaatgtcag   2340
I  N  T  F  K  G  L
tagtttatttttttatttttatcttgttatttggagaaacaaaatttactgttctacata   2400
caccaaaccagatgataccagttacaatactgaattttatgttgtaaaaatatgtttgaa   2460
ccttatagcaatagactgccactaatttaagaagtcacgttgtcaagtgttagtttttaa   2520
ctcactcatcaacagatcaagcaagaggttttttctgttttgaccagAGTGGAGCACTGG   2580
                                                  V  E  H  W
GATGTGAACAACGAGAACCTGCATGGCCAGTGGTACCAGCATCAACTGAATGACAATGGC   2640
 D  V  N  N  E  N  L  H  G  Q  W  Y  Q  H  Q  L  N  D  N  G
TACAACCTGGAACTGTTCCGTATCGCACACGCCGCCGACCCCAACGTCAAACTCTTCCTC   2700
 Y  N  L  E  L  F  R  I  A  H  A  A  D  P  N  V  K  L  F  L
AACGACTACAACGTTGTGTCCAACAGTTATTCAACAAACgtgagtgaatgagaaaacttg   2760
 N  D  Y  N  V  V  S  N  S  Y  S  T  N
cttctacatcacctgagaacaaatgtttgcaagcatcgtatgctttatttctgcctaaaa   2820
tatactcttgttgtctttagcattatagtcttctgtgccttgctatgttgccgaacaatg   2880
caacttgtactgctccatactagtctgtagtgtctatttttttttatctgtaaaaacgaat   2940
atcatcttcacaagtctttgcggtgtgtgacagGACTATCTTCGACAAGGTCAACAGTTT   3000
                                    D  Y  L  R  Q  G  Q   Q  F
AAGGCCGCTAATGTGGGTCTTTACGGTTTGGGTGCTCAATGCCACTTTGGCGACGAAGCC   3060
 K  A  A  N  V  G  L  Y  G  L  G  A  Q  C  H  F  G  D  E  A
GACCCAGAACCCGGTACTAAGgtaagaaaaataatataaacatccttcgcctttctcatt   3120
 D  P  E  P  G  T  K
ccccattttattattattctaaaaaataaaacaaaggtacgcacaacaatgaattagttc   3180
tacaatgcagtaagatacaagatttcagcttctgtccgtatccttttttctatttagtta   3240
tttgcatttttttactattatttgttaatttacgttaaatggaataacatatctttatgt   3300
tgagaaatcaaaaagtcttcattttatagctgtggatgtaaaaaatactcatctaacgtc   3360
ttctatactccaaaagctgttttagctacatgtgggttttttttctaatctcagCAACGTC   3420
                                                      Q  R
TGGATACTTTAGCTCAAGTGGGCGTGCCCATCTGGGCCACTGAGTTGGATGTGGTAGCTT   3480
L  D  T  L  A  Q  V  G  V  P  I  W  A  T  E  L  D  V  V  A
CGGATGAGAACAGACGAGCGGATTTCTACGAGCACGCGCTGACAGTCCTGTACGGCCATC   3540
S  D  E  N  R  R  A  D  F  Y  E  H  A  L  T  V  L  Y  G  H
ATGCCGTGGAGGGCATCCTGATGTGGGGCTTCTGGGACAAGGCCCACTGGCGTGGTGCCA   3600
H  A  V  E  G  I  L  M  W  G  F  W  D  K  A  H  W  R  G  A
