Gene 536 (2014) 207–212

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Short Communication

The *myostatin* gene of *Mytilus chilensis* evidences a high level of polymorphism and ubiquitous transcript expression

Gustavo Núñez-Acuña, Cristian Gallardo-Escárate *

Laboratory of Biotechnology and Aquatic Genomics, Interdisciplinary Center for Aquaculture Research (INCAR), University of Concepción, PO Box 160-C, Concepcion, Chile

ARTICLE INFO

Article history:
Accepted 30 November 2013
Available online 12 December 2013

Keywords:
Myostatin
Mytilus chilensis
SNPs
GDF-8
Muscle
qPCR

ABSTRACT

Myostatin (MSTN) is a protein of the Transforming Growth Factor-β (TGF-β) superfamily and plays a crucial role in muscular development for higher vertebrates. However, its biological function in marine invertebrates remains undiscovered. This study characterizes the full-length sequence of the *Mytilus chilensis myostatin* gene (*Mc-MSTN*). Furthermore, tissue transcription patterns and putative single nucleotide polymorphisms (SNPs) were also identified. The *Mc-MSTN* cDNA sequence showed 3528 base pairs (bp), consisting of 161 bp of 5′ UTR, 2110 bp of 3′ UTR, and an open reading frame of 1257 bp encoding for 418 amino acids and with an RXXR proteolytic site and nine cysteine-conserved residues. Gene transcription analysis revealed that the *Mc-MSTN* has ubiquitous expression among several tissues, with higher expression in the gonads and mantle than in the digestive gland, gills, and hemolymph. Furthermore, high levels of polymorphisms were detected (28 SNPs in 3′-UTR and 9 SNPs in the coding region). Two SNPs were non-synonymous and involved amino acid changes between Glu/Asp and Thr/Ile. Until now, the *MSTN* gene has been mainly related to muscle growth in marine bivalves. However, the present study suggests a putative biological function not entirely associated to muscle tissue and contributes molecular evidence to the current debate about the function of the *MSTN* gene in marine invertebrates.

1. Introduction

Myostatin (MSTN), also known as growth differentiation factor 8 (GDF-8), is a protein that belongs to the Transforming Growth Factor-β (TGF-β) superfamily. Reported evidence shows that this protein has a pivotal role in muscle growth and differentiation, mainly acting as a negative regulator of the cell cycle in this tissue (McPherron et al., 1997). To inhibit muscle protein synthesis, MSTN has a major role in different signaling pathways including the Akt, MAPK, and Smad cascades (Li et al., 2008; Trendelenburg et al., 2009; Zhu et al., 2004). Generally *MSTN* cDNA encodes for proteins of around 375 amino acids in aquatic vertebrates (Roberts and Goetz, 2001), but recently *MSTN* cDNAs encoding for longer proteins have been found in marine invertebrates (Guo et al., 2012; Hu et al., 2010). Moreover, characterized sequences of *MSTN* present conserved sites such as a proteolytic processing site (RXXR) at the carboxy-terminal region and nine conserved cysteine residues after this site (McPherron et al., 1997).

The *MSTN* gene has been characterized in a wide variety of vertebrate species, including mammals, fish, and birds (Rodgers and Garikipati, 2008). However, this gene has been characterized in few invertebrate species. Regarding marine species, an important difference between the *MSTN* of fish and invertebrates has been observed in regards to the number of presented isoforms. In fish species, such as salmonids and other teleost, two different isoforms of the same gene have been identified (Roberts and Goetz, 2001), while in marine invertebrates, such as pectinids, just one isoform has been identified (Guo et al., 2012; Hu et al., 2010). Another major difference in marine species is in relation to the expression patterns that *MSTN* exhibited in different tissues. For example, in humans *MSTN* is almost exclusively expressed in muscle tissue, and so it is thought to function more specifically in the development of these tissues (Carlson et al., 1999). In contrast, for some marine invertebrates, such as the crustacean *Gecarcinus lateralis*, it is expressed in a wider array of tissues than in higher vertebrates (Covi et al., 2008), and therefore its potential function could not be precisely specified.

