

RESEARCH ARTICLE

Open Access



Comparative profiling of hepatopancreas transcriptomes in satiated and starving *Pomacea canaliculata*

Lei Yang^{1*}, Tian-yin Cheng² and Fei-yan Zhao¹

Abstract

Background: Although *Pomacea canaliculata* is native to South and Central America, it has become one of the most abundant invasive species worldwide and causes extensive damage to agriculture and horticulture. Conventional physical and chemical techniques have been used to eliminate *P. canaliculata*, but the effects are not ideal. Therefore, it is important to devise a new method based on a greater understanding of the biology of *P. canaliculata*. However, the molecular mechanisms underlying digestion and absorption in *P. canaliculata* are not well understood due to the lack of available genomic information for this species, particularly for digestive enzyme genes.

Results: In the present study, hepatopancreas transcriptome sequencing produced over 223 million high-quality reads, and a global de novo assembly generated a total of 87,766 unique transcripts (unigenes), of which 19,942 (22.7%) had significant similarities to proteins in the UniProt database. In addition, 296,675 annotated sequences were associated with Gene Ontology (GO) terms. A Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment was performed for the unique unigenes, and 262 pathways (p -value $< 10^{-5}$) in *P. canaliculata* were found to be predominantly related to plant consumption and coarse fiber digestion and absorption. These transcripts were classified into four large categories: hydrolase, transferase, isomerase and cytochrome P450. The Reads Per Kilobase of transcript per Million mapped reads (RPKM) analysis showed that there were 523 down-regulated unigenes and 406 up-regulated unigenes in the starving apple snails compared with the satiated apple snails. Several important genes are associated with digestion and absorption in plants: endo-beta-1, 4-glucanase, xylanase, cellulase, cellulase EGX1, cellulase EGX3 and G-type lysozyme genes were identified. The qRT-PCR results confirmed that the expression patterns of these genes (except for the longipain gene) were consistent with the RNA-Seq results.

Conclusions: Our results provide a more comprehensive understanding of the molecular genes associated with hepatopancreas functioning. Differentially expressed genes corresponding to critical metabolic pathways were detected in the transcriptome of starving apple snails compared with satiated apple snails. In addition to the cellulase gene, other genes were identified that may be important factors in plant matter metabolism in *P. canaliculata*, and this information has the potential to expedite the study of digestive physiology in apple snails.

Keywords: *Pomacea canaliculata*, Hepatopancreas, Transcriptome, Digestion

* Correspondence: zhazha_yanglei@163.com

¹Laboratory of Molecular Physiology, College of Veterinary Medicine, Hunan Agricultural University, Changsha, Hunan Province 410128, People's Republic of China

Full list of author information is available at the end of the article



Background

Pomacea canaliculata (apple snail) is a member of Gastropoda and originates from South and Central America, and it is beginning to emerge in locations worldwide, including China, representing one of the 100 most invasive species in the world [1–4]. *P. canaliculata* features high reproductive capacity, strong adaptability, large food intake and a lack of effective predators, which indicate that it is a considerable threat to the ecosystem balance of fields and ponds [5–7]. This species has rapidly adapted to new environments and is now found in at least 11 provinces in southern China [8].

Presently, medication and predation approaches are applied to eradicate *P. canaliculata*. However, the results are not satisfactory. Moreover, these approaches are associated with drug residues and animals such as geese and ducks, which exert a negative influence on crops. In addition, the control of *P. canaliculata* in wetland ecosystems has received limited attention [9–15].

The hepatopancreas is an important digestive organ in *P. canaliculata* [16], and the expression of digestive enzyme genes in *P. canaliculata* may be closely related to nutrient metabolism and energy balance. The relative lack of information on enzyme genes has impeded research on the digestive physiology of *P. canaliculata*. Because of the advancements in DNA sequencing technology, generating a large amount of sequence data from non-model organisms such as the apple snail has become increasingly affordable. Although transcriptomic analyses of the hepatopancreas have been performed in several snail species [17, 18] using RNA-Seq, limited information is available on the *P. canaliculata* hepatopancreas transcriptome [19]. Moreover, the available reports were confined to topics such as invasion prevention, population control, species discrimination, etc. [20, 21]. Therefore, a dearth of research has been conducted on the digestive physiology of *P. canaliculata*. A transcriptome analysis can enrich the supply of genetic information on the apple snail, explore its digestion and the molecular mechanisms of its growth and development, and provide a reference for research on targeted drugs for snail control.

In this study, we divided snails into two groups, starving apple snails and satiated apple snails; collected total hepatopancreatic RNA from both groups; and constructed cDNA libraries. The cDNAs were sequenced using an Illumina/Solexa platform. Approximately eight gigabytes of high-quality sequencing data were generated. For the functional annotation of the assembled gene transcripts (contigs), we searched against the NCBI nonredundant (NR) and UniProt databases using BlastX. In addition, we compared the gene expression in the hepatopancreatic tissue between satiated apple snails and starving apple snails. This strategy can help further our understanding of

the biological responses of *P. canaliculata* and the hepatopancreas transcriptome dynamics of plant consumption in *P. canaliculata*.

Methods

Sample collection and RNA extraction

P. canaliculata specimens (15–20 mm shell length; 21.11 ± 0.23 g; 20 individuals) were collected from the Hunan Academy of Agricultural Sciences experimental paddy and then sent to the laboratory. The specimens were divided into hunger and satiety feeding groups. After 7 days of rearing in the aquarium, total RNA was extracted from two sets of snail hepatopancreases. The RNA quality was verified with an absorbance microplate reader (260/280 ratio > 2.0) and by agarose gel electrophoresis. Spare samples were stored at -80 °C.

Construction of a cDNA library for sequencing

The cDNA libraries were constructed via PCR amplification using random hexamer priming (the process was conducted in strict accordance with the manual). After library construction, cluster formation and Illumina sequencing were conducted on a HiSeq 2000 platform (Illumina, SY-401-2501) by Bohao Biotechnology Corporation (Shanghai, China) using the methods previously reported by our group [22].

Data preprocessing and de novo assembly

For the preprocessing, all the adaptor sequences, low-quality reads, ambiguous bases and sequences containing fewer than 20 nucleotides were removed from further analysis prior to contig assembly.

The sequencing data from the two samples were merged into the pooled reads, and de novo assembly was performed using CLC Genomics Workbench (version 6.0.4, CLC Bio, Aarhus, Denmark). De novo sequence assembly was performed on all remaining reads using a contig scaffolding algorithm with a minimum contig length of 200 bp or longer and a word size of 24 characters. The sequences resulting from this phase were designated contigs. Then, we applied CAP3 EST to the spliced primary unigenes to obtain the final unigenes.

