Molecular Characterization and Expression of the LAP3 Gene and Its Association with Growth Traits in the Blood Clam *Tegillarca granosa*

Hanhan Yao 1, Fuzhen Ren 1, Yongbo Bao 1, Yinghui Dong 1, and Zhihua Lin 2

1 Zhejiang Key Laboratory of Aquatic Germplasm Resources, Zhejiang Wanli University, Ningbo 315100, China; yaohanhan1020@126.com (H.Y.); renpiaoyu@126.com (F.R.); bobbao2001@gmail.com (Y.B.)
2 Ninghai Institute of Mariculture Breeding and Seed Industry, Zhejiang Wanli University, Ningbo 315604, China; zhihua9988@126.com
* Correspondence: dongyinghui118@126.com; Tel.: +86-0574-88222015

Abstract: Leucine aminopeptidase 3 (LAP3) is a metalloproteinase that cleaves N-terminal residues and is involved in protein maturation and degradation. In this study, we characterized the *leucine aminopeptidase 3* (LAP3) gene from *Tegillarca granosa* (*Tg*-LAP3 for short), which appeared to consist of 15,731 nucleotides encoding 530 amino acids. We identified 12 introns and 13 exons in the *Tg*-LAP3 gene, suggesting a highly conserved genomic structure. The proximal promoter sequence consists of 1922 bps with a typical TATA box structure, which is the general structural characteristic of core promoters in eukaryotes. We found two functional domains in the *Tg*-LAP3 protein, including an N-terminal domain (41–174aa) and a peptidase_M17 catalytic domain (209–522aa). Multiple alignment showed that *Tg*-LAP3 shares 73.4% identity with LAP3 of *Mizuhopecten yessoensis* and 55.2–70.7% identity with LAP3 of other species. Quantitative analysis of *Tg*-LAP3 in embryos/larvae and adult tissues indicated that the highest expression occurred in eyebot larva, with limited expression in other stages; among tissues, the highest expression was found in the liver (*p* < 0.05). Association analysis found that three single-nucleotide polymorphisms (SNPs) (g.-488A > G, g.-1123C > T, and g.-1304C > A) in the proximal promoter were successfully typed, but there was no significant difference in growth traits (body weight, shell length, shell width, and shell height) among these genotypes. The results of our study demonstrate the functional roles of the *Tg*-LAP3 gene and provide valuable information for molecular marker-assisted selection (MAS) of the blood clam.

Keywords: blood clam; growth traits; leucine aminopeptidase 3; single nucleotide polymorphisms (SNPs)

1. Introduction

Leucine aminopeptidase 3 (LAP3) has been shown to catalyze the hydrolysis of leucine residues from the amino termini of proteins or peptide substrates [1]. LAP3 is a zinc-containing enzyme found in many tissues, including lens, kidney, pancreas, muscle, liver, mammary gland, and in subcellular locations in a diversity of species [2]. As a cell maintenance enzyme, LAP3 has different functions in mammals, invertebrate, microbes, and plants. In mammals, it processes peptides for MHC I antigen presentation and bioactive peptides (oxytocin, vasopressin, and enkephalins) and is involved in vesicle trafficking to the plasma membrane [3,4]. In microbes, LAP3 plays a role in proteolysis and contributes to promoting DNA binding ability [4]. Additionally, in plants, it has roles in defense, membrane transport of auxin receptors, meiosis, and osmoregulation [4,5].

However, the major roles of LAP3 are protein maturation and degradation, which are essential for metabolism, development, adaptation, and repair [6]. The level of protein maturation and degradation varies greatly, depending on the stages of development, environmental factors, and genotypes [7]; therefore, variations in the LAP3 gene are likely...
to be responsible for some of the variations observed in growth rates in some species [8] and in genetic polymorphism significantly associated with traits of interest in breeding. The potential applications of the LAP3 gene in animal breeding as well as its single-nucleotide polymorphisms (SNPs) and their associations with growth traits have been investigated in mammals, such as cattle [9] and sheep [10]. Three SNPs associated with high milk production in cow and two SNPs associated with high birth weight in sheep were investigated and could be used in marker-assisted selective breeding [9,10]. In marine mollusks, LAPs are also known to play an important role in growth and development, stress responses, and adaptation to changing environmental conditions [11,12], and SNPs associated with growth traits have been found in several aquatic species [13,14]. For example, 23 SNPs of the LAP gene were found to be linked to growth rates in the Pacific oyster, Crassostrea gigas, and three SNPs of the LAP3 gene were identified in the razor clam, Sinonovacula constricta; these SNPs could be used for genetic breeding markers in aquaculture. Meanwhile, gene transcripts and enzyme activities of LAP were also found to play important roles throughout early embryonic development in some bivalves, such as C. gigas [15] and Meretrix meretrix [16].

