GROWTH-RELATED EXPRESSED SEQUENCE TAG - SAMPLE SEQUENCE REPEATS (EST-SSRS) SCREEN OF MUD IVORY WHELK (*BABYLONIA LUTOSA*) THROUGH TRANSCRIPTOME SEQUENCING

XIONG, G.[1*] – WANG, P.[2] – KANG, L.[1] – CHEN, Z. N.[2] – ZHOU, X. W.[3] – WU, Q. S.[4] – WANG, X. Q.[2*]

[1]*Hunan Biological and Electromechanical Polytechnic, Changsha 410127, China*

[2]*College of Animal Science and Technology, Hunan Agriculture University, Changsha 410128, China*

[3]*Aquatic science extension station, Xiangxi Autonomous State of Tujia and Miao Minorities, Jishou, 416000 Hunan, China*

[4]*Fisheries Research Institute of Fujian Province, Xiamen 361013, China*

[*]*Corresponding authors*
e-mail: xionggang709@126.com (Xiong, G.), wangxiao8258@126.com (Wang, X. Q.)

(Received 15th Nov 2019; accepted 30th Jan 2020)

Abstract. Mud ivory whelk (*Babylonia lutosa*) is one of the most economically important seawater mollusks in China. However, currently, there is no report on the molecular markers related to the growth of *Babylonia lutosa*, which hinders the its genetic breeding with dominant growth traits. The present study was the first to sequence the transcriptomes of *B. lutosa*, and then searched for EST-SSR in the growth-related genes. Wild *B. lutosa* samples were collected from the Lianjiang River, Fujian Province of China. The RNA of gastropods and hepatopancreas tissues from three *B. lutosa* were measured and mixed, and sequenced by Illumina HiSeq sequencer. A total of 24,829,460 reads containing 5,015,236,064 bp were obtained through sequencing. Eight growth-related genes (TBC-1-A, TBC-1-2B, FGFR-1, GFRBP2, GFR-15, FGFR-1-A, SMAD6, and TGFβR-1) containing EST-SSR sequences were detected in the *B. lutosa* transcriptomes. Our results would provide theoretical basis for *B. lutosa* breeding based on growth-related EST-SSRs.
Keywords: *aquaculture, Gastropoda, growth-related gene, insulin-like growth factors, microsatellites*

Introduction

Growth-related genes of aquatic animals have always been the focus of researchers. For instance, insulin-like growth factors (IGFs) of aquatic animals are a class of proteins with insulin-like metabolism and mitotic function. IGFs participate in the embryonic development and reproduction of vertebrates and are the main determinants of embryonic and postnatal growth (Reinecke et al., 1997; Upton et al., 1998; Zhang et al., 2011). In addition, IGFs mediate growth-promoting effects of growth hormones (Moriyama et al., 2000). Aquatic animal IGF system includes IGF-I, IGF-IR, IGF-II, IGF-IIR, and IGFBPs (Li et al., 2011), which play important roles in inducing myopoietin (MyoG) gene expression (Florini et al., 1991; Yano et al., 1999), promoting myocyte differentiation (Hembree et al., 1991; Le Roith et al., 2001; Buckingham et al., 2003), regulating skeletal muscle growth (Doumit et al., 1993; Lamberson et al., 1995; Davis et al., 2002), and other aspects. Growth-related genes have genetic polymorphism, and different genotypes correspond to different growth traits.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1802_34713482

Single nucleotide polymorphism (SNP) loci and microsatellites are important molecular markers commonly used in population genetic diversity research. Ruan et al. (2011) identified three SNPs by comparing the sequences of exon 1, exon 3, intron 2, and intron 4 of the IGF-1 gene in Giraffe tilapia (*Oreochromis niloticus*). Then they detected intron genotypes of 121 *G. tilapia* from five genetic families by PCR-RFLP, and screened two SNPs that correlated with weight gain. Therefore, the two SNPs could be used as molecular assisted markers for breeding. Microsatellites have been widely used in population genetic diversity analysis, because of their wide distribution in genome, good polymorphism, genetic stability, and easy detection (Nie, 2014). For instance, the population genetic diversities of blacklip abalone (*Haliotis rubra*) (Huang et al., 2000), scallop (*Pectinidae* sp.) (Feng et al., 2014), marine shrimp (*Fenneropenaeus chinensis*) (Zhang et al., 2005), whiteleg shrimp (*Litopenaeus vannamei*) (Tong et al., 2009), and Japanese pufferfish (*Fugu rubripes*) (Lu et al., 2013) have been studied using microsatellite markers. High-throughput sequencing technology provides a convenient way to transcriptome sequence and large-scale screen SNPs and microsatellite markers. Hou et al. (2011) sequenced transcriptomes of the larvae of scallop (*Patinopecten yessoensis*) at different development stages by high-throughput sequencing technology, and screened a large number of expressed sequence tag - sample sequence repeats (EST-SSRs) and SNPs that were correlated to growth, reproduction, and immune recognition.

