Peer

Molecular characterization of the giant freshwater prawn (*Macrobrachium rosenbergii*) *beta-actin* gene promoter

Junjun Yan^{1,2}, Qiang Gao³, Zongbin Cui¹, Guoliang Yang⁴ and Yong Long¹

¹ State Key Laboratory of Freshwater Ecology and Biotechnology, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan, China

² University of the Chinese Academy of Sciences, Beijing, China

³ Zhejiang Institute of Freshwater Fisheries, Huzhou, China

⁴ Huzhou University, Huzhou, China

ABSTRACT

Constitutive promoters are important tools for gene function studies and transgenesis. The *Beta-actin* (*actb1*) gene promoter has been isolated from many species but remains to be cloned from the giant freshwater prawn (*Macrobrachium rosenbergii*). In this study, we cloned and characterized the *Mractb1* gene promoter. Two alternative promoters were identified for the *Mractb1* gene, which direct the generation of two transcripts with different 5' untranslated regions. Three CpG islands were predicted in the upstream sequence, which are intimately related to transcription initiation and promoter activity. In addition to the CCAAT-box and the CArG-box, molecular dissection of the flanking sequence revealed the existence of one negative and two positive elements in the upstream region and the first intron. Finally, the *Mractb1* promoter. Our investigations provide a valuable genetic tool for gene function studies and shed light on the regulation of the *Mractb1* gene.

Subjects Aquaculture, Fisheries and Fish Science, Genetics, Molecular Biology **Keywords** Giant freshwater prawn, Beta-actin gene, Promoter, Transcriptional activity

INTRODUCTION

Giant freshwater prawn is the biggest freshwater prawn in the world. The males can grow up to 320 mm in body length and weigh over 200 g (*Ling*, 1969). It has a wide distribution throughout the world and is particularly favored for farming in the tropical and subtropical developing countries in Asia, such as China, India and Thailand. Owing to its fast growth rate and higher profit realization per unit area, it has now become the most widely aquacultured freshwater prawn species and contributed 51.7% to the total global production of freshwater prawns (*Banu & Christianus*, 2016; *New*, 2005; *New & Nair*, 2012). However, the outbreak of diseases is a major limitation facing the aquaculture industry, and several diseases such as white spot disease and white tail disease have been reported to cause mass mortality for giant freshwater prawn farming worldwide (*Bonami & Widada*, 2011; *Saurabh & Sahoo*, 2008).

Submitted 9 July 2018 Accepted 6 September 2018 Published 23 October 2018

Corresponding authors Guoliang Yang, ygl0572@163.com, 02462@zjhu.edu.cn Yong Long, longyong@ihb.ac.cn

Academic editor Vladimir Uversky

Additional Information and Declarations can be found on page 14

DOI 10.7717/peerj.5701

Copyright 2018 Yan et al.

Distributed under Creative Commons CC-BY 4.0

OPEN ACCESS

Transgenic technology has been considered as a powerful and effective way to improve economic traits of aquaculture species (*Bennett & Jennings, 2013; Rasmussen & Morrissey, 2007; Sin, 1997*), through which pathogen-resistant strains of prawn could be generated to reduce or even avoid the disease-related economic losses. So far, successful gene transfer in freshwater prawn and several marine shrimp species have been reported (*Arenal et al., 2000; Liu et al., 2001; Preston et al., 2000*). However, most of the researches used expression vectors based on heterologous promoters such as the CMV (cytomegalovirus) and SV40 (simian vacuolating virus 40) promoter, the hobo transposable element (HFL1) and the carp *beta-actin* gene promoter. Studies about fish transgenesis have shown that promoters of fish origin were more efficient in driving expression of the target genes than those with mammalian or viral origins (*Alam et al., 1996; Hanley et al., 1998*). Moreover, negative concerns about the use of viral promoters to express transgenes have led to the generation of "all-fish" transgene constructs (*Rasmussen & Morrissey, 2007*). In accordance with the "all-fish" concept, promoters originated from prawn species have strong potential for use in prawn transgenesis and gene function studies.

Beta-actin is a housekeeping gene with ubiquitous expression. The promoter of *beta-actin* gene has been isolated from and characterized in a wide variety of species such as human (*Gunning et al., 1987*), chicken (*Kosuga et al., 2000*), amphioxus (*Feng et al., 2014*), teleosts (*Barman et al., 2015*; *Hwang et al., 2003*; *Kong et al., 2014*; *Liu et al., 1991*; *Noh et al., 2003*) and shrimps (*Shi et al., 2015*; *Shi et al., 2016*). Because the *beta-actin* gene promoter confers a high level of constitutive transcriptional expression, it is widely used for transgenesis in both plants and animals (*Cho et al., 2011*; *Hong et al., 2016*; *Kosuga et al., 2000*). Although the cDNA of the *beta-actin* gene has been cloned from giant freshwater prawn (*Zhu et al., 2005*), its promoter remains to be characterized. Here, we report the isolation and characterization of the giant freshwater prawn beta-actin (Mractb1) gene promoter (Mbap). Our results provide insights into the regulation mechanisms of the *Mractb1* gene and indicate the potential usage of Mbap in gene function and transgenesis studies for giant freshwater prawn and closely related species.

MATERIALS AND METHODS

Animals

The animal protocol of this study was approved by the Institutional Animal Care and Use Committee of Institute of Hydrobiology (Approval ID: Y341131501). Adult giant freshwater prawns of both sexes (body weight 20–30 g) were acquired from Zhejiang Institute of Freshwater Fisheries, Huzhou, China. After transportation, the prawns were acclimated to laboratory conditions for 2 weeks in a circulating fresh water system at 28 °C.

Total RNA and genomic DNA extraction

The prawns were immersed in eugenol solution (0.125 mL/L in phosphate buffered saline) until loss of consciousness and then placed into ice-slurry for a 2-Step euthanasia procedure (*Leary et al., 2013*). After that their bodies were dissected, and total RNA was extracted from the gill, liver, muscle and intestine of both sexes, and from testis and ovary, using TRIZOL (Invitrogen) according to the manufacturer's instructions. For genomic DNA

extraction, about 100 mg muscle tissue was put into a 1.5 mL Eppendorf tube with 0.5 mL lysis buffer (10 mM Tris-Cl, 1 mM EDTA, 0.1% SDS, 200 μ g/mL proteinase K, pH 8.0). The samples were incubated in 58 °C water bath for 6 h with gentle vibration. After that, the lysate was subjected to a traditional phenol/chloroform extraction procedure. The concentration of RNA and DNA samples was measured by a NanoDrop 8000 from Thermo Scientific and the quality of RNA and DNA samples was assessed by agarose gel electrophoresis.

