

文章编号:1000-0615(2009)01-0001-08

大口黑鲈 *MyoD* 基因结构和单核苷酸多态性位点的筛选

于凌云^{1,2}, 白俊杰¹, 叶 星¹, 李胜杰¹, 李小慧¹

(1. 中国水产科学研究院珠江水产研究所, 中国水产科学院热带亚热带鱼类选育
与养殖重点开放实验室, 广东 广州 510380;
2. 上海海洋大学水产与生命学院, 上海 201306)

摘要:采用 PCR 技术和基因组步移技术从大口黑鲈基因组 DNA 中扩增得到 *MyoD* 基因及其 5' 调控区序列。该基因序列全长 3 797 bp, 其中 5' 调控区长 1 077 bp, *MyoD* 基因转录区由 3 个外显子(分别为 591、81 和 78 bp)和 2 个内含子(分别为 1 077 和 486 bp)组成。5' 调控区含有与肌肉特异性基因转录密切相关的转录调控元件 E box、肌细胞增强因子 2(MEF2)、肌肉特异性金属硫蛋白结合位点(MTBF)及一些转录反应调控元件(TATA box、OCAAT box、OCT1、PRE、AP4、Pit1)。运用 PCR-SSCP 技术和直接测序法进行大口黑鲈 *MyoD* 基因 SNP 位点筛选, 结果表明 *MyoD* 基因序列中存在 7 个突变点, 均位于内含子上。养殖群体中这 7 个突变点分析结果显示突变比例范围在 0.042~0.353 之间。本研究结果为 SNPs 位点与大口黑鲈生长性能关联分析奠定了基础。

关键词:大口黑鲈; *MyoD*; SNPs; 筛选

中图分类号:Q 579.1; S 917

文献标识码:A

MyoD 基因是生肌调节因子 MRFs(myogenic regulatory factors) 家族的主要成员之一^[1]。MRFs 家族基因包括 *MyoD*、*MyoG*、*Myf5* 和 *MRF4* 4 种转录调节因子, 它们都具有一个保守的碱性螺旋-环-螺旋结构(bHLH)^[2]。这些转录调节因子在肌肉发生过程中具有不同的时空表达模式, 各自发挥不同作用, 个体出生后 *MyoD*、*Myf-5* 和 *MyoG* 基因仅在肌卫星细胞中表达, 并影响肌卫星细胞的活性。其中 *MyoD* 基因是脊椎动物胚胎期肌肉发育的主导调控基因之一, 对骨骼肌的形成和分化起主要作用, 可通过多个途径激活肌肉基因的转录, 从而促进成肌细胞的分化^[3-4]。相对其它生肌调节因子, *MyoD* 主要在成肌过程中起作用, 其表达对维持肌细胞分化有重要作用, *MyoD* 缺失可导致成肌细胞的增殖和分化无法进行^[5-6]。相反, *MyoD* 基因的过度表达会抑制成肌细胞的增殖过程, 并促进成肌细胞分化形成成熟的肌纤维细胞^[7]。此外, *MyoD* 基因还会通过影

响 *MyoG* 基因的活性来间接影响肌细胞的终端分化过程^[8]。目前有关人和小鼠 *MyoD* 基因的 cDNA 结构和功能研究较多^[9-11], 而对 *MyoD* 基因的 5' 调控区序列及基因全序列的结构和功能研究较少。

SNP(single nucleotide polymorphism), 即单核苷酸多态性, 是指在染色体基因组水平上单个核苷酸的变异引起的 DNA 序列多态性, 它包括单碱基的转换, 颠换、插入及缺失等形式^[12]。SNP 作为一类遗传标记以其信息含量丰富、遗传稳定等特性得以广泛应用^[13]。候选基因法是一种常用的从 DNA 水平寻找与数量性状相连锁的方法之一。许多肉质性状为数量性状, 其遗传基础受到多基因控制, 存在主基因效应^[14]。在畜禽方面, 把 *MyoD* 基因作为与肌肉性能相关的功能基因已经进行了研究。如, 朱砾^[15] 在猪上发现 *MyoD* 基因内含子 1 上存在 *Dde* I 酶切位点, 关联分析表明突变基因具有使肌纤维生长更充分, 肌

收稿日期:2008-01-12 修回日期:2008-02-08

资助项目:国家科技支撑项目(2006BAD01A1209); 国家科技基础条件平台工作项目(2005DKA21103); 广东省重大科技兴渔项目(B200701A06); 广东省自然科学基金项目(7301732)

