Molecular characterization and differential expression of the myostatin gene in channel catfish (*Ictalurus punctatus*)

Arif M. Kocabas, Huseyin Kucuktas, Rex A. Dunham, Zhanjiang Liu*

Fish Molecular Genetics and Biotechnology Laboratory, Department of Fisheries and Allied Aquacultures and Program of Cell and Molecular Biosciences, Auburn University, 201 Swingle Hall, Auburn, AL 36849, USA

Received 15 October 2001; received in revised form 22 January 2002; accepted 25 January 2002

Abstract

Myostatin is a recently discovered gene that inhibits muscle growth. In the present study, we characterized the myostatin locus and its expression in channel catfish (*Ictalurus punctatus*). The genomic DNA and cDNA encoding the channel catfish myostatin were cloned and sequenced. The myostatin gene has three exons encoding a protein of 389 amino acids. Comparison of the genomic sequences with those of the cDNA revealed that the myostatin cDNA was 1673 base pair (bp) long with a 5'-untranslated region (UTR) and 3'-UTR of 180 and 323 bp, respectively. The deduced amino acid sequences of the catfish myostatin is highly conserved with those of other organisms. The myostatin locus is highly polymorphic in channel catfish because of the presence of several microsatellites and single nucleotide polymorphic sites. The myostatin gene was expressed in various tissues and developmental stages at differential levels, suggesting complex regulation of this gene and perhaps roles for myostatin in addition to those originally suggested. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Catfish; Myostatin; Muscle; Differential expression; Gene; cDNA; Microsatellite; Polyglutamine expansion; Marker; Polymorphism; SNP

1. Introduction

Myostatin is a recently discovered gene that inhibits muscle growth [1]. It is a member of the transforming growth factor β (TGF-β) superfamily [2] although it shares only low levels of identity with the well-known members of the superfamily such as the inhibins, the TGF-βs, and the bone morphogenetic proteins [1]. In developing mouse embryos, the myostatin gene is expressed only in cells forming the myotome compartment of the somites, and skeletal muscles of the developing and adult mice [1]. However, it has been reported to be expressed by cells of different origins in mature sheep and fish [3–5].

Myostatin is a negative regulator of muscle deposition. Direct evidence came from knockout experiments in mice that had a significant increase in skeletal muscle mass. Supporting this discovery was the natural mutations of the myostatin gene in two breeds of cattle, the Belgian Blue and the Piedmontese, both of which possess a double muscling phenotype [6–10]. Such a major effect of a single gene on processing yields was surprising to many animal breeders, but it opened a potential channel for improving processing yields of animals using knockout technology. Therefore, molecular characterization of the myostatin locus from important agricultural species is important to produce genomic resources for development of knockout technology as well as for understanding the structure, function, and evolution of the gene.

Channel catfish (*Ictalurus punctatus*) is the major aquaculture species in the United States [11]. As part of our catfish genome research, here we report molecular cloning, sequencing, and differential expression of the myostatin gene from channel catfish.

2. Materials and methods

2.1. Animals and tissue collection

All experimental fish were raised in troughs located in the hatchery of the Auburn University Fish Genetics Facility and conducted under a protocol approved by the Institutional Animal Care and Use Committee (IACUC) at Auburn...
University. Unfertilized eggs were collected by stripping mature female channel catfish. After fertilization, samples of 1-, 3- (with eye spot), 7-, 14-, and 21-day-old embryos and larvae were collected. Channel catfish generally hatch 6 days after fertilization. Within 14 days after fertilization, whole eggs, embryos, and larvae were used since dissecting muscles was difficult. Both the whole larvae and the dissected muscle samples were used from 14- and 21-day-old fish; only skeletal muscle tissues were used from fish older than 21 days including 9-month and 1-year, 2-, 3-, 4-, 5-, and 7-year-old fish. To determine the expression of the myostatin gene in various tissues, samples were also collected from brain, heart, eye, spleen, liver, stomach, swimbladder, head kidney, trunk kidney, intestine, and gill, all from the 5-year-old fish. Multiple individuals were used to collect the tissues and the tissue samples were mixed to avoid major bias, if any, that may exist from a single individual. All samples were kept in liquid nitrogen until isolation of RNA.