Mud ivory whelk (*Babylonia lutosa*) is one of the most economically important seawater mollusks in China (Xiong et al., 2015). In 2018, the output of *B. lutosa* is approximately 4,000 tons in China. Although aquaculture output of marine snails has decreased by 6.17% compared with 2017 in China, the cultivation of *B. lutosa* increases (Fishery and Fishery Administration of the Ministry of Agriculture and Rural Areas of China et al., 2019). It is widely distributed in southeast coast of China, southeast Asia and Japan (Xiong et al., 2015). *B. lutosa* inhabits in the sediment bottom of several meters to tens of meters deep in the subtidal zone, mainly feeds on oyster eggs, organic debris, and protozoa in the juvenile stage, and small fish, shrimp, crab, and shellfish in the adult stage (Yin et al., 2007). Currently, studies on *B. lutosa* are mainly focused on morphological comparison (Wang et al., 2007; Pan et al., 2010; Huang et al., 2010; Qi et al., 2011), metamorphosis (Yang et al., 2008), environmental adaptation (Lin, 2012), genetic characteristics (Chen et al., 2011; Qin et al., 2014), and culture technologies (Yin et al., 2007). Molecular breeding is very important for the cultivation of snails with high growth characteristics. However, there is no report on the molecular markers related to the growth of *B. lutosa*, which hinders the genetic breeding of *B. lutosa* with dominant growth traits. The present study was the first time that sequenced the transcriptomes of *B. lutosa*, and then searched for EST-SSR in the growth-related genes. Our results would provide theoretical basis for *B. lutosa* breeding based on growth-related EST-SSRs.

Materials and Methods

Sample collection, RNA extraction, cDNA library construction and sequencing analysis

Wild *B. lutosa* samples were collected from the estuary of Lianjiang River, Fujian Province of China (26°20′ N, 119°42′ E) in 2015 (*Fig. 1*). Three living *B. lutosa* were dissected and their gastropods and hepatopancreas were quickly cut out and used to extract total RNA by a TRIzol reagent kit (TAKARA, China) (Xiong et al., 2015, 2019).

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1802_34713482

The RNA concentration and purity were detected by an Ultrospec^{TM} 2100UV/Visible spectrophotometer, and the RNA integrity was detected by 1% agarose gel electrophoresis. The RNA of gastropods and hepatopancreas tissues from the three *B. lutosa* were measured and mixed. After enrichment, purification, interruption and cDNA synthesis, the constructed cDNA library was sequenced by Illumina HiSeq sequencer at Beijing Biomarker Technologies Co., China.

Figure 1. *Map shows the sampling location*

Sequence assembly, annotation, and growth-related EST-SSR screening

The original reads were assembled to transcripts using Trinity software. Several sliceable transcripts were clustered into one gene, and finally obtained the unigene library. SSRs in the cDNA sequence of unigenes were searched using MISA software (http://pgrc.ipk-gatersleben.de/misa/) (Kanehisa et al., 2008). The searching parameters were set as a single base with a minimum repetition of more than 10 times, two bases with a minimum repetition of more than 6 times, and 3 - 6 bases with a minimum repetition of more than 5 times. The distance between two repetitions should not be greater than 100 bp. The ORFs were predicted using Getorf software based on the unigene sequences (Mortazavi et al., 2008). The software parameters were set to predict the forward and reverse directions of three bases at both ends of unigenes, and the longest sequence was as the ORF of unigenes. The reads obtained by the sequencing were aligned to unigenes library, and calculated the expression abundance (RPKM) of each unigene (Mortazavi et al., 2008). The unigenes were annotated using BlastX program referenced to the protein sequence databases of Swiss-Prot (Apweiler et al., 2004), TrEMBL (Apweiler et al., 2004), GenBank Nr (Deng et al., 2006), Gene Ontology (GO) (Ashburner et al., 2000), COG (Tatusov et al., 2000), and KEGG (Kanehisa et al., 2004). The parameter E value was set to $< 1e^{-6}$, and the maximum similar sequence was selected as the annotated gene. According to GO annotation information, functions of the unigenes were classified to molecular function, cellular component, and biological process (Ashburner et al., 2000). KEGG analysis was implemented by KEGG Automatic Annotation Service (KAAS) (Masoudi-Nejad et al., 2007). Growth-related genes were screened from the assembled annotated genes, and EST-SSR markers were excavated for the selected growth-related genes by SSR analysis.

The EST-SSRs were submitted to NCBI GenBank database with accession number from KM972681 to KM972698, and from KT763437 to KT784812.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1802_34713482

Results

Transcriptome characteristics and EST-SSR analysis of B. lutosa

A total of 24,829,460 reads containing 5,015,236,064 bp were obtained by sequencing. The GC content of the DNA sequences was 50.54%. CycleQ20 was 100%. A total of 119,113 unigenes that contained 16,877 unigenes with more than 1 kb were obtained through *de novo* assembling. N50 was 835 bp, and the average length of unigenes was 627.3 bp (*Table S1*).

A total of 13,962 EST-SSRs were found from 8,177 unigenes, which total length was 262,687 bp. We also found a large number of mixed repetitive nucleotide sequences. These mixed repetitive nucleotide sequences were distributed in 2,272 unigenes (*Table 1*). The frequency of EST-SSRs in the transcriptome of *B. lutosa* was 11.72% (ratio of the number of EST-SSRs detected to the total number of unigenes); and the frequency of EST-SSR occurrence was 6.86% (ratio of the number of unigenes with EST-SSR to the total number of unigenes). In the transcriptome of *B. lutosa*, an SSR occurred on average every 350.78 kb.

Table 1. *Occurrence of different types of EST-SSR in the B. lutosa transcriptome*

Repeat type	Number	Proportion	Frequency (%)	Average distance (kb)	Total length (bp)	Average length (bp)
Mononucleotide	5,536	39.65	4.65	884.70	75,641	13.7
Dinucleotide	3,906	27.98	3.28	1253.89	68,946	17.7
Trinucleotide	2,141	15.33	1.80	2,287.57	40,314	87.6
Tetranucleotide	100	0.72	0.08	48,976.91	2,536	25.4
Pentanucleotide	7	0.05	0.006	699,670.14	175	25
Compound SSR	2,272	16.27	1.907	2,155.67	75,075	75.5
Total	13,962	100	11.72	350.79	262,687	18.4