通讯作者:白俊杰, Tel:020-81616129, E-mail: baijj2005@21cn.com

纤维变粗,面积增大的作用。田璐等^[16]在肉牛上发现 *MyoD* 基因内含子 2 上不同基因型对肉牛的宰前活重、胴体重、净肉重、高档肉重、眼肌面积等性状影响极显著或显著,而在水产动物方面把 *MyoD* 基因作为主效功能基因来研究其与水产动物肌肉生长性能的关联研究,国内、外未见有相关报道,并且再利用第三代分子标记 SNPs 技术结合与肌肉相关的主效基因——*MyoD* 基因去研究其与生产性能关联的更少。本研究旨在通过大口黑鲈 (*Micropterus salmoides*) *MyoD* 基因的克隆与单核苷酸多态性的筛选,来分析 *MyoD* 基因的结构及在该基因上寻找单核苷酸多态性,为进一步分析 *MyoD* 基因的结构变异与肌肉生长性状的关联奠定基础。

1 材料与方法

1.1 材料

样品来源 克隆基因所用的大口黑鲈购自广州黄沙市场;进行 SNP 位点筛选的大口黑鲈 24 尾及群体分析的大口黑鲈 90 尾均来自广东省大口黑鲈良种场。

主要试剂 *Taq* DNA 聚合酶体系为上海申能博彩公司产品,pMD18-T vector system 购自大连宝生生物公司;胶回收 (E. Z. N. A Gel Extraction Kit) 试剂盒为美国 OMEGA 公司产品。GenomeWalker Universal Kit 是 Clontech 公司产品。DNA 提取试剂盒购自北京天根时代公司。大肠杆菌 DH5 α 由本实验室保存。

引物合成 根据本实验室分离到的大口黑鲈 *MyoD* 基因的 cDNA 序列 (GenBank 序列号: EU367961) 设计引物 R1 和 R2 用于 5' 调控区扩增,同时设计了一对引物 F1 和 R3 用来扩增 *MyoD* 基因组序列。最后根据获得的大口黑鲈 *MyoD* 基因组序列设计了 13 对引物 (P1 ~ P12), 用于 SNPs 筛选。各引物信息见表 1。

1.2 方法

大口黑鲈基因组 DNA 的提取 采用 ACD 抗凝剂,实验鱼尾静脉活体取血,按天根离心柱型基因组 DNA 提取试剂盒介绍的方法提取样品基因组 DNA,取 100 μ L 双蒸灭菌水溶解,0.8% 的琼脂糖凝胶电泳检测 DNA 质量和浓度。检测完

毕后取 20 μ L 基因组 DNA 保存于 4 $^{\circ}$ C 供使用,余 80 μ L 基因组 DNA 保存 -20 $^{\circ}$ C 备用。

PCR 扩增

5' 调控区序列的扩增: 以特异引物 R1 和试剂盒提供的接头引物 AP1 为引物采用基因组步移法进行第一次 PCR 扩增,反应条件:94 $^{\circ}$ C、25 s,72 $^{\circ}$ C、3 min,7 个循环,接着 94 $^{\circ}$ C、25 s,67 $^{\circ}$ C、3 min,32 个循环,最后 67 $^{\circ}$ C、7 min;第二次扩增反应中利用嵌套特异引物 R2 和锚式引物 AP2 验证扩增片断大小。

***MyoD* 基因序列的扩增:** 以大口黑鲈基因组 DNA 为模板,用特异引物 F1 与 R3 扩增基因组内含子序列。反应体系按上海申能博采公司推荐的反应体系与条件进行,反应条件:预扩增 94 $^{\circ}$ C、3 min,然后 32 个循环 94 $^{\circ}$ C、50 s,57 $^{\circ}$ C、50 s,72 $^{\circ}$ C、2 min,72 $^{\circ}$ C 再延伸 7 min。

筛选 SNPs 位点的 PCR 扩增: 以广东省大口黑鲈良种场 48 尾大口黑鲈的基因组 DNA 为模板,分别利用设计的 12 对特异引物为引物 (表 1)。反应程序为:94 $^{\circ}$ C 预变性 3 min 后进入循环体系 94 $^{\circ}$ C 变性 30 s,X $^{\circ}$ C 退火 30 s,72 $^{\circ}$ C 延伸 30 s,40 个循环,最后 72 $^{\circ}$ C 延伸 7 min。

SSCP 分析 将 5 μ L PCR 产物与 9 μ L 上样缓冲液 (95% 甲酰胺,10 mmol/L EDTA,0.09% 二甲苯青,0.09% 溴酚蓝,pH 8.0) 混匀后,100 $^{\circ}$ C 变性 10 min 后迅速置于冰上冷却 5 min,160 V 电压,4 $^{\circ}$ C 下 12% 非变性聚丙烯酰胺凝胶电泳 16 ~ 22 h,银染显色。

克隆测序 对所获得的目的片断经低熔点琼脂糖凝胶电泳回收后连接到 pMD-T 载体上,筛选阳性转化子进行序列测定,测序由上海英骏生物技术有限公司完成。

2 结果与分析

2.1 大口黑鲈 *MyoD* 基因 PCR 扩增结果

用引物 R1 和 GenomeWalker Universal Kit 中的引物 AP1 进行 5' 调控区序列扩增获得了一条大小约 1 400 bp 的条带 (图 1-a:1),用引物 R2 和锚式引物 AP2 验证。用引物 F1 和 R3 进行了 *MyoD* 基因内含子序列的扩增,获得了约为 2 100 bp 的特异条带 (图 1-b:1)。