Functional annotation and classification

The final unigene sequences were compared with the NR database and the UniProt database (filter: E-value < $1e-5$). The NR annotations of the resulting unigenes were searched with Blast2GO Gene Ontology (GO) functional classification algorithms, and the unigenes were functionally classified using cluster of orthologous groups (COG). The unigenes were searched in the CDD using rpstblastn, and all results with E-values < $1e-5$ were assigned KOG functional classification predictions and mapped to each COG level of

classification, whereas the online KAAS pathway alignment analysis tool was applied to the unigene splices for the KEGG mapping analysis.

Screening of differentially expressed unigenes

Taking Final Unigene as the reference, respectively, reads each sample was reference mapping, to obtain a sample of each coverage in each Unigene reads the application RPKM (Reads Per Kilobase of transcript per Million mapped reads), which was calculated via the following formula:

$$\text{RPKM} = \frac{\text{transcription}_{\text{reads}}}{\text{transcription}_{\text{length}} \times \text{total}_{\text{assembly}}\text{reads}_{\text{in}}\text{run}} \times 10^9$$

where *transcription_reads* is the number of reads throughout the unigene coverage, *transcription_length* is the length of the unigene, and *total_assembly_reads_in_run* is the total number of reads involved in splicing for all samples. Each RPKM result was normalized to the total expression of the corresponding sample as an internal standard. Fisher's exact test was used for the statistical analysis. Differentially expressed genes (DEGs) were identified using the analysis package DEGseq under the following conditions for at least one sample in the group: FDR < 0.05; fold change ≥ 1; and an average RPKM ≥ 3.

Putative molecular markers and open reading frames (ORFs)

The flanking sequences of a particular microsatellite are usually highly conserved, and the repeat unit is species specific. Using this characteristic, an SSR analysis was conducted on the unigene sequences with MISA software (<http://pgrc.ipk-gatersleben.de/misa/>). SSR markers can be used to develop STMS, one of the most commonly used microsatellite markers, using the principle of simple sequence repeat length polymorphisms (SSLPs). EMBOSS (6.4.0)-getorf [23] was used to generate ORF predictions for all unigenes, and the ORF greater than 300 bp was considered the ORF of the unigene.

Differentially expressed gene validation by real-time quantitative RT-PCR (qRT-PCR)

We screened the protein-coding genes associated with plant digestion and absorption using RT-PCR validation. The following six genes were selected: cellulase, cellulase EGX3, cellulase EGX1, endo-beta-1, 4-glucanase, xylanase and G-type lysozyme. The primer sequences (Additional file 1: Table S1) for the qRT-PCR were designed utilizing the Roche Universal Probe Library Assay Design Center (https://lifescience.roche.com/en_cn.html) and synthesized by Sangon Biotechnology Corporation

(Shanghai, China). The cDNA synthesis and RT-PCR processes were conducted according to the manufacturer's instructions. A reaction volume of 20 μL was used. The amplification program was as follows: 95 °C for 30 s, followed by 34 cycles of 95 °C for 10 s and annealing at 60 °C for 30 s.

Results and discussion

RNA sequencing and de novo assembly of the hepatopancreas transcriptome

After pre-processing the sequencing data of the two samples, we obtained 95,126,010 and 128,258,226 clean reads. The clean ratios (clean ratio = clean reads/raw reads) (%) were 94.9 and 94.6%.

The high-quality reads were assembled de novo using CLC Genomics Workbench v6.0.4. A total of 90,141 contigs were generated, with an average length of 108,487 bp and an N50 of 1,710 bp. To obtain the unigenes, we applied CAP3 EST to the secondary splices, which were then assembled into 87,766 unigenes, with a mean unigene size of 110,818 bp, a total length of 97,260,581 bp and an N50 of 1,772 bp. The results of the assembly are listed in Table 1. The majority of the unigenes were between 400 and 600 bp (30.37%). The length distribution of the assembled unigenes is shown in Fig. 1.

Functional annotation and classification

A total of 16,384 unigenes were identified by comparing our sequences with the NCBI NR database. The unigenes were searched in a protein sequence database and subjected to a UniProt comparative analysis, with the filtering criterion: E-value < 1e-5 (Table 2). The results show that of all unigenes, only 18.7% (*n* = 16,384) had significant similarities to proteins in the NR database (Additional file 2: Table S2), and a total of 19,942 unigenes had significant similarities to proteins in the UniProt database (Additional file 3: Table S3). However, other unigenes did not return any results in either of the two databases, which indicates that our study revealed a number of new genes and non-coding RNA sequences and enriched the snail hepatopancreas transcriptome gene data resources. Overall, the transcriptome sequencing yielded a great number of unique genes in the species, which is consistent with similar results reported in other species [20].

The Blast2GO algorithm was used to determine the GO functional classifications of the NR results. All sequences were classified according to the three major GO categories, with 23 in "biological process," 17 in "cellular component" and 14 in "molecular function". The distribution of the terms in each of the major GO categories is shown in Fig. 2 (Additional file 4: Table S4). In the biological process category, unigenes related to metabolic

Table 1 Statistical summary of the SG transcriptome for assembly

Statistics	Counts (number)	Total length (bp)	N50 (bp)	Average length (bp)	Longest (bp)	GC%
Contig	90,141	97,790,982	1,710	1,084.87	24,789	40.50
Unigene	87,766	97,260,581	1,772	1,108.18	24,789	40.45

processes (4131, 27.68%) and cellular processes (3538, 23.69%) were predominant, indicating that certain basic processes occurred in the hepatopancreas. A high percentage of genes were classified as “cell” and “cell part” under the cellular components category. Moreover, a COG classification of all of the unigenes indicated that 13,060 unigenes were assigned to 25 COG categories (Fig. 3) (Additional file 5: Table S5). The top 3 groups were “signal transduction mechanisms” (2,600, 19.9%), “general function prediction only” (2,052, 15.7%) and “posttranslational modification, protein turnover and chaperones” (1,205, 9.2%). “cell motility” (25, 0.02%) had the least representation.

KAAS was used for the KEGG map analysis of 20,885 unigenes using comparative analysis tools. These unigenes were mapped to a 262-article pathway, and most were concentrated in the following categories: metabolic: biosynthesis of secondary metabolites; microbial metabolism in diverse environments; and protein processing in the endoplasmic reticulum.