GAGCTGCTCTTGTTGTCGGAGACAACCTGCAGgtgtgttgggttacaattaacctaatgg   3660
R  A  A  L  V  V  G  D  N  L  Q
gaatcttacaggtattatagcacagccagcatccagacagacatttaggaatcactgtca   3720
tgatgagaacacattggattcttcttttattagaatttctacccagaactacccttgact   3780
tcaacttttctttttccttataagagacatctgacaaatgcttactcagtgtacctatgg   3840
ttagtccagattatgcgtgtatgtttttccattttagaattactttagcacaaacacgac   3900
caaaacgattaccatctatgtgaggacaaagatacgtgttctgcttgtttatatgaata   3960
atattgttatgataaaactgcaaaatactcagtttgaagtagacatttttttccatgaaaa   4020
ttcgcaaaaaaaagaaagaccggaaatgaacccatttgaacgagtttaaaaaattgttta   4080
cagcaaagtccttttatctctgctataaatgtgttgaaggtaaatgtatattcacatcagt   4140
attatttgactgagaatgttttcggctgttttcttcatatataatcaaaagtatcacaat   4200
gttcattgaagaaacgaggacacctggttgtagccaacgacattgtttgacattaatgca   4260
aataataacgtgttacaacaatctaagcggtactttgacagaagtgtaaaaattaatatt   4320
gagacttcatcaacgcttaaacaattccaagactttaaaattgcaagaatcaagcttgcg   4380
cgctttttttcacatcgataaggatgatgttataatgtaaccagctggtaaagtgaaaca   4440
ctgtttaaattactattggttgcggcagaagagttatgacgttgtaaaggacattgtctt   4500
acttaaggattgcttcattcgaaatacttctgatagtgaatgcgaaacgcacattccatc   4560
aataataaaagtataagatatgagaattgttactgaaatatgtctgttttagtctccatg   4620
attagccaaatgattgctttcgttaaagttcatgaaacaccgggtgtttgtggtagacat   4680
aaaataaatgtttgattctatagCTGACGGCGGCCGGACGTCGCGTGCTGGAGCTCTTTG   4740
                        L  T  A  A  G  R  R  V  L  E  L  F
AGCACAGGTGGATGACAGACGAGACGCACAACCTGGCAGCGGGCACCCAGTTCACAGTAC   4800
E  H  R  W  M  T  D  E  T  H  N  L  A  A  G  T  Q  F  T  V
GCGGTTTCCATGGCGACTACGAGGTGCAAGTCATCGTCCAGGGTCAAGAGCACACCAACC   4860
R  G  F  H  G  D  Y  E  V  Q  V  I  V  Q  G  Q  E  H  T  N
TGAGGCAGACGTTCTCGTTGGGCAACGGTCCCCACACCGTCAACATTAACATTAGCTAGA   4920
L  R  Q  T  F  S  L  G  N  G  P  H  T  V  N  I  N  I  S  *
GCGACACTCAGAGGGCA                                           4937
```



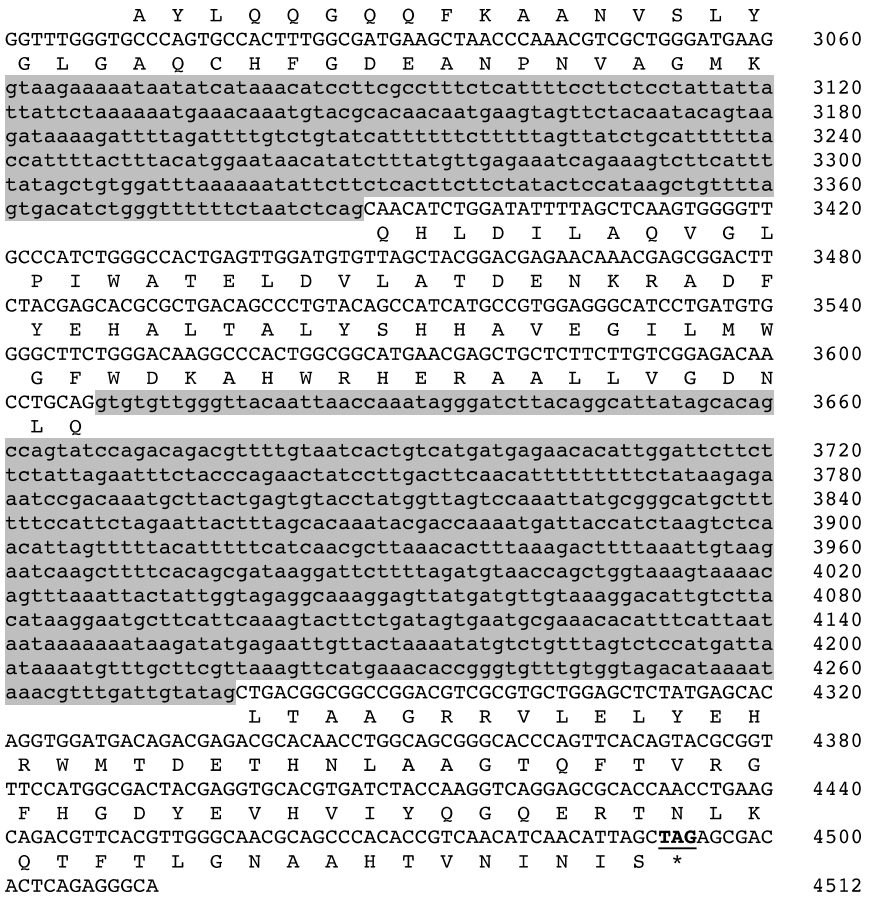


Figure 3. Genomic structures of GHF10-*Pc*1 (A) and GHF10-*Pc*3 (B). Coding nucleotides and deduced amino acids of each exon are capitalized. Introns are shaded and illustrated with lower letters.