表 1 大口黑鲈 *MyoD* 基因引物信息
Tab. 1 The primers of *MyoD* gene of largemouth bass

引物 primer	碱基组成(5'-3') nucleotide constitutes	位置 sites	长度 (bp) length	扩增区域 amplification region	产物长度 (bp) length of product	最适合退火 温度 T_M ($^{\circ}C$) optimum annealing temperature	用处 purpose
R1:	CGTCGTTACCTTGCTTAGTCGCCG	extron 1	25	1 462		58	扩增 5'调控区 amplify 5'-flanking region
AP1:	GTAATACGACTCACTATAGGGC						
R2:	GGAGGAGTCGTCCGGCTTCAGTAGG	extron 1	25	1 224		60	扩增 5'调控区 amplify 5'-flanking region
AP2:	ACTATAGGGCACGCGTGGT						
F1:	TTAACACCAGCGACATGCACTTC	extron 1	23	1 142			扩增内含子 amplify introns
R3:	CTAGAGATTCCGCTCCACGATGCT	extron 3	21	3 394	2 100	60	
P1:	F TGAACCATGGAGCTGTCCGAT	promoter	21	1 072			筛选 SNPs 位点 identify site of SNPs
	R GAGGACGGGGAGAAGGAAGAG	extron 1	21	1 274	203 bp	58	
P2:	F CTCCTGCATCTCCATTACCAC	extron	21	1276			筛选 SNPs 位点 identify site of SNPs
	R; GGCAGCCGCTGGTTGGGGTTG	extron 1	21	1 514	239 bp	54	
P3:	F; ACCCCAACCAGCGGCTGCCCA	extron 1	21	1 496			筛选 SNPs 位点 identify site of SNPs
	R; CATGCCGTCGGAGCAGITGGA	extron 1	21	1 668	173 bp	58	
P4:	F; TTTTAAGCATTTCGTTGTT	extron 2	19	2 702			筛选 SNPs 位点 identify site of SNPs
	R; CAGGTTTCTTACCGTTTGG	intron 2	19	2 856	155 bp	52	
P5:	F; GCAAGGCGGCTCTGAAAAGTGA	extron 3	21	3 326			筛选 SNPs 位点 identify site of SNPs
	R; GTGGAATACGAGTCCCTCGGCT	extron 3	21	3 522	197 bp	58	
P6:	F; TGCTCCGACGGCATGGTGAGT	extron 1	21	1 654			筛选 SNPs 位点 identify site of SNPs
	R; TAGGTGAAGTTTAGTTTCCGTGTTTA	intron 1	27	1 928	274 bp	58	
P7:	F; GCTTTTTTTGTTTCTATTGTGGGC	intron 1	25	1 933			筛选 SNPs 位点 identify site of SNPs
	R; AGAACTCTTGIGATGAAAGGAATA	intron 1	25	2 208	276 bp	56	
P8:	F; GTAGTCCCTTCATCACAAGAGTTC	intron 1	25	2 183			筛选 SNPs 位点 identify site of SNPs
	R; CATCTATTATTATGCTGATTGTTT	intron 1	25	2 394	212 bp	54	
P9:	F; GAAACAAATCAGCATAATAATAAGA	intron 1	25	2 368			
	R; CAAAACATCAAAAGTAAAGAAAATA	intron 1	25	2 615	248 bp	52	
P10:	F; TTTCITTACTTTTGATGTTTGA	intron 1	25	2 594			筛选 SNPs 位点 identify site of SNPs
	R; CTTCIGTCATAACTTCCCTCTCIGT	extron 2	25	2 818	225 bp	52	
P11:	F; GAGACTCCAAACGGTAAGAAAC	extron 2	22	2 815			筛选 SNPs 位点 identify site of SNPs
	R; TTTTATIGITTTTATGCTGCTTTTA	intron 1	25	3 000	186 bp	54	
P12:	F; AGACAAACAGGACAAGCAATAA	intron 2	22	2 957			筛选 SNPs 位点 identify site of SNPs
	R; GACACAAAAGGAACTGAAAAT	extron 3	22	3 220	263 bp	54	
P13:	F; AGGGTTCGCTCAGTATTGGGGTGT	promoter	25	18			
	R; AGACAAATCCAGCGAAAGAGGC	promoter	22	1 052	1 035	58	筛选 SNPs 位点 identify site of SNPs

