The Israeli Journal of Aquaculture – Bamidgeh • ISSN 0792-156X • IJA.74.2022.1783855, 15 pages
CCBY-NC-ND-4.0 ● https://doi.org/10.46989/001c.38101

The *IJA* is a peer-reviewed open-access, electronic journal, freely available without charge to users
-
Produced by the AquacultureHub non-profit Foundation
Sale of *IJA* papers is strictly forbidden

Molecular cloning, expression pattern of Trypsin gene and association analysis with growth traits in *Penaeus monodon*

Yun-Dong Li[1,2,3], Fa-Lin Zhou[1,2], Qi-Bin Yang[2], Song Jiang[1], Li-Shi Yang[1], Jian-Hua Huang[1], Xu Chen[4], Qianlin Mai[1], Shi-Gui Jiang[1,2,4*]

[1] *South China Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Key Laboratory of South China Sea Fishery Resources Exploitation and Utilization, Ministry of Agriculture and Rural Affairs, Guangzhou 510300, China*

[2] *Hainan Yazhou Bay Seed Laboratory, Sanya 572025, China*

[3] *Key Laboratory of Tropical Hydrobiology and Biotechnology of Hainan Province, Hainan Aquaculture Breeding Engineering Research Center, College of Marine Sciences, Hainan University, Haikou, Hainan 570228, China*

[4] *Tropical Fishery Research and Development Center, South China Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Key Laboratory of Efficient Utilization and Processing of Marine Fishery Resources of Hainan Province, Sanya 572018, China*

Keywords: Gene clone; Growth traits-related; *Penaeus monodon*; Single nucleotide polymorphism (SNP); Trypsin gene (*PmTry*)

Abstract

A novel TRY homolog was cloned in *Penaeus monodon* by RACE technology, named *PmTry* (GenBank: KP998480). The *PmTry* cDNA was 916 bp, which encodes 266 amino acids with a predicted molecular weight of 28.38 KDa and an isoelectric point of 4.58. Homologous analysis indicated that *PmTry* shared 42%~91% similarity with other species. The phylogenetic tree showed that *PmTry* was closely related to *Fenneropenaeus chinensis*. Tissue expression profiles showed that *PmTry* was most expressed in the hepatopancreas and the lowest in the eyestalk nerve. It was expressed in the whole growth stage of *P. monodon*, but the relative expression level of each stage was significantly different. In addition, *PmTry*-524 and *PmTry*-798 were particularly related to growth traits of *P. monodon* by genotype. The SNP markers may provide a basis for genetic selection and breed improvement studies in *P. monodon*.

* Corresponding author. Shi-Gui jiang, jiangsg@21cn.com

Introduction

Trypsin (TRY) is a kind of enzyme belonging to the serine protease family. Trypsin is also a major alkaline proteolytic enzyme in animals' digestive tracts. Trypsin has endopeptidase activity which can specifically hydrolyze the peptide bond formed by the lysine and arginine shuttle group (Wang et al., 2011). It has two significant roles in digestion: digesting protein food and activating other proteases such as chymotrypsinogen, carboxypeptidase, and elastase to perform digestive functions (Muhlia-Almazán et al., 2008).

Bovine trypsin crystals were the first to be purified. The complete amino acid sequence and three-dimensional structure of trypsin were subsequently determined (Walsh et al., 1964, Huber et al., 1974). There are also many reports on the isolation and purification of trypsin from aquatic animals, such as *Engraulis encrasicholus* (Huber et al., 1974), *Salmo salar* (Outzen et al., 1996), *Oncorhynchus keta* (Sekizaki et al., 2000), *Gadus morhua* (Beirao et al., 2001), etc. In recent years, many studies have related to the function of trypsin in larval development in aquatic animals (Pan et al., 2006, Wang, 2004). Studies in *Sparus aurata* have shown that trypsin and chymotrypsin are the most abundant proteolytic enzymes in its digestive tract and play an essential role in protein hydrolysis and utilization (Alarcón et al., 1998). Studies on *Oncorhynchus solar* found that trypsin activity is a good indicator of nutritional condition parameters, and it was significantly positively correlated with its specific growth rate (Ueberschr, 1995). Using qPCR and enzyme activity assay methods to study the expression and enzyme activity of TRY mRNA in the early developmental stages of three species of fish, *Culter alburnus*, *Ctenopharyngodon idellus*, and *Elopichthys bambusa* (Ruan, 2010). The change process was compared, and the results showed that the activity of trypsin before the opening of larvae was a programmed expression, which could meet the specific nutritional and physiological needs in the early development of larvae. The activity of trypsin would increase with the feeding rate while rising (Infante et al., 1996). In shrimp studies (Wang et al., 2011, Yang et al., 2005), it was found that the activity of trypsin was vitality gradually increased with the larval stage increase in larvae of *L. vannamei*, *Fenneropenaeus chinensis*, *Marsupenaeus japonicus* and *Macrobrachium rosenbergii*.

With the development of molecular biology and the advancement of biotechnology, domestic and foreign research on trypsin is not only limited to the field of enzymology but also has been studied in detail from the perspective of genes. In invertebrates, trypsin from various insects has been cloned (Dong and Liu, 1998). The trypsin gene from a few crustaceans like *L. vannamei*, *F. chinensis*, *Procambarus claikii*, *Paralithodes camtschaticus*, *Penaeus merguiensis*, *M. japonicus*, *Oratosguilla oratoria*, and *Neocaridina denticulate sinensis*, was also cloned (Wang, 2004). In addition, studies on trypsin gene polymorphisms in aquatic animals have also been reported. For example, previous researchers used the direct sequencing method of PCR products to analyze the single nucleotide polymorphism of the trypsin gene of *Siniperca chuatsi.* They found a G169A polymorphism in the third exon. The G169A polymorphism was further analyzed in 312 samples, and three genotypes, GG, GA, and AA, were found, and the genotype frequencies were 0.26, 0.54, and 0.20, respectively (Yu et al., 2011). Additional studies on *L. vannamei* have shown that the trypsin gene plays a vital role in shrimp molt and that TRY is thought to be a gene associated with shrimp growth and development (Klein et al., 1998, Klein et al., 1996). Studies on *TRY* polymorphism in *P. monodon* have not been reported.

