

基于转录组测序对翘嘴鲌(*Siniperca chuatsi*) 2种肌球蛋白重链基因的克隆与分析*

陈之航^{1,2} 董浚键¹ 孙成飞¹ 田园园¹ 叶星^{1,2①}

(1. 中国水产科学研究院珠江水产研究所 广州 510380;
2. 上海海洋大学水产与生命技术学院 上海 201306)

摘要 肌球蛋白重链(Myosin heavy chain, MYH)是骨骼肌粗肌丝的重要组成单位,其表达量高低影响肌纤维的组成和肌肉生长。为了解其在翘嘴鲌(*Siniperca chuatsi*)早期生长过程中的作用,本研究从前期生长差异显著的2组(快长组和慢长组)翘嘴鲌幼鱼转录组差异表达 Unigene 中筛选出2个 MYH 基因, RACE (Rapid amplification of cDNA ends)克隆到其全长 cDNA(MYH-7a, MYH-7b)。MYH-7a 全长为 6071 bp, 开放阅读框为 5820 bp; MYH-7b 全长为 5896 bp, 开放阅读框为 5745 bp。序列分析显示, 2个 MYH 均有 Loop1、Loop2 环、ATP 结合位点等关键结构域; 进化树聚类分析显示, MYH-7a 与 MYH-7b 均属于慢肌球蛋白。实时荧光定量 PCR 验证发现, 其在快长组样本中表达量显著高于慢长组, 与转录组测序结果一致; 检测其在翘嘴鲌心肌、红白肌和皮肤等 14 种组织的表达水平, 结果显示, MYH-7a 主要在心肌中表达, 而 MYH-7b 主要在红肌中表达; 在胚胎发育不同阶段, 二者随着胚胎发育的进行, 表达量不断增加; 在幼鱼早期生长过程(孵化出膜后 15 dph、30 dph 和 60 dph)的翘嘴鲌白肌中, 在 15 dph 和 30 dph 快长组的表达量显著高于慢长组, 而到 60 dph 时快长组的表达量均显著低于慢长组。翘嘴鲌 MYH-7a、MYH-7b 在快长组与慢长组鱼中的差异表达提示它们在翘嘴鲌胚胎及其早期生长发育过程中发挥重要作用。

关键词 翘嘴鲌; 转录组测序; 肌球蛋白重链; cDNA 克隆; 表达分析

中图分类号 S917.4 **文献标识码** A **文章编号** 2095-9869(2017)03-0051-11

肌球蛋白重链(Myosin heavy chain, MYH)的结构分为头区和尾区。常规 MYH 的 N 端头区通常与 1 个调节型肌球蛋白轻链(Regulatory myosin light chain, RLC)和 1 个必需轻链(Essential myosin light chain, ELC)结合, 折叠盘绕形成球状头部作为“马达”功能区; 2 个 MYH 的 C 端棒状区盘绕形成卷曲螺旋, 并聚集成束, 形成粗肌丝束主体。其中, 头部“马达”功能区有 ATP 结合位点, 可结合 ATP, 将其水解为

ADP 释放能量, 为肌肉收缩提供动力而影响运动、摄食等活动, 从而直接影响鱼类生长(Campion *et al.*, 2012)。MYH 也广泛存在于非肌细胞中, 是细胞骨架的组成成分, 参与细胞内的细胞器运动与有丝分裂, 细胞外的吞噬、运动、受精和吸收等生理过程, 充当非肌细胞生命活动的不同层次的调节者, 包括细胞间的信号传递、细胞形状改变等较高级的调节(Sebé-Pedrós *et al.*, 2014)。

* 国家引智项目(国家外专局)(20140326016)、广东省海洋与渔业科技推广(A201601A06)、广东省科技计划项目(2015A020209034)和广州市科技计划项目(2014J4100087)共同资助[This work was supported by State Administration of Foreign Experts Affairs (20140326016), Ocean Fisheries Science and Technology Promotion Project of Guangdong Province (A201601A06), Science and Technology Project of Guangdong Province (2015A020209034), and Science and Technology Project of Guangzhou City (2014J4100087)]. 陈之航, E-mail: kevinchenaddress@163.com

① 通讯作者: 叶星, 研究员, E-mail: gzyexing@163.com

收稿日期: 2016-04-14, 收修改稿日期: 2016-04-27

哺乳动物肌肉生长只在未成年阶段涉及肌纤维增生和增粗,成体只有增粗而不再增生;而鱼类肌纤维始终存在增生和增粗,鱼类这种生长现象被称为非限制性生长。因 *MYH* 基因表达量高低影响肌纤维的增生,被认为与鱼类的非限制性生长有关(Biga *et al.*, 2006)。在鱼类中发现的 *MYH* 亚型远远多于哺乳动物,这也被认为是鱼类基因组倍化的可靠证据(Yasmin *et al.*, 2011; Asaduzzaman *et al.*, 2016)。在虹鳟(*Oncorhynchus mykiss*)、大西洋鲑(*Salmo salar*)和大菱鲆(*Scophthalmus maximus*)等鱼类中发现, *MYH* 的表达量与鱼生长速度有显著相关性(Overturf *et al.*, 2001; Dhillon *et al.*, 2008; Churova *et al.*, 2015)。Hevrøy 等(2006)发现,投喂不同膳食水平 68 d 后,大西洋鲑肌肉 *MYH* 的表达量与肌肉蛋白总量和体重显著相关, *MYH* 的表达量可作为表征鱼类生长的重要标记。

为了解影响翘嘴鲌(*Siniperca chuatsi*)早期生长的关键基因,本研究从同批孵化、生长存在显著差异的翘嘴鲌的转录组差异表达谱数据中,筛选差异表达 *MYH* 基因,克隆获得 2 个表达量差异显著的 *MYH* 全长 cDNA,进行序列结构分析,并分析其组织分布及胚胎发育的表达规律;检测这 2 个基因在孵化出膜后不同阶段的(15、30、60 days post hatching, dph)快长组和慢长组翘嘴鲌白肌中的表达量,探讨其与生长的相关性,为后续翘嘴鲌快长品系的分子选育提供候选基因与位点。

1 材料与方法

1.1 材料

实验时间为 2014 年 5 月–2015 年 11 月,实验用翘嘴鲌取自广东省清新县宇顺农牧渔业科技服务有限公司养殖基地。转录组测序用翘嘴鲌选择同批孵化、培育 30 dph 的翘嘴鲌,从中选取生长快和生长慢的个体各 12 尾,取白肌组织;不同胚胎发育时期的翘嘴鲌源自同一对亲本的受精卵,显微镜下观察胚胎发育时期(刘希良等, 2013),分别取未受精卵、受精卵、囊胚期、神经胚期、体节出现期、尾芽期、肌肉效应期和出膜幼鱼各期样品 10 枚卵(尾)/组,各 3 组;随机选取 3 尾健康成鱼,取眼、鳃、皮肤、白肌、红肌、心肌、胃、肠、肝、脾、肾、性腺、脑及垂体共 14 个组织作为组织分布样品;取同批 15 dph、30 dph 和 60 dph 的快长组和慢长组的翘嘴鲌各 5 尾,测体重,取白肌组织作为不同生长时期快长组和慢长组样品(平均体重及差异见表 1)。上述样品置于盛有 1 ml TRIzol (Invitrogen, 美国)的 1.5 ml 离心管中并匀浆,随即液氮冻存。