Genome walking

Genome walking was performed as previously described (*Zhou et al.*, 2015) to clone the 5' flanking sequence of the *Mractb1* gene. Each walking step contains three rounds of nested PCR using different gene specific primers (GSP). Before genome walking, primers Mractb1-F and Mractb1-R (sequences of all the primers used in this study are listed in Table 1) designed according to the cDNA sequence deposited in GenBank (AY626840) were used to amplify partial genomic sequence. A 1,293-bp genomic DNA fragment was obtained and sequenced. The gene specific primers from Mractb1-GSP1 to Mractb1-GSP3 were subsequently designed according to this known genomic sequence. These gene specific primers were sequentially mated with the four degenerate primers included in the genome walking kit (Takara, Kusatsu, Japan) to amplify the unknown 5' flanking sequence. Similarly, gene specific primers from Mractb1-GSP4 to Mractb1-GSP6 based on the product of the first walking experiment were used for the second walking step. A total of three walking experiments were conducted to identify the 5' flanking sequence of the *Mractb1* gene.

5' RACE

5' RACE (rapid amplification of cDNA ends) was performed using the Takara 5'-Full RACE Kit according to the manufacturer's instructions. Briefly, the RNA sample was first treated with CIAP (calf intestine alkaline phosphatase) to remove the naked phosphorous from the RNA molecules, followed by TAP (tobacco acid pyrophosphatase) treatment for mRNA decapitation. The decapped mRNA was ligated to the 5' RACE adaptor using T4 RNA ligase. Reverse transcription was performed using M-MLV (moloney murine leukemia virus) reverse transcriptase and 9-mer random primer. After that, two rounds of nest PCR using 5'-RACE- outer-primer/GSP1 and 5'-RACE-inner-primer/GSP2 primer pairs were conducted sequentially to clone the 5' cDNA ends. Product of the second round PCR was purified using the Biospin Gel Extraction Kit (BioFlux), subcloned into the pMD18-T vector (TaKaRa) and then sequenced by the Tsingke Biological Technology (Wuhan, China) Co., Ltd.

Quantitative real-time PCR

Quantitative real-time PCR (qPCR) was performed as previously described (*Long et al.*, 2013) to determine the expression level of the *Mractb1* transcripts. Primers for qPCR were designed using the Primer Premier 6.0 software. First-strand cDNA was synthesized using total RNA extracted from different tissues by the RevertAid First Strand cDNA Synthesis

Table 1 Sequences of primers.		
Primer name	Sequence (5'-3')	Note
Mractb1-F	TTCCCATCCATTGTCGGCAG	cDNA amplification
Mractb1-R	GCATTCTGTCAGCGATTCCTGG	cDNA amplification
GSP1	TGACCCATACCAACCATCACAC	5'-RACE
GSP2	ACCGGAGCCATTGTCTACAACCAAC	5'-RACE
Exon1-F	ACTCGCTCTTCGACATC	cDNA cloning
Exon1'-F	AGTCTGTCACTTGCTCC	cDNA cloning
Exon3-R	TGTCTGTAGAAATGAATTTATTC	cDNA cloning
Mractb1-SP1	GTACTGATATGAAGGCGTGTTCAG	Genome walking
Mractb1-SP2	GATAATGATGGTAAGGCAAACATTG	Genome walking
Mractb1-SP3	TACACCTGGAGTGTCTAAGCAG	Genome walking
Mractb1-SP4	CTACTCCTGAAGATGTCGAAGAGCGAGTG	Genome walking
Mractb1-SP5	TAACATCTGAAATGAAAGCGGACGAAACTG	Genome walking
Mractb1-SP6	CAAACGTCTTGCCTTATATGGACATGGAG	Genome walking
Mractb1-SP7	GTAGAAAGACCGGGATTTCTTTCGGT	Genome walking
Mractb1-SP8	ACTGGGCGTAACTACTATGCCTCTAA	Genome walking
Mractb1-SP9	GTTGAAGGGAAATGTACTGAGAACA	Genome walking
QMractb1-F1	TCAGGAGTAGCACGTACAC	qPCR
QMractb1-F1'	ATCACTGGTGCTCGTTG	qPCR
QMractb1-R0	ACAATGGATGGGAACAC	qPCR
QMr18S-rRNA-F	TAGTTGGAGGTCAGTTCC	qPCR
QMr18S-rRNA-R	ATTCCAGAGTAGCCTGC	qPCR
Mractb1P1-F	ACT <u>GAGCTC</u> TCCCGAAGTGATCACTG	Promoter analysis
Mractb1P1-R	AGACT <u>GCTAGC</u> TTTGTATTAGCTGCAAGAGAAAG	Promoter analysis
Mractb1P2-F	ACT <u>GAGCTC</u> TCATTTAGTAAGTAGGAGAG	Promoter analysis
Mractb1P2-R	AGACT <u>GCTAGC</u> AGCAAGTGACAGACTGAAC	Promoter analysis
Mractb1P3-F	ACT <u>GAGCTC</u> AGGTAAGTACACGTTGGC	Promoter analysis
Mractb1P3-R	AGACT <u>GCTAGC</u> TGAGATAAATTATGGAAC	Promoter analysis
Mractb1P4-F	ACT <u>GAGCTC</u> TTTGCCCTTCCGCGAAATTTAC	Promoter analysis
Mractb1P4-R	AGACT <u>GCTAGC</u> GGTGAGAGTGTACGTGCTAC	Promoter analysis
Mractb1P5-F	ACT <u>GAGCTC</u> ACATTATGGAAACATTTC	Promoter analysis
Mractb1P5-R	AGACT <u>GCTAGC</u> ATTCTTTGCATGTGACGAG	Promoter analysis
Mractb1P6-F	ACT <u>GAGCTC</u> ATGATGAGGGTCACGCGTTAAG	Promoter analysis
Mractb1P6-R	AGACT <u>GCTAGC</u> TTCGGCGAACTGGGCGT	Promoter analysis
Mractb1P7-F	ACT <u>GAGCTC</u> TGCGCTGCTTTTACCAAATAC	Promoter analysis
Mractb1P7-R	AGACT <u>GCTAGC</u> TCATCATCTTTACCATC	Promoter analysis
Mractb1P42-F	TACACTCTCACCTCATTTAGTAAG	Promoter analysis
Mractb1P24-R	CTTACTAAATGAGGTGAGAGTGTA	Promoter analysis
Carpactb1-F1	ACG <u>AAGCTT</u> TCAAACTGTGGCACCATC	Promoter analysis
Carpactb1-R1	ACG <u>CCATGG</u> CTGAACTGTAAATGAATG	Promoter analysis