P. monodon is one of the world's three most economically farmed shrimp and a crucial traditional resource of high-quality shrimp in China and Southeast Asia. Although the technology of artificial breeding of *P. monodon* has long been overcome, there is still a lack of information on the molecular regulation mechanism of growth and development of *P. monodon* (Dai et al., 2015, Zhou et al., 2014, Huang et al., 2005). Therefore, in this study, we cloned the *TRY* gene from *P. monodon*, analyzed its structure and expression characteristics, and

screened its SNPs to find SNP markers related to the growth traits of *P. monodon*, which provided scientific data for marker-assisted selection of *P. monodon*.

Materials and Methods

Animal sample and traits data collection

The individuals of *P. monodon* were taken from the Shenzhen Experimental Base of the South China Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, to explore the expression of PmTry in different tissues and different growth processes. Three male and female individuals were randomly selected. Hepatopancreas, gill, heart, intestine, stomach, lymph, muscle, eyestalk, ovary, testis, and thoracic, abdominal nerve samples were collected, respectively. Three same organs from different *P. monodons* were collected in the same tube. Referring to the research on the life history of *P. monodon* (Jiang et al., 2013), the nauplius, zoea, mysis, larvae, postlarva, juvenile, subadult, and adult samples were collected respectively for gene expression experiments at various growth stages. Sample tissue is stored in RNAalater® solution. Store overnight at 4 °C, then reserve for long-term storage at -80 °C.

The geographic populations in 4 regions of *P. monodon*: Africa, Indonesia, Sanya, and Thailand stock were used in resequencing experiments and SNP screening. There are 8 samples in each population for a total of 32 samples. The genotyping and detection of SNP used the population of offspring produced by the parents of the African *P. monodon*, from now on referred to as the African stock, with a total of 300 samples. Each shrimp's bodyweight (BW) was measured using an electronic balance (accurate to 0.01 g). Body length (BL), carapace length (CL), carapace width (CW), carapace height (CH), first pleon segment length (FSL), second pleon segment length (SSL), and telson length (TL), a total of 7 growth traits, were measured by a ruler (accurate to 0.1 mm). All the data was collected and recorded. For DNA extraction, take a small amount of muscle tissue stored in a 2 mL centrifuge tube (have added 95% ethanol).

Total RNA extraction and cDNA preparation

Total RNA was extracted from tissue expression and samples of different growth periods under the guidance of HiPure Fibrous RNA Plus Kit (Magen, Guangzhou, China). The integrity of each RNA was checked by 1.5% agarose gel electrophoresis, and the RNA concentration was checked by NanoDrop 2000. The samples for quantitative real-time PCR (qPCR) were reverse-transcribed according to the instructions of the PrimeScript™ RT reagent Kit with gDNA Eraser Kit (TaKaRa, Dalian, China). The obtained cDNA was detected with *EF* primers and stored at –80 °C for later use.

The full-length cDNA of PmTry

The EST sequences were obtained from the *P. monodon* transcriptome library. Primer Premier 5.0 was used to design the specific primers *PmTry*-F1 and *PmTry*-F2. Rapid amplification of cDNA end (RACE), touchdown PCR, and semi-nested PCR were used to amplify the 3′ ends of the target genes. Reaction procedures and conditions are based on laboratory research literature (Li et al., 2016). PCR amplification products were detected by 1.2% agarose gel electrophoresis and sent to Beijing Ruibo Xingke Biotechnology Co., Ltd. Guangzhou Branch for sequencing. The sequencing results were compared with the target gene to obtain the full-length cDNA of the *PmTry* gene. The primers used are listed in **Table 1**.

Table 1 Primers sequences used in experiments

Primers	Sequence (5'–3')	Application
PmTry-F1	TCCACCAGTGACCAGCCATG	3'RACE-PCR
PmTry-F2	CTGTGCCCGTCCCAACTACCC	
PmTry(RT)-F	CAGGGCGATGACTTTGATAAT	Real-time PCR
PmTry(RT)-R	ACTGATGGTGAAGCCGTTGTA	
PmTry-SF1	TTGTGTGTATGAACATAATTGTTGAC	SNPs screening
PmTry-SR1	ACCATATCTCCATATTTGTTCATACTG	
PmTry-SF2	ACGACTCAGAGGTCGAGGAGTCATTC	
PmTry-SR2	GCAAGAGCCACACTGGCCCTGGTC	
PmTry-SF3	AGAACAACCGTGCTTACACCGGAG	
PmTry-SR3	AGGCATTGTCCATGAGTCCACCG	
PmTry-SF4	TGCCTTGGAGGGCCAGACATTC	
PmTry-SR4	AAAGAGAAGAGTACTAATATCAAC	
PmTry-SF5	CACTGAAGAAGGCTGTTGCCACCG	
PmTry-SR5	GGCCATCTTAACGTAGCCCTGATCAC	

Sequence analysis of PmTry

DNAMAN was used to assemble the sequenced sequences and obtain the full length of the *PmTry* gene. The software DNAstar was used to predict the open reading frames and peptides. The BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi) tool was used to analyze the similarity between predicted amino acid sequences and protein databases. Multiple sequence alignments were performed using Clustal X software. Amino acid isoelectric points and theoretical molecular weights were predicted using ExPASy ProtParam (https://web.expasy.org/protparam/). Protein domain analysis was performed using the NCBI structural analysis tool (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) and SMART 4.0 (http://smart.embl-heidelberg.de/smart/set_mode.cgi?GENOMIC=1). Glycosylation sites were predicted using the NetNGlyc 1.0 Server (http://www.cbs.dtu.dk/services/NetNGlyc/). Phosphorylation sites were predicted using the NetPhos 3.1 Server (http://www.cbs.dtu.dk/services/NetPhos/). Phylogenetic trees were constructed using Clustal x 2.0 and MEGA 5.0.5 software.