MSTN is considered a highly polymorphic gene at interspecific and intraspecific levels. Previous studies have evidenced this through the identification of a high number of punctual polymorphisms in different species, mainly in mammals and in chicken, and through the association of these mutations to certain phenotypes such as muscle mass and lipid content (Grisolia et al., 2009; Johnson et al., 2009; Xianghai et al., 2007). It has been suggested that punctual mutations in the *MSTN* sequence could affect these phenotypes by inactivating the gene or by decreasing

Abbreviations: MSTN, myostatin; TGF-β, Transforming Growth Factor-β; *Mc-MSTN*, *Mytilus chilensis myostatin*; SNPs, single nucleotide polymorphisms; UTR, untranslated region; bp, base pairs; GDF-8, growth differentiation factor 8; mRNA, messenger RNA; RNA-seq, RNA sequencing; ORF, open reading frame; RACE, rapid amplification of cDNA ends.

* Corresponding author at: Departamento de Oceanografía, Centro de Biotecnología, Universidad de Concepción, P.O. Box 160-C, Concepción, Chile. Tel.: +56 41 2203422; fax: +56 41 2207310.
E-mail address: crisgallardo@udec.cl (C. Gallardo-Escárate).

0378-1119/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.gene.2013.11.067

the activity of the protein. In this context, a single nucleotide polymorphism (SNP) in the promoter region of *MSTN* in spotted halibut, *Verasper variegatus*, has been linked to individual growth performance (Li et al., 2012). Moreover, in the marine invertebrate *Chlamys farreri* an association analysis between growth traits and one non-synonymous SNP in the *MSTN* gene was conducted with significant results (Wang et al., 2010).

Mytilus chilensis is a marine bivalve with distribution along the Chilean coast, and is most abundant in southern Chile (Brattström and Johanssen, 1983). This species is susceptible to variations in this natural distribution as a consequence of local environmental conditions and anthropogenic activities, such as aquaculture which implies the displacement of seeds and breeders from native locations (Núñez-Acuña et al., 2012). Despite the great commercial importance of this species and its relationship with the "*Mytilus* complex," little is known about the implications of specific genes related to growth, cell development, reproduction, and immune response. The aim of this study was to characterize the *MSTN* gene as a potential candidate gene related to muscular growth, and also to identify SNPs-*MSTN*. Overall, this study provides the first report of *MSTN* mRNA in the Mytilidae family and gives relevant molecular evidence towards the understanding of its biological function in mussels.

2. Materials and methods

2.1. Myostatin gene discovery

A transcript annotated as *Myostatin* (*MSTN*) was obtained from a RNA-seq library performed for *Mytilus chilensis* (unpublished data, manuscript in preparation). Briefly, a RNA-seq library was generated through high-throughput sequencing of hemocytes of *M. chilensis* in the HiSeq 2000 platform (Illumina®, San Diego, CA, USA). Around eight hundred million reads were obtained and *De Novo* assembled, yielding over two hundred and fifty thousand contigs. Then, *MSTN* sequences from marine invertebrates were downloaded from the Genbank database in order to perform a MultiBlast analysis against all contigs by using the CLC Genomic Workbench software (version 6.0, CLC bio, Aarhus, Denmark). One contig with a significant e-value (e-value = 1.2E − 120), blasted with the *Crassostrea gigas* MSTN (Genbank accession number: EKC29862), was selected. This contig was 2753 bp in length and was used to obtain the full-length cDNA of *MSTN*.

2.2. Full MSTN cDNA characterization

To verify if the transcript was correctly annotated as *MSTN*, PCR amplifications were performed. Here, total RNA from muscle tissue of *M. chilensis* was extracted using the Trizol Reagent (Ambion®, Life Technologies, Grand Island, NY, USA) following the manufacturer's protocol. Then, from 200 ng of RNA, cDNA was synthetized using the RevertAid H Minus First Strand cDNA synthesis kit (Thermo Scientific, Glen Burnie, MD, USA). In parallel, a pair of primers was designed to flank 719 bp in the C-terminal region of the ORF (Primers Mc-MSTN1-F and Mc-MSTN1-R, see sequences in Table 1) by using the Geneious 6.0.4 software (Biomatters, Auckland, New Zealand). PCR amplification was conducted using 200 ng of cDNA, 10 μM of each primer, 1.5 mM MgCl2, and 0.06U taq DNA polymerase (Thermo Scientific). PCR cycles consisted of initial denaturation at 95 °C for 5 min, 35 cycles of denaturation at 95 °C for 60 s, annealing at 58 °C for 30 s, extension at 72 °C for 60 s, and a final extension at 72 °C for 7 min. PCR products were visualized through electrophoresis in agarose gels stained with GelRed™ (Biotium, Hayward, CA, USA) and then directly sequenced in an ABI 3730xl capillary sequencer.