Kit (Thermo scientific) and used as template for qPCR analysis. *Mr18S-rRNA* was used as the internal reference for gene expression normalization.

Bioinformatic analyses

CpG islands in the upstream region were predicted by the Methprimer software using default parameters (http://www.urogene.org). Core promoter elements were predicted by the YAPP Eukaryotic Core Promoter Predictor (http://www.bioinformatics.org/yapp/cgi-bin/yapp.cgi).

Construction of plasmids

To dissect transcriptional activity of the 5' flanking sequence of the *Mractb1* gene, the cloned upstream sequences were divided into seven fragments according to location of the predicted GC islands and the TSSs (transcriptional start sites) revealed by 5' RACE. The position of the first nucleotide of the initiation codon was regarded as +1. The length and relative location of fragments 1 to 7 are 558 bp (-558 to -1), 523 bp (-1,080 to -558), 240 bp (-1,317 to -1,078), 249 bp (-1,568 to -1,320), 195 bp (-1,763 to -1,569), 357 bp (-2,120 to -1,764) and 501 bp (-2,620 to -2,120), respectively. These fragments were amplified and cloned into the SacI/NheI site of the pGL3- Basic vector (Promega, Madison, WI, USA) separately or in combination. Deletion of fragment 3 in the 5' flanking sequences was conducted via the PCR-driven overlap extension method as previously described (Heckman & Pease, 2007) with primers Mractb1P42-F and Mractb1P24-R. The constructs were named according to the inserted fragments, for example, pGL-Mba7654321 includes the whole sequence, pGL-Mba1 contains only fragment 1 and pGL-Mba21 consists of both fragment 1 and fragment 2. To compare the promoter activity of *Mractb1* gene with that of carp *actb1* gene, the promoter of carp *actb1* gene was amplified by the primer pair Carpactb1-F1/Carpactb1-R1 and inserted into the HindIII/NcoI site of the pGL3-Basic vector.

Luciferase activity assay

The EPC (endothelial progenitor cells) cells (ATCC[®] CRL-2872TM) were maintained at 28 °C with 5% CO₂ in M199 medium (Hyclone, Logan, UT, USA) supplemented with 10% fetal bovine serum, 100 μ g/mL streptomycin, 100 U/mL penicillin and 2 μ g/mL amphotericin B. One day before transfection, cells were seeded into 24-well plates at a density of 2 × 10⁵ cells/well. Transfection was performed using the X-tremeGENE HP Transfection Reagent (Roche). The pRL-TK plasmid from Promega was used as internal control for luciferase activity assays and co-transfected with the promoter-luciferase constructs. The total amount of plasmid used for transfection was 500 ng/well and the ratio between the promoter-luciferase vector and pRL-TK was 5:1. Twenty-four hours after transfection, cells were lysed and subjected to luciferase activity assays using the Dual-Luciferase[®] Reporter Assay System from Promega.

Western blot

EPC cells cultured in six-well plates were transfected with 2 µg plasmid. Twenty-four hours after transfection, cells were lysed by the RIPA (radio immunoprecipitation assay) lysis



Figure 1 Cloning of the 5'-flanking sequence of the *Mractb1* gene. The 5'-flanking sequence of *Mractb1* was cloned by three rounds of genome walking based on asymmetric interlaced PCR. The dashed line represents the assembled 5' flanking sequence. The rectangular blocks and the dotted arrows indicate the cloned sequences and the primers used for each round of genome walking respectively. The electrophore-tograms were displayed accordingly. The numbers above the electrophoretograms indicate different round of PCR for each walking step.

Full-size 🖾 DOI: 10.7717/peerj.5701/fig-1

buffer from beyotime. Then western blot was performed as previously described (*Mo et al., 2010*) and GAPDH (glyceraldehyde-3-phosphate dehydrogenase) was used as loading control. The primary antibodies used for western blot were rabbit anti-Firefly Luciferase (Abcam, #ab185923, 1:5000) and mouse anti-GAPDH (Boster, #BM3876, 1:100). The secondary antibodies were goat anti- rabbit IgG (immunoglobulin) (Boster, #BM3894, 1:5000) and goat anti-mouse IgG (#BM3895, 1:5000; Bosterbio, Pleasanton, CA, USA), respectively. Intensity of the bands in the western blot image was analyzed by the ImageJ software (http://rsb.info.nih.gov/ij/).

Statistical analysis

The results of dual luciferase assays and qPCR are presented as mean \pm standard deviation. Independent samples *t*-tests were performed using SPSS 15.0 (Chicago, IL, USA) to analyze the significant difference between groups.

RESULTS

Cloning of the 5' flanking sequence for the Mractb1 gene

To clone the 5' flanking sequence for the *Mractb1* gene, a 1,293 bp genome sequence was firstly amplified using the primers Mractb1-F/ Mractb1-R designed according to the reported cDNA sequence (*Zhu et al., 2005*). Three rounds of genome walking experiments were subsequently performed to clone the upstream sequence of the *Mractb1* gene. As shown in (Fig. 1), the length of the DNA fragments obtained by the three genome walking



experiments were 2,094, 531 and 713 bp, respectively. Furthermore, PCR using the forward primer Mractb1-F and the reverse primer Exon3-R (matches to the 3' cDNA end) generated a 1,593 bp sequence. Assembling of these fragments resulted in a 4,306-bp long contig, which contains both the 5' flanking and full sequence for the *Mractb1* gene. The obtained upstream sequence counted from the first nucleotide of the initiation codon is 2,620 bp in length, which has been submitted to the GenBank database under the accession number KY038927.