EST-SSR was abundant in the transcriptome of *B. lutosa*, and it could be found from mononucleotide repeat to pentanucleotide repeat types, but no hexanucleotide repeat type was found. The frequencies of occurrence of various EST-SSR types were quite different. It mainly focused on mononucleotide, dinucleotide, and trinucleotide repeats, which accounted for 39.65%, 27.98%, and 15.33% of the total EST-SSRs, respectively. The tetranucleotide and pentanucleodite repeats accounted for 0.72% and 0.05%, respectively. Among them, the largest number of SSR types was: A repeat primitives were 3,534 (25.31%), and T repeat primitives were 2,984 (21.37). The number of EST-SSR loci with different repeats was also quite different (*Fig. 2A*). The major EST-SSRs with more than 10 repeat elements were mononucleotides. The dinucleotide repeats with more than 10 repeat elements were very rare. TG repeats were 1,136 (8.14%), CA repeats were 866 (6.35%), AC repeats were 717 (5.14%), and AC/GT repeats were 3,409 (24.42%). In the trinucleotide repeats, GGA repeat was 151, GAG repeat was 114, GGT repeat was 108, GTG repeat was 104, TTG repeat was 100, AGG/CCT repeat was 560 (0.040%), and ACC/GGT repeat was 508 (0.036%; *Fig. 2B*). According to the overall distribution, polynucleotide repeats predominated in the frequency of six repeat.

Through analyzing the length of 13,962 EST-SSRs, the lengths of microsatellite sequences contained in the EST sequences of *B. lutosa* ranged from 9 to 265 bases with a total length of 262,687 bp. The average length of the microsatellite sequences was

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1802_34713482

18.4 bp. The length of EST-SSRs was mainly in the range of 12 to 25 bp, of which 8928 (63.95%) were in the range of 12 to 20 bp. Number of SSRs with the length between 21 to 25 bp were 1,276 (9.14%), and the number of SSRs with the length of more than 25 bp were 225 (1.60%; *Fig. 2C*).

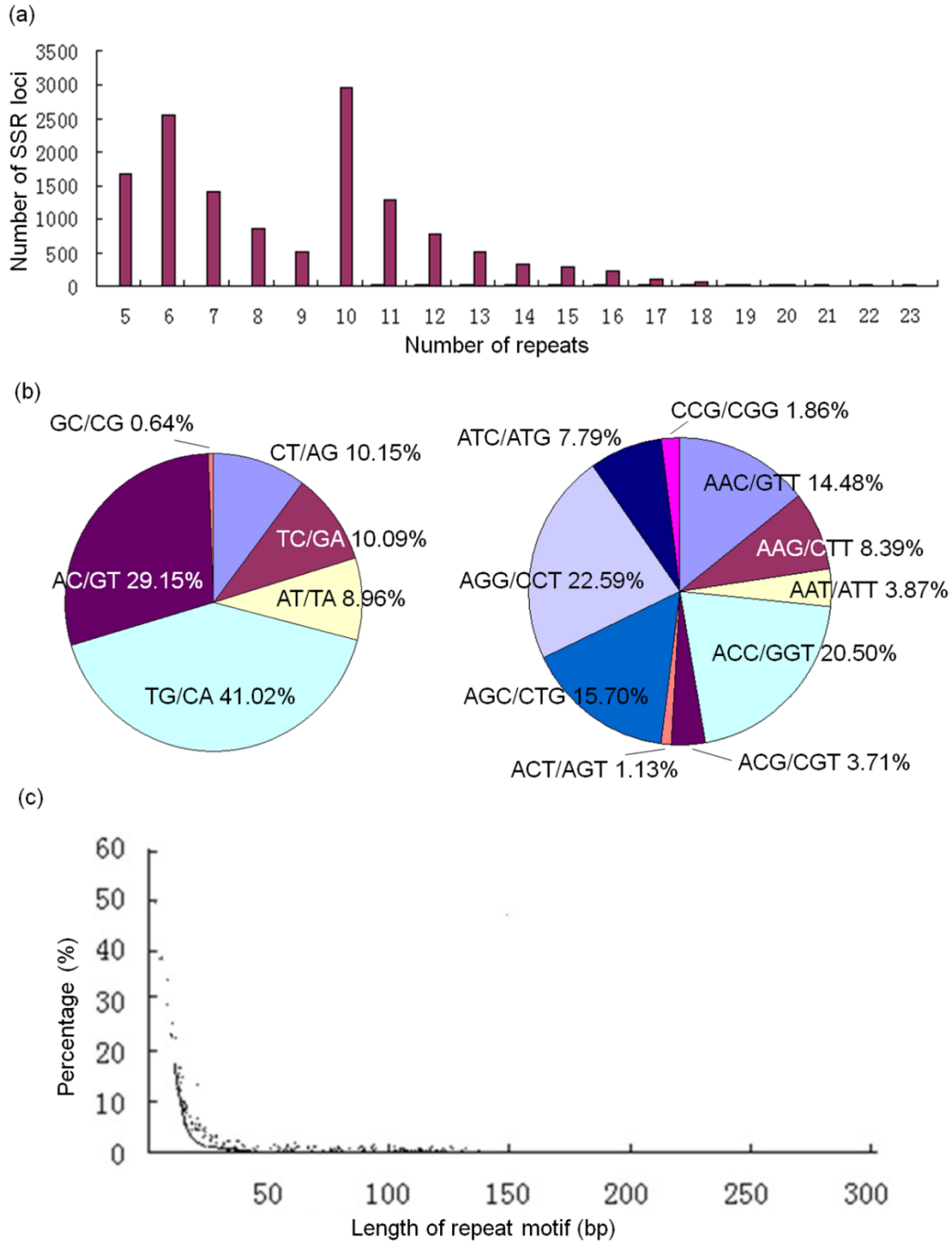


Figure 2. *Distribution of the number of repeats of EST-SSR (A), percentage of different motifs of dinucleotide and trinucleotide SSRs (B), and distribution of the repeat motif length of SSRs in B. lutosa transcriptome (C)*

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1802_34713482

Functional annotation and Growth-related Genes Containing EST-SSR Sequences

Total of 118,608 unigenes were protein encoding unigenes, of which 78.84% were unigenes within 300 bp, and 6.18% were unigenes above 900 bp (*Appendix 1*). A total of 24,829,460 reads were belonged to the protein encoding unigenes, which showed the high reliability of the obtained transcriptome data (*Appendix 1*).