表 1 孵化出膜后各时期翘嘴鲌的平均体重

Tab.1 The average body weight of *S. chuatsi* at different growth stages after hatching (g)

组别 Groups	生长时期 Growth stages (dph)		
	15	30	60
慢长组 Slow growing group	0.09	0.36	21.75
快长组 Fast growing group	0.17**	2.31*	38.60**

*表示差异显著($P<0.05$), **表示差异极显著($P<0.01$), 下同

* denoted significant difference ($P<0.05$), ** denoted highly significant difference ($P<0.01$), the same as below

1.2 RNA 提取及 cDNA 合成

将上述组织样品按照 TRIzol 试剂盒说明书提取总 RNA,总 RNA 用 DNase I (TaKaRa, 日本)去除基因组 DNA 后,用 Biophotometer (Eppendorf, 德国)及琼脂糖凝胶电泳检测 RNA 浓度及纯度,样品浓度 ≥ 400 ng/ μ l,电泳检测条带清晰,样品 -70°C 保存备用。cDNA 合成,使用 Transcriptor First Strand cDNA Synthesis Kit (Roche, 瑞士)反转录试剂盒,取 1 μ g 总 RNA 合成 cDNA 第 1 链, -20°C 保存备用,具体步骤参考说明书。将快长组和慢长组翘嘴鲌 RNA (12 个/组)各取 3 μ l 混合,形成快、慢 RNA 混池,送深圳华大基因公司,经随机打断、反转录、末端修复、加测序接头后,通过 Hiseq2000 (Illumina, 美国)平台,进行 RNA denovo 测序。

1.3 转录组表达量分析

测序后得到 clean reads,用短 reads 组装软件 Trinity 对 clean reads 进行从头组装和过滤后得到高质量的 Unigenes。用 FPKM (Fragments per kilobase of exon per million fragments mapped)法计算基因表达量(Mortazavi *et al.*, 2008)。通过错误发生率(False discovery rate, FDR)多重检验校正 P 值,并根据 FPKM 比较该基因在快长、慢长组中的差异表达倍数。筛选 $\text{FDR} \leq 0.001$ 且差异倍数不低于 2 倍(即 $|\log_2 \text{ration}| \geq 1$) 的基因为差异显著基因。

1.4 翘嘴鲌肌球蛋白重链基因 RACE 克隆

提取健康翘嘴鲌肌肉总 RNA,采用 SMARTer™ RACE cDNA Amplification Kit (Clontech, 美国)合成 5'端 cDNA 第 1 链和 3'端 cDNA 第 1 链。根据转录组差异表达谱数据,从转录组表达量差异显著的 43 个的翘嘴鲌肌球蛋白重链基因中,选出表达量差异最大的 2 条 *MYH* 基因 (CL127.Contig1_All 和 Unigene19394_All),根据其序列片段,用 Primer Premier 5.0 软件设

计 5'、3'RACE 引物(5'-GSP/3'-GSP), 并与 UPM/UPS 通用引物进行 PCR 反应, 引物信息见表 2。反应体系: 5'/3'端 cDNA 2.5 μ l, 10 \times UPM 5 μ l, 5'/3'特异引物(5'/3' GSP-1) 1 μ l, 2 \times SeqAmp Buffer 25 μ l, SeqAmp DNA 酶 1 μ l, ddH₂O 补足至 50 μ l。扩增程序: 94 $^{\circ}$ C 5 min, 5 个循环: 94 $^{\circ}$ C 30 s, 72 $^{\circ}$ C 4 min; 5 个循环: 94 $^{\circ}$ C 30 s, 70 $^{\circ}$ C 30 s, 72 $^{\circ}$ C 4 min; 20 个循环: 94 $^{\circ}$ C 30 s, 68 $^{\circ}$ C 30 s, 72 $^{\circ}$ C 4 min。第 2 轮巢式 PCR 以第 1 轮

PCR 产物稀释 50 倍为模板, 用通用引物 UPS 与 5'/3'GSP-2 进行 PCR 反应, 扩增程序同第 1 轮 PCR。5'和 3'RACE 扩增产物经凝胶电泳检测目的带并切胶回收, PCR 产物克隆至 pEASY blunt zero 载体(全式金)中, 经菌液 PCR 鉴定后, 挑取阳性菌送广州艾基生物技术有限公司测序。测序结果利用 DNAMAN 8.0 软件将 5'RACE 片段、3'RACE 片段和 Unigene 片段序列进行拼接。

表 2 翘嘴鲌肌球蛋白重链基因的克隆及实时荧光定量 PCR 表达分析使用的引物
Tab.2 Primers for cloning and qRT-PCR analysis of *S. chuatsi* myosin heavy chain

目的片段 Target fragment	引物名 Primer name	引物序列 Primer sequence (5'-3')	用途 Purpose
CL127.Contig1_All	5'GSP-1	GATTACGCCAAGCTTAGCAGCCATCTCCTCAGTCAGGTTCTT	5'RACE PCR
	5'NGSP-2	GATTACGCCAAGCTTACCCAGCAATCCAGCCTTGAAGAACA	5'RACE nested PCR
	3'GSP-1	GATTACGCCAAGCTTAGCTCCAAGTTCAAGCGGAGCAAGAT	3'RACE PCR
	3'NGSP-2	GATTACGCCAAGCTTGACTGAGGAGATGGCTGCTTTGGATGAA	3'RACE nested PCR
Unigene19394_All	5'GSP-1	GATTACGCCAAGCTTCTGCTCTCTGCTTCTCCACCTTGCC	5'RACE PCR
	5'NGSP-2	GATTACGCCAAGCTTGCGCTCCTCCAGCTGTTGCTTGTC	5'RACE nested PCR
	3'GSP-1	GATTACGCCAAGCTTGGACTGCTGGGTCAGCTAGAGGAGATG	3'RACE PCR
	3'NGSP-2	GATTACGCCAAGCTTCAGTGGAACATCCGTGCCCTTCATGGG	3'RACE nested PCR
	UPM	CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGT	5'/3'RACE PCR
	UPS	CTAATACGACTCACTATAGGGC	5'/3'RACE nested PCR
	M13-F	GTA AACGACGGCCAG	阳性克隆筛选 Screening positive clones
	M13-R	CAGGAAACAGCTATGACC	
MYH-7a	7a-F	GCTCAAGGCTGAGAGAGATTAC	qRT-PCR (109 bp)
	7a-R	GAAGGCGTAGTCATAGGGATTATT	
MYH-7b	7b-F	AAGCACAGATCCCCGATTTC	qRT-PCR (125 bp)
	7b-R	GCTTCTAGACGCTCCCTATCT	
β -actin	β -actinF	GATCTGGCATCACACCTTCTAC	qRT-PCR (104 bp)
	β -actinR	TCTTCTCCCTGTTGGCTTTG	

GSP: 基因特异引物; NGSP: 基因特异巢式引物; UPM: RACE 通用引物; UPS: RACE 巢式通用引物

GSP: Gene-specific primers; NGSP: Nested gene specific primers; UPM: Universal primer mix; UPS: Universal primer short