Identification of two transcripts for the Mractb1 gene

Before characterizing the promoter of the *Mractb1* gene, 5'-RACE was performed to identify the transcriptional start site (TSS). A clear DNA band in the range of 100 to 250 bp was obtained (Fig. 2). The DNA fragments were purified and cloned into the pMD18-T vector. Subsequent DNA sequencing revealed two 5' cDNA end sequences for the *Mractb1* gene. One is 106 bp (10 clones sequenced) and the other is 145 bp (6 clones sequenced) in length (Fig. 2). Further analysis revealed that the *Mractb1* gene possesses 2 transcript variants (designated Mractb1_tv1 and Mractb1_tv2, respectively), which only differ in the 5' UTR (untranslated region sequence) (Fig. 3A). PCR using different forward primers (Exon1-F and Exon1'-F) and the same reverse primer (Exon3-R) amplified the full-length sequence for the 2 transcripts (Figs. 3A, 3B), which are 1,186 and 1,225 bp in length, respectively. Since the transcripts contain different exon 1 (Fig. 3A), we speculate that they are driven by alternative promoters (promoter-1 and promoter-2).



Figure 3 Transcripts of the *Mractb1* gene. (A) Structure of the *Mractb1* gene and its two transcript variants with different initial exon. The numbers and arrows indicate the relative position of the exons and coding sequence (CDS). The first base of the start codon was defined as position +1. The numbers in the rectangles indicate different exons of the Mractb1 gene. The alternative first exons are designated as exon 1 and exon 1'. (B) Electrophoretogram for the two transcripts. Primers used to amplify the two transcripts are Exon1-F, Exon1'-F and Exon3-R as displayed in (A). M: DNA marker.

Full-size DOI: 10.7717/peerj.5701/fig-3

Expression of the Mractb1 transcripts in different sexes and tissues

To shed light on the activities of the promoters in different sexes and tissues, the expression levels of the two transcript variants were measured by qPCR assays. As shown in Fig. 4, the *Mractb1* gene demonstrate quite different expression levels among tissues and between sexes. The highest overall expression level of the *Mractb1* gene was found in the gonad and gill, followed by the intestine, muscle and liver. Sexual dimorphism could be found for the expression level of the *Mractb1* gene in all the tissues (Fig. 4). The expression levels of both Mractb1_tv1 and Mractb1_tv2 in the female gonad (ovary) and gill is significantly higher than those in the corresponding male tissues, while muscle and intestine of the males demonstrated higher expression than those of the females (Fig. 4). A difference between the expression of the two transcripts in the same tissue was identified in the male gill and the female gonad and muscle. Mractb1_tv2 showed a higher expression in the male gill, but a lower expression in the female muscle and gonad when compared to Mractb1_tv1 (Fig. 4). These findings indicate that expression of the *Mractb1* gene is regulated by tissue and sex specific factors.

Molecular dissection of the Mractb1 gene promoter

Three CpG islands were predicted in the 5' flanking sequence of the *Mractb1* gene. These CpG islands are located at -1,762 to -1,569 (i1, 194 bp), -1,437 to -1,320 (i2, 118 bp) and -747 to -542 (i3, 206 bp), respectively (Fig. 5A). The 2,620 bp flanking sequence was divided into 7 segments (from Mba7 to Mba1) according to the locations of the CpG

Peer



islands and the TSSs (Fig. 5A). Transcriptional activities of the upstream fragments were investigated by dual luciferase assays. When characterized separately, Mba7, Mba6, Mba3 and Mba1 demonstrated no or minimal transcriptional activity, while high promoter activity was detected for Mba5, Mba4 and Mba2, each of them contains a predicted CpG island (Fig. 5B), suggesting importance of the CpG islands in transcription. Based on structure of the Mractb1 transcripts and transcriptional activities of the upstream fragments, we concluded that the two alternative promoters are in Mba54 (promoter-1) and Mba21 (promoter-2), respectively. Activity of the two promoters and regulations by their corresponding upstream regions were further characterized. For promoter-1, combination of Mba5 and Mba4 only slightly increased the promoter activity, Mba6 displayed no effect on Mba54, while fusing Mba7 to Mba654 significantly enhanced the promoter activity (Fig. 5C). As for promoter-2, significant synergistic effect was found between Mba2 and Mba1; however, Mba3 demonstrated negative effect on the activity of Mba21 (Fig. 5D). Serial deletion experiments indicated that Mba654321 had the highest transcriptional activity, followed by Mba7654321 and Mba54321, while Mba4321, Mba321 and Mba21 only displayed basal promoter activity (Fig. 5E). Furthermore, interaction between the two alternative promoters was observed. Mba54 (promoter-1) and Mba21 (promoter-2)



Figure 5 Molecular dissection of the *Mractb1* gene promoter. (A) Schematic diagram of the DNA fragments subjected to promoter activity assays. The 2,620 bp 5'-flanking sequence was divided into 7 fragments (rectangles shaded with different color) according to the location of the TSSs and CpG islands (i1: -1,762/-1,569, i2: -1,437/-1,320, i3: -747/-542) predicted by the Methprimer software (http://www.urogene.org/). The arrowheads indicate TSSs. The dashed rectangles represent first exons of the two transcripts. (B) Promoter activity of the flanking regions characterized individually. (C) Characterization of promoter-1. (D) Characterization of promoter-2. (E) Deletion analysis of the whole sequence. (F) Interaction between the two alternative promoters. The DNA fragments were cloned into the pGL3-basic vector to generate promoter activity analyzing plasmids. The constructs were transfected into EPC cells and the transcriptional activities were analyzed by dual luciferase assays. The relative luciferase activity represents the ratio between the firefly luciferase activity and the corresponding renilla luciferase activity. The bars indicate mean \pm SD (n = 3).