Expression analysis of PmTry gene

The specific primers *PmTry* (RT)-F and *PmTry* (RT)-R were designed based on the cDNA sequence of PmTry. β-actin was chosen as the internal reference gene, and the primer sequences were βactin-F (GGTATCCTCACCCTCAAGTA) and βactin-R

(AAGAGCGAAACCTTCATAGA), as reported in a previous study (Li et al., 2016). The PCR reactions were performed according to the instructions of the TaKaRa SYBR Premix Ex Taq™ Kit (Perfect Real Time) (TaKaRa, Dalian, China). Distilled water was used as a negative control instead of a template, and three replicates were set up for both samples and internal reference. The relative expression of *PmTry* in various tissues and larval developmental stages of *P. monodon* was analyzed by the relative quantification $2^{-\Delta\Delta CT}$ method. The obtained data were analyzed using One-Way ANOVA in SPSS 18.0, and the results were expressed as Mean ± SD. Duncan's test for significant differences was used ($P < 0.05$).

Acquisition of candidate SNP loci and genotyping

Primers were designed on the cDNA sequence of the *PmTry* gene using Premier 5.0 software. The primers specifically used in the experiment are shown in Table 1. Using the designed primers, PCR amplification was carried out with DNA samples from 4 geographic groups in Sanya, Indonesia, Thailand, and Africa as templates. The length and brightness of the products were detected by 1.0% agarose electrophoresis, and the PCR products that met the requirements were sent to the Beijing Genomics Institute for sequencing. The Sequencher 4.1.4 software was used to compare and screen SNP sites, and the SNP sites in the amplified fragments of each set of primers were screened by the peak map of the sequencing results. Genotyping of candidate SNP loci using SnaPshot technology. Non-synonymous mutant loci on exons were screened as candidate SNP loci. The locus information and a population of 300 African *P. monodon* progeny were sent to Generay Biotech Co., Ltd. for typing.

Correlation analysis of candidate SNP loci with growth traits

The SPSS 18.0 software was used to analyze the correlation between the genotype and traits of each locus. The multivariate analysis of the variance module in the general linear model (GLM) was used for correlation analysis (Cao, et al., 2012). The dependent variables were body weight, length, and other morphological traits, and the independent variables were the different genotypes of the screened SNP loci. The model is $Yij=\mu+Bi+eij$, where Yij represents the observed value of the i-th marker of a trait on the j-th individual; μ represents the average value of all individuals observed in the experiment; Bi represents the effect value of the i-th marker; eij represents the random residual effect of the observed value of the corresponding individual.

Results

Sequence analysis of PmTry

Amplify using specific primers, splicing the sequencing results of the RACE amplification product with the EST sequence to obtain the full-length cDNA sequence of the trypsin gene of *P. monodon* (**Figure 1**), named *PmTry* (GenBank ID: KP998480). The full length of the gene is 916 bp, including 35 bp in 5'-end non-coding region, 80 bp of 3'-end non-coding region, and a 831 bp open reading frame. The amino acid sequence analysis showed that the *PmTry* encodes a protein composed of 266 amino acid residues, with a molecular weight of 28.38 kDa and a theoretical isoelectric point of 4.58. Analysis of the secondary structure of PmTry found that there are 3 α-helix and 15 β-sheet. It has phosphorylation sites of 4 series, 2 threonines, and 4 tyrosines. Domain analysis showed that its N-terminal contained a 15-amino acid signal peptide. Trypsin was a serine protease, and the deduced amino acid sequence contained a conserved active site of serine protease, namely His 74, Asp 125, and Ser 218 active triplet.

```
   1 TGTCTGCAAGTGCCTCTCTCCACCAGTGACCAGCCatgaagtccctcgtgctctgcctgc 60
   1                                   M  K  S  L  V  L  C  L  L 9
  61 tcctcgccggggctctcgccgccccctccaggaagcccaccttccgccgcggtctgaaca 120
  10   L  A  G  A  L  A  A  P  S  R  K  P  T  F  R  R  G  L  N  K 29
 121 agatcgtgggaggaacggacgtcactcctggtgagcttccttaccagctcagcttccagg 180
  30   I  V  G  G  T  D  V  T  P  G  E  L  P  Y  Q  L  S  F  Q  D 49
 181 acatctcttggggtaccgaatttcacttctgcggcgcctccatctacaatgaacattggg 240
  50   I  S  W  G  T  E  F  H  F  C  G  A  S  I  Y  N  E  H  W  A 69
 241 ccatctgcgccggtcactgcgtccagggcgatgactttgataatcctaactaccttcagg 300
  70   I  C  A  G  H  C  V  Q  G  D  D  F  D  N  P  N  Y  L  Q  V 89
 301 tcgtggccggagacaataaccttcacattgacgagggcaacgagcagaaggtcgtcctct 360
  90   V  A  G  D  N  N  L  H  I  D  E  G  N  E  Q  K  V  V  L  S 109
 361 ccaagattattcaacacgaggactacaacggcttcaccatcagtaacgacatctccctgc 420
 110   K  I  I  Q  H  E  D  Y  N  G  F  T  I  S  N  D  I  S  L  L 129
 421 tcaagctgtctcagcctctgaccttcaacgacttcgtgcgcgccatcgatattcccgctc 480
 130   K  L  S  Q  P  L  T  F  N  D  F  V  R  A  I  D  I  P  A  Q 149
 481 agggtcacgctgcctcaggtgactgcatcgtgtccggctggggcgctctctctgaaggag 540
 150   G  H  A  A  S  G  D  C  I  V  S  G  W  G  A  L  S  E  G  G 169
 541 gaagctcccccagtgtcctccagaaggtgtccgttcccatcgtgtctgatgacgaatgtc 600
 170   S  S  P  S  V  L  Q  K  V  S  V  P  I  V  S  D  D  E  C  R 189
 601 gtgatgtttatggccagaacgatattgaggactccatgatctgtgccggagtgcctgagg 660
 190   D  V  Y  G  Q  N  D  I  E  D  S  M  I  C  A  G  V  P  E  G 209
 661 gcggcaaggactcgtgccagggtgactctggcggtccccttgcttgctctgacaccggct 720
 210   G  K  D  S  C  Q  G  D  S  G  G  P  L  A  C  S  D  T  G  S 229
 721 ccacctacctggccggcatcgtgtcctggggctacggctgtgcccgtcccaactaccctg 780
 230   T  Y  L  A  G  I  V  S  W  G  Y  G  C  A  R  P  N  Y  P  G 249
 781 gcgtgtacgctgaggtcgcttaccatgtcaattggatcaaggccaatgctgtttaaTCTT 840
 250   V  Y  A  E  V  A  Y  H  V  N  W  I  K  A  N  A  V  *       266
 841 GCTTTGTTTTAACATGCATTTACTTAAATGTGTTAAATAAAGAAGGATATACAAAAAAAA 900
 901 AAAAAAAAAAAAAAAA 916
```