is 1071 bp as compared to that of GHF10-*Pc*3, which is 671 bp.

Expression analysis by reverse transcription-polymerase chain reaction

To examine the expression pattern of GHF10-*Pc*1 and GHF10-*Pc*3 in *P. canaliculata,* RT-PCR was carried out against the first-stranded cDNA synthesized from total RNAs of different ages of juvenile snails. Results showed that none of GHF10-*Pc*1 and GHF10-*Pc*3 transcripts was detected at the egg stage whereas the fragment of approximately 1200 bp RT-PCR product was amplified in 1- and 10-day-old snails (Figure 4A). Elongation factor la generated band of 150 bp (Figures 4B and 5B). The undetectable of mRNA transcripts encoded by the GHF10-*Pc*1 and GHF10-*Pc*3 genes from the eggs suggested that the expression of both genes should start during or after the hatching period.

From sequence analysis, both GHF10-*Pc*1 and GHF10-*Pc*3 cDNA sequences contain *Rsa* I recognition sites at different location. Therefore, this enzyme can be used to differentiate the family of GHF10-*Pc*1 and GHF10-*Pc*3 cDNAs. The *Rsa* I digested GHF10-*Pc*1 showed restriction fragments of approximately 519, 227, 196, 124 and 122 bp, respectively whereas *Rsa* I digested GHF10-*Pc*3 gave fragments of 395, 354, 196, 124 and 122 bp, respectively (Figure 5A). The result showed that the *Rsa* I digested PCR product of 1- day-old snails revealed restriction fragments that corresponded well to both the GHF10-*Pc*1 and GHF10-*Pc*3 pattern. In contrast, *Rsa* I digested PCR product of 10-day-old snails corresponded only with the GHF10-*Pc*1 pattern. The restriction profile indicated that the GHF10-*Pc*1 was expressed in both ages of juvenile *P. canaliculata* whereas GHF10-*Pc*3 was observed in 1-day old snails but not in 10-day-old snails. Our result suggested that the GHF10-*Pc*3 transcripts should have expressed in the early development of juvenile snail. However, the expression of cellulase GHF10-*Pc*1 and GHF10-*Pc*3 transcripts was investigated in pooled cDNA of the whole body of 1 and 10-day-old juvenile *P. canaliculata*. Therefore, it is interesting for further studies to examine the cellular type involved in cellulase GHF10-*Pc*1 and GHF10-*Pc*3 expression using *in situ* hybridisation.

Table II. GC content and length of exons and introns in GHF10-*Pc*1 and GHF10-*Pc*3 genes.

Exon	Genomic DNA (Number of nucleotides)	GC content (%)	Intron	Genomic DNA (Number of nucleotides)	GC content (%)
GHF10-*Pc*1					
1	1–93 (93 bp)	53	1	94–545 (452 bp)	34
2	546–591 (46 bp)	36	2	592–874 (283 bp)	35
3	875–1033 (159 bp)	50	3	1034–1688 (655 bp)	32
4	1689–1744 (56 bp)	35	4	1745–2161 (417 bp)	30
5	2162–2299 (138 bp)	48	5	2300–2567 (268 bp)	27
6	2568–2729 (172 bp)	48	6	2740–2973 (234 bp)	35
7	2974–3081 (108 bp)	48	7	3082–3413 (332 bp)	27
8	3414–3632 (219 bp)	56	8	3633–4703 (1071 bp)	33
9	4704–4937 (234 bp)	56			
GHF10-*Pc*3					
1	1–93 (93 bp)	48	1	94–539 (446 bp)	27
2	540–585 (46 bp)	38	2	586–1003 (418 bp)	33
3	1004–1162 (159 bp)	49	3	1163–1689 (527 bp)	31
4	1690–1745 (56 bp)	31	4	1746–2138 (393 bp)	32
5	2139–2276 (138 bp)	48	5	2277–2545 (269 bp)	28
6	2546–2717 (172 bp)	51	6	2718–2949 (232 bp)	28
7	2950–3060 (111 bp)	46	7	3061–3388 (328 bp)	26
8	3389–3607 (219 bp)	55	8	3608–4278 (671 bp)	32
9	4279–4512 (234 bp)	58			

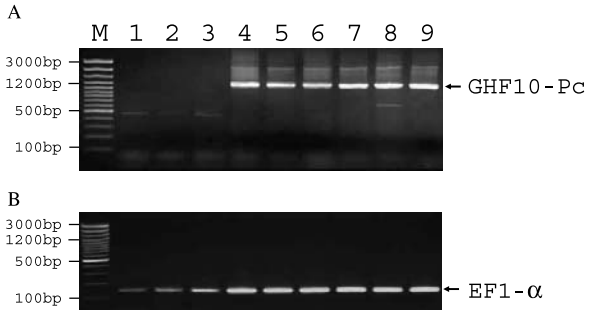


Figure 4. RT-PCR of GHF10-*Pc* (1200 bp; A) and EF1α (150 bp; B) was examined in the egg (lane 1–3), 1- day-old (lane 4–6), and 10-day-old (lane 7–9) juveniles of *P. canaliculata*. Lanes M are 100-bp DNA marker.