The blood clam Tegillarca granosa is a well-known seafood with red blood, special flavor, and rich nutrients and is an economically important aquaculture species in China. As a lower marine invertebrate animal, the blood clam is an oviparous bivalve and reaches sexual maturity at two years of age. Like most bivalve mollusks, it undergoes a metamorphosis from embryo to larva to juvenile clam in its early life, including the stages of fertilized egg, blastomere, blastula, gastrula, trochophore, D-shaped larva, umbo larva, eyebot larva, and juvenile clam. In recent years, selective breeding for rapid growth has been a primary focus of the aquaculture industry. The identification of growth genes and genetic variations associated with growth could benefit genetic improvement. So far, several genes involved in growth have been cloned and analyzed in T. granosa, including Smad3 (Mothers against decapentaplegic homolog 3), BMP7 (Bone morphogenetic protein 7), ERK2 (Extracellular signal-regulated kinase 2) [17], GRB2 (Growth factor receptor-bound protein 2) [18], HDAC1 (Histone deacetylase 1) [19], Smad1/5 [20], BMP2/4 [21], and IGF1R (Insulin-like growth factor 1 receptor) [22], and SNPs associated with growth traits have been identified in Smad1/5 [21] and HDAC1 [23], which might be useful as molecular markers for assisted breeding. However, to date, no studies on the LAP3 gene in this species have been conducted.

In the present study, we characterized the Tg-LAP3 gene from the blood clam, determined the quantitative expression profiles, and analyzed associations between SNPs and growth-related traits. This study improves our understanding of the evolutionary and functional roles of the Tg-LAP3 gene and provides valuable information for molecular breeding programs of this species.

2. Materials and Methods

2.1. Sample Collection

Adult two-year-old blood clams (body weight 8.99 ± 0.48 g, shell length 29.14 ± 1.88 mm, shell width 22.32 ± 0.88 mm, shell height 19.82 ± 0.75 mm, on average) were collected from the Ningbo Yongsheng Shellfish Hatchery in Ningbo, China. Eight tissues from four individuals, including blood, liver, gill, mantle, adductor muscle, foot, testis, and ovary were sampled, immediately preserved in liquid nitrogen, and stored at −80 °C for tissue expression analysis. Samples of nine developmental stages (including mature eggs, 2–4 cells, blastula, gastrula, trochophore, D-shaped larva, umbo larva, eyebot larva, and juvenile clam, n > 500) were collected by artificially induced spawning and preserved at −80 °C for embryo/larva expression analysis. To study SNPs of the Tg-LAP3 gene, 80 two-year-old blood clams were collected, and their shell lengths, shell heights, shell widths, and body weights were measured. The adductor muscles of each individual were collected and preserved in ethanol absolute.
2.2. Cloning and Characterization of Tg-LAP3

To obtain the full-length cDNA, expressed sequence tag (EST) of Tg-LAP3 were retrieved from a 454 cDNA library of T. granosa to design specific primers (Table 1). Total RNAs were extracted by the common method described above, and the integrity and quality of RNAs were assessed using 1.5% agarose gel electrophoresis and ultraviolet spectrophotometry. The 5' RACE and 3' RACE reactions were conducted using a SMART™ RACE cDNA amplification kit (Clontech, Mountain View, CA, USA) according to the manufacturer’s instructions. The obtained PCR products were purified, cloned, and sequenced.

| Primer Name | Sequence from 5’ to 3’ | Information of Primer |
|-------------|------------------------|-----------------------|
| Tg-LAP3-GSP1 | CCCCCGACCTCTCCACGGTACCAATGT | 5’-RACE |
| Tg-LAP3-GSP2 | AAACAGGTCAGAATGCTGGAGAATGCC | 3’-RACE |
| Tg-LN1-F | TTTTATTTTTTATTCAGTTTA | Cloning of intron 1 |
| Tg-LN1-R | TTAGGTTGCTCTTGGCTCAGG | Cloning of proximal promoter sequence |
| Tg-LN2-F | TGGGAGGTCCTGGTTCTTG | HRM primer |
| Tg-LN2-R | SNP-F1 | SNP-R1 |
| Tg-LN3/4-F | SNP-F2 | SNP-R2 |
| Tg-LN3/4-R | SNP-F3 | SNP-R3 |
| Tg-LN5-F | LAP3-real-F | qRT-PCR |
| Tg-LN5-R | LAP3-real-R | qRT-PCR |
| Tg-LN6-F | 18S rRNA-real-F | qRT-PCR |
| Tg-LN6-R | 18S rRNA-real-R | qRT-PCR |