Figure 1 Full cDNA sequence and predicted amino acid sequence of PmTry gene. Start codon (ATG) is marked with bold, and stop codon is marked with asterisk.

Homology analysis of PmTry with other species

The deduced amino acid sequences of *P. monodon* trypsin were compared with those of some other reported species of trypsin. The homology analysis showed that the *PmTry* gene had high homology with all other species (**Table 2**). PmTry shared 91% homology with *F. chinensis* and *L. vannamei*, respectively. The similarity with other crustaceans such as *M. japonicus* and *Eriocheir sinensis* was 83% and 73%, respectively. The amino acid sequence similarity of PmTry with *Danio rerio* and *Homo sapiens* trypsin was 48% and 45%, respectively.

Table 2 Amino acid sequence percent identity of PmTry compared to several other species

Species	GenBank ID	Identity
Penaeus monodon	KP998480	----
Fenneropenaeus chinensis	ACQ45454.1	91%
Litopenaeus vannamei	AEZ67461.1	88%
Palaemon carinicauda	ABQ02534.1	86%
Marsupenaeus japonicus	ACE80257.1	83%
Metapenaeus ensis	ABQ02531.1	80%
Eriocheir sinensis	AKN46052.1	73%
Scylla paramamosain	AGO02163.1	72%
Larimichthys crocea	KKF14922.1	42%
Danio rerio	NM_131708.2	48%
Homo sapiens	M30038.1	45%

The evolutionary tree constructed by the neighbor-joining method is shown in **Figure 2**. The results show that *P. monodon* is closely related to *F. chinensis*. The next closest relatives are the *L. vannamei* and *Palaemon carinicauda*, the *M. japonicus,* and the *M. ensis*, which are clustered in a small branch. They are then clustered in the crustacean branch together with *E. sinensis* and *Scylla paramamosain*.

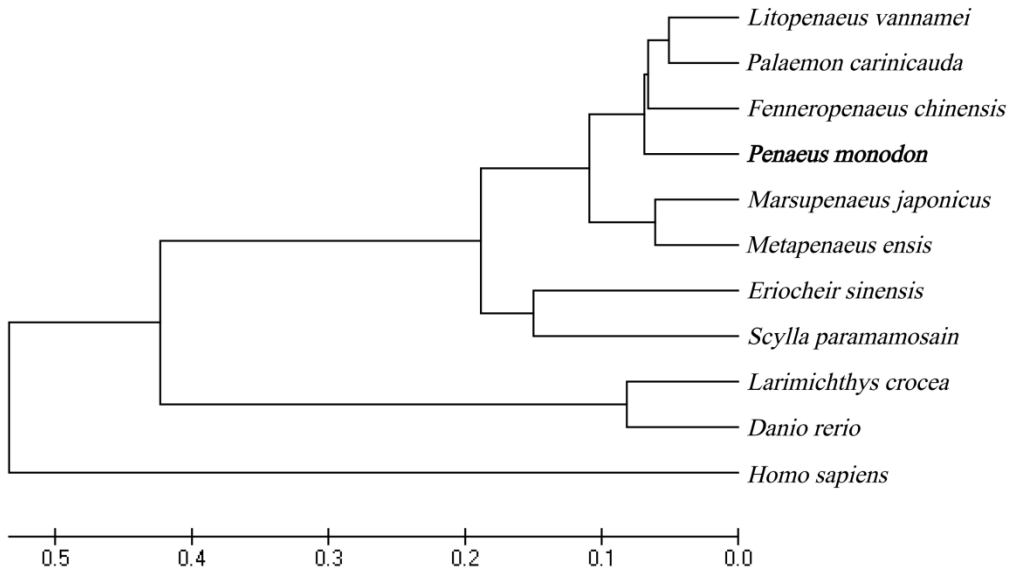

Figure 2 Phylogenetic tree of the TRY amino acid sequence in different groups by Clustal W and MEGA 5.0. The TRY sequence number of different species are listed in **Table 2**

Expression of PmTry in different tissues

The *PmTry* mRNA of *P. monodon* was basically expressed in the whole tested tissues, but the expression levels were significantly different. The expression in hepatopancreas was significantly higher than that in other tissues ($P < 0.05$), followed by the stomach, ovary, and

lymph. The relative expression level is low in the eyestalk nerve and brain tissue. Especially in the eyestalk nerve, it is difficult to observe in the figure (**Figure 3**).