Total of 35,931 unigenes were annotated, which were in corresponding the sequences from sea-slug (*Aplysia californica*), oyster (*Crassostrea gigas*), polychaete worm (*Capitella teleta*), amphioxus (*Branchiostoma floridae*), sea urchin (*Strongylocentrotus purpuratus*), acorn worm (*Saccoglossus kowalevskii*), filarial nematode worm (*Brugia malayi*), eye worm (*Loa loa*), and sea anemone (*Nematostella vectensis*) (*Appendix 2*), and some sequences did not have similar homologous sequences in the databases, which was probably caused by the lack of mollusk EST sequences in these databases. A total of 35,931 annotated functions were obtained from the *B. lutosa* transcriptome annotation analysis (*Appendix 3*).

Total of 109,359 unigenes obtained GO information, of which 60,143 (55.0%) unigenes were belonged to biological process, 31,931 (29.2%) genes were belonged to cellular component, and 17,285 (15.8%) were belonged to molecular function. In the GO classification system, the biological process, cellular component, and molecular function were classified into more detailed 56 sub-categories (*Fig. 3*). In the biological process, there were 3,289 unigenes participated in the reproduction, 7,751 unigenes participated in the metabolic process, 353 unigenes participated in the immune system process, 387 unigenes participated in the cell proliferation, 8,826 unigenes participated in the cellular process, 2,876 unigenes participated in the reproductive process, and 5,562 unigenes participated in the biological regulation.

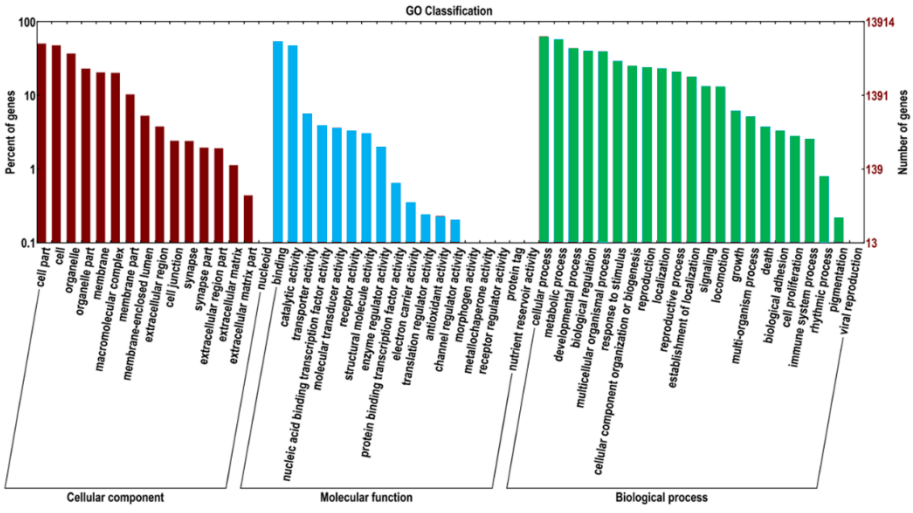


Figure 3. *GO function annotation results of B. lutosa transcriptome library*

Growth-related gene containing EST-SSR sequences were growth hormone-regulated TBC protein 1-A (TBC-1-A), growth hormone-regulated TBC 1 domain family member 2B (TBC-1-2B), fibroblast growth factor receptor 1 (FGFR-1), growth factor receptor-bound protein 2 (GFRBP2), growth factor receptor substrate 15 (GFR-15), fibroblast growth factor receptor 1-A (FGFR-1-A), SMAD family member 6 (SMAD6), and TGF-beta receptor type-1 (TGFβR-1, *Table 2*).