Table 1. Primers and their sequences.
To obtain the intron sequence of Tg-LAP3, genomic DNA was extracted by the phenol/chloroform method and dissolved in sterile water at a concentration of 100 ng/ µL. Eleven pairs of primers were designed according to the cDNA sequence of Tg-LAP3 (Table 1). The PCR conditions were as follows: 5 min at 94 °C; 30 cycles of 45 s at 94 °C, annealing for 45 s and 1 min at 72 °C; a final 10 min extension at 72 °C. A Genome Walker™ Universal Kit (Clontech, Mountain View, CA, USA) was used to generate the proximal promoter sequence of Tg-LAP3 according to the manufacturer’s instructions. Each PCR product was cloned and sequenced following the procedures described above.

2.3. Analysis of Nucleotide and Amino Acid Sequences

The sequences of nucleotides and deduced amino acids were analyzed using DNA-MAN8.0 software (https://dnaman.software.informer.com/8.0/, accessed on 15 August, 2013), and protein domains were predicted by Pfam software (http://pfam.xfam.org/, accessed on 1 March, 2021). TRANSFAC and Alibaba 2.1 (http://www.gene-regulation.com/pub/programs/aliba2/index.html, accessed on 2 January, 2002) were used to analyze transcription binding sites of promoter sequences, and transcription start site analysis was performed by NNPP (http://www.fruitfly.org/seq_tools/promoter.html, accessed on 1 March, 1999). EMBOSS Cpgplot (http://www.ebi.ac.uk/Tools/seqstats/emboss_cpgplot/, accessed on 1 July, 2004) was used to analyze CpG island structures, and putative regulatory elements were identified using MatInspector software (Genomatrix, http://www.genomatrix.de/, accessed on 28 April, 2005) [24]. Multiple alignments of LAP3 proteins from T. granosa with other species (Table 2) were performed using the ClustalW2 multiple alignment program (http://www.ebi.ac.uk/Tools/msa/clustalw2/, accessed on 10 September, 2007) [25], and a phylogenetic tree was constructed by using the neighbor-joining method of MEGA 7.0 [26] and iTOL (https://itol.embl.de/, accessed on 3 March, 2006) software.

Table 2. Sequence characteristics of introns of the Tg-LAP3 gene.

| Intron Name | Starting Position(bp) | Length (bp) | Characteristy |
|-------------|-----------------------|-------------|---------------|
| Tg-LN1      | 111                   | 615         | GT-AG         |
| Tg-LN2      | 233                   | 1040        | GT-AG         |
| Tg-LN3      | 394                   | 540         | GT-AG         |
| Tg-LN4      | 463                   | 864         | GT-AG         |
| Tg-LN5      | 575                   | 1042        | GT-AG         |
| Tg-LN6      | 737                   | 1127        | GT-AG         |
| Tg-LN7      | 899                   | 581         | GT-AG         |
| Tg-LN8      | 1113                  | 1105        | GT-AG         |
| Tg-LN9      | 1216                  | 1853        | GT-AG         |
| Tg-LN10     | 1296                  | 1567        | GC-AG         |
| Tg-LN11     | 1446                  | 378         | GT-AG         |
| Tg-LN12     | 1540                  | 1345        | GT-AG         |