Here we show that in spite of the high similarity of GHF10-*Pc*1 and GHF10-*Pc*3, the two cellulase genes revealed different expression patterns. The differential expression of the two isoforms supports the idea that these forms may play different physiological roles in *P. canaliculata*. However, the functional role of the cellulase gene and its catalytic properties remains to be elucidated.

In summary, we have successfully cloned two full-length cDNAs of cellulase GHF10-*Pc*1 and GHF10-*Pc*3 belonging to GHF10 from the fresh water golden apple snail, *P. canaliculata*. We also report here the gene structures of two endogenous GHF10-*Pc*1 and GHF10-*Pc*3 genes. Both genes contained nine exons and eight introns. The endogenous origin of these cellulase genes in *P. canaliculata* was unambiguously verified in this study. The expression analysis of GHF10-*Pc*1 and GHF10-*Pc*3 during several ages of *P. canaliculata* was analysed by RT-PCR. The result showed that GHF10-*Pc*1 and GHF10-*Pc*3 transcripts were developmentally expressed.

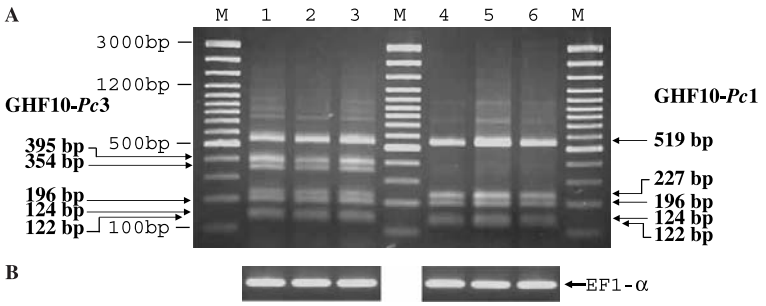


Figure 5. Restriction patterns of the *Rsa* I-digested RT-PCR products of GHF10-*Pc*1 and GHF10-*Pc*3 (A) and EF1α (B). The 1-day old juveniles (lane 1–3) of *P. canaliculata* revealed restriction patterns corresponding to both the GHF10-*Pc*1 (519, 227, 196, 124 and 122 bp) and GHF10-*Pc*3 (395, 354, 196, 124 and 122 bp) whereas the 10-day old juveniles (lane 4–6) corresponded only with the GHF10-*Pc*1 pattern. Lanes M are 100-bp DNA marker.

Acknowledgements

This work was supported by grant from the Thailand Research Fund (TRF), grant number MRG4980168. It was also supported by Ratchadaphiseksomphot Endowment Fund, Chulalongkorn University, awarded to Chanprapa Imjongjirak.

References

Amparyup P, Jitvaropas R, Pulsook N, Tassanakajon A. 2007. Molecular cloning, characterization and expression of a masquerade-like serine proteinase homologue from black tiger shrimp *Penaeus monodon*. Fish Shellfish Immunol 22:535–546.

Beguin P, Aubert JP. 1994. The biological degradation of cellulose. FEMS Microbiol Rev 13:25–58.

Brummell DA, Lashbrook CC, Bennett AB. 1994. Plant endo-l,4-β-D-glucanases. ACS Symp Ser 566:100–129.

Byrne KA, Lehnert SA, Johnson SE, Moore SS. 1999. Isolation of a cDNA encoding a putative cellulase in the red claw crayfish *Cherax quadricarinatus*. Gene 239:317–324.

Gilkes NR, Henrissat B, Kilburn DG, Miller RC, Warren RAJ. 1991. Domains in microbial β-l,4-glycanases: Sequence conservation, function, and enzyme families. Microbiol Rev 55: 303–315.

Henrissat BA. 1991. Classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem J 280:309–316.

Imjongjirak C, Klinbunga S, Sittipraneed S. 2005. Cloning, expression and genomic organization of genes encoding major royal jelly protein 1 and 2 of the Honey 20 Bee (*Apis cerana*). J Biochem Mol Biol 38:49–57.