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1802_34713482

Table 2. *Parameters of primer pairs for growth-relative EST-SSR*

Primer name		Primer sequence (5'-3')	Homologous gene	Core SSR sequence	Amplification location or target fragment length
3'RACE Olig(T)-Adaptor		CTGATCTAGAGGTACCGGATCCTTTTTTTTTTTTTT			3'-terminal amplification
3'RACE Adaptor		CTGATCTAGAGGTACCGGATCC			3'-terminal amplification
5'RACE Olig(T)-Adaptor		GACTCGAGTCGACATCGATTTTTTTTTTTTTTTT			5'-terminal amplification
5'RACE Adaptor		GACTCGAGTCGACATCG			5'-terminal amplification
TBC -1-A	(TG)7P1+	CAAGTGTGTGAATGATGTTGTCTAC	XM_005107837: *Aplysia californica* growth hormone-regulated TBC protein 1-A-like	(TG)7	297 bp
	(TG)7P1-	CACACATTCACATACATGCACA			
	GSPP1+	CCTTCCCAGAGAACATTTACTT			3'-terminal amplification
	NGSPP1+	CAAGTGTGTGAATGATGTTGTCTAC			3'-terminal amplification
	GSPP2-	ACTGATGTAGGACTGCACCGT			5'-terminal amplification
	NGSPP2-	TCCCTTCCTGCAGAACCGCTTGACT			5'-terminal amplification
TBC -1-2B	*P1+	GCATCACTCCGGTCACCAGCTT	XM_013087488: *Aplysia californica* TBC-1 domain family member 2B-like	(ACC)5*	225 bp
	*P1-	CCTGCTGCGGCACGGTCC		(GAG)5	BL-106
FGFR-1	(GGA)5P1+	TTGAGCAAGAGGGAAGAGC	XM_003974594.1: *Takifugu rubripes* fibroblast growth factor receptor 1-A-like	(GGA)5	206 bp
	(GGA)5P1-	ACCAGCGTGGCCTCGTCCCTT			
	(AGA)5P2+	TTCTCCGTGAACCTGACGCT		(AGA)5	250 bp
	(AGA)5P2-	AGGAAGAAGGTGTTGAGGGTGT			
FGFR-15	(CTG)6P1+	TTGACCGGACCTAGACTGC	XM_005110162.1: *Aplysia californica* epidermal growth factor receptor substrate 15-like	(CTG)6	217 bp
	(CTG)6P1-	TCTACGTGGCCCTCAAACT			
FGFR-1-p	(AC)8P1+	CGAGAACAGGCGGAGAAT	XM_005105989.1: *Aplysia californica* FGFR1 oncogene partner-like (LOC101862887)	(AC)8	261 bp
	(AC)8P1-	AAGGACGGAGGGAGACG			
	(GT)*P2+	TGCTGGGACTGACGGGCT		(GT)6*(G)16	245 bp
	(GT)*P2-	TCTTAGCAATTGCTTGGCGC			
GFR BP2	(CA)*P1+	TGCTACTACACACACATACCACACA	NM_001165306.1: *Salmo salar* growth factor receptor-bound protein 2	(CA)7*(TG)6	229 bp
	(CA)*P1-	GTGCTTTGTTTAGAATCTGGGC			
SMAD6	(CCT)5P1+	GCCAGTATTTGCCCAACAGT	XM_011385995.1: *Pteropus vampyrus* SMAD family member 6 (SMAD6)	(CCT)5	231 bp
	(CCT)5P1-	TGGAGGTGGTGGGGAGTGT			BL-120
TGFβR-1	(TG)*P1+	GTATCAACTACAAATGACACAACTT	XM_005103324.1: *Aplysia californica* TGF-beta receptor type-1-like	(TG)6*(TG)6	241 bp
	(TG)*P1-	ACTTGTGAAAGAAAAAGATTATTAT			
	(CA)7P2+	ATCATTACAGACAGGTTCTTAGC		(CA)7	202 bp
	(CA)7P2-	GGTAGGGAGGGCAAGGT			BL-151

* Intermediate non-repetitive sequence fragments of composite microsatellite sequences. Length of core sequence is amplification length of the EST-SSR

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1802_34713482

Discussion

A total of 13,962 EST-SSRs was obtained from the *B. lutosa* transcriptome, which total length was 262,687 bp. The number and frequency of EST-SSRs in the *B. lutosa* transcriptome were more abundant than those in *P. yessoensis* (2,700 EST-SSRs; Hou et al., 2011). Besides the improvement of NGS technology provided more quantity and higher quality sequencing data of *B. lutosa* transcriptome, the reason caused the more SET-SSRs were detected from the *B. lutosa* transcriptome was we used a mixture of *B. lutosa* hepatopancreas and gastropodal muscle samples, which contain abundant transcriptional sequences. These EST-SSR molecular markers had laid an important foundation for large-scale development of *B. lutosa* molecular markers, evaluation of artificial multiplication effect, and population genetic diversity and molecular assisted breeding.

If EST-SSR markers can be directly identified on growth-related genes, it would effectively improve the breeding efficiency of *B. lutosa* with fast-growth traits. In the present study, we found eight growth-related genes that contain EST-SSR sequences in the *B. lutosa* transcriptome. These EST-SSR sequences were likely to participate in the expression and regulation of their representative genes, and affect the growth-related traits of *B. lutosa.*

TCB-1 gene was found in the growth-related genes containing EST-SST sequences. TBC-1 is a kind or nucleic acid protein. Richardson and Zon (1995) first discovered this nucleic acid protein in the mouse labrocyte cDNA library. A total of 29 proteins were found in the TBC-1 family, in which TBC-1-D4 and TBC-l-D1 have been extensive studied. Koumanov and Holman (2007) speculate that TBC-1-D4 and TBC-l-D1 play key roles in cell membrane transport and connection signal transduction. Bryant (2002) report that TBC-1D1 stimulates glucose transporter protein (GLUT4) from a cellular location to the cell surface in muscle and adipocytes, thereby completing signal transduction and substance transport. Hargett et al. (2015) confirm that TBC-1D1 regulates the expression of GLUT4 in skeletal muscle cells through knockout TBC-1D1 gene in mice.

Fibroblast growth factor (FGFR) belongs to tyrosinkinase (TK) receptor family (Kornbluth et al., 1988). Four kinds of FGFRs, i.e. FGFR-1, FGFR-2, FGFR-3, and FGFR-4, have been found. Deng et al. (1994) find that mouse embryos with defective FGFR-1 gene are severely blacked and died during development. Li (2008) shows that the FGFR1 regulates phosphorus metabolism and calcium metabolism in blood and bone through ALK3, and FGFR1 regulates osteoblasts proliferation and apotosis through inhibiting PI3K and promoting MAPK signaling pathway, and inhibits osteoblasts mineralization and proliferation physiologically. Gudernova et al. (2015) find that mutant FGFR3 causes human dwarfism-achondroplasia (ACH) symptoms.