Functional enrichment analysis of DEGs

To better understand the function of the DEGs, GO functional and KEGG pathway enrichment analyses were performed when a *p*-value was less than 0.05 was observed. The up-regulated and down-regulated genes were significantly enriched in 194 GO categories

(Additional file 6: Table S6). The DEGs were mainly concentrated in the functional categories of “carbohydrate metabolic process,” “hydrolase activity,” “hydrolyzing O-glycosyl compounds,” “oxidation-reduction process,” “metabolic process” and “oxidoreductase activity” (Table 3). Our results suggest that the hepatopancreas is closely related to the digestion and metabolism of food and indicated that a variety of hydrolytic enzymes may play an important role in the process of digestion.

A total of 148 pathways were found to be enriched with DEGs, and DEGs were primarily identified in the following GO categories: metabolic pathways (104, 63.4%); chemical carcinogenesis (18, 10.98%); metabolism of xenobiotics by cytochrome P450 (16, 9.76%); and drug metabolism - cytochrome P450 (14, 8.53%). Notably, up-regulated genes were mostly found in the metabolic pathways of starch and sucrose metabolism and glutathione metabolism, and only up-regulated genes were found in the pathways of galactose metabolism, caprolactam degradation, and carbohydrate digestion and absorption (Table 3). The full lists of DEGs is shown in Additional file 7: Table S7 Collectively, these transcriptome sequences and pathway annotations provide an essential resource for further screening and expression analyses of candidate genes related to digestion and absorption in *P. Canaliculata*.

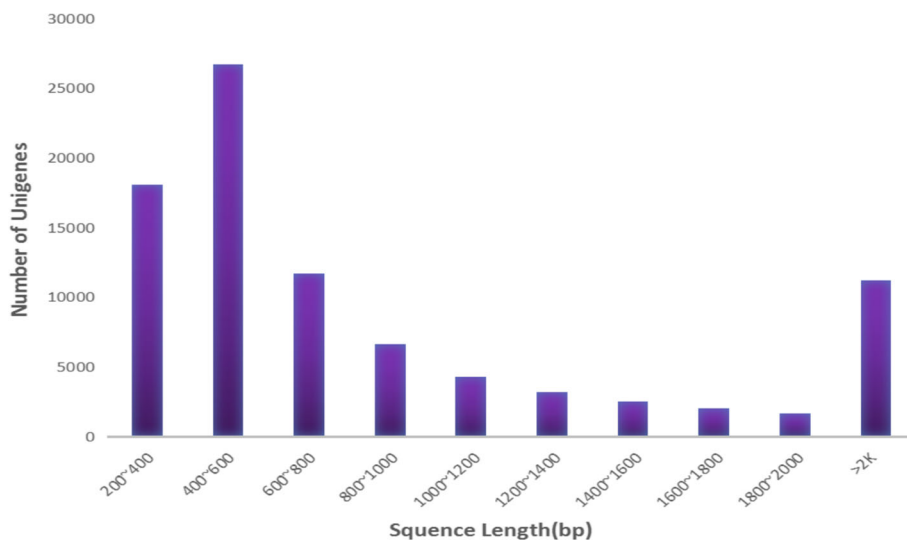


Fig. 1 Length distribution of the final assembled unigenes. The X axis shows the sequence lengths of the unigenes, and the Y axis shows the number of unigenes

Table 2 Results of blasting against the NCBI NR and UniProt protein databases

Database	Total unigene	No. sequence with hits	No. unknown sequence	Percentage of annotation (%)
NR database	87,766	16,384	71,382	18.7
UniProt	87,766	19,942	67,824	22.7

SSR analysis and ORF prediction

Our analyses identified 35,582 SSRs (or microsatellites) in total using MISA (Microsatellite) software, and 3171 of these SSRs (8.91%) were hybrid nucleotide repeats, 21,203 SSRs (59.59%) were mononucleotide repeats, 10,748 SSRs (30.21%) were dinucleotide repeats, and 3139 SSRs (8.82%) were trinucleotide repeats (Additional file 8: Table S8). In addition, 775 ORFs were predicted.

Analysis of digestive enzymes and genes in hepatopancreas

In this study, we focused on CDS (coding sequences) extracted from the putative ORF sequences to classify the functions of the hepatopancreas transcripts. Based on sequence alignment using the NR and UniProt databases and by referencing the results of the functional enrichment analysis of DEGs, the transcripts involved in plant feed digestion were analyzed and presented in Table 4. These transcripts were classified into four large categories as follows: hydrolase, transferase, isomerase and cytochrome P450. Among them, a variety of hydrolase genes were identified throughout our comparative

study of satiety and starvation, including endoglucanase, beta-1,4-endoglucanase, cellulase, family 10 cellulase, cellulase EGX1, cellulase EGX3, and xylanase. These genes are involved in many metabolic processes, such as starch and sucrose metabolism, galactose metabolism, the TCA cycle, secondary metabolite biosynthesis, carbohydrate digestion and absorption, and the pentose phosphate pathway. Moreover, we identified P450 cytochromes (CYPs), which is a major family of enzymes involved in detoxification and metabolism. Previous studies have reported a correlation between increased exposure to metabolic neurotoxic pesticides and over-expression of P450 genes in many pest species [24–31].

Comparative analysis between the two hepatopancreas libraries

To identify the DEGs in the starving apple snails and satiated apple snails, we used the RPKM algorithm to compare the differences in gene expression. Based on an RNA-Seq (quantification) analysis, 929 unigenes in total were differentially expressed between the two samples, among which the number of unigenes with up-regulated