1.5 序列分析

用 BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) 对测序结果进行同源性检索。NCBI ORF Finder (<http://www.ncbi.nlm.nih.gov/projects/gorf/>) 确定其开放阅读框并预测氨基酸序列。用 ProtParam 工具 (<http://web.expasy.org/protparam/>) 对推测蛋白进行基本理化参数分析; 根据 PredictProtein 工具 (<https://www.predictprotein.org/>) 分析蛋白二级结构及功能; interProScan 工具 (<http://www.ebi.ac.uk/interpro/scan.html>) 分析结构域; 用 NLS Mapper (http://nls-mapper.iab.keio.ac.jp/cgi-bin/NLS_Mapper_form.cgi) 在线预测核定位序列。用 ClustalW 软件将翘嘴鲌及其他物种肌球蛋白重链基因的氨基酸头区序列进行多重比对, 用 MEGA 5.0 软件构建 M-L(Maximum Likelihood) 系统发育树(Bootstraps 设为 1000 次)。

分析结构域; 用 NLS Mapper (http://nls-mapper.iab.keio.ac.jp/cgi-bin/NLS_Mapper_form.cgi) 在线预测核定位序列。用 ClustalW 软件将翘嘴鲌及其他物种肌球蛋白重链基因的氨基酸头区序列进行多重比对, 用 MEGA 5.0 软件构建 M-L(Maximum Likelihood) 系统发育树(Bootstraps 设为 1000 次)。

1.6 翘嘴鲌 MYH-7s 基因表达分析

使用合成的 cDNA 第 1 链作为实时荧光定量 PCR (qRT-PCR) 分析的模板, 以 β -actin 为内参基因(Zhou

et al, 2010), 分别设计目的基因及内参基因的 qRT-PCR 引物(表 1)。

qRT-PCR 采用 ABI 7300 实时荧光定量 PCR 仪 (Life Technologies, 美国), 反应体系根据 SYBR Premix Ex Taq (Tli RNaseH Plus) (TaKaRa, 日本)说明书进行设置: cDNA 模板 2 μ l, 上、下游引物(10 μ mol/L) 各 0.4 μ l, 2 \times SYBR Premix Ex Taq 10 μ l, 50 \times ROX Reference Dye 0.4 μ l, ddH₂O 6.8 μ l。反应程序: 95 $^{\circ}$ C 1 min, 40 个循环: 95 $^{\circ}$ C 30 s, 55 $^{\circ}$ C 30 s, 72 $^{\circ}$ C 30 s。熔解曲线: 95 $^{\circ}$ C 15 s, 55 $^{\circ}$ C 15 s, 95 $^{\circ}$ C 15 s。采用 2^{- Δ ACT} 方法计算目的基因相对表达量。用 SPSS 20.0 软件进行单因素方差分析(One-way ANOVA), 利用 Origin

6.0 对统计结果进行作图。

2 结果

2.1 转录组差异表达基因筛选

使用 Hiseq 2000 测序平台分析快长组与慢长组翘嘴鲈的转录组, 共获 41641 条 Unigene (平均长度为 902 nt), 经数字基因表达谱 DGE 分析, 筛选出 549 个表达量有显著差异 Unigene, 其中 310 个上调(图 1)。这 549 个 Unigene 中, 有 43 条目的 annotation 注释为肌球蛋白重链, 其中, 快长组与慢长组相比, 有 32 个显著上调, 11 个显著下调。

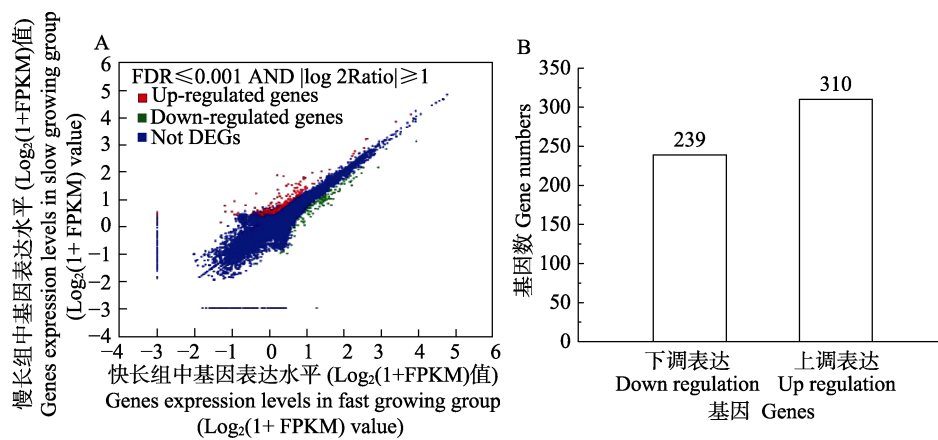


图 1 翘嘴鲈生长快、慢群体中差异表达基因分布

Fig.1 The distribution of differentially expressed genes in fast and slow growing groups of *S. chuatsi*

A: 所有的 Unigene 表达量分布; B: 表达量差异显著基因数统计

FDR: 错误发生率; FPKM: 每千个碱基的转录每百万映射读取的碎片, 表征表达量高低

A. The expression distribution of all unigenes; B. Genes number of significantly different expression

FDR: False discovery rate; FPKM: Fragments per kilobase of exon per million fragments mapped (represent the expression level)

2.2 翘嘴鲈肌球蛋白重链基因的克隆与序列结构分析

以 5'/3'RACE cDNA 为模板分别进行 5'/3'端 RACE, 获得 2 个 *MYH* 基因 cDNA 全长。NCBI BLAST 分析显示, 2 个基因均与 *MYH-7* 基因高度同源, 命名为 *MYH-7a* (以 Unigene9779_All 为模板, GenBank 登录号: KX085225)和 *MYH-7b* (以 CL2161.Contig6_All 为模板, GenBank 登录号: KX085226)亚型。*MYH-7a* 全长为 6071 bp, 5'UTR 为 107 bp, 3'UTR 为 143 bp; ORF 长为 5820 bp (109–5928), 编码 1939 个氨基酸, 推测分子量为 223 kDa, 等电点为 5.57。*MYH-7b* 全长为 5896 bp, 5'UTR 为 95 bp, 3'UTR 为 146 bp; ORF 长为 5745 bp (96–5840), 编码 1914 个氨基酸, 推测分子量为 220 kDa, 等电点为 5.62。二者核酸相似性为 83.71%, 氨基酸相似性为 89.12%。

核苷酸序列同源性比较显示, *MYH-7a* 与深裂眶

锯雀鲷(*Stegastes partitus*)、尼罗罗非鱼(*Oreochromis niloticus*)、斑马宫丽鱼(*Maylandia zebra*)、伯氏朴丽鱼(*Haplochromis burtoni*)、红丽鱼(*Pundamilia nyererei*)和红鳍东方鲀(*Takifugu rubripes*)等的 *MYH-7* 的相似性极高, 均高于 90%。*MYH-7b* 与已报道的翘嘴鲈慢肌 *MYH* (KF601703.1)相似性最高(97%), 与大黄鱼(*Larimichthys crocea*) *MYH-7* 转录剪接体 1 和剪接体 2、布氏新亮丽鲷(*Neolamprologus brichardi*)、深裂眶锯雀鲷、斑马宫丽鱼和红丽鱼等的 *MYH-7* 的相似性均高于 90%。进一步与其他物种的 *MYH-7a*、*MYH-7b* 氨基酸序列推断的头区进行同源比对, 结果显示, 在头区中都有与“马达”功能相关的关键结构域 Loop1 和 Loop2 环。这 2 个 Loop 环区上各有 1 个 ATP 结合位点, 在氨基酸序列起始的 175 aa (*MYH-7b* 是 176)和 403 aa (*MYH-7b* 是 404)处还各有 1 个 ATP 结合位点(图 2)。