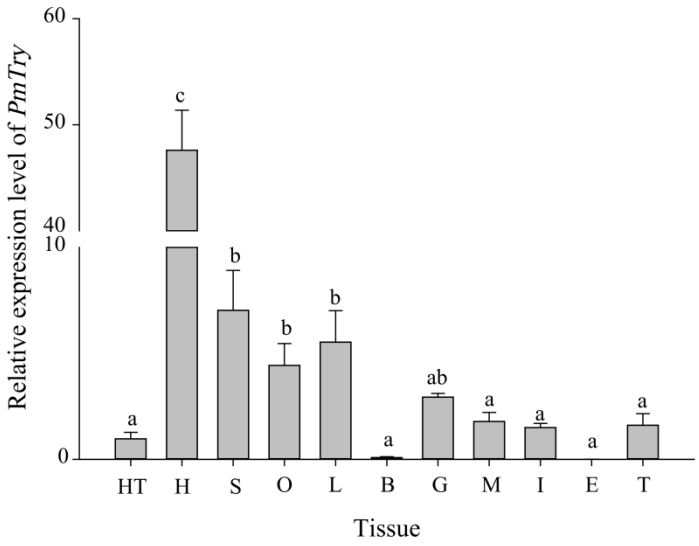

Figure 3 Distribution of *PmTry* gene expression in different tissues of *P. monodon.*
Vertical bars represented the mean ± SD (n=3); Significant differences were indicated with different letters above vertical bars ($P < 0.05$) (L. lymph; H. hepatopancreas; G. gill; I. intestine; O. ovary S. stomach; T. thoracic abdominal nerve; M. muscle; E. eyestalk; B. brain; HT. Heart)

Expression of PmTry in developmental stages

PmTry mRNA of *P. monodon* was expressed in the detection of nauplius, zoea, mysis, larvae, postlarva, juvenile, subadult, and adult in the detection of the whole growth stage. Still, there was a significant difference in the amount of relative expression in each stage. The expression level of nauplius was the highest during larval development and then decreased in the stages of zoea, mysis, and larvae. In the later growth stages, the expression level of juvenile shrimp was significantly lower than that of postlarva, subadult shrimp, and adult shrimp (**Figure 4**).

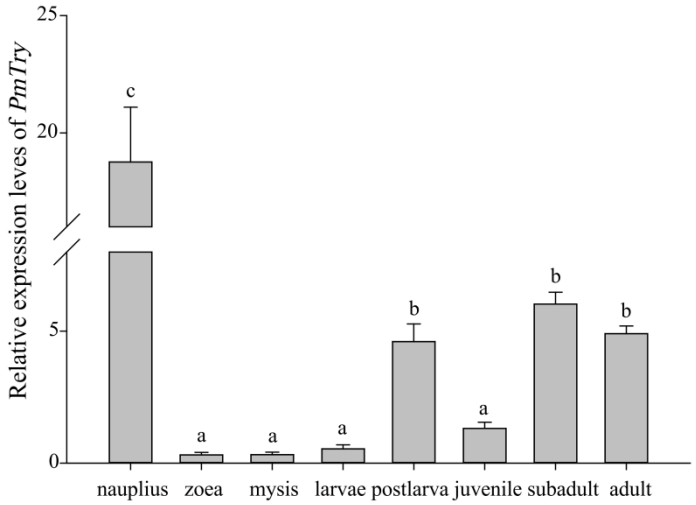

Figure 4 Expression level of *PmTry* in different growth period of *P. monodon*. Vertical bars represented the mean ±SD (n=3); Significant differences were indicated with different letters above vertical bars ($P < 0.05$)

SNP loci screening

The SNP loci were screened in 32 samples of four wild populations, and the amplification and sequencing results were compared and analyzed in Sequencher 4.1.4 software. Eight SNP sites were screened in the *PmTry* gene sequence. Three of them belonged to non-synonymous mutations, and the rest were synonymous mutations. The naming of the SNP site starts from the first base of the cloned gene sequence. The specific information on SNPs is shown in **Table 3**.

Table 3 The SNPs isolated information in PmTry

SNP ID	Location	Frequency	Alleles	Type	Amino acid	S/NS[1]
PmTry-488	488	8/23	CAC/CAT	transition	H	S
PmTry-524	524	7/25	GGA/GGC	transversion	G	S
PmTry-527	527	10/22	GCC/GCT	transition	A	S
PmTry-596	596	11/21	GAG/GAA	transition	E	S
PmTry-623	623	12/20	GAG/GAT	transversion	E/D	NS
PmTry-629	629	5/27	GAT/GAG	transversion	D/E	NS
PmTry-723	723	6/22	CCC/ACC	transversion	P/T	NS
PmTry-798~800	798	9/22	GCC/GCA	transition	A	S

[1] S/NS represents synonymous mutation/Nonsynonymous mutation.

Association analysis of SNPs and growth traits

The screened *PmTry*-524, *PmTry*-629, and *PmTry*-798 were analyzed and detected, and the genotypes of 240 samples in the African population were obtained. The genotype frequencies and gene frequencies are shown in **Table 4**, which are associated with growth traits. The results of the analysis are shown in **Table 5**. The results showed that three genotypes were detected in both loci *PmTry*-524 and *PmTry*-798, and only two genotypes of GT and TT were detected in the locus *PmTry*-629. The correlation analysis between the locus *PmTry*-629 and growth traits showed that there was no significant correlation between GT and TT genotypes and the measured growth traits. The AA genotype of the *PmTry*-524 was significantly smaller than the AC type and CC genotype in the two traits of carapace length and carapace width. In the two indicators of the length of the first abdominal segment and the length of the second abdominal segment, the AA genotype was significantly higher than the CC genotype but not significantly different from the AC genotype. The genotype results of the *PmTry*-798 were found to be correlated with the width and height of the cephalothorax. The AC type was significantly higher than the AA and CC genotypes on the cephalothorax width trait, and the cephalothorax height was significantly higher in the AC genotype. The AC genotype was significantly higher than the CC genotype on the cephalothorax width trait. There were no significant associations found in other traits.