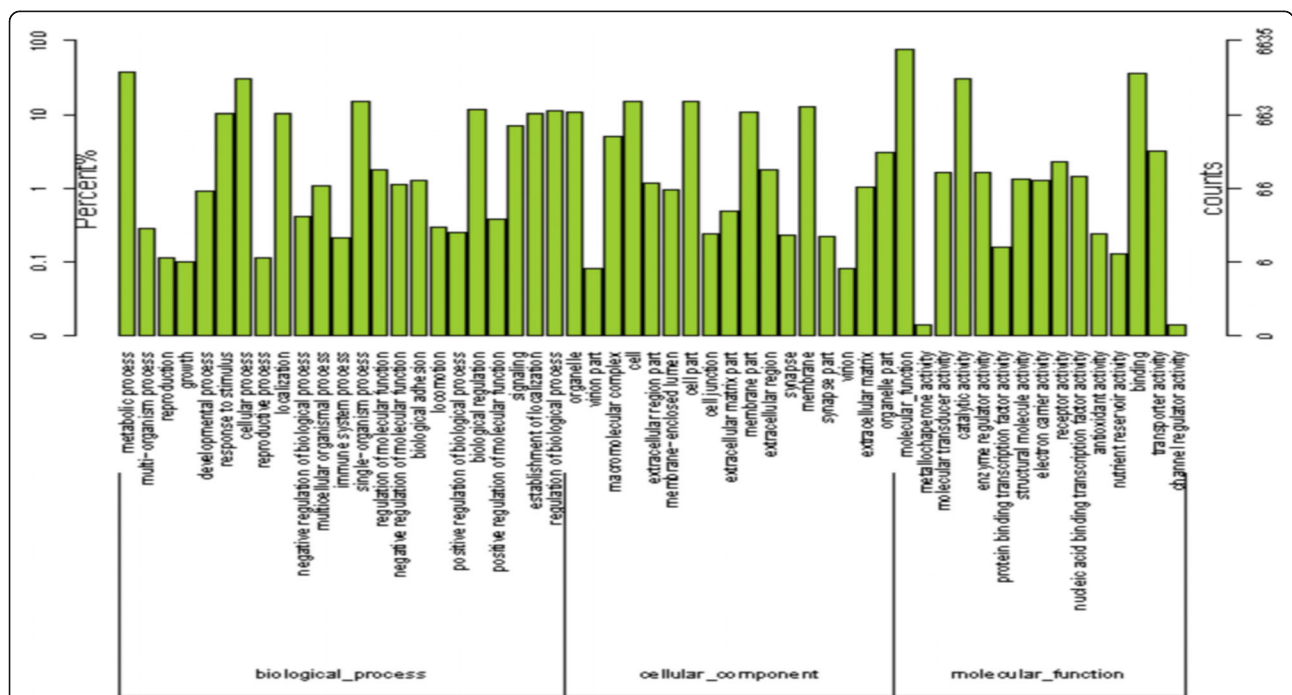
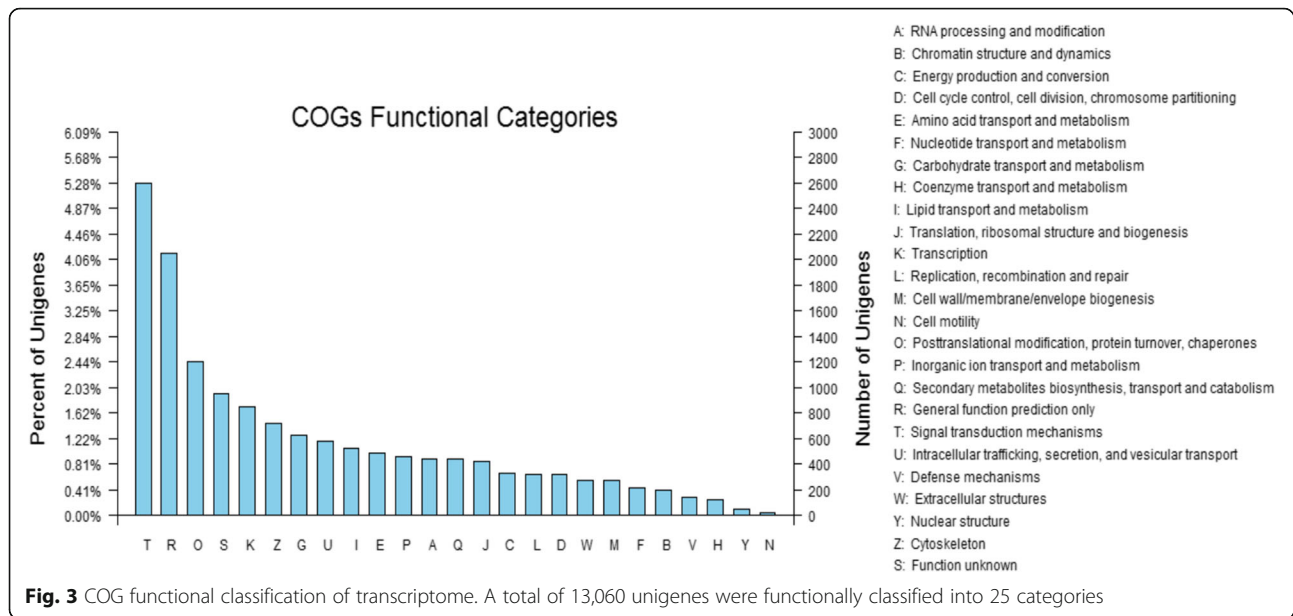


Fig. 2 Gene Ontology analysis of hepatopancreas transcriptome data. The results are summarized for the three main GO categories; red bars refer to “biological process”, blue bars refer to “cellular component” and green bars refer to “molecular function”. The Y axis shows the number of unigenes in each category



expression was 406, and the number with down-regulated expression was 523 in the starving apple snails compared with the satiated apple snails (as shown in Fig. 4). Many of the DEGs had only a 2- to 5-fold change in gene expression between the two samples. (The full list of DEGs is included in Additional file 9: Table S9).

Genes encoding secreted proteins that mediate digestive absorption

We concentrated on the most significant sequenced transcripts encoding secreted proteins. This de novo assembly of the hepatopancreas transcriptome of *P. canaliculata* provided us with valuable raw material for the identification and analysis of *P. canaliculata* hepatopancreas proteins. Similar to other herbivores, *P. canaliculata* produces various secreted proteins that are involved in the digestion of the meal or modulation of host defenses for survival. Among these modulatory molecules, cellulase, cellulase EGX1, cellulase EGX3, endo-beta-1, 4-glucanase, xylanase and G-type lysozyme enzymes are vital for *P. canaliculata* survival and feeding.

We found that only the above six digestive and absorption-related enzymes were differentially expressed in this experiment, which may be related to the characteristics of *P. canaliculata* digestive organs themselves, the distribution of enzymes and the role of symbiotic gastrointestinal bacteria in digestion [32]. Our group is continuing to perform related research.

Quantitative expression and comparative analysis for genes associated with digestive absorption and host defenses

qRT-PCR technology was used for the RNA-Seq screening of the six DEGs to validate the analysis. These genes

include two up-regulated unigenes (contig473, endo-beta-1, 4-glucanase; contig6717, xylanase) and four down-regulated unigenes (contig4257, cellulase; contig1040, cellulase EGX1; contig3816, cellulase EGX3; contig4913, G-type lysozyme) in starving apple snails compared with satiated apple snails. We used the fold change to reflect the differential expression of unigenes. The results of the above two methods were generally consistent (shown in Fig. 5).

In this study, cellulase EGX1, cellulase, cellulase EGX3 and G-type lysozyme were up-regulated in the hepatopancreases of fully fed *P. canaliculata*, and this finding is consistent with previous studies (xylanase [33], cellulase EGX3 [34], endo-beta-1, 4-glucanase [35, 36] and G-type lysozyme [37]), which indicates that these enzymes play an important role in the digestion and absorption of fibrous food sources. Cellulase [38] and cellulase EGX1 [36] may be involved in this process. These results showed that the hepatopancreatic secretion of proteins plays an important role in the regulation of plant digestion and utilization in *P. canaliculata*.