Conclusion

In conclusion, we firstly sequenced the transcriptomes of *B. lutosa*, and detected eight growth-related genes (TBC-1-A, TBC-1-2B, FGFR-1, GFRBP2, GFR-15, FGFR-1-A, SMAD6, and TGFβR-1) that contain EST-SSR sequences in the *B. lutosa* transcriptome. However, which EST-SSR sequences should be used as biomarkers for breeding need to further study.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1802_34713482

Acknowledgments. This research was supported by the National Natural Science Foundation of China [31672640], the Natural Science Foundation of Hunan Province [2017JJ3134], Blue Granary [2018YFD0900200], and the Scientific Research Fund of Hunan Province Education Department [17C0935]. The authors thank anonymous technicians at Guangdong Meilikang Bio-Science Ltd., China for assistance with data analysis.

REFERENCES

[1] Apweiler, R., Bairoch, A., Wu, C. H., Barker, W. C., Boeckmann, B., Ferro, S., Gasteiger, E., Huang, H., Lopez, R., Magrane, M., Martin, M. J., Natale, D. A., O'Donovan, C., Redaschi, N., Yeh, L. S. L. (2004): UniProt: the universal protein knowledgebase. – Nucleic Acids Research 32: D115-D119.

[2] Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., Sherlock, G. (2000): Gene ontology: tool for the unification of biology. – Nature Genetics 25: 25-29.

[3] Bryant, N. J., Govers, R., James, D. E. (2002): Regulated transport of the glucose transporter GLUT4. – Nature Review Molecular Cell Biology 3: 267-277.

[4] Buckingham, M., Bajard, L., Chang, T., Daubas, P., Hadchouel, J., Meihac, S., Montarras, D., Rocancourt, D., Relaix, F. (2003): The formation of skeletal muscle: from somite to limb. – Journal of Anatomy 202: 59-68.

[5] Chen, F., Luo, X., Shen, M., Huang, M., Ke, C. (2011): Karyotypes of two species of the genus *Babylonia*. – Journal of Xiamen University (Natural Science) 50: 789.

[6] Davis, T. A., Fiorotto, M. L., Burrin, D. G., Vann, R. C., Reeds, P. J., Nguyen, H. V., Beckett, P. R., Bush, J. A. (2002): Acute IGF-I infusion stimulates protein synthesis in skeletal muscle and other tissues of neonatal pigs. – American Journal of Physiology-Endocrinology and Metabolism 283: E638-E647.

[7] Deng, C. X., Wynshaw-Boris, A., Shen, M. M., Daugherty, C., Ornitz, D. M., Leder, P. (1994): Murine FGFR-1 is required for early postimplantation growth and axial organization. – Genes & Development 8: 3045-3057.

[8] Deng, Y. Y., Li, J. Q., Wu, S. F., Zhu, Y. P., He, F. C. (2006): Integrated nr database in protein annotation system and its localization. – Computer Engineering 32: 71-74.

[9] Doumit, M. E., Cook, D. R., Merkel, R. A. (1993): Fibroblast growth factor, epidermal growth factor, insulin-like growth factors, and platelet-derived growth factor-BB stimulate proliferation of clonally derived porcine myogenic satellite cells. – Journal of Cellular Physiology 157: 326-332.

[10] Feng, L., Hu, L., Fu, X., Liao, H., Li, X., Zhan, A., Zhang, L., Wang, S., Huang, X., Bao, Z. (2014): An integrated genetic and cytogenetic map for Zhikong scallop, Chlamys farreri, based on microsatellite markers. – PLoS One 9: e92567.

[11] Fishery and Fishery Administration of the Ministry of Agriculture and Rural Areas of China, National Aquatic Technology Promotion Terminal of China, Chinese Society of Aquaculture. (2019): China Fisheries Statistics Yearbook. – China Agriculture Press, Beijing.

[12] Florini, J. R., Ewton, D. Z., Roof, S. L. (1991): Insulin-like growth factor-I stimulates terminal myogenic differentiation by induction of myogenin gene expression. – Molecular Endocrinology 5: 718-724.

[13] Gudernova, I., Vesela, I., Balek, L., Buchtova, M., Dosedelova, H., Kunova, M., Pivnicka, J., Jelinkova, I., Roubalova, L., Kozubik, A., Krejci, P. (2015): Multikinase activity of fibroblast growth factor receptor (FGFR) inhibitors SU5402, PD173074, AZD1480, AZD4547 and BGJ398 compromises the use of small chemicals targeting FGFR catalytic activity for therapy of short-stature syndromes. – Human Molecular Genetics 25: 9-23.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1802_34713482