2.2. Genomic library screening, gene and cDNA cloning, and sequencing

A channel catfish genomic DNA library [12] was screened with a heterologous cDNA probe from zebrafish. The cDNA probe was made by reverse transcriptase-PCR (RT-PCR) and labeled by α32P-dCTP using random primer methods [13]. The recombinant clones containing the myostatin gene were identified by repeated plaque purification [13]. The positive clones were isolated, mapped by restriction endonuclease digestions, and subcloned into the pBlue vector [13]. The recombinant clones containing the myostatin gene were identified by repeated plaque purification [13]. The positive clones were isolated, mapped by restriction endonuclease digestions, and subcloned into the pBlue vector [13]. The recombinant clones containing the myostatin gene were identified by repeated plaque purification [13]. The recombinant clones containing the myostatin gene were identified by repeated plaque purification [13].

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Primers and their sequences used for primer walking sequencing of the channel catfish myostatin locus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primer name</td>
<td>Primer sequence</td>
</tr>
<tr>
<td>GDF 2.2 kb-700</td>
<td>5'-AGGGGGAGAGACGGAGGAG-3'</td>
</tr>
<tr>
<td>GDF 2.2 kb-800</td>
<td>5'-CGTCTGGGACAGTTCAAAACAG-3'</td>
</tr>
<tr>
<td>GDF-GSP#1-700</td>
<td>5'-TTTCCAATCCAGCTCCCG-3'</td>
</tr>
<tr>
<td>GDF-GSP#2-700</td>
<td>5'-ATAATCTTTTATATAXAAT-3'</td>
</tr>
<tr>
<td>GDF-GSP#4-700</td>
<td>5'-CTAAATCATGATGACGCGGT-3'</td>
</tr>
<tr>
<td>GDF-GSP#5-700</td>
<td>5'-CCGAGATGCCACACGCCAG-3'</td>
</tr>
<tr>
<td>GDF-GSP#6-700</td>
<td>5'-CCATGCTCATGAGCGGTCTCGG-3'</td>
</tr>
</tbody>
</table>

2.3. Southern blot analysis

Genomic DNA (∼20 μg) was isolated from blood of channel catfish as previously reported [16], digested with four restriction enzymes: HindIII, PstI, EcoRI and SacI, and analyzed on a 0.7% agarose gel. After electrophoresis, DNA was transferred to a nylon membrane (Roche Diagnostics, Indianapolis, IN) by capillary transfer with 0.4 N NaOH overnight [17]. The transferred DNA was cross-linked to the membrane by a UV linker (Stratagene, CA). The membrane was washed with 0.5% SDS (w/v) at 65 °C for 25 min and then pre-hybridized in a pre-hybridization solution [13] overnight at 42 °C in a Hy-Blot 12S hybridization oven (Phoenix Research Products, CA). Hybridization was conducted in a hybridization solution similar to that of the pre-hybridization, but excluding 5'× Denhart's. The probe used in hybridization was a 2.2-kb genomic fragment of catfish myostatin gene (i.e. the 2.2-kb EcoRI fragment from pA7-2). The probe was labeled with 32P-dCTP by using the High-Prime random primer labeling kit (Roche Diagnostics). The hybridization was incubated at 42 °C overnight in the hybridization oven. After hybridization, the membrane was washed first in 100-μl 2× SSC for 5 min at room temperature followed by three washes in 0.2× SSC and 0.2% SDS (w/v) at 60 °C for 10 min each. The membrane was wrapped with Saran wrap and exposed to Kodak BioMax MS film overnight.
Fig. 1. Genomic sequences of the channel catfish myostatin locus. Exons are shown as black boxes, whereas the upstream region, the introns, and the downstream region are shown as gray boxes. The deduced amino acid sequences are shown by single letter code of amino acids below the exons. Translational start site is in bold letters and termination site is indicated by an asterisk. Simple sequence repeats (microsatellites) are underlined.
Fig. 2. Alignments of amino acid sequences of the myostatin gene from various organisms. Polyglutamine tracts are indicated by a rectangular box. The accession numbers for these sequences are: human, NM_005259; rat, NM_019151; chicken, AF019621; Atlantic salmon isoform I, AJ297267; Atlantic salmon isoform II, AJ344158; tilapia, AF197193; seabream, AF258448; and catfish, AF396747.
2.4. RT-PCR