Conclusions

In this study, to explore the digestive function of *P. canaliculata* as a starting point, the de novo hepatopancreatic transcriptome of *P. canaliculata* was initially characterized using NGS technology and bioinformatics. The transcriptomic approach provides opportunities to identify pharmacologically active proteins and further insights into the digestive and metabolic functions of *P. canaliculata*. In total, 90,141 contigs and 87,766 unigenes have been discovered via high-throughput sequencing of the hepatopancreas. By comparing unigenes with the NCBI NR and UniProt databases,

Table 3 Top 30 functional classifications of DEGs determined by comparisons with the GO and KEGG databases

GO term/pathway	ID	Counts of DEGs	Up-regulated	Down-regulated	p-value
Based on the GO database					
Carbohydrate metabolic process	GO:0005975	27	21	6	1.02E-11
Hydrolase activity, hydrolyzing O-glycosyl compounds	GO:0004553	22	16	6	2.08E-11
Oxidation-reduction process	GO:0055114	45	29	16	5.83E-11
Metabolic process	GO:0008152	38	28	10	6.64E-09
Oxidoreductase activity	GO:0016491	34	24	10	9.51E-09
Pyrimidine nucleoside metabolic process	GO:0006213	5	0	5	4.29E-07
Pyrimidine nucleobase metabolic process	GO:0006206	5	0	5	4.29E-07
transferase activity, transferring pentosyl groups	GO:0016763	5	0	5	4.29E-07
Phosphorylase activity	GO:0004645	5	0	5	4.29E-07
Pyrimidine-nucleoside phosphorylase activity	GO:0016154	5	0	5	4.29E-07
Respiratory chain	GO:0070469	6	3	3	4.52E-07
ATP binding	GO:0005524	2	1	1	5.54E-07
Electron transport chain	GO:0022900	6	3	3	1.11E-06
NADH dehydrogenase (ubiquinone) activity	GO:0008137	5	2	3	2.59E-06
ATP synthesis coupled electron transport	GO:0042773	4	2	2	4.24E-06
Pentose-phosphate shunt	GO:0006098	4	4	0	8.27E-06
Mitochondrion	GO:0005739	10	5	5	1.78E-05
Enzyme regulator activity	GO:0030234	4	4	0	3.93E-05
RNA binding	GO:0003723	3	0	3	4.27E-05
Serine-type endopeptidase inhibitor activity	GO:0004867	8	5	3	4.69E-05
Nucleic acid binding	GO:0003676	4	2	2	5.15E-05
Regulation of catalytic activity	GO:0050790	4	4	0	5.98E-05
Peptidoglycan catabolic process	GO:0009253	3	0	3	8.32E-05
Hydrolase activity, acting on glycosyl bonds	GO:0016798	10	8	2	0.00010235
Glutathione transferase activity	GO:0004364	4	4	0	0.00012431
Respiratory electron transport chain	GO:0022904	3	2	1	0.00014634
Catalytic activity	GO:0003824	31	24	7	0.00021324
Intracellular	GO:0005622	4	3	1	0.00023455
Carbohydrate binding	GO:0030246	14	9	5	0.00038725
Based on the KEGG database					
Metabolic pathways	ko01100	104	64	40	1.02E-31
Chemical carcinogenesis	ko05204	18	12	6	1.21E-24
Metabolism of xenobiotics by cytochrome P450	ko00980	16	11	5	1.57E-22
Drug metabolism - cytochrome P450	ko00982	14	10	4	8.19E-20
Starch and sucrose metabolism	ko00500	17	16	1	2.43E-18
Steroid hormone biosynthesis	ko00140	12	6	6	1.62E-14
Biosynthesis of secondary metabolites	ko01110	32	21	11	3.94E-14
Glutathione metabolism	ko00480	11	10	1	5.47E-12
Retinol metabolism	ko00830	11	4	7	2.68E-11
Arachidonic acid metabolism	ko00590	10	6	4	1.06E-10
Lysosome	ko04142	16	11	5	1.19E-10
Phenylpropanoid biosynthesis	ko00940	6	5	1	1.20E-10
Pentose phosphate pathway	ko00030	9	9	0	4.07E-10

Table 3 Top 30 functional classifications of DEGs determined by comparisons with the GO and KEGG databases (Continued)

Ascorbate and aldarate metabolism	ko00053	7	7	0	4.19E-10
Proteoglycans in cancer	ko05205	16	9	7	1.18E-09
Degradation of aromatic compounds	ko01220	5	5	0	1.27E-09
Microbial metabolism in diverse environments	ko01120	19	13	6	2.79E-09
Cyanoamino acid metabolism	ko00460	5	4	1	6.96E-09
Galactose metabolism	ko00052	8	8	0	2.02E-08
Caprolactam degradation	ko00930	5	5	0	7.53E-08
Carbon metabolism	ko01200	14	10	4	7.69E-08
Glycine, serine and threonine metabolism	ko00260	9	6	3	1.12E-07
Carbohydrate digestion and absorption	ko04973	6	6	0	2.12E-07
Glycosaminoglycan degradation	ko00531	6	5	1	2.12E-07
Linoleic acid metabolism	ko00591	5	2	3	4.02E-07
Oxidative phosphorylation	ko00190	13	6	7	4.07E-07
Cardiac muscle contraction	ko04260	7	4	3	4.62E-07
Phenylalanine metabolism	ko00360	5	2	3	2.51E-06
Drug metabolism - other enzymes	ko00983	6	3	3	3.10E-06