Full-size DOI: 10.7717/peerj.5701/fig-5

demonstrated significant synergistic effects when combined (Mba5421). Both Mba3 and Mba6 can enhance activity of Mba5421 and the interaction between Mba3 and Mba6 is necessary for the full activity of the *Mractb1* gene promoter (Fig. 5F). Taken together, the upstream fragments demonstrate different effects under different circumstances. For example, Mba7 enhances activity of the promoter-1 but inhibits activity of the whole promoter (Figs. 5C, 5E). On the contrary, Mba3 inhibits activity of the promoter-2 but is indispensable for the full activity of the whole promoter (Figs. 5D, 5F). These findings

```
-1766 acattatgga aacatttcaa ccgaaagaaa tcccggtctt tctactattc
-1716 ccgcctcgca ttcaacgtcc tgctc<u>gacca atcag</u>cggcc gagatcgtct
                                     CCAAT-box
                             acgtttg
-1666 ccatgtccat ataaggcaag
                                     tet gggcggggee gaegaggeea
              CArG-box
                                          GC-box
-1616 ttgcggcaaa
                ataatagggg
                             agttcactcg
                                         tcacatgcaa agaattttgc
-1566 ccttccgcga
                aatttacgcc
                             gcagtttcgt
                                         ccgctttcat ttcagatgtt
-1516 aatcgcgagg
                tttgtgtcaa
                             cgggtcatcg
                                         tgaaatattc ctttaatata
-1466 cgggaaactc
                ggtgaaagat
                             aaagtg<mark>atat</mark>
                                         attgccctcc ccccacctgg
-1416 aaagtcagag agacggtcgg tgtttggagt
                                         ggcgcgctgc tccttct<u>ctc</u>
-1366 ACTCGCTCTT CGACATCTTC AGGAGTAGCA CGTACACTCT CACCTGAGgt
    InR
-1316 aagtacacgt
                tggctgttac
                             caagtatctt cgtcaatgtg
                                                      tagttttatt
-1266 ggaagattta
                tgatgtggtg
                             ttgtgtagag
                                         gtatggtttt
                                                      cgtttgggat
-1216 attggttagt
                ttttggggt
                                         ttcaggctta
                                                      acttgattag
                             tatggaaggc
-1166 aattgcagtg
                tgtatgagtt
                             attggtttgg
                                         ttatatgcaa
                                                      agtgtcgaga
                gtattatagt
-1116 gtattataga
                             tccataattt
                                         atctcattta
                                                      gtaagtagga
                tttttaata
                             tgcttttaa
                                                      tctttaaaac
-1066 gaggccgtga
                                         ccccagttat
                             gtttagtagt
-1016
   ttagccggtt
                attgtgtatg
                                         taatacagtc
                                                      attttacat
-966 cgcgtataag
                taaaggaaac
                             ttgcttttcc
                                         aaccataagt
                                                      taggtcactt
-916
    cccgaaaccg
                ccattgaaca
                             atgtactttc
                                         accggtaact
                                                      ttacttcctt
-866
   tcccctcgtt
                aagttattt
                             caccagagtc
                                         gtttaagtga
                                                      ccgattttgt
-816 gaatattgct
                gttttggga
                             gaatttttg
                                         aagtctgaga
                                                      cgataaaatc
-766 cttctaaaat
                gaaagactcg
                             tggttaggga
                                         gtgtcaagcg
                                                      catggtgacg
-716
   tcacgtggca
                atctggcggt
                             cacggggctg
                                          acacccccc
                                                      cacccccca
-666
   ctcgacggcc
                tcaccgccgt
                             accacctcat
                                         tgacggtgga
                                                      ggaaaa<u>ggg</u>
                                                               GC-
                                                      cAGTCTGTCA
-616
   <u>gcggg</u>gccct ttatatgcta gggccgttcc ggtgttgg<u>tt</u>
    box
                                                       InR
    CTTGCTCCCG
                AAGTGATCAC
                             TGGTGCTCGT
                                         TGGGCCTCTT
                                                      GTCCCATTCC
-566
-516 ATTTAGTCAT
                TATCTGTGAT
                             AATATCTGg t
                                                      tggtttttt
                                         aaatatcttg
-466 gtgcagttac
                agatataact
                             tttacttgtg
                                         aaatttcttg
                                                      gttcaatgga
-416 gtgttgtgtt
                gaaacgatgt
                             gccaccaacc
                                         ggtttggaag
                                                      ttgtgatggc
                                         atatttgaca
-366 atcgttgtgg
                tttaactttg
                             tatataaatt
                                                      actttggatg
-316 gaaacttact
                tgcacactgg
                             catgtaaagt
                                         gaatgtctag
                                                      cattgatttc
                acaggtttcg
-266 agtgttggat
                             attccttaga
                                         atccttctga
                                                      acaattttc
                                         tagcttttg
-216 gtgtaaattg
                ttgccggatt
                             tcgaggcttc
                                                      accatcaaat
-166 acacatgttg
                aagcttattt
                             tgaaaacttg
                                         aagcttagtt
                                                      tgaaagcttg
                                         gttttactca
-116 aagggtttag
                acctgaattt
                             cattgatgaa
                                                      agttttatat
                tgttaatttt
                             aaatttactt
                                         aacactatta
 -66
   ccaagcagtg
                                                      ctttctcttg
 -16 c a g C T A A T A C
                A A A <u>A T G</u>
```

Figure 6 The conserved genetic elements in the *Mractb1* 5'-flanking region. The conserved genetic elements in the *Mractb1* 5'-flanking region. The CpG islands are shown in red. The exons are shown in uppercase and the introns are shown in lowercase. The numbers represent relative position of the nucleotides, and the first base of the start codon is regarded as +1. The donor/acceptor sites for splicing are shown in red italics. The arrows indicate transcription start sites. The CCAAT-box, CArG-box, GC-box and InR are underlined.

Full-size 🖾 DOI: 10.7717/peerj.5701/fig-6

suggest complex interactions among genetic elements in regulating expression of the *Mractb1* gene.