根据肌球蛋白重链头区氨基酸序列构建的系统



图 2 翘嘴鲌与其他鱼类 MYH-7s 的头区保守氨基酸位点的比对

Fig.2 Amino acid sequence alignments of the myosin head region of MYH-7s of *S. chuatsi* and other fish species

灰色区为 ATP 结合位点; 方框为 P-loop 环; 方框上为 loop 环的保守序列; Sc-s: 翘嘴鲌慢肌球蛋白重链亚型 1, AHB33632.1; Dr-v: 斑马鱼慢肌球蛋白亚型 1, AAI63721.1; Sp-MYH7: 深裂眶锯雀鲷 MYH-7, XP_008303539.1; On-MYH7: 尼罗罗非鱼 MYH-7, XP_013131789.1; Mz-MYH7: 斑马宫丽鱼 MYH-7, XP_012778396.1; Tr-MYH7: 红鳍东方鲀 MYH-7, XP_011614688.1

The gray regions represented ATP binding sites; the black box represented P-loop; the amino acids sequence above the box represented the conserved sequence of the P-loop; Sc-s: *S. chuatsi* slow skeletal muscle myosin heavy chain isoform 1, AHB33632.1; Dr-v: *D. rerio* slow myosin heavy chain 1, AAI63721.1; Sp-MYH7: *S. partitus* MYH-7, XP_008303539.1; On-MYH7: *O. niloticus* MYH-7, XP_013131789.1; Mz-MYH7: *M. zebra* MYH-7, XP_012778396.1; Tr-MYH7: *T. rubripes* MYH-7, XP_011614688.1

进化树, 形成快肌肌球蛋白重链组和慢肌(或心肌)肌球蛋白重链组两大分支。MYH-7a 先与红鳍东方鲀的 MYH-7 聚类, MYH-7b 则与鳊鱼、大黄鱼、短鳍花鲈(*Poecilia mexicana*)的慢肌球蛋白重链聚类, 二者同归慢肌球蛋白重链分支(图 3)。

2.3 蛋白理化性质及结构分析

采用 InterPro 分析蛋白结构域发现, MYH-7a 与 MYH-7b 主要结构域相似, 均可分为头区和尾区。头区均有一种选择性结合多脯氨酸配体的结构域 Src 同源区 3 (Src homology 3, SH3), 以及 MLC、钙调蛋白或 EF 手相蛋白超家族结合位点 IQ 基序(IQ motif)。

Predict Protein 分析发现, MYH-7a 和 MYH-7b 有相似的结构组成, 均有 77.2%的螺旋结构, 5.0%的直链, 17.8%的 loop 环(Loop), 未发现二硫键形成位点。2 个 MYH-7s 均在各自的 178–185 aa 处有 8 个氨基酸(GESGAGKT)为 ATP/GTP 结合活性位点 A [ATP/GTP-binding site motif A (P-loop)]; 942–971 aa 处 30 个氨基酸(KKRKLEDECSSELKDDIDLELTLAKV-EKEK)为核定位序列(Nuclear localization sequence,

NLS)。此外, MYH-7b 还在其 632–639 aa 处一段 8 个氨基酸(GGKGGKGGG)为核定位序列(图 4)。

进一步用 NLS Mapper 在线预测与 MYH-7a 相似性较高的伯氏朴丽鱼(XP_005938260.1)、斑马宫丽鱼(XP_012778396.1)和尼罗罗非鱼(XP_013131789.1)等的 MYH-7 序列, 以及与 MYH-7b 相似性较高的斑马宫丽鱼(XP_014266369.1)、非洲齿鲤(*Nothobranchius furzeri*)(XP_015812773.1)和斑马鱼(*Danio rerio*)(XP_009292203.1)等的 MYH-7 序列, 均在其 C 端对应位置找到 NLS 序列。

2.4 MYH-7s 基因的表达量变化分析

2.4.1 翘嘴鲌 MYH-7s 的表达验证 根据转录组数据计算 MYH-7a、MYH-7b 的 FPKM 值, 在快长组 2 个基因的表达量均显著高于慢长组($P < 0.01$)(图 5-A)。以用于转录组测序的同批 cDNA 为模板, qRT-PCR 检测验证这 2 个基因的表达量差异, 显示快长组表达量均显著高于慢长组(图 5-B)。qRT-PCR 结果与转录组差异表达谱分析结果一致。

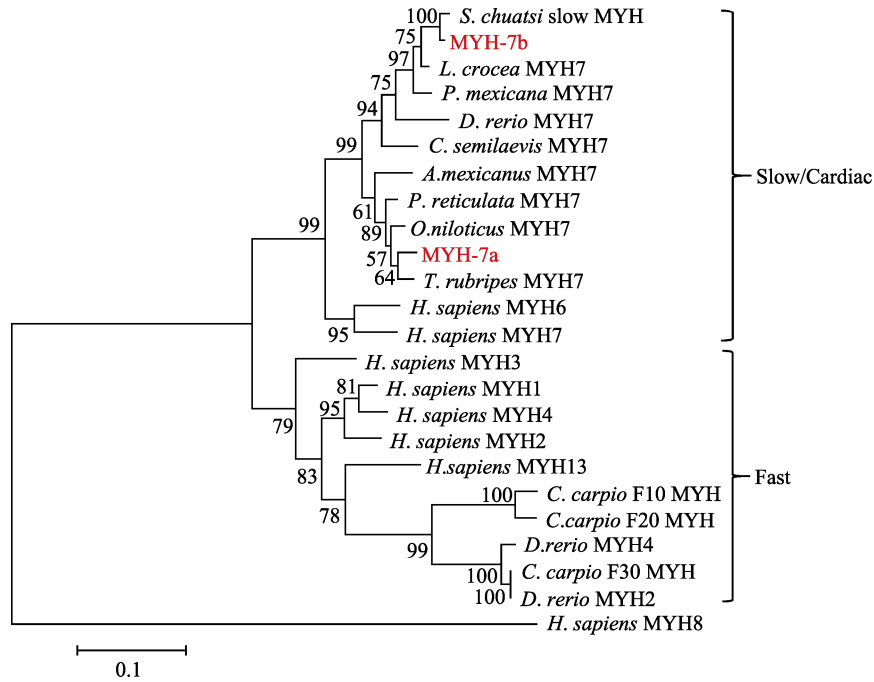


图 3 基于翘嘴鲌及其他物种 MYH-7s 头区序列的系统进化树

Fig.3 Phylogenetic tree based on head domain amino acids of myosin of *S. chuatsi* and other species