Full-length cDNA of *MSTN* was obtained by the rapid amplification of cDNA ends (RACE) technique using the SMARTer™ RACE cDNA Amplification kit (Clontech Laboratories Inc., Mountain View, CA, USA). Primers for amplifying the 5′UTR and 3′UTR ends were designed according manufacturer's instructions (Table 1). Amplicons were purified and cloned using the TOPO TA cloning kit (Invitrogen, Life Technologies) and then transformed into *E. coli* JM109 in LB/amp/IPTG/Xgal. Positive clones were selected by galactosidase reaction visualization, and then its plasmids were purified using the E.Z.N.A® Plasmid DNA Mini kit II (Omega Bio-tek, Doraville, GA, USA). Purified plasmids were sequenced in both reverse and forward directions using M13 universal primers. Sequence analyses were carried out in the Geneious software, and consisted of quality visualizations, assembly, and BLASTn against the Genbank non-redundant database.

Table 1
Primers used in the present study. The "Specification" column shows the reaction in which each primer was used.

Primer name	Sequence (5′ to 3′)	Specification
Mc-MSTN1-F	TCAACTTCCTCCAACGGAAC	end-point PCR
Mc-MSTN1-R	TCGGCAAGGTTGTGAATTGA	
Mc-MSTNq-F	AACAAGTGGATGCGATACCC	real-time PCR
Mc-MSTNq-R	TCGGCAAGGTTGTGAATTGA	
α-tubulin-F	GAGCCGTCTGCATGTTGAGC	endogenous control (qPCR)
α-tubulin-R	TGGACGAAAGCACGTTTGGC	
Mc-MSTN_R5-i	CTCGCCTCAACTGCTCGCTAG	RACE 5′ inner
Mc-MSTN_R5-o	ACTCGAGCAGGATTGCCGTCC	RACE 5′ outer
Mc-MSTN_R3-i	GGTGCACAACTAGTGCTG	RACE 3′ inner
Mc-MSTN_R3-o	TAGTATAAATGAGAAATAGACC	RACE 3′ outer

2.3. MSTN gene transcription analysis

Total RNA extractions were performed from different tissue samples of mussels ($N = 10$); the adductor muscle, mantle, gills, gonad, digestive gland, and hemolymph. Extractions were performed using the Trizol Reagent® (Invitrogen, USA) with the same conditions as in the previous section. Regarding hemolymph, minor changes to the protocol were made. Specifically, centrifugation at 1200 ×*g* for 30 min at 4 °C after hemolymph collection and homogenization, after the Trizol was added, in the Mixer Mill Retsch MM200 (Retsch Inc., Düsseldorf, Germany) at 20 Hz for 5 min were performed.

Primers for qPCR analysis were designed by Geneious software (Table 1). End-point PCR reactions were conducted to standardize the optimal conditions for the amplification by using the same reagents described above but with a new PCR program consisting of initial denaturation at 95 °C for 2 min, 35 cycles denaturation at 95 °C for 30 s, annealing at 57 °C for 30 s, extension at 72 °C for 30 s, and a final extension at 72 °C for 5 min. Then, a dynamic range analysis was conducted to obtain qPCR efficiencies and optimal conditions for qPCR runs. For this, five serial dilutions of cDNA were prepared starting with 80 ng and with a serial factor of 1:5 in order to generate an amplification curve with both pairs of primers designed for qPCR analysis. The qPCR runs were assayed in an ABI StepOnePlus™ Mastercycler (Applied Biosystems®, Life Technologies) with the Maxima Kit® SYBR Green/ROX qPCR Master Mix (Thermo Scientific®) following manufacturer's instructions. The run method comprised of a holding stage at 95 °C for 10 min to activate the enzyme, then 40 cycles of 95 °C for 15 s, 57 °C for 30 s, and 72 °C for 30 s. After the amplification stage, a melting curve was performed from 57 °C to 95 °C, and data was collected every 0.3 °C to determine the existence of a sole amplification product and to verify the inexistence of contaminations and primer dimers. Efficiencies were calculated in the StepOnePlus™ Software (version 2.2, Applied Biosystems®). To quantify mRNA relative levels, the $\Delta\Delta C_T$ method was used adding the α-tubulin gene as an endogenous control (primers in Table 1). Finally, statistical analyses were performed in Microsoft Excel and GraphPad Prism (GraphPad Software, Inc., La Jolla, CA, USA).

2.4. SNP identification

Three pools of total RNA ($n = 10$ mussels for each pool) were used to identify SNPs associated to the *MSTN* gene. SNP identification was