注: F,正向引物;R,反向引物

Notes: F, forward primer; R, reverse primer

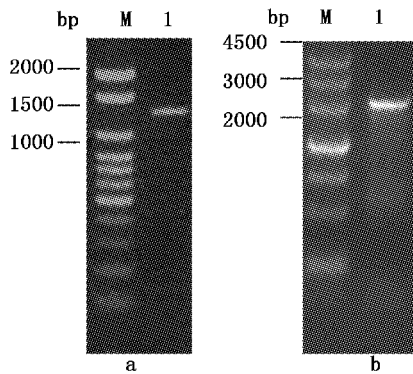


图1 *MyoD* 基因 5'调控区序列扩增结果(a)和内含子序列扩增结果(b)

M:核算分子量; a:1,5'调控区的扩增结果;
b:1,*MyoD* 基因内含子的扩增

Fig.1 PCR amplified result of 5'-flanking region (a) and introns (b) of the largemouth bass *MyoD* gene

M:Marker; a:1, amplified PCR product of 5'-flanking region;
b:1, amplified PCR product of introns of the largemouth bass *MyoD* gene

2.2 *MyoD* 基因序列分析

运用 Vector NTI 8.0 软件将所得的序列进行对比拼接成一条完整的 DNA 序列,即为大口黑鲈 *MyoD* 基因序列。大口黑鲈 *MyoD* 基因序列全长 3 797 bp,其中 5'调控区序列长度为 1 077 bp。*MyoD* 基因含三个外显子和两个内含子。三个外显子中第一个外显子 591 bp,编码 197 个氨基酸残基,第二个外显子 81 bp,编码 27 个氨基酸残基,第三个外显子 78 bp,编码 26 个氨基酸残基;而两个内含子大小分别为 1 077、486 bp,内含子序列都起始于 GT,终止于 AG,遵循 GT-AG 法则。该基因的序列如图 2。

通过 MatInspector 软件对大口黑鲈 *MyoD* 基因的 5'调控区域进行潜在的转录因子结合位点的分析(图 2),结果表明与肌肉特异性基因转录密切相关的转录调控元件 E box 有 13 个,肌细胞增强因子 2 (myocyte-specific enhancer factor 2, MEF2)、肌肉特异性金属硫蛋白结合位点 (muscle-specific Mt binding site, MTBF) 各 2 个。此外还有其他多个转录调控元件如孕激素受体结合位点 (progesterone receptor binding site, PRE)、八聚体结合因子 1 (octamer-binding factor 1, OCT 1)、激动蛋白 4 (activator protein 4, AP4)、垂体特异性转录因子 (pituitary-specific transcription factor 1, PIT 1) 和 3 个转录反应起始框 TATA box

及 1 个 OCAAT BOX 等。

2.3 大口黑鲈 *MyoD* 基因 PCR-SSCP 结果

用 13 对引物对来自广东省大口黑鲈良种场的 24 个样品进行 *MyoD* 基因序列的 SNPs 位点筛选,其中 P7、P8、P12 三对引物扩增的 PCR 产物表现出单链构象多态性,P7、P8 引物扩增的片段位于内含子 1,P12 引物扩增的片段位于内含子 2 上(图 3)。

对 PCR-SSCP 分析表现为多态性的样品进行双向测序,结果表明在内含子 1 上有 5 个 SNP 位点,分别为 2014(A→T)、2137(T→A)、2138(T→G)、2181(T→C)、2315(G→A),内含子 2 上有两个 SNP 位点,分别为 2976(T→C)、2992(A→G),测序结果如图 4。

2.4 大口黑鲈 *MyoD* 基因突变点频率分析

对大口黑鲈群体 90 尾鱼进行 *MyoD* 基因 7 个 SNPs 位点的群体分析结果如表 2。

表 2 大口黑鲈 *MyoD* 基因突变点频率分析

Tab. 2 SNPs frequency analysis of *MyoD*

gene in the largemouth bass

突变点 SNP site	群体数(尾) no. of sample	基因频率 allele frequency		突变率(%) percentage of SNPs
		A	B	
A2014T	90	0.647	0.353	35.3
T2137A	90	0.824	0.176	17.6
T2138G	90	0.824	0.176	17.6
T2181C	90	0.882	0.118	11.8
G2315A	90	0.912	0.088	8.8
T2976C	90	0.958	0.042	4.2
A2992G	90	0.958	0.042	4.2

根据表 2 分析:这 7 个 SNPs 位点在大口黑鲈群体中的突变比例范围为 4.2% ~ 35.3% 不等,其中 A2014T 所占的比例为 35.3%,而 T2976C、A2992G 所占的突变率为 4.2%。突变点 A2014T 比突变点 T2976C 和 A2992G 在大口黑鲈群体的突变比例高 8.4 倍。