图中氨基酸序列的物种名和相应的登录号分别为(The species and GenBank accession numbers of amino acid sequences were listed as follows): 翘嘴鲌慢肌 MYH (*S. chuatsi* slow MYH), KF601703.1; 大黄鱼 MYH-7 (*L. crocea* MYH-7), KKF11324.1; 短鳍花鲈 MYH-7 (*P. mexicana* MYH-7), XP_014864995.1; 半滑舌鳎 MYH-7 (*C. semilaevis* MYH-7), XP_008332930.1; 墨西哥脂鲤 MYH-7 (*A. mexicanus* MYH-7), XP_015456340.1; 网纹鲈 MYH-7 (*P. reticulata* MYH-7), XP_008431929.1; 尼罗罗非鱼 MYH-7 (*O. niloticus* MYH-7), XM_013276335.1; 红鳍东方鲀 MYH-7 (*T. rubripes* MYH-7), XM_011616386.1; 鲤 10℃、20℃、30℃ 适应型 MYH (*C. carpio* 10℃, 20℃, 30℃ acclimated MYH), D89990、D89991、D89992; 斑马鱼 MYH2、4、7 (*D. rerio* MYH2, 4, 7), NM_152982、AY921650、XP_009292203.1; 人 MYH 1-4、6-8、13 (*H. sapiens* MYH 1-4, 6-8, 13), AF111785、AF111784、NM_002470.3、AF11783、NM_002471、NM_000257、NM_002474、AF111782

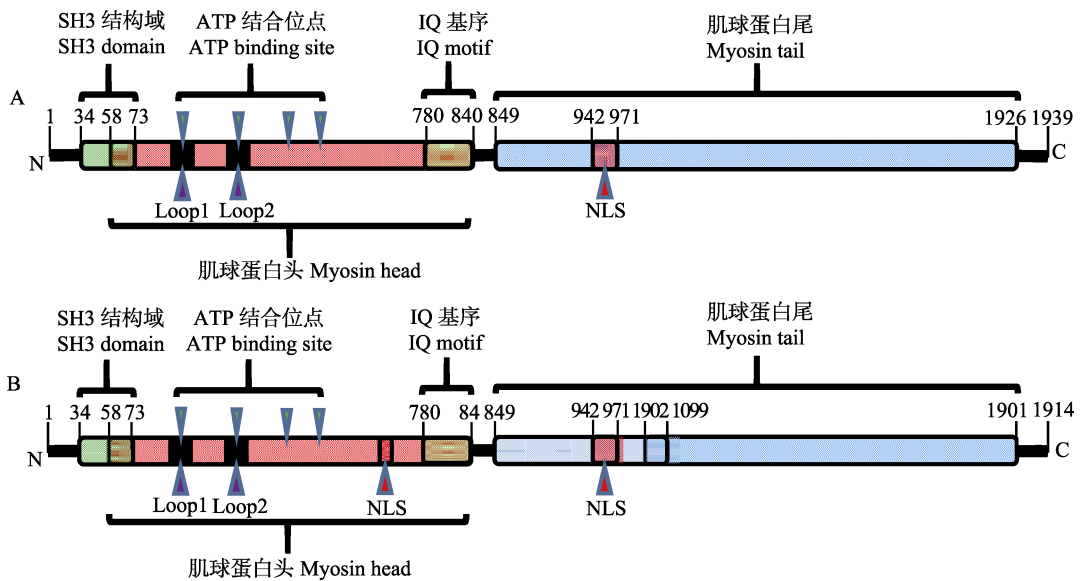


图 4 2 个翘嘴鲌 MYH-7s 氨基酸序列结构模式

Fig.4 Structure models of the two MYH-7s amino acid sequence of *S. chuatsi*

A: MYH-7a; B: MYH-7b

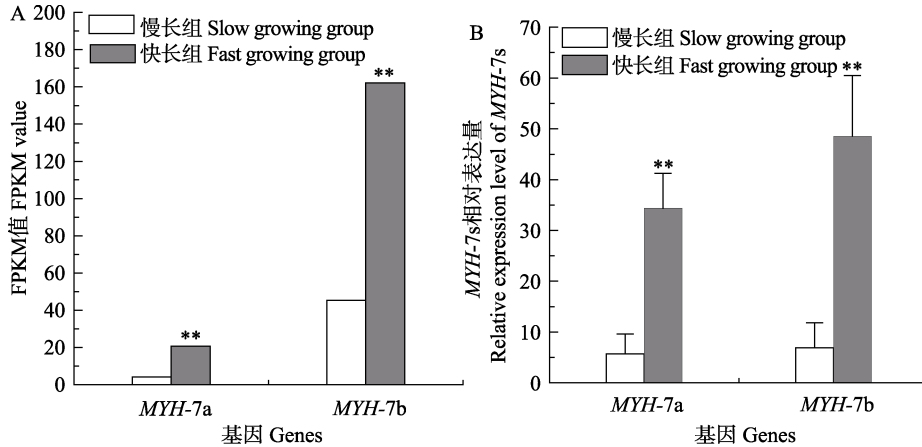


图 5 翘嘴鲌快长组和慢长组样本的 MYH-7s 表达量
Fig.5 Expression level of two MYH genes of *S. chuatsi* in fast and slow growing groups measured by transcriptome sequencing and qRT-PCR

A: 转录组测序; B: 实时荧光定量 PCR; **表示差异极显著($P<0.01$), 下同

A: Transcriptome sequencing; B: qRT-PCR; ** denoted highly significant difference ($P<0.01$), the same as below

2.4.2 翘嘴鲌 MYH-7s 在成鱼各组织中的表达 qRT-PCR 分析 2 个 MYH 基因在翘嘴鲌成鱼 14 个组织中的表达, 结果显示, MYH-7a 在心肌中表达量最高, 其次是红肌、眼和皮肤, 在肝、脾、肠和胃中低表达; MYH-7b 在红肌中表达量最高, 其次是皮肤、垂体和眼, 在鳃、性腺和心肌中低表达(图 6)。

况。MYH-7a 在未受精卵中有微量表达, 在尾芽期表达量明显上调, 并随着孵化出膜, 表达量不断增加。MYH-7b 在未受精卵中检测到表达, 从受精开始表达, 随着胚胎发育整体呈上升趋势(差异均不显著)。

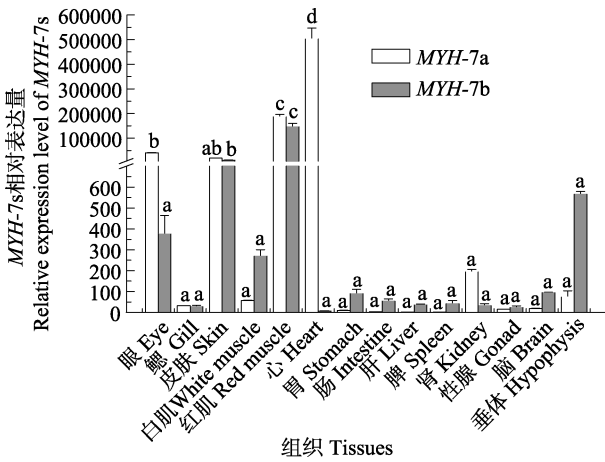


图 6 qRT-PCR 检测翘嘴鲌不同组织中 MYH-7s 的表达
Fig.6 qRT-PCR analysis of MYH-7s expression in tissues of adult *S. chuatsi*

同一基因中不同字母表示差异显著($P<0.05$), 相同字母表示差异不显著($P>0.05$), 下同

In the same gene expression feature, different letters represented significant difference ($P<0.05$), the same letter showed no significant difference ($P>0.05$), the same as below