The poly (A)+RNA was isolated from various samples of catfish using a poly (A) pure mRNA kit (Ambion, TX). The first strand cDNA was made in a 25 µl reverse transcriptase (RT) reaction containing the following: 1 µg mRNA, 1× first-strand buffer (Life Technologies, Bethesda, MD), 20 units of RNase Block Ribonuclease Inhibitor (Stratagene), 500 ng of synthetic oligo (dT)$_{18}$ primer, 0.4 mM each of dATP, dCTP, dGTP, dTTP, and 200 units of the M-MLV reverse transcriptase (Life Technologies). The reaction was incubated at 37 °C for 1 h and then at 70 °C for 15 min to stop the reaction. The RT reaction was treated with *Escherichia coli* RNAse H for 20 min at 37 °C. After the first strand cDNA synthesis, equal amounts of the first strand reaction were amplified by the gene-specific primers. The RT-PCR primers were designed from the coding regions of myostatin (RT-GDF-8 upper primer: 5’ CTGTGAGATATCGCGACTCA 3’ which is located within the second exon; RT-GDF-8 lower primer: 5’ CCTGATTCTCTCCTGGTTCG 3’ which is located in the third exon). The reaction also included the primers of β-actin (upper primer: 5’-AGAGAGAAATTGTCCGTGACATC-3’ and lower primer: 5’-CTCCGATCCAGACAGTATTTG-3’), serving as an internal control. After an initial incubation at 94 °C for 4 min, the RT-PCR was carried out at 94 °C for 1 min, 58 °C for 1 min, 72 °C for 1 min for either 25 or 35 cycles, as specified. Recombinant Taq DNA polymerase was used (Life Technologies) to amplify the myostatin cDNAs. Upon the completion of PCR, the reaction was incubated at 72 °C for an additional 10 min. The RT-PCR products were electrophoresed on a 1.5% agarose gel and documented with a Gel Documentation System (Nucleotech, CA). The results of the RT-PCR were quantified by a Kodak 1D Image Analysis Software (Eastman Kodak Co., NY). The intensity of the bands were expressed as units of the gene expression (i.e. the net intensity that is the sum of the background subtracted pixels within the band). Relative gene expression is calibrated by the intensity of β-actin in the same sample.

3. Results

3.1. Molecular characteristics of the myostatin gene

Two recombinant λ clones, λ7-2 and λ5, were obtained after screening the genomic library with the myostatin
Restriction mapping plus partial sequencing and Southern blot analysis indicated that E5 contained the upstream portion of the gene, while E7-2 contained downstream portion of the gene. The relevant segments of the myostatin gene included in these two phage clones were subcloned into plasmid vector pBluescript SK+. The complete nucleotide sequences of the catfish myostatin gene was obtained by primer walking sequencing, which is shown in Fig. 1, and has been deposited in the GenBank with an accession number of AF396747. Comparison of the genomic DNA sequences with cDNA sequences indicated that the channel catfish myostatin gene contains three exons and two introns. The gene encodes a protein of 389 amino acids with 180 bp 5’-UTR and 323 bp 3’-UTR. The deduced amino acid sequences of the channel catfish myostatin gene are highly conserved with those of other organisms, especially at the carboxyl terminus (Fig. 2).

Channel catfish has a single copy of the myostatin gene as revealed by Southern blot analysis. A single band was detected with all four restriction enzymes tested (Fig. 3). The catfish myostatin gene contains multiple microsatellites (Fig. 1). The upstream UTR contained highly repetitive sequences including several hundred base pairs of simple sequence repeats with a simplified consensus of TGGTAG.

One (CAG) repeat was found within the first exon encoding a polyglutamine tract (Fig. 2). Three microsatellite sequences existed in the first intron with (GTTT)$_7$, (TG)$_{11}$, and (TA)$_{38}$ repeats. One microsatellite repeat of (GAA)$_7$ existed in the second intron. In the 3’-UTR, one (AC) repeat was present.

Analysis of the myostatin cDNA sequences indicated that the channel catfish myostatin gene is highly polymorphic

<table>
<thead>
<tr>
<th>Position</th>
<th>Coordinate</th>
<th>Base change</th>
<th>Amino acid change</th>
</tr>
</thead>
<tbody>
<tr>
<td>5’UTR</td>
<td>1507</td>
<td>G to A</td>
<td>N/A</td>
</tr>
<tr>
<td>First exon</td>
<td>1691</td>
<td>T to C</td>
<td>no change</td>
</tr>
<tr>
<td>First exon</td>
<td>1796</td>
<td>T to C</td>
<td>no change</td>
</tr>
<tr>
<td>First exon</td>
<td>1820</td>
<td>A to C</td>
<td>Q to H</td>
</tr>
<tr>
<td>First exon</td>
<td>1854</td>
<td>T to C</td>
<td>no change</td>
</tr>
<tr>
<td>First exon</td>
<td>1931</td>
<td>A to C</td>
<td>no change</td>
</tr>
<tr>
<td>Second exon</td>
<td>3796</td>
<td>A to G</td>
<td>K to R</td>
</tr>
<tr>
<td>Second exon</td>
<td>3801</td>
<td>T to C</td>
<td>C to R</td>
</tr>
<tr>
<td>Second exon</td>
<td>3865</td>
<td>A to G</td>
<td>H to R</td>
</tr>
<tr>
<td>Second exon</td>
<td>3875</td>
<td>G to A</td>
<td>no change</td>
</tr>
<tr>
<td>Second exon</td>
<td>3988</td>
<td>G to A</td>
<td>S to N</td>
</tr>
<tr>
<td>Third exon</td>
<td>4658</td>
<td>A to G</td>
<td>no change</td>
</tr>
<tr>
<td>3’UTR</td>
<td>5043</td>
<td>T to C</td>
<td>N/A</td>
</tr>
<tr>
<td>3’UTR</td>
<td>5080</td>
<td>G to A</td>
<td>N/A</td>
</tr>
<tr>
<td>3’UTR</td>
<td>5135</td>
<td>C to T</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Table 2: Single nucleotide polymorphic sites within the channel catfish myostatin gene