3 讨论

大口黑鲈 *MyoD* 基因 5'调控区 1 077 bp 长度中存在 13 个 E box 和两个肌细胞增强因子 2 (MEF2) 及其它一些与肌肉转录调控有关的转录因子。E box 是基因转录调控区 DNA 序列中一段 CANNTG(N 代表任何一种碱基)序列^[17-18]。MFRs 与 E box 结合是激活肌肉基因转录的重要

```

1 tagogattta cctttaaagg gttctgctca gtattgggtt gtggagtctt ttgctactga atgacaacgt cgtgatcgtc attctagatt ctagtgatc
101 ttatttcaagt tgcgctgtgc gtccgttggg tatacagat gtttggttca gtgactcacc atttcagctt ttatttattt tactttcacc tgcactgattt
201 cactctgaagg aacttcagct gcgacaacat cacagccaat agcgtttatg aaggaggaaa octgottaag ttacaactgt acataaatat aatatcaaat
OCCAAT BOX MEF 2 TATA BOX
301 aaatgtgaaa tttcgttctt tatacttata cttttatttt aactgtattt ggaaatttatt tgtttccacc cctttattct tttcaggttt atgtatcgtt
MPBI
401 ttattgtcac ttcagtgat ttgttttttg taacttttgg taattttaaag tcttcttctt tttttaattc attttttctt cagtgtattc taattatata
CCT 1 TATA BOX
501 gatcaatata ttttaattat ctctttcaaa tcggtcaaat attgcaagc aagaactttc catccaata ccatttttca ttgtgtactt tacattaaaa
MPBI
601 tgttattagt caacaagtaa cttgcacaga acataatgaa taagaagaaa ggcaatgaaa aagacaaaa ctacaaata tctctgatatt tcttcaagtg
Pit 1 MEF 2
701 taccctccc caatctgtc agcagatata tatcctgctg ttaactgctt gttacgggta attgtacact aattaacgtg aagttgtaaa ccctctctgc
OCT 1 TATA BOX
801 tgcacagctca ttggtcagac atcagttggc atcagttggtc atcaccgtcc tgaccocgoc cactgogaat gtcccaggtc taaggagatc ctggtcaaca
901 gctgaggaca caacagtttg tgacaggtact ctacattccc ccgaaaacac tgccttca catctgagttcct ctctctctct gttccctctt tctttctctt
AP 4 PRE
1001 tcttcttctc tgtaaacagt ttttgtttg gcctctttcg ctggatttgt cttgtttggt gtgtttggga ctgaaaccatg GAGCTGTGGG ATATCTCTTT
1101 CCCATCCCT GCCTGATG APTTCTATGCA GACCCCTGCT TTTAACACCA GCGATGCA CTTCTTCGAG GACCTGGADC CCGGGCTGGT CCATGTGGGC
120 CTACTGAAGC GGGAGACTC CTCCTCTTCG GCTTCACCTT CCCCCTCCFC CTCCTCTTCC TTCTCCCTGT CCTCCCTCTT GCATCTCAT TACCACGGAG
1301 AGGGGAGGA GACGAGCAC GTCCGCTGCC CCAGCGGCA TCACAGGCG GCGCGTGC TGCTCTGGCT CTGCAAGGC TGCAAGCGGA AGCACACCA
1401 GCGATATCG GGAAGGCG CCACCTGCG GAAAGCGCG GCACATGCA AGGTGACGA CCGCTTCGAG ACCCTGAGC GCTGCACGTC GGCACACCC
1501 AACGAGCGC TGCCAAAGT GGAGATCTG GCACAGCCA TCAGTACAT CGATCCCTG CAGCCCTGCT TGCGCGCGG GCAGGACGAC GGCCTCTAC
1601 CGTGTCTGA GCATACAGC GGGACTCAG AGGCTCCAG CCCCGCTCC AACTGCTCC AGGCATGGT GAGTTCAGG ACTGACAGt gttaaagagg
1701 gcaagagggg ggggtccgtg gagaagagtc tggctgtgcy ctttgggtgc gtcctccag tgcgtctta cgcaggtta tatctcactt catttgattt
1801 aatgtaaat gatttatttg atttactttg gttttgttta atgagtgatg atgacgactg caattcaaat tctgatttta acaattctcg atgcaaat
1901 ttaaacacag gaaactaaac ttcaactatt aagctttttt tgtttctatt gttgggccc atttaactct tttctctctt catctcatgt aaagtctgtt
2001 tttgtgtcca aactatttta caactccaag cagaagagct aaagtgcag actttgaat taattgtaat tatctgtagg cctgttaatt ctcaagatga
2101 aaagaaaaa ctactaagca ggtogaatac aagttttttc tgggggaaag aaccttccg cgtgcaatca tctctgtttt ttgtagttcc tttcatca
2201 agagttctga ctaataaaa taatgagctg atttataat attttaaac agyagctgtt ttatgttta agaactcagtt cgggtttttg aaacytaaaa
2301 acatgcyaca caactgttag tgcgacata tagatogaac aaatatctg aatttaoaga atgtgttga acaaatcagc ataaataaa gatgtataat
2401 tacgttttgt gacacttgat agttctgtag ttttctgtca tttttgata tgaattgact ttgttggctc gcagtaatta ataaaaata aacatttat
2501 ttttaaaaa aagttgaatt tctataagac acagcttagt agtgttgtt gaatttattt aaatatctg ccaatttcta cagtttcaat tatttcttt
2601 acttttgatg ttttgatcat atgtctgcat catataagct gcaagtttcc tctaaaacta atagygact gataaactgg ctacttttaa goatttctgt
2701 gttgtaaaaa tctaatttct ttataaattt tttttccctc tgcagAGGA TTTTAAAGC CCGACTGTC ACTCAACAG AAGAGGAAGT TATCACAGAA
2801 GCTCTTATTT CTCAGAGACT CCAACGgta agaaacctga aaagcttta ataatacaat ttaogatgat attagcagct agcaggccaa ttggtctaca
2901 aacaacacaga ctttttttgc tcccgttag taaaacaaaa gacatttagc agggaaagac aaacaggaca agcaatataa gcaataaaa adcaataaaa
3001 ttaagttttt ttttaactg tcaatataa ataaaaatt ttgtataaaa agaactcagc ttttaagtagc ctagtatttt tatataat: cagcaaaaa
3101 atctttatga aaacatctca ttagttaac ccagagataa caaacacttt tatttgatta aactaattta asataaaacc aaacaaagga ccaataagat
3201 tttcagttcc ttttgggtgc ttttgaaagt gtcccgactg tttttccga tcaatttga gtttaacaag tcaataaatt aatattaac tctgatgttt
3301 ggtttcctgc aagGCGTCT GAAAGTGAC CGGACTCCG TGGTCTCCAG TCTGACTGT CTGTCCAGCA TCGTGGAGG GAATCTCTAG taagcacaga
3401 tccagcgtgt ccccccgg gcgacagcgt ggttcccgc ggcctggtat cgcctccca acagctcta cgaaccgctc tgaactgat ggcagcaggt
3501 actcgtatc caaaactta tacagacttt ttctgaagt ttcttcagac tgcgtctttg gtttttatcc tccgaaaca actgaagact gtttccatt
3601 aatctcagca gaacaggac cttttgttt gtaataaga gctatttctt cacagtacc tctcacaca acacacac acgtcacat atacagacac
3701 aaggatggtt gggattgttg tgcataatg ttctatatat tatttatgag tgaatgacat ttaataaag aaatatttat attctgaaaa aaaaaa