Genetic elements in the Mractb1 gene promoter

As shown in (Figs. 5A 5B), the CpG islands in the promoter of the *Mractb1* gene are intimately associated with the promoter activity. It is interesting that despite no TSS was identified in the first CpG island (i1, Mba5), it still directs high level of transcription (Fig. 5B). Further investigation indicated that this region contains a CCAAT box, a CArG box and a GC box (Fig. 6). Both CCAAT box and CArG box were reported to be important for the constitutive expression of carp and human *beta-actin* gene (*Liu et al., 1990*; *Quitschke et al., 1989*). The second CpG island (i2) encompasses the first TSS and an initiator element (InR) (Fig. 6). The third CpG island (i3) contains the second TSS, a GC box and an InR



Figure 7 *Mractb1* promoter demonstrates comparative activity with carp *actb1* gene promoter. (A) Relative luciferase activity. (B) Western blot of firefly luciferase. (C) Relative levels of the firefly luciferase protein. The *Mractb1* promoter (Mbap) and carp *actb1* promoter (Cbap) were cloned into the pGL3-basic vector and the resulted constructs were transfected into EPC cells. Dual luciferase assays and western blot were performed to analyze the activities and protein levels of the luciferase reporter, respectively. GAPDH was used as loading control for western blot. The relative intensity represents ratio of the brightness of the firefly luciferase bands to those of GAPDH. The bars indicate mean \pm SD (n = 3). Full-size \supseteq DOI: 10.7717/peerj.5701/fig-7

(Fig. 6). Since no TATA box was identified in the core promoters for the *Mractb1* gene, the InRs may play important roles in transcription initiation.

Mractb1 promoter demonstrates comparative activity with the carp *actb1* promoter

To justify whether the *Mractb1* promoter cloned in this study has the potential to be used as a genetic tool in gene function and transgenesis studies, the activity of which was compared with that of carp *beta-actin (actb1)* gene promoter, a promoter which is commonly used in transgenic fish. The fragment with the highest promoter activity (Mba654321) was designated as Mbap. The promoter of the carp *actb1* gene (designated as Cbap) was cloned into the pGL3-basic vector as well. Dual luciferase assays in EPC cells indicated that the Mbap demonstrated comparative activity with the Cbap (Fig. 7A). Western blot using antibody against the firefly luciferase detected similar expression level of firefly luciferase in the EPC cells transfected with pGL-Mbap and pGL-Cbap (Figs. 7B, 7C); however, the empty vector pGL-basic demonstrated no luciferase expression at all. These results indicate

that the *Mractb1* promoter is as active as the carp *actb1* promoter and which can be used for transgenesis and gene function studies for prawn and related species in the future.

DISCUSSION

As a house-keeping gene, the promoter of *actb1* is cloned from a wide range of organisms and extensively used for gene function and transgenesis studies (*Barman et al., 2015*; *Cho et al., 2011*; *Hwang et al., 2003*; *Kong et al., 2014*; *Liu et al., 1991*; *Noh et al., 2003*). Although the CDS (coding sequence) and 3' UTR sequence of the *Mractb1* gene have been cloned previously (*Zhu et al., 2005*), the promoter remains to be characterized. In this study, we cloned the upstream sequence of the *Mractb1* gene, identified the TSSs, investigated expression of the transcripts in different tissues of both sexes and characterized transcriptional activity of the *Mractb1* gene promoter. Both luciferase assays and western blot indicated that the *Mractb1* promoter in the EPC cells, suggesting that it is a valuable tool for gene function and transgenesis studies for *Macrobrachium rosenbergii*.

Identification of two TSSs and the results of promoter activity assays indicated that the *Mractb1* gene possesses two alternative promoters which initiate transcription from different TSSs and lead to the generation of two transcript variants. These two transcript isoforms contain different initial exons but share the same open reading frame (ORF) and 3' UTR, indicating that they only differ in the 5' UTR. Existence of alternative promoters is common for mammalian genes. It was reported that 18% of human genes have evidence of alternative promoter usage (*Landry, Mager & Wilhelm, 2003*). Alternative promoters may differ in tissue and developmental stage specificity and transcriptional activity (*Landry, Mager & Wilhelm, 2003*). The usage of alternative promoters may account for distinct mRNA levels of the *Mractb1* gene transcripts in different tissues and sexes, and different 5' UTR of the two transcript variants may affect their translation efficiency. Moreover, molecular dissection for the activities of different upstream regions revealed significant synergistic effects between the two alternative promoters. Although the *actb1* gene promoters of human and carp were well characterized, it is unknown whether they possess alternative promoters as well.

Three CpG islands were identified in the 5' flanking sequence of the *Mractb1* gene. These CpG islands are closely related to the activity of the *Mractb1* gene promoter according to luciferase assays. All of them demonstrated transcriptional activity and both of the identified TSSs are located in the two proximal CpG islands, consistent to the discovery that most CpG islands function as sites of transcription initiation by destabilizing nucleosomes and attracting proteins that create a transcriptionally permissive chromatin state (*Deaton & Bird*, 2011). CpG islands are prevalent in the promoter of housing keeping (HK) genes and more than three quarters (78.7%) of human HK (housing keeping) genes predominantly have a CpG+/TATA- core promoter (*Zhu et al.*, 2008). CpG-islands usually contain multiple GC-boxes, and GC-boxes coupled with InR elements are essential to initiate transcription in the absence of TATA-box (*Butler & Kadonaga*, 2002). Consistent with this notion, both GC-box and InR, but no TATA-box were found in the core promoters of the *Mractb1* gene. Genetic elements including the CCAAT-box and CArG-box were reported to be important for the constitutive expression of human and carp *actb1* gene (*Liu et al., 1990*; *Quitschke et al., 1989*). These elements were also identified in the *Mractb1* promoter. Fusing the DNA fragment containing these elements (Mba5) to the proximal promoter sequence (Mba4321) dramatically increased the promoter activity (Figs. 5A, 5E), suggesting the function of these elements in regulating the *Mractb1* gene. As in the carp *actb1* gene promoter, a distal upstream region (Mba7) was found to negatively regulates the *Mractb1* promoter activity. Moreover, the upstream fragment Mba6 and 5' end of the first intron Mba3, and the interaction between these two regions are necessary for the constitutive transcriptional activity of the *Mractb1* gene promoter. These results indicate that in addition to the conservative elements, other unknown elements and factors are involved in regulating the *Mractb1* gene.

CONCLUSIONS

The *beta-actin* (*Mractb1*) gene promoter of *M. rosenbergii* was cloned and characterized. Two alternative promoters were identified for the *Mractb1* gene, which direct the generation of two transcripts with different 5' UTR. Genetic dissection of the upstream sequence of the *Mractb1* gene revealed one distal negative element and two proximal positive elements regulating the activity of the *Mractb1* gene promoter. Finally, the *Mractb1* promoter demonstrated comparative activity to the carp (Cyprinus carpio) *actb1* promoter. Our investigations provide a valuable genetic tool for gene function studies and shed light on the regulation of the *Mractb1* gene.