Fig. 5. Tissue distribution of myostatin gene expression. RT-PCR was conducted for 25 cycles (panel A), and 35 cycles (panel B) using mRNA isolated from various tissues of mature fish; PCR products were analyzed by electrophoresis. The designations of the lanes are as follows: M, 1 kb ladder molecular weight markers; 1, brain; 2, eye; 3, swimbladder; 4, gill; 5, head kidney; 6, heart; 7, intestine; 8, liver; 9, muscle; 10, spleen; 11, stomach; 12, trunk kidney; c, negative control. Molecular weight standards are indicated on the left margin of the gel. RT-PCR products of myostatin and of the internal control beta-actin are indicated by the arrows on the right margin. A quantitative presentation of the RT-PCR data of panel A is shown in panel C. Obviously, band intensities shown in panel B is no longer proportional to those in panel A because of high number of cycles used in amplification. The worth noting point is the detection of low expression in the stomach. Relative gene expression was determined as the expression of myostatin as compared to that of beta-actin in the same sample.
because of the microsatellites. Among the four clones of cDNAs sequenced from the upstream, three clones contained six CAG repeat, while one contained five CAG repeat in the coding region of the first exon (Fig. 4), making their encoded amino acids differ in length of polyglutamine tract. Similarly, among the six clones sequenced from the downstream of the cDNA, three contained nine (AC) repeats, three contained 10 (AC) repeats, while the genomic clone contained seven (AC) repeats. We did not determine the level of polymorphism caused by microsatellites in the upstream and in the introns. In addition to the microsatellites, we identified many single nucleotide polymorphism (SNP) within the myostatin cDNA (Table 2). Many of these SNPs were neutral without causing changes in amino acid sequences, but SNPs at 1820, 3796, 3801, 3865, and 3988 (coordinates are shown in Fig. 1) caused changes in amino acid sequences (Table 2).

3.2. Differential expression of the myostatin gene

To determine the expression of the myostatin gene in various tissues, RT-PCR was conducted. The gene had a wide distribution of expression in all 12 tested tissues. However, its expression levels varied greatly among tissues; expression was strongest in the muscle; intermediate in brain, eye, intestine, and trunk kidneys; but low in gill, spleen, heart, and liver; the lowest expression was observed in the stomach, head kidney, and swimbladder (Fig. 5). Expression in the stomach was so low that it was not detectable with 25 cycles of RT-PCR (Fig. 5A), but was visible after 35 cycles of RT-PCR (Fig. 5B).

Similarly, expression was variable during development and at various ages (Figs. 6 and 7). Myostatin was not detected in unfertilized eggs even after 35 cycles of RT-PCR (Fig. 6A and B). However, its expression was detected in 1-day-old embryos although expression level at this stage of development is quite low, requiring 35 cycles of RT-PCR for detection (Fig. 6B). Myostatin expression gradually increased as the development progressed reaching high levels of expression in about 2 weeks after fertilization (Fig. 6C).

It appeared that myostatin expression was lower in older fish (Fig. 7). Repeated experiments indicated that expression was high before channel catfish reach 3 years of age, but

![Fig. 6. Expression of myostatin gene during development of channel catfish. RT-PCR was conducted for 25 cycles (panel A), and 35 cycles (panel B) using mRNA isolated from various developmental stages; PCR products were analyzed by electrophoresis. Loading orders of the gel were as the following: MW, 1 kb molecular weight standards with sizes as indicated on the left margin; 0d, 0 day non-fertilized eggs; 1d, 1 day after fertilization; 3d, 3 days pre-hatched larvae; 7d, 1-week-old larvae; 14d, 2-week-old whole larvae; 21d, 3-week-old whole larvae; 14dM, muscle of 2-week-old larvae; 21dM, muscle of 3-week-old larvae; c, negative control with no DNA template. A quantitative presentation of the RT-PCR data of panel A is shown in panel C. Obviously, band intensities shown in panel B is no longer proportional to those in panel A because of high number of cycles used in amplification. The worth-noting point is the detection of low expression in 1-day-old embryos. Relative gene expression was determined as the expression of myostatin as compared to that of beta-actin in the same sample.](image-url)
levels of expression dropped three- to five-fold at ages of 4–7 years (Fig. 7A and B).