```

图2 大口黑鲈 *MyoD* 基因序列
 5'调控区和内含子用小写字母表示,外显子用大写字母表示,5'调控区中的转录因子结合位点用下划线和方框标记,名称标在下面。CaNNtg表示 E box, N表示 SNP 位点

Fig. 2 Genomic sequence of the largemouth bass *MyoD* gene
 The intron and 5'-flanking region are shown in lowercase letters, Nucleotide sequences in exon are shown in uppercase. The putative transcription factor binding sites in 5' -flanking region are underlined or in frame and under them is indicated the transcription factor that binds to it. the initiation codon is indicated by an arrow.
 CaNNtg means E box, N means sites of SNP

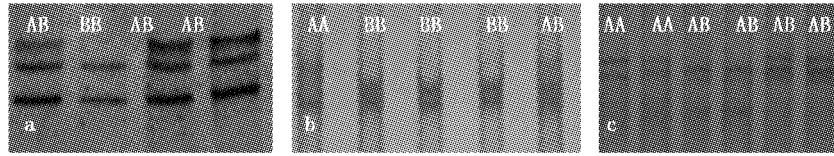


图3 加州鲈 *MyoD* 基因 SSCP 电泳结果

a. P7 引物 SSCP 结果 ; b. P8 引物 SSCP 结果 ; c. P12 引物 SSCP 结果

Fig. 3 SSCP electrophoretic result of the largemouth bass *MyoD* gene

a. SSCP result of P7 primer; b. SSCP result of P8 primer; c. SSCP result of P12 primer

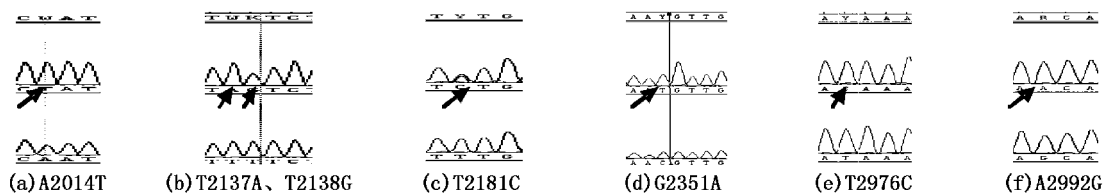


图4 加州鲈 *MyoD* 基因 SNP 测序图

Fig. 4 The largemouth bass *MyoD* gene SNP sequenced result

途径^[19-20]。Claude 等^[21]通过对鸡的 *MyoD* 基因 5'调控区启动子中的各个 E box 分析,结果在 *MyoD* 基因的启动子中存在 17 个 E box 和 2 个 MEF2,其中的 E1、E5、E9、E10 和 E13 能有效的激活 *MyoD* 基因的转录与表达,特别是 E9 box。在牛的肌肉生肌抑止因子 Myostatin 基因的启动子中,E6 box 中碱基突变造成了双肌牛的出现,Crisa 等^[22]论证了 E6 box 是 Myostatin 基因转录过程中最关键的一个 E box。大口黑鲈 *MyoD* 基因 5'调控区中的 E8 box 与鸡 *MyoD* 基因的启动子中 E9 及牛的 Myostatin 基因启动子中的 E6 box 所在的位置相当。因此我们推测 E8 box 是大口黑鲈 *MyoD* 基因的转录与表达的重要元件。