2.4.3 翘嘴鲌 MYH-7s 在胚胎发育期的表达 图 7 显示在翘嘴鲌胚胎发育各时期 2 种 MYH-7s 的表达情

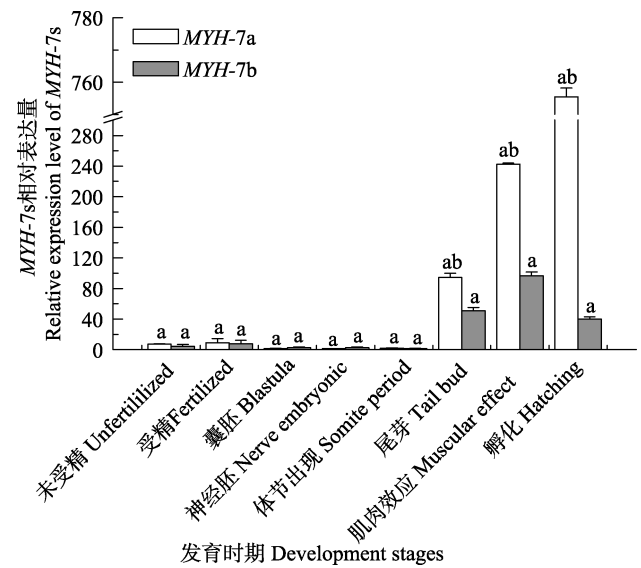


图 7 MYH-7s 在翘嘴鲌各发育时期中的表达水平
Fig.7 The expression levels of MYH-7s in *S. chuatsi* at different developing stages

2.4.4 MYH-7s 在翘嘴鲌早期生长阶段的表达情况 定量检测 2 个 MYH-7s 在孵化后 15、30 和 60 dph 快长组和慢长组的翘嘴鲌白肌中的表达, 2 个 MYH-7s 在 15 和 30 dph 快长组白肌中表达量均显著高于慢长组, 而到 60 dph 时, 快长组中的表达量均显著低于慢长组(图 8)。

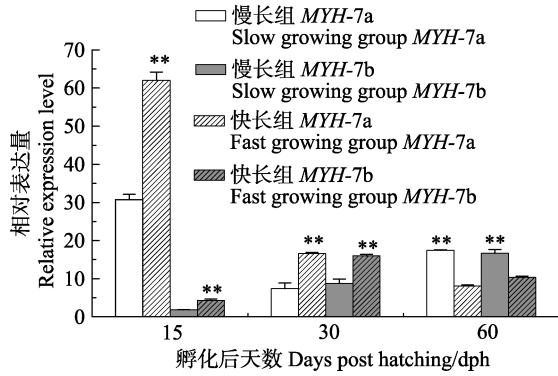


图8 孵化后不同生长时期的翘嘴鲌白肌中 MYH-7s 的表达

Fig.8 The expression pattern of MYH-7s in white muscle of *S. chuatsi* during different growth periods

3 讨论

3.1 翘嘴鲌 MYH-7s 蛋白结构分析

本研究克隆获得的翘嘴鲌 MYH-7a、MYH-7b 均具有 MYH 家族蛋白的典型结构:由头部和尾部组成,头部有 ATP 结合位点、Loop1 和 Loop2 环区以及头部末端的 IQ 基序等关键结构。MYH 主要由 N 端球状盘绕的头部和 C 端 α -螺旋型的尾部组成。头区是 Myosin 动力产生的主要区域,具有 ATP 酶活性位点,可与 ATP 结合,水解 ATP 为 ADP,释放能量,供头部迁移、滑动,是肌球蛋白“分子马达”功能的关键结构区,也是 MYH 家族中的重要保守区(Sebé-Pedrós *et al.*, 2014);头区的 2 个功能性表面环状结构 Loop1 和 Loop2, Loop1 表面环结构靠近 ATP 结合位点,决定肌纤维收缩速率和动力持续;Loop2 则可与肌动蛋白(actin)的非极性末端反应,影响 MYH 上由 actin 激活的 ATP 酶水解活性(Holmes *et al.*, 2002);IQ 基序是 MYH 与 MLC 以及钙调蛋白结合区,因其保守氨基酸 IQxxxRGxxxR(x 指任何一种氨基酸残基)的起始 2 个残基得名;Ca²⁺与 MYH 上 IQ 基序的结合可调节细胞定向运动(如肌细胞收缩),IQ 基序也是 MYH 家族蛋白的重要特征结构之一(Bähler *et al.*, 2002)。此外,在 2 个 MYH-7s 还发现 SH3 结构域和核定位序列等特殊结构。

SH3 是一种选择性结合多脯氨酸配体的结构域,在多种生物进程中有重要作用,包括分子内相互作用调节酶活性、调节信号通路中某些成员的局部浓度和亚细胞定位以及介导一些大型多重复合蛋白的组装,在低等的酵母菌到人等物种的数百种蛋白中广泛存在(Mayer, 2001)。含 SH3 结构的 MYH 可通过 SH3 与含有 Arp2/3 复合体和富脯氨酸区的蛋白结合,通

过其上的 Arp2/3 介导肌动蛋白丝的聚合,对细胞骨架维持和细胞的内吞作用等有重要作用(Okrut *et al.*, 2015)。酵母的 2 个 MYH1 亚型: Myo3p 和 Myo5p 的 SH3 结构域的缺失可导致肌动蛋白细胞骨架的严重缺陷(Evangelista *et al.*, 2000);鸡(*Gallus gallus*)胸肌中一种 MYH 的 SH3 可促进肌 ELC 的 N 端延伸与肌动蛋白结合,调节肌纤维运动(Lowey *et al.*, 2007)。SH3 是只在 MYH-1、2、4、7、15 等亚型中发现的特征结构之一(Sebé-Pedrós *et al.*, 2014)。本研究获得的 2 个翘嘴鲌 MYH-7s 均有 SH3 结构,其对细胞骨架维持和肌纤维运动等的功能尚待进一步研究。

核定位序列(NLS)作为蛋白质的一个结构域,与入核载体相互作用,使蛋白能被运进细胞核的特定位置发挥作用。已报道具有 NLS 的 MYH 有 NM1(myosin I C)、myosin VI、myosin X VI B 和 myosin X VIII B (Sobczak *et al.*, 2009)。本研究通过软件预测,在大多数鱼的 MYH-7 上也发现 NLS。Dzijak 等(2012)发现,在 NM1 的第 2 个 IQ 基序中,13 aa 的 GRRKAAKRK-WAAQ(754-766)是其必需的 NLS, NM1 可通过此结构进入细胞核,参与核内转录、染色质移动和染色质重塑等关键核运动。本研究所获得的 2 个 MYH-7s 的 C 端均有核定位序列,推测 MYH-7a、MYH-7b 蛋白也可能通过 NLS 入核参与核内运输与转录调控等过程,但具体作用还有待进一步研究验证。

3.2 翘嘴鲌 MYH-7s 基因的表达分析

MYH 具有组织和时期表达特异性。研究发现,MYH-7 主要在哺乳动物和鱼的胚胎、幼鱼期表达,在成体则主要是心室肌和慢肌中表达;而快肌中受甲状腺激素信号通路的抑制,表达量低(van Rooij *et al.*, 2009; Schiaffino *et al.*, 2015)。Spurgeon 等(2008)检测发现,人的 MYH-7 主要在心肌和慢肌中表达,在肠、脑、前列腺和睾丸中有一定表达。Ikeda 等(2007)将红鳍东方鲀基因组 DNA 序列与人 DNA 序列比较发现,红鳍东方鲀与人同源的 MYH-7 主要在慢肌、心肌和皮肤中高表达,此外也在眼、胚胎(168 hours post fertilization, hpf)和快肌中有表达。而 Akolkar 等(2010)在红鳍东方鲀成鱼快肌中未检测到 MYH-7 的表达,在外侧浅肌和深层的竖肌与降肌 2 种慢肌中有表达,在心肌中高表达。本研究中,MYH-7a、MYH-7b 在心肌、红肌等部位高表达,与上述 MYH-7 的组织分布情况相符。