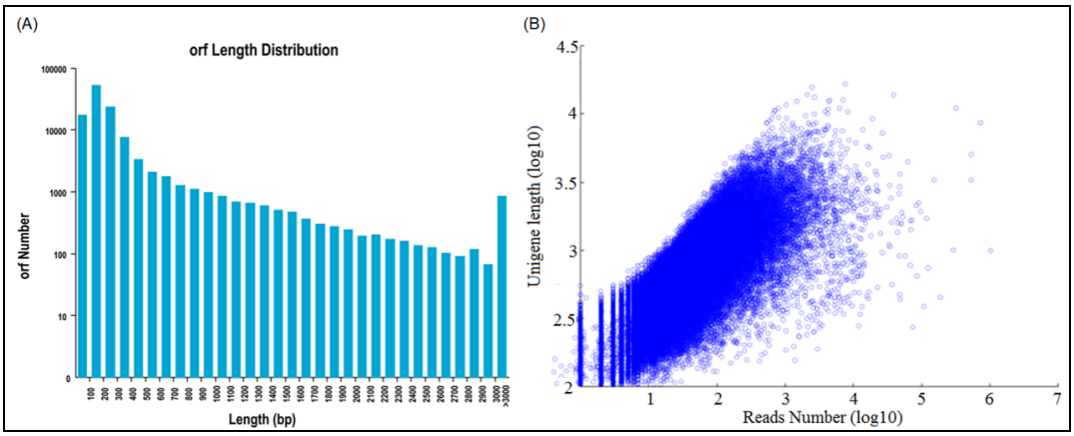
[14] Hargett, S. R., Walker, N. N., Hussain, S. S., Hoehn, K. L., Keller, S. R. (2015): Deletion of the Rab GAP Tbc1d1 modifies glucose, lipid, and energy homeostasis in mice. – American Journal of Physiology Endocrinology and Metabolism 309(3): E233-E245.
[15] Hembree, J. R., Hathaway, M. R., Dayton, W. R. (1991): Isolation and culture of fetal porcine myogenic cells and the effect of insulin, IGF-I, and sera on protein turnover in porcine myotube cultures. – Journal of Animal Science 69: 3241-3250.
[16] Hou, R., Bao, Z., Wang, S., Su, H., Li, Y., Du, H., Hu, J., Wang, S., Hu, X. (2011): Transcriptome sequencing and de novo analysis for Yesso scallop (Patinopecten yessoensis) using 454 GS FLX. – PLoS One 6: e21560.
[17] Huang, B. X., Peakall, R., Hanna, R. J. (2000): Analysis of genetic structure of blacklip abalone (Haliotis rubra) populations using RAPD, minisatellite and microsatellite markers. – Marine Biology 136: 207-216.
[18] Huang, R., Huang, B., Tang, W., Chen, Z. (2010): Morphological observation of the early developmental stages of *Babylonia areolata*. – Journal of Oceanography in Taiwan Strait 29: 380-388.
[19] Kanehisa, M., Goto, S., Kawashima, S., Okuno, Y., Hattori, M. (2004): The KEGG resource for deciphering the genome. – Nucleic Acids Res 32: D277-D280.
[20] Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., Katayama, T., Kawashima, S., Okuda, S., Tokimatsu, T., Yamanishi, Y. (2008): KEGG for linking genomes to life and the environment. – Nucleic Acids Res 36: D480-D484.
[21] Kornbluth, S., Paulson, K. E., Hanafusa, H. (1988): Novel tyrosine kinase identified by phosphotyrosine antibody screening of cDNA libraries. – Molecular and Cellular Biology 8: 5541-5544.
[22] Koumanov, F., Holman, G. D. (2007): Thrifty Tbc1d1 and Tbc1d4 proteins link signalling and membrane trafficking pathways. – Biochemical Journal 403: e9-e11.
[23] Lamberson, W. R., Sterle, J. A., Matteri, R. L. (1996): Relationships of serum insulin-like growth factor I concentrations to growth, composition, and reproductive traits of swine. – Journal of Animal Science 74: 1753-1756.
[24] Le Roith, D., Scavo, L., Butler, A. (2001): What is the role of circulating IGF-I? – Trends in Endocrinology and Metabolism 12: 48-52.
[25] Li, F. (2008): Role of FGFR1 signaling in bone development and remodeling in mice. – Chongqing: Third Military Medical University.
[26] Li, Y., Zhong, Y., Lu, L., Chen, C., Duan, C. (2011): Research progress of insulin-like growth factor binding protein IGFBP-3. – Periodical of Ocean University of China 41: 41-47.
[27] Lin, G. (2012): Acute toxicity tests of four heavy metal salts to juvenile snail of *Babylonia lutosa*. – Fujian Journal of Agricultural Sciences 27: 232-236.
[28] Lu, L., Ma, A., Wang, X., Yue, L., Meng, X., Zhai, J., Tan, L., Hou, S. (2013): Polymorphisms analysis of five populations of Takifugu rubripes with microsatellite. – Progress in Fishery Sciences 34: 27-33.
[29] Masoudi-Nejad, A., Goto, S., Jauregui, R., Ito, M., Kawashima, S., Endo, T. R., Moriya, Y. (2007): EGENES: transcriptome-based plant database of genes with metabolic pathway information and expressed sequence tag indices in KEGG. – Plant Physiol 144: 857-866.
[30] Moriyama, S., Ayson, F. G., Kawauchi, H. (2000): Growth regulation by insulin-like growth factor-I in fish. – Bioscience, Biotechnology and Biochemistry 64: 1553-1562.
[31] Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., Wold, B. (2008): Mapping and quantifying mammalian transcriptomes by RNA-Seq. – Nature Methods 5: 621-628.
[32] Nie, Z. (2014): Transcriptome analysis and study on genetic diversity of different geographic populations in Megalobrama terminalis. – Wuhan: Huazhong Agricultural University.
[33] Pan, Y., Li, B., Li, G. (2010): Morphological comparison on radula from different populations of *Hemifusus*. – Journal of Tropical Oceanography 29: 125-129.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1802_34713482