Table 4 Genotype of SNP site of PmTry gene in *P. monodon* and its allele frequency

SNP ID	Number	Ne[1]	Genotypic frequency (%)			Allelic frequency (%)	
			AA	AB	BB	A	B
PmTry-524	240	1.64	6.7	62.9	30.4	38.2	61.8
PmTry-629	240	1.49	55	45		77.5	22.5
PmTry-798	240	1.56	48.3	46.3	5.4	71.4	28.6

[1]Ne represents an effective number of alleles.

Table 5 Correlation of PmTry genotypes with growth traits

SNP ID	Genotype	BL (mm)	CL (mm)	CW (mm)	CH (mm)	FSL (mm)	SSL (mm)	TL (mm)	W (g)
PmTry-524	AA	140.05±11.95[a]	40.31±3.38[a]	18.43±2.33[a]	23.12±2.02[a]	16.50±1.36[b]	15.87±1.62[b]	21.12±1.82[a]	39.69±9.95[a]
	AC	140.48±11.74[a]	43.15±4.67[b]	20.57±2.36[b]	23.23±2.18[a]	16.11±1.66[a,b]	15.41±1.60[a,b]	21.54±2.42[a]	39.52±9.46[a]
	CC	140.00±10.78[a]	41.42±4.09[a,b]	21.02±2.56[b]	23.76±2.82[a]	15.67±1.36[a]	15.08±1.54[a]	21.10±1.83[a]	39.82±9.55[a]
PmTry---62	GT	139.10±11.19[a]	41.94±4.32[a]	19.79±2.44[a]	23.28±2.12[a]	16.17±16.1[a]	15.44±1.60[a]	21.35±2.51[a]	38.59±9.28[a]
	TT	141.31±11.56[a]	42.85±4.64[a]	21.21±2.34[a]	23.48±2.59[a]	15.86±1.60[a]	15.26±1.59[a]	21.40±1.96[a]	40.48±10.10[a]
PmTry-798	AA	139.04±11.06[a]	42.12±4.36[a]	19.85±2.40[a]	23.28±2.08[a,b]	16.18±1.55[a]	15.44±1.60[a]	21.31±2.46[a]	38.59±9.15[a]
	AC	141.94±12.00[a]	42.94±4.74[a]	21.38±2.38[b]	23.63±2.65[b]	15.86±1.63[a]	15.25±1.64[a]	21.46±2.00[a]	41.00±9.49[a]
	CC	137.69±7.72[a]	41.00±3.43[a]	20.00±2.12[a]	22.23±2.38[a]	15.53±0.96[a]	15.23±1.01[a]	21.30±1.75[a]	37.13±7.17[a]

Discussion

The *PmTry* gene cloned in this study is predicted to encode a polypeptide sequence of 266 amino acids, similar to the trypsin gene of shrimp reported from other similar species, such as the *TRY* gene of *L. vannamei*, *M. japonicus*, and *F. chinensis*. Homology and domain analysis showed that the similarity with other species was between 42% and 91%. It had many conserved sequences and active sites for serine proteases.

The study of the expression distribution of the *PmTry* gene in *P. monodon* tissues found that the expression level in hepatopancreas was significantly higher than that in other tested tissues. This may be related to the source of shrimp trypsin. In fish or vertebrates, trypsin is secreted by the pancreas and then exists in the caecum, intestine, and other tissues to play a role, while in *P. monodon*, there is no such thing as the spine. Animals are specialized in pancreatic organs for secreting trypsin, and the hepatopancreas undertakes this secretory function. The hepatopancreas of *P. monodon* is the main organ of material metabolism, which can synthesize various digestive enzymes and is also the primary organ for digesting and absorbing various nutrients. The specificity of the trypsin gene in the hepatopancreas has a particular relationship with its function.

In this study, the relative expression level of the *PmTry* gene in the nauplius of *P. monodon* was significantly higher than that in the zoea, mysis, and larvae shrimp. Changes in the expression of *PmTry* in different developmental stages of larvae are thought to have a significant relationship with the diet switch and self-development of *P. monodon* during this process. During the breeding process of *P. monodon*, the nauplii are not fed with bait. During this period, no exogenous food enters, but their nutrients are used for energy. Currently, the gastric glands are not fully formed to function, and the digestive tract lacks the activity of many proteases. Therefore, the extracellular digestion of food at this stage may mainly depend on the secretion and synthesis of trypsin from the hepatopancreas.

As the larvae of *P. monodon* continue to develop, the zoea open their mouths to feed and are fed with algae, rotifers, and nauplius of fairy shrimp. After that, the expression level of the *PmTry* gene was low in zoea, mysis, and larvae stages. Until the postlarvae stage, the expression level of *PmTry* increased again and continued to the adult stage. Such expression patterns may be related to exogenous feeding, maturation of tissues and organs, and the co-action of maturation of various enzymes. The study of larval development in other animals may indicate changes in *PmTry* expression in *P. monodon*. Studies on the tryptase activity of larvae of *L. vannamei* believed that the change of tryptase activity of *L. vannamei* with the development of larvae was consistent with the changes in the diet of shrimp (Yang et al., 2005). Studies on *E. sinensis* show that protein in the diet significantly inhibits the hepatopancreas active enzymes of *E. sinensis* juvenile crabs. This inhibitory effect is significant in the early stage of artificial diet feeding and declines in the later period (Liu et al., 2006). However, there are also studies showing no relationship between feeding feed ingredients and digestive enzyme activities. For example, different protein sources in the feed have no significant effect on the amylase activity of *M. rosenbergii* (Pan et al., 2006). Studies on many fish (*Ctenopharyngodon idella*, *Culter alburnus*, etc.) found that trypsin activity was detected early in development, even before opening (Ruan, 2010). These results suggest that before the emergence of a functional stomach, the hepatopancreas was the only organ in larvae that expressed and secreted proteolytic enzymes, including trypsin, which played an essential role in digestion. The expression of the trypsin gene suggested it may be programmed to be expressed during early development (Infante and Cahu, 2001). During the period of exogenous feeding from larvae to adult shrimp, the change of *PmTry* expression is related to the perfection and growth of the digestive system of *P. monodon*, such as the development and maturity of the reproductive system, the formation of the digestive tract, gastric glands, and the secretion and synthesis of other enzymes, etc (Gisbert et al., 2009; Zouiten et al., 2008).