基因组中的 SNPs 绝大多数位于内含子区域,而外显子相对较保守。Halushka 等^[23]通过对人的 75 个基因进行检测后,推测人类基因组有数百万个 SNP 位点,其中大约有 50 万个在非编码区,估计有 24 ~ 40 万个在编码区,且与蛋白质的功能有关。Rafalsk 等^[24]利用直接测序法对美国优良玉米品种进行多态性分析发现在非编码区平均每 48 bp 出现 1 个 SNP,在编码区每 131 bp 有 1 个 SNP。Nie 等^[25]在鸡中对与生长相关的 12 候选基因进行 SNPs 分析,结果 68% 的突变在内含子上,而外显子的突变只有 16.6%。本研究在大口黑鲈 *MyoD* 基因上筛选的 7 个 SNPs 位点出现在内含子上,外显子和 5'调控区上均未发现

SNPs 位点。这与 Knoll 等^[26]、朱砺^[15]只在猪的 *MyoD* 基因内含子上发现 *Dde* I 酶切多态性位点,而在外显子上不存在多态性一样,这个结果符合基因组 SNPs 突变的基本规律,也说明我国养殖大口黑鲈的遗传多样性不高,这与我们实验室应用微卫星标记对养殖大口黑鲈的遗传多样性分析结果相一致。遗传多样性不高与我国引进大口黑鲈种质来源单一和多年来不同程度的近亲繁殖有关。

参考文献:

- [1] 王立新,白俊杰,叶星,等.草鱼 *MyoD* cDNA 的克隆和序列分析[J].中国农业科学,2005,38(10):2134-2138.
- [2] Rawls A, Valdez M R, Zhang W, et al. Overlapping functions of myogenic bHLH genes MRF4 and *MyoD* revealed in double mutant mice [J]. Development, 1998, 125(13):2349-2353.
- [3] Lee H, Habas R, Abate-Shen C, et al. Mx1 cooperates with histone H1b for inhibition of transcription and myogenesis [J]. Science, 2004, 304(5677):1675-1678.
- [4] Rocky S T. Cellular and molecular regulation of embryonic skeletal development and morphogenesis [J]. Cells and Materials, 1998,(8):3-18.
- [5] Alves H J, Alvares L E, Gabriel J E, et al. Influence of the neural tube/notochord complex on *MyoD* expression and cellular proliferation in