Spurgeon 等(2008)发现,MYH-7 在人胎儿期的肝脏和肌肉中均有表达,且肌肉中表达量显著高于肝脏。在大鼠(*Rattus norvegicus*)和小鼠(*Mus musculus*)

胚胎发育中主要表达胚胎型的 *MYH-3*、神经型的 *MYH-8* 和慢肌型的 *MYH-7* (Schiaffino *et al*, 2015)。在小鼠胚胎发育中, 其 *MYH-7a* 在受精后第 9 天的胚胎期即见表达; 鸡的 *MYH-7* 则最早出现在神经分化早期(Schiaffino *et al*, 2015; Nguyen *et al*, 2015)。用 qRT-PCR 检测, 红鳍东方鲀 4–8 dph 的胚胎发育期均可见 *MYH-7* 表达, 幼鱼期的 8–16 dph 也有表达(Asaduzzaman *et al*, 2013)。在鲤中, 心肌型 *MYH* (也归为 *MYH-7*) 最早在 61 hph 检测到表达, 且随着胚胎发育表达量逐渐升高(Nihei *et al*, 2006)。本研究中, *MYH-7a*、*MYH-7b* 在未受精卵即有表达, 且随着胚胎发育, 表达量呈上升趋势, 提示 2 种 *MYH-7* 在翘嘴鲌胚胎发育过程具有重要作用。

MYH 基因从早期的慢型逐渐被快型 *MYH* 替代, 与动物早期生长有很大相关性。在大鼠早期发育中, 在出生几天后才能在腿部骨骼肌中检测到成体高表达的快肌型 *MYH* (*MYH-1*、2 和 4), 而在后续的几周内, 快肌型 *MYH* 大量表达并迅速替代胚胎型和慢型 *MYH* (*MYH-3*、7) (Schiaffino *et al*, 2015)。在小鼠 C2C12 成肌细胞系, 用低血清浓度诱导细胞分化, 首先表达 *MYH-7*、*MYH-3* 和 *MYH-8*, 3 种 *MYH* 在第 1 天即见表达, 到 2–4 d 达到峰值, 之后降低; 而快肌型的 *MYH-1* 和 *MYH-4* 在 2–4 d 才检测到表达, 但表达量不断增加(Brown *et al*, 2012)。鼠机体和细胞中这种 *MYH* 由慢型到快型的表达转变受甲状腺激素的调节(Soukup *et al*, 2015a)。在鱼类中, 特别是幼鱼生长阶段, *MYH* 快慢亚型表达量变化可能与肌肉类型在白肌(快肌)、红肌(慢肌)间的转变有关, 在斑马鱼、河鲈(*Perca fluviatilis*)幼鱼中均有发现(McGuigan *et al*, 2004)。Soukup 等(2015b)认为, 慢肌球蛋白在鱼胚胎期和幼鱼期表达量较高, 而随着鱼类生长, 身体其他部位的慢型 *MYH* 表达量逐渐降低, 被快型 *MYH* 取代, 而慢型 *MYH* 在成体慢肌存在的部位如心脏、红肌表达量持续增加。在孵化后 1 个月和 7 个月的鲤躯干肌肉中也能检测到 *MYH-7* 表达, 但 7 个月时的表达量低于 1 个月的(Nihei *et al*, 2006)。本研究结果显示, *MYH-7a*、*MYH-7b* 在孵化后 30–60 dph 的生长时期, 在白肌中表达量有明显的降低, 与鱼类中肌纤维由慢型向快型转变相符, 在对应时期中慢长组 *MYH-7* 表达量高于快长组, 说明 2 个 *MYH-7* 表达水平的变化与鱼生长进程以及生长速度有极大的相关性。

4 结论

本研究从翘嘴鲌快长组和慢长组的转录组差异

表达基因中筛选到 2 个 *MYH* Unigenes (*MYH-7a*、*MYH-7b*), 进行了 cDNA 序列克隆、结构分析以及 qRT-PCR 表达水平分析。这 2 个基因具有不同的组织分布, 在胚胎发育过程早期均有表达且随着发育的推进表达水平明显提高, *MYH-7a* 的表达量明显高于 *MYH-7b*; 在不同生长阶段的快长组与慢长组翘嘴鲌白肌中, 2 个基因的表达量在 15 dph 和 30 dph 快长组均显著高于慢长组, 而 60 dph 快长组显著低于慢长组。研究结果显示, 这 2 个基因在翘嘴鲌早期发育过程中发挥着重要的作用, 其表达量高低与翘嘴鲌生长快慢相关。

参 考 文 献

- Akolkar DB, Kinoshita S, Yasmin L, *et al*. Fibre type-specific expression patterns of myosin heavy chain genes in adult torafugu *Takifugu rubripes* muscles. *Journal of Experimental Biology*, 2010, 213(1): 137–145
- Asaduzzaman M, Ahammad AKS, Asakawa S, *et al*. 5'-flanking sequences of zebrafish fast myosin heavy chain genes regulate unique expression in the anterior, medial subsection and posterior tail somites of the skeletal muscle. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, 2016, 191: 1–12
- Asaduzzaman MD, Akolkar DB, Kinoshita S, *et al*. The expression of multiple myosin heavy chain genes during skeletal muscle development of torafugu *Takifugu rubripes* embryos and larvae. *Gene*, 2013, 515(1): 144–154
- Bähler M, Rhoads A. Calmodulin signaling via the IQ motif. *FEBS Letters*, 2002, 513(1): 107–113
- Biga PR, Goetz FW. Zebrafish and giant danio as models for muscle growth: Determinate vs. indeterminate growth as determined by morphometric analysis. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 2006, 291(5): R1327–R1337
- Brown DM, Parr T, Brameld JM. Myosin heavy chain mRNA isoforms are expressed in two distinct cohorts during C2C12 myogenesis. *Journal of Muscle Research and Cell Motility*, 2012, 32(6): 383–390
- Campion LA, Choi S, Mistry HL, *et al*. Myosin heavy chain and parvalbumin expression in swimming and feeding muscles of *centrarchid* fishes: The molecular basis of the scaling of contractile properties. *Comparative Biochemistry and Physiology Part A: Molecular and Integrative Physiology*, 2012, 163(2): 223–230
- Churova MV, Meshcheryakova OV, Veselov AE, *et al*. Activity of enzymes involved in the energy and carbohydrate metabolism and the level of some molecular-genetic characteristics in young salmon (*Salmo salar* L.) with