Table 4 Categories of transcripts potentially involved in plant feed digestion

Transcript	Putative functions	Best match to NR database	E-value	GenBank
Hydrolase				
Contig_6437	Endoglucanase	Endoglucanase-like isoform x3 [<i>Biomphalaria glabrata</i>]	2E-45	XM_013220130.1
Contig_2881	Beta-1,4-endoglucanase	Endo-beta-1,4-glucanase [<i>Ampullaria crossean</i>]	3.00E-76	DQ367350.1
Contig_1140	Cellulase	Endoglucanase-like isoform x3	7.00E-120	
Contig_576	Family 10 cellulase	Family 10 cellulase (EGXA) [<i>Ampullaria crossean</i>]	3.00E-19	FJ183727.1
Contig_1040	Cellulase EGX1	Cellulase EGX1 [<i>Pomacea canaliculata</i>]	2E-40	DQ848667.1
Contig_3816	Cellulase EGX3	Cellulase EGX3 [<i>Pomacea canaliculata</i>]	5.00E-73	DQ848668.1
Contig_3668	Xylanase	Xylanase [<i>Ampullaria crossean</i>]	1.00E-73	AY941794.1
Contig_896	Probable beta-D-xylosidase	Probable beta-D-xylosidase 7 [<i>Aplysia californica</i>]	4E-42	XM_005105779.2
Contig_4457	Probable beta-D-xylosidase 5	Probable beta-D-xylosidase 7-like [<i>Aplysia californica</i>]	2E-148	XM_005100399.2
Contig_8505	Maltase-glucoamylase	Maltase-glucoamylase, intestinal-like [<i>Aplysia californica</i>]	3.00E-100	XM_013089164.1
Contig_4793	Endo-1,3-beta-D-glucanase	Beta-1,3-glucan-binding protein-like [<i>Biomphalaria glabrata</i>]	1.00E-51	NM_001311278.1
Transferase				
Contig_10465	Glucose-6-phosphate-1-dehydrogenase	Glucose-6-phosphate 1-dehydrogenase-like [<i>Crassostrea gigas</i>]	0	XM_011413892.1
Contig_20042	Thymidine phosphorylase	Thymidine phosphorylase [<i>Chelonia mydas</i>]	9.00E-53	XM_007066781.1
Isomerase				
Contig_6615	Aldose 1-epimerase	Aldose 1-epimerase-like [<i>Aplysia californica</i>]	7.00E-58	XM_005089091.2
Cytochrome P450				
Contig_7990	Cyp3A	Cytochrome P450 3A16-like [<i>Aplysia californica</i>]	1.00E-59	XM_013090539.1
Contig_37274	Cytochrome P450 II f2-like protein II	Heat shock protein 60 (HSP60) gene [<i>Pomacea canaliculata</i>]	4.00E-22	KM504522.1

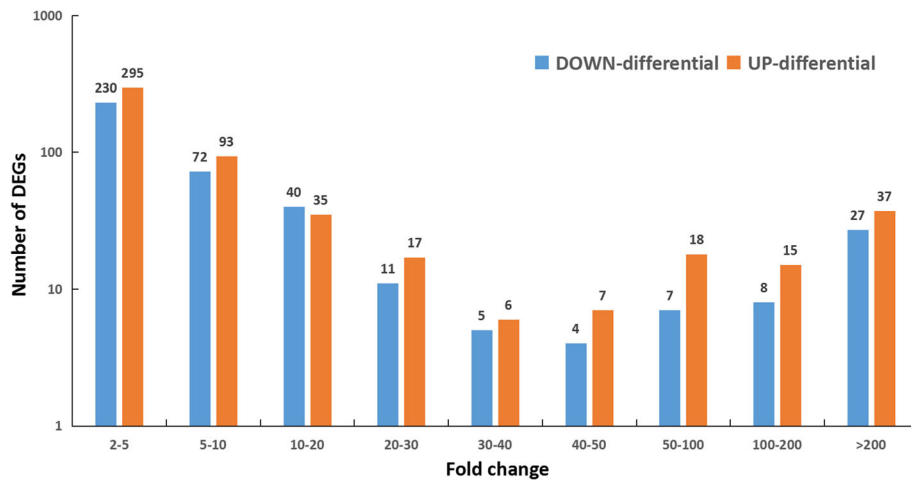


Fig. 4 Histogram of the multiples and *p*-values between the two groups of samples. *Blue* pillars refer to down-regulated DEGs and *Orange* pillars refer to up-regulated DEGs in satiated compared with starving snails. The X axis show the fold change of DEGs and the Y axis shows the number of DEGs. DEG: differentially expressed genes. (The references to colors in this figure legend are provided in the web version of this article.)

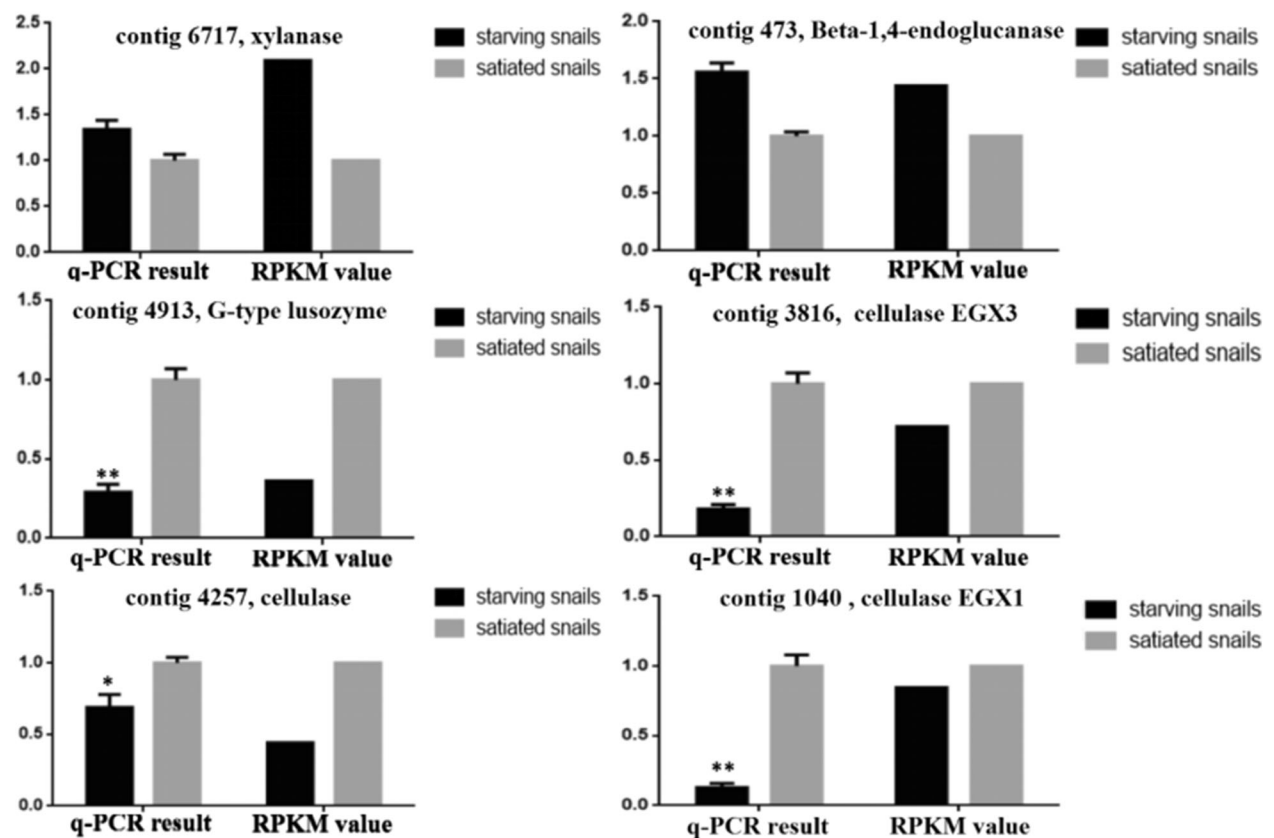


Fig. 5 qRT-PCR validation of the differentially expressed genes analyzed by RNA-seq. qRT-PCR was performed for nine genes that were identified as differentially expressed between the starving apple snails and satiated apple snails. The expression level of each gene was normalized to the level in the satiated snails. The Y axis shows the relative mRNA expression levels. $P^* < 0.05$, $P^{**} < 0.01$