ACKNOWLEDGEMENTS

The authors would like to thank Dr. Bolan Zhou from the Hunan University of Arts and Science for technical assistance.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This work was supported by the Youth Innovation Promotion Association of Chinese Academy of Sciences (No.2016305) and the Major Project for New Agricultural Varieties Breeding of Zhejiang Province (No.2012C12907-2). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors: Youth Innovation Promotion Association of Chinese Academy of Sciences: 2016305. Major Project for New Agricultural Varieties Breeding of Zhejiang Province: 2012C12907-2.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Junjun Yan performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper.
- Qiang Gao performed the experiments.
- Zongbin Cui and Guoliang Yang conceived and designed the experiments, approved the final draft.
- Yong Long conceived and designed the experiments, performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, prepared figures and/or tables, authored or reviewed drafts of the paper, approved the final draft.

Animal Ethics

The following information was supplied relating to ethical approvals (i.e., approving body and any reference numbers):

The animal protocol of this study was approved by the Institutional Animal Care and Use Committee of Institute of Hydrobiology (Approval ID: Y341131501).

DNA Deposition

The following information was supplied regarding the deposition of DNA sequences:

The obtained upstream sequence counted from the first nucleotide of the initiation codon is 2,620 bp in length, is available at GenBank, accession number KY038927.

Data Availability

The following information was supplied regarding data availability:

Long, Yong; Yan, Junjun (2018): Molecular characterization of the giant freshwater prawn (*Macrobrachium rosenbergii*) beta-actin gene promoter. figshare. Fileset. https://doi.org/10.6084/m9.figshare.7076543.v3.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/ peerj.5701#supplemental-information.

REFERENCES

Alam MS, Lavender FL, Iyengar A, Rahman MA, Ayad HH, Lathe R, Morley SD,

Maclean N. 1996. Comparison of the activity of carp and rat β -actin gene regulatory sequences in tilapia and rainbow trout embryos. *Molecular Reproduction & Development* **45**:117–122

DOI 10.1002/(SICI)1098-2795(199610)45:2<117::AID-MRD2>3.0.CO;2-W.

- Arenal A, Pimentel R, Guimarais M, Rodríguez A, Martínez R, Abad Z. 2000. Gene transfer in shrimp (*Litopenaeus schmitti*) by electroporation of single-cell embryos and injection of naked DNA into adult muscle. *Biotecnologia Aplicada* 17:247–250.
- **Banu R, Christianus A. 2016.** Giant freshwater prawn *Macrobrachium rosenbergii* farming: a review on its current status and prospective in Malaysia. *Journal of Aquaculture Research & Development* 7:Article 1000423 DOI 10.4172/2155-9546.1000423.

- Barman HK, Mohanta R, Patra SK, Chakrapani V, Panda RP, Nayak S, Jena S, Jayasankar P, Nandanpawar P. 2015. The β-actin gene promoter of rohu carp (*Labeo rohita*) drives reporter gene expressions in transgenic rohu and various cell lines, including spermatogonial stem cells. *Cellular & Molecular Biology Letters* 20:237–247 DOI 10.1515/cmble-2015-0010.
- **Bennett DJ, Jennings RC. 2013.** Successful agricultural innovation in emerging economies: transgenic marine algae for aquaculture: a coupled solution for protein sufficiency. New York: Cambridge University Press, 233–246.
- Bonami JR, Widada JS. 2011. Viral diseases of the giant fresh water prawn *Macrobrachium rosenbergii*: a review. *Journal of Invertebrate Pathology* 106:131–142 DOI 10.1016/j.jip.2010.09.007.
- **Butler JEF, Kadonaga JT. 2002.** The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes & Development* **16**:2583–2592 DOI 10.1101/gad.1026202.
- **Cho YS, Sang YL, Kim YK, Dong SK, Nam YK. 2011.** Functional ability of cytoskeletal β -actin regulator to drive constitutive and ubiquitous expression of a fluorescent reporter throughout the life cycle of transgenic marine medaka *Oryzias dancena*. *Transgenic Research* **20**:1333–1355 DOI 10.1007/s11248-011-9501-2.
- Deaton AM, Bird A. 2011. CpG islands and the regulation of transcription. *Genes & Development* 25:1010–1022 DOI 10.1101/gad.2037511.
- Feng J, Li G, Liu X, Wang J, Wang YQ. 2014. Functional analysis of the promoter region of amphioxus *beta-actin* gene: a useful tool for driving gene expression in vivo. *Molecular Biology Reports* 41:6817–6826 DOI 10.1007/s11033-014-3567-x.
- Gunning P, Leavitt J, Muscat G, Ng SY, Kedes L. 1987. A human *Beta-actin* expression vector system directs high-level accumulation of antisense transcripts. *Proceedings* of the National Academy of Sciences of the United States of America 84:4831–4835 DOI 10.1073/pnas.84.14.4831.
- Hanley S, Smith TJ, Muller F, Maclean N, Uzbekova S, Prunet P, Breton B. 1998. Isolation and functional analysis of the histone H3 promoter from atlantic salmon (*Salmo salar L*). *Molecular Marine Biology & Biotechnology* 7:165–172.
- Heckman KL, Pease LR. 2007. Gene splicing and mutagenesis by PCR-driven overlap extension. *Nature Protocols* 2:924–932 DOI 10.1038/nprot.2007.132.
- Hong JK, Suh EJ, Kwon SJ, Lee SB, Jin AK, Lee SI, Lee YH. 2016. Promoter of chrysanthemum actin confers high-level constitutive gene expression in Arabidopsis and chrysanthemum. *Scientia Horticulturae* 211:8–18 DOI 10.1016/j.scienta.2016.08.006.
- Hwang G-L, Rahman MA, Razak SA, Sohm FDR, Farahmand H, Smith A, Brooks C, Maclean N. 2003. Isolation and characterisation of tilapia β -actin promoter and comparison of its activity with carp β -actin promoter. *Biochimica et Biophysica Acta* 1625:11–18 DOI 10.1016/S0167-4781(02)00534-1.
- Kong HJ, Kim JL, Kim WJ, Kim HS, Yeo SY, Park JY, An CM. 2014. Genomic cloning and promoter analysis of the *beta-actin* gene from Korean rose bitterling (*Rhodeus uyekii*). *Genes & Genomics* 36:861–869 DOI 10.1007/s13258-014-0221-x.