- different age and weight. *Russian Journal of Developmental Biology*, 2015, 46(5): 254–262
- Dhillon RS, Wang Y, Tufts BL. Using molecular tools to assess muscle growth in fish: Applications for aquaculture and fisheries management. *Comparative Biochemistry and Physiology Part C: Toxicology and Pharmacology*, 2008, 148(4): 452
- Dzijak R, Yildirim S, Kahle M, *et al.* Specific nuclear localizing sequence directs two myosin isoforms to the cell nucleus in calmodulin-sensitive manner. *PLoS One*, 2012, 7(1): e30529
- van Rooij R, Quiat D, Johnson BA, *et al.* A family of microRNAs encoded by myosin genes governs myosin expression and muscle performance. *Developmental Cell*, 2009, 17(5): 662–673
- Evangelista M, Klebl BM, Tong AHY, *et al.* A role for myosin- I in actin assembly through interactions with Vrp1p, Bee1p, and the Arp2/3 complex. *Journal of Cell Biology*, 2000, 148(2): 353–362
- Hevrøy EM, Jordal AEO, Hordvik I, *et al.* Myosin heavy chain mRNA expression correlates higher with muscle protein accretion than growth in Atlantic salmon, *Salmo salar*. *Aquaculture*, 2006, 252(2–4): 453–461
- Holmes JM, Whiteley NM, Magnay JL, *et al.* Comparison of the variable loop regions of myosin heavy chain genes from Antarctic and temperate isopods. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, 2002, 131(3): 349–359
- Ikeda D, Ono Y, Snell P, *et al.* Divergent evolution of the myosin heavy chain gene family in fish and tetrapods: Evidence from comparative genomic analysis. *Physiological Genomics*, 2007, 32(1): 1–15
- Liu XL, Bin SY, Wang KZ, *et al.* Artificial propagation and embryonic development observation of mandarin fish. *Journal of Guangxi Normal University: Natural Science Edition*, 2013, 31(2): 100–106 [刘希良, 宾石玉, 王开卓, 等. 翘嘴鲈的人工繁殖与胚胎发育观察. *广西师范大学学报: 自然科学版*, 2013, 31(2): 100–106]
- Lowey S, Saraswat LD, Liu HJ, *et al.* Evidence for an interaction between the SH3 domain and the N-terminal extension of the essential light chain in class II myosins. *Journal of Molecular Biology*, 2007, 371(4): 902–913
- Mayer BJ. SH3 domains: Complexity in moderation. *Journal of Cell Science*, 2001, 114(7): 1253–1263
- McGuigan K, Phillips PC, Postlethwait JH. Evolution of sarcomeric myosin heavy chain genes: Evidence from fish. *Molecular Biology and Evolution*, 2004, 21(6): 1042–1056
- Mortazavi A, Williams BA, McCue K, *et al.* Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods*, 2008, 5(7): 621–628
- Nguyen K, Hall AL, Jones J M. Expression of myosin VIIA in the developing chick inner ear neurons. *Gene Expression Patterns*, 2015, 19(1): 36–44
- Nihei Y, Kobiyama A, Ikeda D, *et al.* Molecular cloning and mRNA expression analysis of carp embryonic, slow and cardiac myosin heavy chain isoforms. *Journal of Experimental Biology*, 2006, 209(1): 188–198
- Okrut J, Prakash S, Wu Q, *et al.* Allosteric N-WASP activation by an inter-SH3 domain linker in Nck. *Proceedings of the National Academy of Sciences*, 2015, 112(47): E6436–E6445
- Overturf K, Hardy RW. Myosin expression levels in trout muscle: A new method for monitoring specific growth rates for rainbow trout *Oncorhynchus mykiss* (Walbaum) on varied planes of nutrition. *Aquaculture Research*, 2001, 32(4): 315–322
- Schiaffino S, Rossi AC, Smerdu V, *et al.* Developmental myosins: Expression patterns and functional significance. *Skeletal Muscle*, 2015, 5(1): 1–14
- Sebé-Pedrós A, Grau-Bové X, Richards TA, *et al.* Evolution and classification of myosins, a pan-eukaryotic whole-genome approach. *Genome Biology and Evolution*, 2014, 6(2): 290–305
- Sobczak M, Majewski Ł, Redowicz MJ. Myosins in nucleus. *Postepy Biochemii*, 2009, 55(2): 239–246
- Soukup T, Diallo M. Proportions of myosin heavy chain mRNAs, protein isoforms and fiber types in the slow and fast skeletal muscles are maintained after alterations of thyroid status in rats. *Physiological Research*, 2015a, 64(1): 111–118
- Soukup T, Smerdu V. Effect of altered innervation and thyroid hormones on myosin heavy chain expression and fiber type transitions: A mini-review. *Histochemistry and Cell Biology*, 2015b, 143(2): 123–130
- Spurgeon SL, Jones RC, Ramakrishnan R. High throughput gene expression measurement with real time PCR in a microfluidic dynamic array. *PLoS One*, 2008, 3(2): e1662
- Yasmin L, Kinoshita S, Asaduzzaman M, *et al.* A 5'-flanking region of embryonic-type myosin heavy chain gene, *MYH_{M743-2}*, from torafugu *Takifugu rubripes* regulates developmental muscle-specific expression. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, 2011, 6(1): 76–81
- Zhou RX, Meng T, Meng HB, *et al.* Selection of reference genes in transcription analysis of gene expression of the mandarin fish, *Siniperca chuatsi*. *Zoological Research*, 2010, 31(2): 141–146

cDNA Cloning and Analyses of Two Myosin Heavy Chain Isoforms of Mandarin Fish (*Siniperca chuatsi*) Based on Transcriptome Sequencing

CHEN Zhihang^{1,2}, DONG Junjian¹, SUN Chengfei¹, TIAN Yuanyuan¹, YE Xing^{1,2}^①

(1. Pearl River Fisheries Research Institute, Chinese Academy of Fishery Sciences, Guangzhou 510380;

2. College of Fisheries and Life Science, Shanghai Ocean University, Shanghai 201306)

Abstract Myosin heavy chain (MYH) is the main component of the skeletal muscle thick filament. The expression levels of MYHs have significant influences on the formation of thick filaments and muscle development. To understand the roles of MYHs on growth during the early development of mandarin fish (*Siniperca chuatsi*), the differentially expressed unigenes from the transcriptome sequencing data of the fast and slow growing fish were analyzed, and two highly expressed unigenes of MYH were screened. The full length cDNAs of these two MYH-7s, MYH-7a and MYH-7b were obtained by RACE method; they were 6070 bp and 5896 bp in length, with coding regions of 5820 bp and 5745 bp, respectively. Sequencing analysis showed that both MYH-7s contain the conserved regions such as Loop1 and Loop2, and ATP binding sites. The evolutionary tree clustering analysis showed that both MYH-7s genes belong to the slow muscle type MYH. The expression patterns of the two MYH-7s in different tissues were detected by qRT-PCR. The highest expression of MYH-7a was found in heart muscle while MYH-7b was highly expressed in red muscle. The expression levels of both MYH-7a and MYH-7b increased gradually in accordance with the embryonic development stages. The expression levels of the MYH-7s in the fast growing group were significantly higher than in the slow growing groups in 15, 30-day post hatching (dph), while both genes showed lower expression levels in fast growing group than in slow growing group in 60 dph. The different expression levels in tissues and early developmental stages of these two genes in the fast and slow growing groups suggested that they play important roles in the embryonic and early development of *S. chuatsi*.

Key words Mandarin fish (*Siniperca chuatsi*); Transcriptome sequencing; Myosin heavy chain; cDNA cloning; Expression analysis

① Corresponding author: YE Xing, E-mail: gzyexing@163.com