2.4. Gene Expression Analysis of Tg-LAP3

Expression of the Tg-LAP3 transcript in eight tissues (blood, liver, gill, mantle, adductor muscle, foot, testis, and ovary; n = 4) and nine developmental stages (mature eggs, 2–4 cells, blastula, gastrula, D-shaped larva, umbo larva, eyespot larva, and juvenile clam; n > 500) was analyzed using qRT-PCR. One pair of 18S rRNA primers (18S rRNA-real-F and 18S rRNA-real-R; see Table 1) of T. granosa was chosen to amplify the 18S rRNA gene of 200 bp as a reference gene for an internal control, and LAP3-real-F and LAP3-real-R (Table 1) were designed to amplify a product of 192 bps. The cDNA from tissues and larvae were diluted to 1:10 for the template for SYBR Green Fluorescent quantitative real-time RT-PCR using an ABI 7500 Fast Real-time PCR System (Applied Biosystems, Foster City, CA, USA). PCR amplifications were performed in a 20 µL volume with 10 µL of iTaq Universal SYBR Green Supermix (Bio-Rad, Hercules, CA, USA), 7.2 µL of deionized water, 0.8 µL of first-strand cDNA, and 1 µL of each primer. The conditions
for the PCR were as follows: incubation at 94 °C for 20 s, 40 cycles of 94 °C for 3 s, 60 °C for 15 s, and 72 °C for 10 s.

2.5. Association Analysis of SNPs with Growth Trait in T. granosa

To collect the SNPs in the Tg-LAP3 proximal promoter, the DNA of 80 individuals was extracted using the phenol–chloroform method. High-resolution melting (HRM) primers (Table 1) were designed to validate the SNPs, and primer annealing temperatures were optimized by temperature-gradient PCR. The PCR products were detected using 8% polyacrylamide gel electrophoresis to visualize the results. HRM genotyping was performed using the 7500 Fast Real-time PCR System, and PCR was performed in a 20 µL volume containing 10 µL of Melt Doctor HRM Master Mix (Applied Biosystems), 10 µM forward and reverse primers, 20 ng/µL of template DNA, and 5.4 µL of deionized water. The PCR was conducted as follows: an initial denaturation at 95 °C for 10 min, then 40 cycles of denaturation at 95 °C for 15 s and annealing for 1 min, followed by melting curve conditions consisting of denaturation at 95 °C for 10 s and annealing for 1 min, and then high-resolution melting at 95 °C for 15 s and annealing for 15 s. The amplification results were analyzed by the HRM system software, and the genotype of each individual was determined through melting curve alignment. To determine the mutant genotype, the genotyped primers were used to amplify the target fragments using the genomic DNA of different genotypes (at least three individuals for each genotype). Each PCR product was purified, recovered, and sequenced following the procedures described above. The association of SNP with growth traits was examined. For each SNP locus, 80 individuals were grouped based on their genotypes, and the mean value and standard deviation of each measured trait were calculated for each genotype group.

2.6. Statistical Analysis

The expression levels of Tg-LAP3 were analyzed using the method of $2^{-\Delta\Delta ct}$, and the results were expressed as the mean ± S.E. For descriptive statistics, the data were assessed using one-way ANOVA, and a Tukey’s honest significant difference analysis was performed on each data set. The effects of SNPs on growth traits were analyzed using the aforementioned method. A comparison of the means among genotypes was performed for each trait using one-way ANOVA. All statistical analyses were performed using SPSS statistical software version 20.0 (SPSS, Chicago, IL, USA), and the level of significance was assumed at $p < 0.05$.

3. Results

3.1. Gene Structure and Characterization of Tg-LAP3

The full-length cDNA of the Tg-LAP3 gene appeared to be 1769 bp (GenBank Accession No. JX103498), including an open reading frame (ORF) of 1593 bp encoding 530 amino acids. The Tg-LAP3 gene contains 12 introns and 13 exons, and all introns were located in the ORF (Figure 1A). Eleven intron–exon junctions conform to the -GT/AG rule, while the intron–exon boundary of Tg-LN10 (intron 10) is -GC–AG- (Table 2). The maximum length of Tg-LAP3 introns is 1853 bp, while the minimum intron length is 378 bp.

The Tg-LAP3 proximal promoter sequence was found to consist of 1922 bp (Figure 1B) without a CpG island, and has three potential transcription start sites at -1641 bp, -512 bp, and -347 bp (ATG, A in the initiation codon is considered as +1). A typical TATA box structure, six E-boxes (CAXXTG), a GATA frame, and octamer transcription factors (Oct-1) binding sites were found, as well as 209 potential transcription factor binding sites, e.g., for TBP, SP1, AP1, and C/EBPα.