Imjongjirak C, Amparyup P, Tassanakajon A, Sittipraneed S. 2007. Antilipopolysaccharide factor (ALF) of mud crab *Scylla paramamosain*: Molecular cloning, genomic organization and the antimicrobial activity of its synthetic LPS binding domain. Mol Immunol 44:3195–3203.

Itakura S, Masuta T, Tanaka H, Enoki A. 2006. Identification of two subterranean termite species (Isoptera: Rhinotermitidae) using cellulase genes. J Econ Entomol 99:123–128.

Keawjam RS, Upatham ES. 1990. Shell morphology, reproductive anatomy and genetic patterns of three species of apple snails of the genus *Pomacea* in Thailand. J Med Appl Malacol 2:45–57.

Moriya S, Ohkuma M, Kudo T. 1998. Phylogenetic position of symbiotic protist *Dinemympha exilis* in the hindgut of the termite *Reticulitermes speratus* inferred from the protein phylogeny of elongation factor lα. Gene 210:221–227.

Nathan L, Watanabe H, Sugimura M. 2003. Evidence for the presence of a cellulase gene in the last common ancestor of bilaterian animals. Proc R Soc Lond B (Suppl.) 270:S69–S72.

Smant G, Stokkermans JPWG, Yan Y, De Boer JM, Baum TJ, Wang X, Hussey RS, Gommers FJ, Henrissat B, Davis EL, Helder J, Schots A, Bakker J. 1998. Endogenous cellulases in animals: Isolation of β-l,4-glucanase genes from two species of plant-parasitic cyst nematodes. Proc Natl Acad Sci USA 95: 4906–4911.

Suzuki M, Watanave Y, Tanaka A, Ohnishi T. 1986. Bull Tokai Reg Fish Res Lab 120:53.

Suzuki K, Ojima T, Nishita K. 2003. Purification and cDNA cloning of a cellulase from abalone *Haliotis discus hannai*. Eur J Biochem 270:771–778.

Thaewnon-ngiw B, Klinbunga S, Phanwichien K, Sangduen N, Lauhachinda N, Menasveta P. 2004. Genetic diversity and molecular markers in introduced and Thai native apple snails (*Pomacea* and *Pila*). J Biochem Mol Biol 37:493–502.

Tokuda G, Lo N, Watanabe H, Slaytor M, Matsumoto T, Noda H. 1999. Metazoan cellulase genes form termites: Intron/exon structure and sites of expression. Biochem Biophys Acta 1447: 146–159.

Tomme P, Warren RAJ, Gilkes NR. 1995. Cellulose hydrolysis by bacteria and fungi. Adv Microbiol Physiol 37:1–81.

Wang H, Jones RW. 1995a. Cloning, characterization and functional expression of an endoglucanase-encoding gene from the phytopathogenic fungus *Macrophomina phaseolina*. Gene 158: 125–128.

Wang H, Jones RW. 1995b. A unique endoglucanase-encoding gene cloned from the phytopathogenic fungus *Macrophomina phaseolina*. Appl Environ Microbiol 61:2004–2006.

Wang J, Ding M, Li YH, Chen QX, Xu GJ, Zhao FK. 2003. Isolation of a multi-functional endogenous cellulase gene from mollusc, *Ampullaria crossean*. Acta Biochim Biophys Sin 35: 941–946.

Watanabe H, Nakamura M, Tokuda G, Yamaoka I, Scrivener AM, Noda H. 1997. Site of secretion and properties of endogenous endo-β-l,4-glucanase components from *Reticulitermes speratus* (Kolbe), a Japanese subterranean termite. Insect Biochem Mol Biol 27:305–313.

Xue XM, Anderson AJ, Richardson NA, Anderson AJ, Xue GP, Mather PB. 1999. Characterisation of cellulase activity in the digestive system of the redclaw crayfish (*Cheraxquadricarinatus*). Aquaculture 180:373–386.

Xu B, Ersson B, Hellman U, Janson JC. 2000. Purification, characterization and amino acid sequence analysis of a thermostable, low molecular mass endo-β-1,4-glucanase from blue mussel, *Mytilus edulis*. Eur J Biochem 267:4970–4977.

Xu B, Janson JC, Sellos D. 2001. Cloning and sequencing of a molluscan endo-β-1,4-glucanase gene from the blue mussel, *Mytilus edulis*. Eur J Biochem 268:3718–3727.

Yan Y, Smant G, Stokkermans J, Qin L, Helder J, Baum T, Schots A, Davis E. 1998. Genomic organization of four β-l,4-endoglucanase genes in plant-parasitic cyst nematodes and its evolutionary implications. Gene 220:61–70.