analyzing NR results via GO functional classification analysis and analyzing unigenes via KOG and KEGG pathways, we have shown that the hepatopancreas of *P. canaliculata* is an essential digestive organ that plays a significant role in the digestion of plant matter. In addition, a comparative analysis of the hepatopancreas transcriptomes of the two groups indicated six genes that may be associated with plant feed digestion, and the qRT-PCR verification, RNA-Seq analysis and qRT-PCR results are consistent. Our research has enriched the genome database of *P. canaliculata* hepatopancreas and provided a reference for environmentally significant research on the digestive physiology of *P. canaliculata*. Moreover, our results provide additional opportunities to screen and clone functional genes.

Additional files

- Additional file 1: Table S1.** Primers for the qRT-PCR experiment. (XLSX 14 kb)
- Additional file 2: Table S2.** The results of the NR annotation. (XLS 1811 kb)
- Additional file 3: Table S3.** The results of the UniProt annotation. (XLS 3010 kb)
- Additional file 4: Table S4.** Gene Ontology analysis of the SG transcriptome data. (XLSX 18 kb)
- Additional file 5: Table S5.** Statistics on the COG functional classification of the transcriptome. (XLS 62 kb)
- Additional file 6: Table S6.** Statistical analysis of the GO functional enrichment of DEGs. (XLS 30 kb)
- Additional file 7: Table S7.** Statistical analysis of the KEGG pathway enrichment of DEGs. (XLS 29 kb)
- Additional file 8: Table S8.** Information on the identified SSRs. (XLS 812 kb)
- Additional file 9: Table S9.** Full list of the DEGs between the two samples. (XLS 3542 kb)

Abbreviations

COG: Cluster of orthologous groups; FDR: False discovery rate; GO: Gene ontology; MISA: Microsatellite; NGS: Next-generation sequencing; NR: Nonredundant; *P. canaliculata*: *Pomacea canaliculata*; RPKM: Reads per kb per million reads; WEGO: Web gene ontology annotation plot

Acknowledgements

The authors thank the Shanghai Bohao Biotechnology Corporation for the assistance with data processing and bioinformatic analysis. We thank Dr. XingLi Xu and Dr. Lei Liu for the valuable comments and revision of the manuscript. In addition, we also thank American Journal Experts for his help to improve our language.

Funding

Hunan agricultural university postdoctoral fund (NO.540390205004).

Availability of data and materials

The data supporting the results of this article are included within the article and in its supplementary files.

Authors' contributions

Conceived and designed the experiments: TYC and LY. Performed the experiments: FYZ and LY. Analyzed the data: LY. Contributed reagents/materials/analysis tools: FYZ and LY. Wrote the paper: FYZ and LY. All authors have read and approved the final manuscript.

Competing interests

I confirm that I have read BioMed Central's guidance on competing interests and have included a statement in the manuscript on any competing interests. The authors declare that they have no competing interests.

Consent for publication

Not applicable.

Ethics approval and consent to participate

Not applicable.

Author details

¹Laboratory of Molecular Physiology, College of Veterinary Medicine, Hunan Agricultural University, Changsha, Hunan Province 410128, People's Republic of China. ²Laboratory of Molecular Pathology, College of Veterinary Medicine, Hunan Agricultural University, Changsha 410128, People's Republic of China.

Received: 18 September 2016 Accepted: 11 February 2017

Published online: 22 February 2017

References

- Pimentel D, Zuniga R, Morrison D. Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecol Econ.* 2005;52:273–88.
- Simberloff D. The politics of assessing risk for biological invasions: the USA as a case study. *Trends Ecol Evol.* 2005;20:216–22.
- Carey MP, Wahl DH. Native fish diversity alters the effects of an invasive species on food webs. *Ecology.* 2010;91:2965–74.
- Lowe MR, Wu W, Peterson MS, Brown-Peterson NJ, Slack WT, Schofield PJ. Survival, growth and reproduction of non-native Nile tilapia II: fundamental niche projections and invasion potential in the northern Gulf of Mexico. *PLoS One.* 2012;7, e41580.
- Lach L, Britton DK, Rundell RJ, Cowie RH. Food preference and reproductive plasticity in an invasive freshwater snail. *Biol Invasions.* 2000;2:279–88.
- Kwong KL, Dudgeon D, Wong PK, Qiu J-W. Secondary production and diet of an invasive snail in freshwater wetlands: implications for resource utilization and competition. *Biol Invasions.* 2010;12:1153–64.
- Kyle CH, Plantz AL, Shelton T, Burks RL. Count your eggs before they invade: identifying and quantifying egg clutches of two invasive apple snail species (*Pomacea*). *PLoS One.* 2013;8, e77736.
- Lv S, Zhang Y, Liu HX, Hu L, Liu Q, Wei FR, et al. Phylogenetic evidence for multiple and secondary introductions of invasive snails: *Pomacea* species in the People's Republic of China. *Divers Distrib.* 2013;19:147–56.
- Litsinger JA, Estano DB. Management of the golden apple snail *Pomacea canaliculata* (Lamarck) in rice. *Crop Protect.* 1993;12:363–70.
- Halwart M. The golden apple snail *Pomacea canaliculata* in Asian rice farming systems: present impact and future threat. *Int J Pest Manag.* 1994;40:199–206.
- San MR. Recent developments in the use of botanical molluscicides against golden apple snails (*Pomacea canaliculata*). In: Joshi RC, Sebastian LC, editors. *Global advances in ecology and management of golden apple snails*. Muñoz: Philippine Rice Research Institute; 2006. p. 393–403.
- Teo SS. Evaluation of different duck varieties for the control of the golden apple snail (*Pomacea canaliculata*) in transplanted and direct seeded rice. *Crop Protect.* 2001;20:599–604.
- Yusa Y, Sugiura N, Wada T. Predatory potential of freshwater animals on an invasive agricultural pest, the apple snail *Pomacea canaliculata* (Gastropoda: Ampullariidae), in Southern Japan. *Biol Invasions.* 2006;8:137–47.
- Ip KKL, Liang Y, Lin L, Wu H, Xue J, Qiu J-W. Biological control of invasive apple snails by two species of carp: effects on non-target species matter. *Biol Contr.* 2014;71:16–22.
- Wong PK, Liang Y, Liu NY, Qiu J-W. Palatability of macrophytes to the invasive freshwater snail *Pomacea canaliculata*: differential effects of multiple plant traits. *Freshwat Biol.* 2010;55:2023–31.
- Choubisa SL, Sheikh Z, Jaroli VJ. Histopathological effects of larval trematodes on the digestive gland of freshwater snail species, *Vivipara bengalensis* and *lymnaea acuminata*. *J Parasit Dis.* 2012;36:283–6.
- Zhao M, Wang T, Adamson KJ, Storey KB, Cummins SF. Multi-tissue transcriptomics for construction of a comprehensive gene resource for the terrestrial snail *Thebapissana*. *Sci Rep.* 2016;6:20685.