- chicken embryos Brazilian [J]. *Journal of Medical and Biological Research*, 2003, 36(2): 191–197.
- [6] Rudinick B T, Hinuma S, Jaenish R. *MyoD* or *Myf-5* is required for the formation of skeletal muscle [J]. *Cell*, 1993, 75: 1351–1359.
- [7] Te Pas M F, Verbung F J, Gerritsen C L, *et al.* Messenger ribonucleic acid expression of the *MyoD* gene family in muscle tissue at slaughter in relation to selection for porcine growth rate [J]. *Journal of Animal Science*, 2000, 78(1): 69–77.
- [8] Kablar B, Krastel K, Ying C H, *et al.* *MyoD* and *Myf-5* differentially regulate the development of limb versus trunk skeletal muscle [J]. *Development*, 1997, 124 (23): 4729–4738.
- [9] Braun T, Grzeschik K H, Bober E, *et al.* The MYF genes, a group of human muscle determining factors, are localized on different human chromosomes [J]. *Cytogenet Cell Genet*, 1989, 51 (1): 969–978.
- [10] Scrabble H J, Johnson D K, Rinchik E M, *et al.* Rhabdomyosarcoma-associated locus and *MyoD1* are syntenic but separate loci on the short arm of human chromosome 11 [J]. *Proc Nat Acad Sci USA*, 1990, 87(6): 2182–2186.
- [11] Guttridge D C, Mayo M W, Madrid L V, *et al.* NF-kappaB-induced loss of *MyoD* messenger RNA: possible role in muscle decay and cachexia [J]. *Science*, 2000, 289(5488): 2363–2366.
- [12] 吴听彦, 张庆华, 陈 竺, 等. 单核苷酸多态性研究及应用 [J]. *中华遗传学杂志*, 2000, 17(1): 57–60.
- [13] Rafalski A. Application of single nucleotide polymorphism in crop genetics [J]. *Current Opinion in Plant Biology*, 2002, 5(2): 94–100.
- [14] 田 璐, 李俊雅, 许尚忠, 等. 动物生肌决定因子的研究进展 [J]. *黄牛杂志*, 2005, 31(1): 43–46.
- [15] 朱 砺, 李学伟. *MyoD* 基因在不同猪种中的 PCR-RFLP 遗传多态性及其遗传效应研究 [J]. *畜牧兽医学报*, 2005, 36(8): 761–766.
- [16] 田 璐, 许尚忠, 岳文斌, 等. *MyoD* 基因对肉牛胴体性状影响的分析 [J]. *Hereditas (Beijing)*, 2007, 3: 29(3): 313–318.
- [17] Malik S, Huang C F, Schmidt J. The role of the CANNTG promoter element (E BOX) and the myocyte-enhancer-binding-factor-2 (MEF-2) site in the transcriptional regulation of the chick myogenin gene [J]. *Eur J Biochem*, 1995, 230(1): 88–96.
- [18] Heidt A B, Rojas A, Harris I S, *et al.* Determinants of myogenic specificity within *MyoD* are required for noncanonical E box binding [J]. *Mol Cell Biol*, 2007, 27(8): 5910–5920.
- [19] Lassar A B, Buskin J N, Lockshon D, *et al.* *MyoD* is a sequence specific DNA binding protein requiring a region of myc homology to bind to the muscle creatine kinase enhancer [J]. *Cell*, 1989, 58(5): 823–831.
- [20] Lin H, Yutzey K, Konieczny S F. Muscle-specific expression of the troponin I gene requires interactions between helix-loop-helix-loop-helix muscle regulatory factors and ubiquitous transcription factors [J]. *Mol Cell Biol*, 1991, 11(1): 267–280.
- [21] Dechesne C A, Qin W, Eldridge J, *et al.* E box and MEF-2-independent muscle-specific expression, positive autoregulation, and cross-activation of the chicken *MyoD* (CMD1) promoter reveal an indirect regulatory pathway [J]. *Mol Cell Biol*, 1994, 14(8): 5474–5486.
- [22] Crisà A, Marchitelli C, Savarese M C, *et al.* Sequence analysis of myostatin promoter in cattle [J]. *Cytogenet Genome Res*, 2003, 102(1–4): 48–52.
- [23] Halushka M K, Fan J B, Bentley K, *et al.* Patterns of single-nucleotide polymorphisms in candidate genes for blood-pressure homeostasis [J]. *Nat Genet*, 1999, 22(3): 239–247.
- [24] Rafalski J A. Novel genetic mapping tools in plants: snp and LD-based approaches [J]. *Plant Science*, 2002, 162(36): 329–333.
- [25] Nie Q, Lei M, Quyang J, *et al.* Identification and characterization of single nucleotide polymorphisms in 12 chicken growth-correlated genes by denaturing high performance liquid chromatography [J]. *Genet Sel Evol*, 2005, 37(3): 339–360.
- [26] Knoll A, Nebola M, Dvorak J, *et al.* Detection of a Dde I PCR-RFLP within intron 1 of the porcine *MyoD1* (*Myf_3*) locus [J]. *Animal Genetics*, 1997, 28(4): 308–322.

Molecular cloning and identification of single nucleotide polymorphisms of *MyoD* in largemouth bass (*Micropterus salmoides*)

YU Ling-yun^{1,2}, BAI Jun-jie¹, YE Xing¹, LI Sheng-jie¹, LI Xiao-hui¹

(1. Key Laboratory of Tropical & Subtropical Fish Breeding & Cultivation, Pearl River Fisheries Research Institute, Chinese Academy of Fishery Sciences, Guangzhou 510380, China;

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

Abstract: In present research, the sequences of *MyoD* gene and 5'-flanking region are amplified from DNA isolated from the largemouth bass by using PCR and GeneWalking technique. A total of 3797 bp sequence of *MyoD* gene and 5'-flanking region was obtained. The *MyoD* encoding gene consisted of three exons (591 bp, 81 bp and 78 bp respectively) and two introns (1077 bp and 486 bp). Sequence analysis of 5'-flanking region of *MyoD* gene revealed that it contained putative muscle growth response elements such as E-boxes, myocyte enhancer factor 2 (MEF2), muscle-specific Mt binding site (MTBF), TATA boxes, Octamer-binding factor 1 (OCT1), activator protein 4 (AP4) and OCAAT box were also detected. 7 SNP sites in introns were identified from *MyoD* gene by using PCR-SSCP technique and sequencing, however no polymorphism existed in exon and 5'-flanking region. The percentage of SNPs is between 0.042 – 0.353 in largemouth bass. These SNPs might be useful for a further study on *MyoD* gene's relation to production performance in largemouth bass.

Key words: *Micropterus salmoides*; *MyoD*; SNPs; identification