In this study, eight SNP loci were screened on the *TRY* gene of *P. monodon*. Two loci, *PmTry*-524 and *PmTry*-798, were found to be significantly associated with growth traits of *P. monodon* by genotyping. The AA genotype of the *PmTry*-524 was significantly smaller than the AC and CC genotypes in the length of the carapace and the width of the carapace. In both the length of the first abdominal segment and the second abdominal segment, the AA genotype was significantly higher than the CC genotype but not significantly different from the AC genotype. This implied that the gene locus is pleiotropic, that is, a gene or locus may play a certain regulatory role on multiple traits, which leads to the AA genotype of this locus being respectively affected in cephalothorax length, cephalothorax width, first abdominal segment length and second abdominal segment length on traits. For the locus *PmTry*-798, the AC genotype was significantly higher than the AA and CC genotypes on the cephalothorax width trait. The AC genotype was only significantly higher than the CC type on the cephalothorax height trait. Although these two loci were found to have certain associations with growth traits, none of the three tested loci were found to be associated with the three traits of body length, caudal length, and body weight. It indicated that the three SNP loci detected and typed by *PmTry* gene had limited effects on the growth traits of *P. monodon*, and there may be other loci or even other multiple genes that jointly regulate the growth traits of *P. monodon*.

In conclusion, the trypsin gene of *P. monodon* was cloned, and its sequences were compared and analyzed, and the phylogenetic tree was constructed in this study. The qPCR method was used to detect that *PmTry* was distributed in many tissues of *P. monodon*, and it was highest expressed in the hepatopancreas. *PmTry* plays an important role in the larval development of *P. monodon*, and its SNP loci were significantly associated with growth traits of *P. monodon*. This provides important molecular markers and genetic information for molecular marker-assisted breeding of prawns.

Acknowledgments

This research was supported by the China Agriculture Research System of MOF and MARA (CARS-48); Hainan Yazhou Bay Seed Laboratory (Project of B21Y10701 and B21HJ0701); Central Public-Interest Scientific Institution Basal Research Fund, South China Sea Fisheries Research Institute, CAFS (2020ZD01, 2021SD13); Hainan Provincial Natural Science Foundation of China (320QN359, 322RC806, 320LH008); Guangdong Basic and Applied Basic Research Foundation (2020A1515110200); Guangzhou Science and Technology Planning Project (202102020208).

References

Alarcón, F. J., Díaz, M., Moyano, F. J. and Abell'N, E. (1998). Characterization and functional properties of digestive proteases in two sparids; gilthead seabream (*Sparus aurata*) and common dentex (*Dentex dentex*). Fish Physiology & Biochemistry. 19(3), 257-267. https://doi.org/10.1023/A:1007717708491

Beirao L. H., Mackie I. M., Teixeira E., Damian C. (2001). Purification and characterisation of trypsin-like enzyme from the pyloric caeca of cod (*Gadus morhua*) II. Brazilian Archives of Biology and Technology. 44(1), 33-40. https://doi.org/10.1590/S1516-89132001000100005

Cao, T. T., Bai, J. J., Yu, Y. L., Fan, J. J. (2012). single Nucleotide Polymorphisms (SNPs) of Carboxpeptidase A1 gene (*CPA1*) Segments and Their Association with the Growth Traits of Grass Carp (*Ctenopharyngodon idella*). Journal of Agricultural Biotechnology. 20(3), 301-307. https://doi.org/10.3969/j.issn.1674-7968.2012.03.011

Dai, W. T., Fu, M. J., Zhao, C., Zhou, F. L., Yang, Q. B., Wang, Y., Shi, J. X., Qiu, L. H. (2015). Molecular cloning and expression analysis of *CDK2* gene from black tiger shrimps (*Penaeus monodon*). South China Fisheries Science. 11(2), 1-11. https://doi.org/10.3969/j.issn.2095-0780.2015.02.001