- Kosuga M, Enosawa S, Li XK, Suzuki S, Matsuo N, Yamada M, Roychowdhury J, Koiwai O, Okuyama T. 2000. Strong, long-term transgene expression in rat liver using chicken *beta-actin* promoter associated with cytomegalovirus immediate-early enhancer (CAG promoter). *Cell Transplantation* **9**:675–680 DOI 10.1177/096368970000900513.
- Landry JR, Mager DL, Wilhelm BT. 2003. Complex controls: the role of alternative promoters in mammalian genomes. *Trends in Genetics* 19:640–648 DOI 10.1016/j.tig.2003.09.014.
- Leary S, Underwood W, Anthony R, Cartner S, Corey D, Grandin T, Greenacre CB, Gwaltney-Bran S, McCrackin MA, Meyer R. 2013. *AVMA guidelines for the Euthanasia of animals: 2013 Edition*. Schaumburg: University of Alaska Anchorage.
- **Ling SW. 1969.** The general biology and development of the *macrobrachium rosenbergii* (de man). *FAO Fisheries Report* **57**:589–606.
- Liu ZJ, Moav B, Faras AJ, Guise KS, Kapuscinski AR, Hackett PB. 1990. Functionalanalysis of elements affecting expression of the *Beta-actin* gene of carp. *Molecular and Cellular Biology* 10:3432–3440 DOI 10.1128/MCB.10.7.3432.
- Liu Z, Xiang J, Zhou G, Gong Z. 2001. Foreign gene transfer into Chinese shrimps (*Penaeus chinensis*) with gene gun. *Chinese Science Bulletin* 46:766–769 DOI 10.1007/BF03187219.
- Liu ZJ, Zhu ZY, Roberg K, Faras A, Guise K, Kapuscinski AR, Hackett PB. 1991. Isolation and characterization of *beta-actin* gene of carp (*Cyprinus carpio*). *Mitochondrial DNA* 1:125–136.
- Long Y, Song G, Yan J, He X, Li Q, Cui Z. 2013. Transcriptomic characterization of cold acclimation in larval zebrafish. *BMC Genomics* 14:612 DOI 10.1186/1471-2164-14-612.
- Mo S, Wang L, Li Q, Li J, Li Y, Thannickal VJ, Cui Z. 2010. Caveolin-1 regulates dorsoventral patterning through direct interaction with β -catenin in zebrafish. *Developmental Biology* **344**:210–223 DOI 10.1016/j.ydbio.2010.04.033.
- New MB. 2005. Freshwater prawn farming: global status, recent research and a glance at the future. *Aquaculture Research* 36:210–230 DOI 10.1111/j.1365-2109.2005.01237.x.
- New MB, Nair CM. 2012. Global scale of freshwater prawn farming. *Aquaculture Research* 43:960–969 DOI 10.1111/j.1365-2109.2011.03008.x.
- Noh JK, Cho KN, Han EH, Kim AR, Lee JS, Dong SK, Kim CG. 2003. Genomic cloning of mud loach *Misgurnus mizolepis* (*Cypriniformes, Cobitidae*) β-actin gene and use-fulness of its promoter region for fish transgenesis. *Marine Biotechnology* **5**:244–252 DOI 10.1007/s10126-002-0066-1.
- Preston NP, Baule VJ, Leopold R, Henderling J, Atkinson PW, Whyard S. 2000. Delivery of DNA to early embryos of the Kuruma prawn, *Penaeus japonicus*. *Aquaculture* 181:225–234 DOI 10.1016/S0044-8486(99)00236-7.
- **Quitschke WW, Lin ZY, Deponti-Zilli L, Paterson BM. 1989.** The beta actin promoter. High levels of transcription depend upon a CCAAT binding factor. *Journal of Biological Chemistry* **264**:9539–9546.

- **Rasmussen RS, Morrissey MT. 2007.** Biotechnology in aquaculture: transgenics and polyploidy. *Comprehensive Reviews in Food Science and Food Safety* **6**:2–16 DOI 10.1111/j.1541-4337.2007.00013.x.
- Saurabh S, Sahoo PK. 2008. Major diseases and the defence mechanism in giant freshwater prawn, *Macrobrachium rosenbergii* (de Man). *Proceedings of the National Academy of Sciences, India—Section B: Biological Sciences* 78:103–121.
- Shi Y, Soderlund M, Xiang J, Lu Y. 2015. Function and regulation domains of a newly isolated putative β -actin promoter from pacific white shrimp. *PLOS ONE* 10:517–528 DOI 10.1371/journal.pone.0122262.
- Shi Y, Xiang J, Zhou G, Ron TB, Tong H-I, Kang W, Sun S, Lu Y. 2016. The Pacific white shrimp β-actin promoter: functional properties and the potential application for transduction system using recombinant baculovirus. *Marine Biotechnology* 18:349–358 DOI 10.1007/s10126-016-9700-1.
- Sin FYT. 1997. Transgenic fish. *Reviews in Fish Biology and Fisheries* 7:417–441 DOI 10.1023/A:1018452214763.
- Zhou B, Long Y, Song G, Li Q, Cui Z. 2015. Molecular characterization of the lgals1 gene in large scale loach *Paramisgurnus dabryanus*. *Gene* 577:65–74 DOI 10.1016/j.gene.2015.11.025.
- Zhu XJ, Dai ZM, Liu J, Yang WJ. 2005. Actin gene in prawn, Macrobrachium rosenbergii: characteristics and differential tissue expression during embryonic development. Comparative Biochemistry & Physiology Part B Biochemistry & Molecular Biology 140:599–605 DOI 10.1016/j.cbpc.2004.12.010.
- Zhu J, He FH, Hu SN, Yu J. 2008. On the nature of human housekeeping genes. *Trends in Genetics* 24:481–484 DOI 10.1016/j.tig.2008.08.004.