4. Discussion

In this study, we isolated and sequenced the channel catfish myostatin gene. The deduced amino acid sequences of the catfish gene were highly conserved through evolution with those of other organisms. It is striking that not only the sequence identity was conserved, but also the structure and organization of the gene. To date, genomic sequences are publicly available only from chicken, bovine, and sea bream (GenBank accession numbers of AF346599, AF320998, and AF258447, respectively); all had three exons and two introns. Such high levels of conservation through evolution suggest high levels of evolutionary constraints, and the importance of its function.

The major role of myostatin is to negatively regulate muscle growth. The availability of the channel catfish genomic sequences makes it possible to explore the potential for genetic improvement in "muscling" through the development of knockout technology. For that purpose, the present study indicated that the catfish myostatin was encoded by a single copy gene, making the knockout less complicated.

High levels of polymorphism in the catfish gene should facilitate genomic mapping of the myostatin gene. It is highly likely that one of the variable sites, including the microsatellites and the SNPs, is polymorphic in specific resource families currently used in catfish genome research [16,18]. One interesting variable site was the CAG repeat in the first exon. There is no polyglutamine tract in the myostatin gene of the higher vertebrates such as in chickens, bovine, rat, and human (Fig. 2). However, in channel catfish, a stretch of CAG repeat resulted in a noticeable polyglutamine tract. Alleles with five and six glutamines were observed. In another teleost fish (sea bream), the polyglutamine tract was expanded to harbor 12 CAG/CAA repeats [19]. This raises the question of the possibility that long expansions of the polyglutamine tract would cause loss of function or gain of new functions for this gene, as reported for several human genetic disease genes such as the Huntington gene involved in Huntington’s disease (HD), and SCA1 gene involved in Spinocerebellar ataxia type 1 [20].

As expected, myostatin is differentially expressed throughout the development of catfish. It was absent in unfertilized eggs, and expressed weakly in 3-day embryos and 7-day-old larvae. Its expression was considerably increased in the muscle after 14 days through the age of two. Myostatin was detectable as early as 1 day after fertilization in catfish although its detection requires extra rounds of PCR (Fig. 6). This is much earlier than in tilapia where myostatin expression was reported after hatching, but not in fertilized eggs and in pre-hatching larvae with eye spots [4]. Interestingly, myostatin was expressed at lower levels in older fish after 2 years of age. Myostatin is known to suppress muscle growth by inhibiting proliferation of myoblasts [21], which dominates in early development in teleosts [22]. Unlike in embryos and developing youngs, myoblasts are terminally differentiated to myocytes which rarely divide in mature organisms [23]. In the mature catfish after 2 years of age, it is likely that the rarely dividing myocytes no longer requires stringent negative regulation of muscle growth. As a result, the demand for myostatin gene expression decreases.

Myostatin was initially reported to be expressed only in skeletal muscle [1]. However, recent reports indicated its...
presence in various tissues [3–5]. Our results confirm the wide tissue distribution of the myostatin gene. As a matter of fact, it was found in all 12 tissues tested, although at different levels. This is consistent with the results of a recent report from Atlantic salmon that myostatin was detected in multiple tissues using Western blot and immunohistochemical analysis [24]. It is possible that myostatin is expressed in a wide variety of other tissues, suggesting other potential roles of myostatin in addition to being a negative regulator of muscle deposition, as originally suggested [1]. Further research is required to elucidate other possible biological functions of myostatin.

Acknowledgements

This project was supported by a grant from USDA NRICGP (#98-35205-6738) and by the Auburn University Competitive BioGrant (BIOGRANTJLIU99) and, in part, by a grant from the US-Israel Bi-National Agricultural Research and Development (BARD, US-2954-97). We appreciate the support of Auburn University Department of Fisheries and Allied Aquacultures, College of Agriculture, and the Vice President for Research for their matching funds to the USDA NRICGP Equipment Grants to ZL (98-35208-6540 and 99-35208-8512). We thank Ping Li for her technical assistance.

References