18. Kang SW, Patnaik BB, Hwang HJ, Park SY, Chung JM, Song DK, et al. Transcriptome sequencing and de novo characterization of Korean endemic land snail, *Koreanohadra kurodana* for functional transcripts and SSR markers. *Mol Genet Genomics*. 2016;291:1999-2014.
19. Wang L, Feng Z, Wang X, Wang X, Zhang X. DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics*. 2010;26:136-8.
20. Mu X, Hou G, Song H, Xu P, Luo D, Gu D, et al. Transcriptome analysis between invasive *Pomacea canaliculata* and indigenous *Cipangopaludina cahayensis* reveals genomic divergence and diagnostic microsatellite/SSR markers. *BMC Genet*. 2015;16:12.
21. Adamson KJ, Wang T, Zhao M, Bell F, Kuballa AV, Storey KB, et al. Molecular insights into land snail neuropeptides through transcriptome and comparative gene analysis. *BMC Genomics*. 2015;16:308.
22. Xu XL, Cheng TY, Yang H, Yan F, Yang Y. De novo sequencing, assembly and analysis of salivary gland transcriptome of *haemaphysalis flava* and identification of sialoprotein genes. *Infect Genet Evol*. 2015;32:135-42.
23. Godoy MS, Castro-Vazquez A, Castro-Vasquez A, Vega IA. Endosymbiotic and host proteases in the digestive tract of the invasive snail *Pomacea canaliculata*: diversity, origin and characterization. *PLoS One*. 2013;8, e66689.
24. Anzenbacher P, Anzenbacherová E. Cytochromes P450 and metabolism of xenobiotics. *Cell Mol Life Sci*. 2001;58:737-47.
25. Bogwitz MR, Chung H, Magoc L, Rigby S, Wong W, O'Keefe M, et al. CYP12A4 confers lufenuron resistance in Anatural population of *Drosophila melanogaster*. *Proc Natl Acad Sci U S A*. 2005;102:12807-12.
26. Joußen N, Heckel DG, Haas M, Schuphan I, Schmidt B. Metabolism of Imidacloprid and DDT by P450 CYP6G1 expressed in cell cultures of *Nicotiana tabacum* suggests detoxification of these insecticides in Cyp6g1-overexpressing strains of *Drosophila melanogaster*, leading to resistance. *Pest Manag Sci*. 2008;64:65-73.
27. Baldwin WS, Marko PB, Nelson DR. The cytochrome P450 (CYP) gene superfamily in *daphnia pulex*. *BMC Genomics*. 2009;10:169.
28. Jones RT, Bakker SE, Stone D, Shuttleworth SN, Boundy S, McCart C, et al. Homology modelling of *drosophila* cytochrome P450 enzymes associated with insecticide resistance. *Pest Manag Sci*. 2010;66:1106-15.
29. Zhu F, Parthasarathy R, Bai H, Woithe K, Kaussmann M, Nauen R, et al. A brain-specific cytochrome P450 responsible for the majority of deltamethrin resistance in the Qtc279 strain of *Tribolium castaneum*. *Proc Natl Acad Sci U S A*. 2010;107:8557-62.
30. Bass C, Carvalho RA, Oliphant L, Puinean AM, Field LM, Nauen R, et al. Overexpression of a cytochrome P450 monooxygenase, CYP6ER1, is associated with resistance to Imidacloprid in the brown planthopper, *Nilaparvata lugens*. *Insect Mol Biol*. 2011;20:763-73.
31. Mitchell SN, Stevenson BJ, Müller P, Wilding CS, Egyir-Yawson A, Field SG, et al. Identification and validation of a gene causing cross-resistance between insecticide classes in *Anopheles gambiae* from Ghana. *Proc Natl Acad Sci U S A*. 2012;109:6147-52.
32. Rice P, Longden I, Bleasby A. EMBOSS: the European molecular biology open software suite. *Trends Genet*. 2000;16:276-7.
33. Imjongjirak C, Amparyup P, Sittipraneed S. Cloning, genomic organization and expression of two glycosyl hydrolase family 10 (GHF10) genes from golden apple snail (*Pomacea canaliculata*). *DNA Seq*. 2008;19:224-36.
34. Li Y, Yin Q, Ding M, Zhao F. Purification, characterization and molecular cloning of a novel endo-beta-1, 4-glucanase AC-EG65 from the mollusc *Ampullariacrossean*. *Comp Biochem Physiol B: Biochem Mol Biol*. 2009;153:149-56.
35. Mishima T, Wada N, Iwata R, Anzai H, Hosoya T, Araya K. Super-protective child-rearing by Japanese Bess beetles, *Cylindrocaulus patalis*: adults provide their larvae with chewed and Predigested Wood. *Insects*. 2016;7:18.
36. Wang J, Ding M, Li YH, Chen QX, Xu GJ, Zhao FK. Isolation of a multi-functional endogenous cellulase gene from mollusc, *Ampullaria crossean*. *Sheng Wu Hua Xue Yu Sheng Wu Wu Li Xue Bao (Shanghai)*. 2003;35:941-6.
37. Guo Y, He H. Identification and characterization of a goose-type lysozyme from sewage snail *Physa acuta*. *Fish Shellfish Immunol*. 2014;39:321-5.
38. Maeda I, Shimohigashi Y, Kihara H, Ohno M. Purification and characterization of a cellulase from the giant snail *Achatina fulica*. *Biosci Biotechnol Biochem*. 1996;60:122-4.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