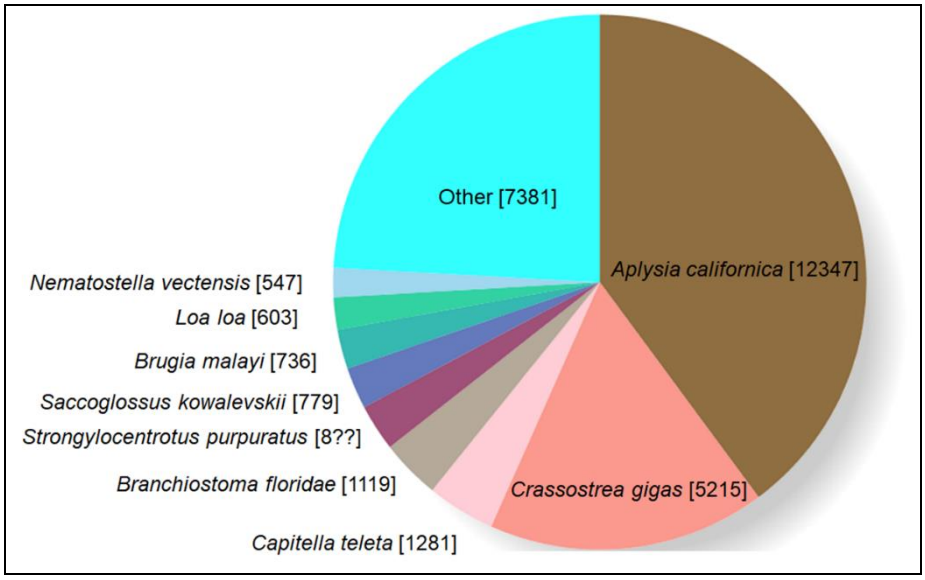
[34] Qi, Z., Wang, Y., Chen, D., Li, Y. (2011): Effect of morphometric traits on body weight in clam *Nerita yoldi*. – Fisheries Science 30: 505-508.
[35] Qin, Q., Wang, X., Zeng, Z., Zhu, D., Xiong, G., Wu, Q., Ning, Y. (2014): Genetic distance in four populations of *Babylonia lutosa* (Lamer) assessed by AFLP makers. – Journal of Human Agricultural University (Natural Sciences) 40: 299-304.
[36] Reinecke, M., Schmid, A., Ermatinger, R., Loffing-Cueni, D. (1997): Insulin-like growth factor I in the teleost Oreochromis mossambicus, the tilapia: gene sequence, tissue expression, and cellular localization. – Endocrinology 138: 3613-3619.
[37] Richardson, P. M., Zon, L. I. (1995): Molecular cloning of a cDNA with a novel domain present in the tre-2 oncogene and the yeast cell cycle regulators BUB2 and cdc16. – Oncogene 11: 1139-1148.
[38] Ruan, R., Yu, J., Li, H., Li, J., Tang, Y. (2011): Effects of IGF-1 genotype on weight gain and body shape in GIFT strain *Oreochromis niloticus*. – Journal of Fishery Sciences of China 18: 682-688.
[39] Tatusov, R. L., Galperin, M. Y., Natale, D. A., Koonin, E. V. (2000): The COG database: a tool for genome-scale analysis of protein functions and evolution. – Nucleic Acids Research 28: 33-36.
[40] Tong, X., Gong, S., Yu, D., Huang, G., Du, B., Li, S. (2009): Genetic diversity of cultured pacific white shrimp (Litopenaeus vannamei) stocks of different generations in China. – Oceanologia et Limnologia Sinica 40: 214-220.
[41] Upton, Z., Yandell, C. A., Degger, B. G., Chan, S. J., Moriyama, S., Francis, G. L., Ballard, F. J. (1998): Evolution of insulin-like growth factor-I (IGF-I) action: in vitro characterization of vertebrate IGF-I proteins. – Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology 121: 35-41.
[42] Wang, W., Cai, L., Liu, W. (2007): Morphological classification of nassariids in Fujian coast. – Journal of Xiamen University (Natural Science) 46: 171-175.
[43] Xiong, G., Ma, X., Wang, X. Q., Zeng, Z. N., Zhu, D. L., Kang, L., Wu, Q. S. (2015): Development and characterization of 18 microsatellite markers for the *Babylonia lutosa*. – Conservation Genetics Resources 7: 471-472.
[44] Xiong, G., Wang, X. Q., Wang, P., Chen, Z. N., Zhou, X. W., Kang, L., Zeng, Z. N. (2019): Development and genetic diversity analysis of Babylonia lutosa with EST-SSR markers. – Progress in Fishery Sciences 40: 1-8.
[45] Yang, Z., Zheng, Y., Li, Z., Zheng, Y. (2008): The effect of KCl on induced metamorphosis of planktonic larvae of *Babylonia aerolato*. – Marine Sciences 32: 6-9.
[46] Yano, K., Bauchat, J. R., Liimatta, M. B., Clemmons, D. R., Duan, C. (1999): Down-regulation of protein kinase C inhibits insulin-like growth factor I-induced vascular smooth muscle cell proliferation, migration, and gene expression. – Endocrinology 140: 4622-4632.
[47] Yin, S., Liao, J., Huang, H., Gao, L. (2007): Research advancement on biology and cultural ecology in *Babylonia* sp. – Fisheries Science 26: 632-636.
[48] Zhang, T., Liu, P., Li, J., Kong, J., Wang, Q. (2005): Genetic diversity of cultured populations of Fenneropenaeus chinensis shrimp using microsatellites. – Journal of Fisheries of China 29: 6-12.
[49] Zhang, J., Shi, Z., Fu, Y., Cheng, Q. (2011): Gene expression and thyroid hormone regulated transcript of IGF-I during metamorphosis of the flounder, Paralichthys olivaceus. – Acta Hydrobiologica Sinica 35: 355-359.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1802_34713482

APPENDIX



Appendix 1. *Length distribution of ORF in B. lutosa transcriptome (A) and the correlation between reads number and unigene length (B)*



Appendix 2. *Distribution of species that were matched by unigenes in B. lutosa transcriptome*

Appendix 3. *Size distribution of Unigenes used for functional annotation of logarithmic transcriptome*

Annotated databases	All sequence	≥300 bp	≥1000 bp
NR	31,076	15,327	10,953
NT	17,033	7,483	7,093
SwissProt	18,413	7,892	8,409
TrEMBL	30,485	15,043	10,772
GO	14,106	6,025	6,300
KEGG	8,390	3,536	3,960
COG	7,191	2,852	3,685
Total	35,931	18,222	11,564

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1802_34713482