Dong, Y. X. and Liu, H. X. (1998). Insect trypsin and its gene expression. Chinese Journal of Applied Entomology. 4, 249-251. https://doi.org/CNKI:SUN:KCZS.0.1998-04-022
Gisbert, E., Giménez, G., Fernández, I., Kotzamanis, Y. and Estévez, A. (2009). Development of digestive enzymes in common dentex *Dentex dentex* during early ontogeny. Aquaculture. 287(3-4), 381-387. https://doi.org/10.1016/j.aquaculture.2008.10.039
Huang, J. H., Ma, Z. M., Zhou, F. L., Ye, L., Jiang, S. G. (2005). Study on anatomy structure and histology of the ovary of wild *Penaeus monodon* (Fabricius) from the north of South China Sea. South China Fisheries Science. (3), 49-53. https://doi.org/10.3969/j.issn.2095-0780.2005.03.008
Huber, R., Kukla, D., Bode, W., Schwager, P., Bartels, K., Deisenhofer, J., et al. (1974). Structure of the complex formed by bovine trypsin and bovine pancreatic trypsin inhibitor. II. Crystallographic refinement at 1.9 A resolution. Journal of Molecular Biology. 89(1), 73-101. https://doi.org/10.1016/0022-2836(74)90163-6
Infante, J. L. Z. and Cahu, C. L. (2001). Ontogeny of the gastrointestinal tract of marine fish larvae. Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology. 130(4), 477-487. https://doi.org/10.1016/s1532-0456(01)00274-5
Infante, J. L. Z., Cahu, C. L., Pères, A., Quazuguel, P. and Gall, M. M. L. (1996). Sea bass (*Dicentrarchus labrax*) larvae fed different Artemia rations: growth, pancreas enzymatic response and development of digestive functions. Aquaculture. 139(1-2), 129-138. https://doi.org/10.1016/0044-8486(95)01149-8
Jiang, S. G., Yang, C. H. and Zhou, F. L. (2013). *Penaeus monodon* breeding techniques.
Klein, B., Sellos, D., Wormhoudt, A. V. (1998). Genomic organisation and polymorphism of a Crustacean trypsin multi-gene family. Gene. 216(1), 123-129. https://doi.org/10.1016/S0378-1119(98)00331-X
Klein, B., Moullac, G. L., Sellos, D. and Wormhoudt, A. V. (1996). Molecular cloning and sequencing of trypsin cDNAs from *Penaeus vannamei* (Crustacea, Decapoda): use in assessing gene expression during the moult cycle. International Journal of Biochemistry & Cell Biology. 28(5), 551-563. https://doi.org/10.1016/1357-2725(95)00169-7
Li, Y. D., Zhou, F. L., Huang, J. H., Ma, Z. H., Yang, Q. B., Qiu, L. H., et al. (2016). Cloning and expression analysis of Cathepsin L cDNA of *Penaeus monodon*. South China Fisheries Science. 12(3), 58-66.
Liu, L. H., Chen, L. Q., Zhou, Y. K. and Jiang, H. B. (2006). Research on crustacean digestive enzymes. Feed Industry. 27(18), 56-62. https://doi.org/10.3969/j.issn.1001-991X.2006.18.018
Muhlia-Almazán, A., Sánchez-Paz, A. and García-Carreño, F. L. (2008). Invertebrate trypsins: a review. Journal of Comparative Physiology B Biochemical Systemic & Environmental Physiology. 178, 655-672. https://doi.org/10.1007/s00360-008-0263-y
Outzen, H., Berglund, G. I., Smalås†, A. O. and Willassen, N. P. (1996). Temperature and pH sensitivity of trypsins from Atlantic salmon (*Salmo salar*) in comparison with bovine and porcine trypsin. Comparative Biochemistry & Physiology Part B Biochemistry & Molecular Biology. 115(1), 33-45. https://doi.org/10.1016/0305-0491(96)00081-8
Pan, L. Q., Liu, H. Y. and Xiao, G. Q. (2006). A review on digestive enzyme of crustacean larvae. Journal of Fishery Sciences of China. 13(3), 492-501. https://doi.org/10.3321/j.issn:1005-8737.2006.03.026
Ruan, G. L. (2010). Early development of trypsins and amylases of grass carp *Ctenopharygodon idellus*, yellowcheck carp *Elopichthys bambusa* and topmouth culter *Culter alburnus*. Huazhong Agricultural University.
Sekizaki, H., Itoh, K., Murakami, M., Toyota, E. and Tanizawa, K. (2000). Anionic trypsin from chum salmon: activity with p-amidinophenyl ester and comparison with bovine and *Streptomyces griseus* trypsins. Comparative Biochemistry & Physiology Part B Biochemistry & Molecular Biology. 127(3), 337-346. https://doi.org/10.1016/S0305-0491(00)00267-4

Ueberschr, B. (1995). The use of tryptic enzyme activity measurement as a nutritional condition index: Laboratory calibration data and field application. Ices Marine Science Symposia. 201, 119-129.

Walsh, K. A., Kauffman, D. L., Kumar, K. S. V. S. and Neurath, H. (1964). ON THE STRUCTURE AND FUNCTION OF BOVINE TRYPSINOGEN AND TRYPSIN. Proceedings of the National Academy of Sciences. 51(2), 301-308. https://doi.org/10.1073/pnas.51.2.301

Wang, N., (2004). Study on the Crustacean Trypsin gene. Ocean University of China.

Wang, P., Wu, Y. Y., Li, L. H., Yang, Q. X. and Diao, S. Q. (2011). Research Progress on Shrimp Trypsin. Biotechnology Bulletin. 2, 42-47+97. https://doi.org/10.13560/j.cnki.biotech.bull.1985.2011.02.036

Yang, Q. H., Zhou, Q. C., Ma, L. S. and Chi, S. Y. (2005). Pepsin and trypsin-like enzyme activities in larvae of *Litopenaeus vannamei* Boone. Marine Science. 5, 7-10. https://doi.org/CNKI:SUN:HYKX.0.2005-05-001

Yu, H. J., Liang, X. F., Fang, R., Peng, M. Y., Zhu, T. (2011). The comparison of four different methods used in genotyping G169A in the exon 3 of trypsin gene in *Siniperca chuatsi*. Freshwater Fisheries. 41(6), 9-14. https://doi.org/10.3969/j.issn.1000-6907.2011.06.002

Zhou, F. L., Huang, J. H., Qiu, L. H., Yang, Q. B., Jiang, S. G. (2014). Construction of lentiviral vector for overexpression of *Penaeus monodon* Cyclin B and its influence on Hela cells proliferation. South China Fisheries Science.10(4), 16-21. https://doi.org/10.3969/j.issn.2095-0780.2014.04.003

Zouiten, D., Khemis, I. B., Besbes, R. and Cahu, C. (2008). Ontogeny of the digestive tract of thick lipped grey mullet (*Chelon labrosus*) larvae reared in "mesocosms." Aquaculture. 279(1), 166-172. https://doi.org/10.1016/j.aquaculture.2008.03.039