3.2. Analysis of the Deduced Amino Acid Sequence of Tg-LAP3

The calculated molecular mass of the deduced mature Tg-LAP3 protein is 57.90 kDa, and the theoretical isoelectric point is 8.80. The mature Tg-LAP3 protein possesses two putative conserved domains: an N-terminal domain (cytosol aminopeptidase family, Pepti-
Figure 1. Sequence characteristics the Tg-LAP3 gene. (A). Genomic DNA structure of the Tg-LAP3 gene. The boxes and lines represent exons and introns, respectively. The numbers on the boxes and below the lines represent the size of each exon and intron. The letters E and I indicate the exons and introns, respectively. (B). Proximal promoter sequence of the Tg-LAP3 gene. The A site of ATG is designated as +1. Partial putative cis elements are indicated in boxes, and their names are listed below each element. Capital letters indicate the transcription start site, and the shaded portion is the predicted promoter region.
The phylogenetic tree showed that all LAP3 of selected animals mainly clustered into three groups (Figure 2), including one group comprising mollusks, one group containing mammals, amphibians, and fish, and one group comprising other invertebrates. Tg-LAP3 in the mollusk group clustered with LAP3 from Yesso scallop *Mizuhopecten yessoensis*, and then with LAP3 from other mollusks. Multiple alignments showed that Tg-LAP3 shares the highest sequence identity (73.4%) with LAP3 of *M. yessoensis*, and 55.2–70.7% identity with LAP3 of other species (Table 3).

![Figure 2. Phylogenetic tree constructed using the neighbor-joining method by MEGA4.0 and iTOL software based on the amino acid sequences of LAP3 from *T. granosa* and other species (Table 3).](image)

| Common Name       | Species                | GenBank No. | Size (aa) | Identity (%) |
|-------------------|------------------------|-------------|-----------|--------------|
| Blood clam        | *Tegillarca granosa*   | XP_021371955| 529       | 73.4         |
| Yesso scallop     | *Mizuhopecten yessoensis* | XP_022345091| 529       | 70.7         |
| Eastern oyster    | *Crassostrea virginica* | MF687490    | 519       | 70.4         |
| Razor clam        | *Sinonovacula constrita* | ACJ05337    | 529       | 71.7         |
| Pacific oyster    | *Crassostrea gigas*    | KU678388    | 417       | 58.0         |
| Hard clam         | *Meretrix meretrix*    | XP_00213156 | 529       | 55.2         |
| Golden apple snail| *Pomacea canaliculata* | XP_025095362| 523       | 69.2         |
| Sea hare          | *Aplysia californica*  | XP_012937771| 527       | 64.4         |
| Freshwater snail  | *Biomphalaria glabrata*| XP_013075022| 497       | 62.2         |
| Sea squid         | *Ciona intestinalis*   | XP_00213156 | 529       | 55.2         |
| Sea cucumber      | *Apostichopus japonicas* | PIK48812    | 547       | 59.8         |
| Sea urchin        | *Strongylocentrotus purpuratus* | XP_030841514| 527       | 66.1         |
| Amphioxus         | *Branchiostoma belcheri* | XP_019624998| 531       | 63.5         |
| Channel catfish   | *Ictalurus punctatus*  | XP_017316877| 518       | 62.4         |
| Zebrafish         | *Danio rerio*          | CAM13051    | 517       | 62.3         |
| Half-smooth tongue sole | *Cynoglossus semilaevis* | XP_008315686| 520       | 61.4         |
| Western clawed frog | *Xenopus tropicalis*   | AAH84523    | 520       | 60.8         |
| African clawed frog | *Xenopus laevis*       | NP_001084691| 495       | 61.6         |
| Human             | *Homo sapiens*         | CAG33409    | 519       | 61.6         |
| Chimpanzees       | *Pan troglodytes*      | JAA04696    | 519       | 61.6         |
| Mouse             | *Mus musculus*         | CAJ18509    | 519       | 61.4         |
3.3. Quantitative Expression Analysis of Tg-LAP3

Among the nine developmental stages, Tg-LAP3 was expressed at low levels from mature eggs to trochophores (Figure 3A). After that, the expression rose from D-shaped larva to eyebot larva and reached the highest level in eyebot larva than in other developmental stages ($p < 0.05$). Afterward, the expression subsided to low levels in juvenile clams (Figure 3A). In adult clams, Tg-LAP3 mRNA was found to be expressed in all eight tissues, and the highest levels were found in the liver ($p < 0.05$). The lowest expression among adult tissues was detected in the ovary, and no statistically significant difference in the expression was detected among the ovary, blood, gill, mantle, and testis (Figure 3B).

Figure 3. Relative expression of the Tg-LAP3 gene in nine early developmental stages ($n > 500$) (A) and eight adult tissues ($n = 4$) (B). Superscript letters represent statistically significant differences at $p < 0.05$.

3.4. Association Analysis of SNPs with Growth Traits

When genotyping the Tg-LAP3 proximal promoter of 80 individuals by HRM, the results showed that only three SNPs (g.-488A > G, g.-1123C > T, and g.-1304C > A) were successfully typed (Table 4). The blood clams with the genotype GG at position g.-488A > G appeared larger (body weight, shell length, shell width, and shell height) than those with AG or AA, and clams with the genotype CC at position eg.-1304C > A tended to be smaller than those with CA or AA. However, the results of one-way ANOVA showed that there was no significant difference in the growth traits with respect to the genotypes of the three SNPs.

Table 4. Growth traits of T. granosa with different genotypes.

| Locus       | Genotype | Number | Frequency (%) | Body Weight (g) | Shell Length (mm) | Shell Width (mm) | Shell Height (mm) |
|-------------|----------|--------|---------------|-----------------|-------------------|------------------|------------------|
| g.-488A > G | AA       | 6      | 10.34         | 8.51 ± 1.91     | 28.37 ± 2.36      | 21.96 ± 1.82     | 19.61 ± 1.45     |
|             | AG       | 28     | 48.28         | 9.16 ± 2.36     | 29.35 ± 2.68      | 22.37 ± 2.12     | 19.93 ± 1.65     |
|             | GG       | 24     | 41.34         | 9.74 ± 2.75     | 29.86 ± 2.60      | 22.97 ± 2.00     | 20.14 ± 1.54     |
|             | CC       | 30     | 49.18         | 8.17 ± 2.45     | 27.88 ± 2.88      | 21.72 ± 2.21     | 19.23 ± 1.78     |
| g.-1123C > T| CT       | 26     | 42.62         | 9.42 ± 1.84     | 29.71 ± 1.98      | 22.60 ± 1.45     | 20.16 ± 1.29     |
|             | TT       | 5      | 8.20          | 8.83 ± 1.88     | 29.20 ± 2.21      | 22.22 ± 1.76     | 19.84 ± 1.68     |
|             | CC       | 39     | 72.22         | 7.78 ± 1.75     | 27.66 ± 2.36      | 21.21 ± 1.66     | 19.11 ± 1.38     |
| g.-1304C > A| CA       | 13     | 22.22         | 9.64 ± 2.23     | 29.75 ± 2.46      | 22.89 ± 2.07     | 20.09 ± 1.55     |
|             | AA       | 3      | 5.56          | 9.66 ± 0.46     | 30.48 ± 0.99      | 22.95 ± 0.38     | 20.26 ± 0.65     |

4. Discussion

4.1. Tg-LAP3 Gene and Amino Acid Sequence Features

Leucine aminopeptidase 3 (LAP3) is an exopeptidase belonging to the M17 family, which cleaves N-terminal residues from proteins and peptides [1,4]. LAP3 is highly conserved at the amino acid level, containing two M17 Pfam domains and lacking the metallopeptidase HEXXH motif [27]. In the current study, the Tg-LAP3 protein was found to include the N-terminal domain (41–174aa) of the peptidase_M17 superfamily and the
peptidase_M17 catalytic domain (209–522aa). It had no HEXXH-Zn combination motif, a common feature for M17 family members, which proved that the Tg-LAP3 protein is highly conserved [4].

The introns and the proximal promoter sequence of Tg-LAP3 gene were characterized and analyzed. The results showed that Tg-LAP3 contains 12 introns and 13 exons with a total sequence of 12,057 bp. By comparison of intron numbers from different species retrieved from GenBank, the number of introns of the LAP3 gene appeared slightly different, varying from 12 to 14. For example, there are 12 introns in human (Homo sapiens), chicken (Gallus gallus), zebrafish (Danio rerio) LAP3 gene, while there are 13 introns in chimpanzee (Pan troglodytes) and 14 introns in great tit (Parus major) LAP3 gene. In addition, 11 intron–exon boundaries (-GT-AG-) and the location of introns are well conserved between T. granosa and vertebrate LAP3 genes, suggesting a highly conserved genomic structure [28]. However, the intron–exon boundary of Tg-LN10 is -GC–AG-, indicating the presence of a C/T mutation [29]. Three possible core promoters were predicted in the 5'-UTR of Tg-LAP3. Obviously, there are conserved TATA boxes upstream of the transcription initiation site, which has the general structural characteristics of eukaryotes’ core promoters [30]. Several putative regulatory elements including GATA-box and E-box were identified in promoter regions of Tg-LAP3. The GATA-box is a binding site for the trans-acting factor ASF-2 and exists in the promoter of genes with tissue-specific expression [31], suggesting that the Tg-LAP3 gene is a gene with tissue-specific expression, which is consistent with results of its specific expression in adult tissues. Six E-boxes (CAXXTG) were found in the proximal promoter of Tg-LAP3. The E-box is one of the sequence motifs characterizing the basic helix–loop–helix myogenic regulatory factors [32], and many genes with tissue-specific expression have multiple E-boxes in their promoter region to cooperatively regulate gene transcription [33,34]; therefore, further studies should be performed to identify how these E-boxes regulate Tg-LAP3 expression.

4.2. Quantitative Expression Analysis of Tg-LAP3

LAP3 plays crucial roles in cell maintenance, growth and development, and defense [13]. In marine mollusks, LAPs are believed to be of selective importance during adaptation to changes in salinity as a result of tides and varying freshwater input, which is the main function in marine organisms [11,35]. Furthermore, LAP3 is mainly involved in protein maturation and degradation, which are essential for metabolism, development, adaptation, and repair [6]. Tg-LAP3 mRNA was found to be expressed at the highest level in the liver, followed by the adductor muscle and foot. The liver is the main digestive and metabolic organ in bivalves, where protein maturation and degradation is very active. This suggests that LAP3 is the main exopeptidase in T. granosa. In bivalves, protein maturation and degradation enable exceptional changes in body morphology during larval development, from the free-living planktonic stage to sessile juvenile and calcified adults [15]. In C. gigas, gene transcripts of LAP were detected throughout larval development, which reflects the essential role of LAP in protein maturation and degradation during all development stages. The peak in LAP enzyme activity corresponds to the onset of eye development, suggesting LAP helps in the major reorganization of tissue structures that typically occurs during this stage [15]. In M. meretrix, the LAP3 gene is expressed at the highest level in D-shaped larva, which suggests that LAP3 might be involved in tissue formation during early development [16]. Among the developmental stages, Tg-LAP3 mRNA expression increases as the larvae develop, is low from the stage of mature egg to the trochophore stage, and increases immediately from the stage of D-shaped larva to that of eyebot larva. The highest expression level was in the eyebot larva stage (p < 0.05); therefore, the Tg-LAP3 gene might play an important role in larval attachment and metamorphosis, as observed in C. gigas [15].
4.3. Association Analysis between SNPs of Tg-LAP3 Gene and Growth Traits

LAP3 plays an important role in protein maturation and degradation, which varies greatly depending on the development stage, environment factors, and genotype [36]. Variations of the LAP3 gene are therefore likely to be responsible for differences of production traits among different individuals, and genetic diversity and correlation analysis between SNP and production traits have been performed in some mammals [9,10] and bivalves [13,14]. For example, in mammals, three SNPs (g.24794T > G, g.24803T > C, and g.24846T > C) of the LAP3 gene were linked to milk production in Chinese Holstein cows, and two SNPs were significantly correlated with birth weight in sheep. In bivalves, 23 SNPs in C. gigas and 3 SNPs in S. constricta were identified to be associated with growth traits, which might be useful for applications in genetic-based stock enhancement and breeding design. In our SNPs typing, however, there was no significant difference in the growth traits among the genotypes of the three SNPs (g.-488A > G, g.-1123C > T, and g.-1304 > A). The blood clams with the genotype GG at position g.-488A > G are larger (body weight, shell length, shell width, and shell height) than those with AG or AA, but we did not find any significant difference in growth traits. It is generally known that the growth traits are controlled by multiple genes [37], indicating that the mutation effects of g.-488A > G in Tg-LAP3 gene may be regulated by other genes and are too small to affect growth traits in T. granosa.

5. Conclusions

In this study, the Tg-LAP3 gene was characterized. It consists of 12 introns and 13 exons. The proximal promoter sequence of Tg-LAP3 consists of 1922 bps without a CpG island, with three potential transcription start sites, a typical TATA box structure, six E-boxes (CAXXTG), a GATA frame, and octamer transcription factors’ binding sites (Oct-1). Tg-LAP3 appeared to be most highly expressed in the eyelite larva when examining embryos/larvae in nine developmental stages, as well as in the liver among eight adult tissues. Three SNPs (g.-488A > G, g.-1123C > T, and g.-1304C > A) in the proximal promoter were successfully typed, but there was no significant difference in growth traits associated with these SNPs. The results of our study provide valuable information for functional studies of Tg-LAP3 and molecular marker-assisted selection of the blood clam.

Author Contributions: Conceptualization, Y.D. and Z.L.; Methodology and software, Y.B. and Y.D.; Validation, Y.D. and H.Y.; Formal analysis and investigation, Y.B.; F.R. and H.Y.; Resources and data curation, Y.D. and H.Y.; Writing—original draft preparation, F.R. and H.Y.; Writing—review and editing, H.Y.; Y.D.; and Z.L.; Visualization, supervision, project administration, and funding acquisition, Y.D. and Z.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported financially by National Key Research and Development Program of China (2018YFD0901405), Zhejiang Major Program of Science and Technology (2021C02069-7), National Marine Genetic Resource Center Program and China Agriculture Research System of MOF and MARA.

Institutional Review Board Statement: In the present study, Tegillarca granosa clams were collected from the genetic breeding research center of Zhejiang Wanli University, China. All experimental procedures were approved by the Institutional Animal Care and Use Committee (IACUC) of Zhejiang Wanli University, China (Approval code: 20210903001).

Data Availability Statement: The sequence of LAP3 gene of Tegillarca granosa was deposited in GenBank, and the accession number is JX103498.

Acknowledgments: We sincerely thank Ningbo Yongsheng Shellfish Hatchery for sample collecting, and thank International Science Editing (http://www.internationalscienceediting.com, accessed on 20 October 2021) for editing our manuscript.

Conflicts of Interest: The authors declare no conflict of interest.
28. Sharp, P.A. Speculations on RNA splicing. Cell 1981, 23, 643–646. [CrossRef]
29. Shi, H.; Maier, S.; Nimmrich, I.; Yan, P.S.; Caldwell, C.W.; Olek, A.; Huang, T.H. Oligonucleotide-based microarray for DNA methylation analysis: Principles and applications. J. Cell Biochem. 2003, 88, 138–143. [CrossRef]
30. Suzuki, Y.; Tsunoda, T.; Sese, J.; Taira, H.; Mizushima-Sugano, J.; Hata, H.; Ota, T.; Isogai, T.; Tanaka, T.; Nakamura, Y.; et al. Identification and characterization of the potential promoter regions of 1031 kinds of human genes. Genome Res. 2011, 11, 677. [CrossRef]
31. Lam, E.; Chua, N.H. ASF-2: A factor that binds to the cauliflower mosaic virus 35S promoter and a conserved GATA motif in cab promoters. Plant Cell 1989, 1, 1147–1156. [PubMed]
32. Ceccarelli, E.; McGrew, M.J.; Nguyen, T.; Grieshammer, U.; Horgan, D.; Hughes, S.H.; Rosenthal, N. An E box comprises a positional sensor for regional differences in skeletal muscle gene expression and methylation. Dev. Biol. 1999, 213, 217–229. [CrossRef]
33. Rao, M.V.; Donoghue, M.J.; Merlie, J.P.; Sanes, J.R. Distinct regulatory elements control muscle-specific, fiber-type-selective, and axially graded expression of a myosin light-chain gene in transgenic mice. Mol. Cell Biol. 1996, 16, 3909–3922. [CrossRef]
34. Wang, D.; Sul, H.S. Upstream stimulatory factor binding to the E-box at-65 is required for insulin regulation of the fatty acid synthase promoter. J. Biol. Chem. 1997, 272, 26367–26374. [CrossRef]
35. Moore, M.N.; Koehn, R.K.; Bayne, B.L. Leucine aminopeptidase (aminopeptidase-I), N-acetyl-b-hexosaminidase and lysosomes in the mussels, Mytilus edulis, in salinity changes. J. Exp. Zool. 1980, 214, 239–249. [CrossRef]
36. Hawkins, A.J.S. Protein turnover: A functional appraisal. Funct. Ecol. 1991, 5, 222–233. [CrossRef]
37. Hu, X.L.; Guo, H.H.; He, Y.; Wang, S.; Zhang, L.L.; Wang, S.; Huang, X.T.; Roy, S.W.; Lu, W.; Hu, J.J.; et al. Molecular characterization of myostatin gene from Zhikong scallop Chlamys farreri (Jones et Preston 1904). Genes Genet. Syst. 2010, 85, 207–218. [CrossRef] [PubMed]