

鲮 *Myf5* 基因克隆及其 SNPs 分析

钟茂春^{1,2}, 郑光明¹, 赵建¹, 朱新平¹, 马丽莎¹, 潘德博¹, 陈昆慈¹, 谢文平¹, 史燕¹

(1. 中国水产科学研究院 珠江水产研究所, 广东 广州 510380; 2. 上海海洋大学 生命科学与技术学院, 上海 200090)

摘要: 提取新鲜鲮 (*Cirrhinus molitorella*) 肌肉总 RNA, 采用 RT-PCR 技术分段克隆 *Myf5* 基因核心序列, 获得 1 104 bp 的目的片段, 其中开放阅读框为 723 bp, 编码 240 个氨基酸。分析表明, 鲮 *Myf5* 蛋白具有 MRF 家族基因的典型性碱性 bHLH 螺旋-环-螺旋 (Basic helix-loop-helix, bHLH) 结构。设计特异性引物扩增获得 2 个内含子, 大小分别为 1 311 bp 和 90 bp。用单链构象多态性 SSCP 的方法分析了 *Myf5* 基因在珠江水系 3 个不同江段鲮中的遗传变异, 分布及群体杂合性等群体遗传信息。发现其编码区非常保守, 仅在外显子 412 位点处发生突变 C→T。对该位点进行基因型分析, 结果表明: 珠江北江段鲮中该位点符合 Hardy-Weinberg 平衡, 而东江和西江段种群不符合 Hardy-Weinberg 平衡, 3 个江段种群总体上符合 Hardy-Weinberg 平衡, 并存在较大基因流。以 *Myf5* 基因该位点的多态性分析表明, 3 个种群未出现明显分化; 通过该位点的遗传结构分析表明: 3 个江段种群杂合度都较高, 中度信息含量 (PIC) 及多态信息指数, 即该位点有较好的变异性及遗传多态性。3 个江段种群总体上该位点在不符 Hardy-Weinberg 的情况下 AA>BB, 并由该位点群体杂合性及中性分析表明该位点突变群体具体良好的多态性且不遵循中性选择模式, 说明 BB 型个体受到自然选择。本研究结果表明, 鲮 *Myf5* 基因的 C412T 与个体生长有相关性, 可作为候选基因标记, 用于鲮生长相关分子标记辅助育种研究。[中国水产科学, 2010, 17 (4): 681–688]

关键词: 鲮鱼; *Myf5* 基因; SSCP; 遗传结构

中图分类号: Q959; S96

文献标识码: A

文章编号: 1005-8737-(2010)04-0681-08

鲮 (*Cirrhinus molitorella*) 属鲤形目、鲤科、野鲮亚科、鲮属。俗称土鲮鱼、鲮公、雪鲮、花鲮。分布于珠江、闽江、韩江、海南岛、台湾岛、元江及澜沧江水系, 国外见于东南亚。鲮在珠江流域地区与草鱼、鲢、鳙合称四大家鱼, 是江河捕捞和池塘养殖的主要对象之一^[1]。

Myf5 (myogenic factor 5) 基因是生肌调节因子 MRFs 家族的成员之一, 该家族包括: *Myf5*、*MyoG*、*MyoD* 和 *MRF6* 4 种, 基本功能是促进成肌细胞的增殖和分化, 其中 *MyoD*、*Myf5* 决定着肌卫星细胞是否能启动成为具有成肌特性的肌肉干细胞, 而 *MyoG* 和

MRF6 则发挥着调节肌肉干细胞终末分化为肌管肌纤维的功能^[2-4]。通过 MRFs 家族基因的典型性的碱性 bHLH 螺旋-环-螺旋结构它们能够与 E 蛋白形成二聚体, 再与肌肉特异性基因如肌球蛋白轻链基因和肌酸激酶等的上游调控序列 E-box 结合, 从而激活这些基因的表达^[5-6]。*Myf5* 基因已经在多种哺乳动物中被鉴定克隆出来, 如小鼠 (*Mus musculus*)、牛 (*Bos taurus*)、猪 (*Sus scrofa*)、绵羊 (*Ovis ammon*)、鸡 (*Gallus gallus*) 等, 但在鱼类中研究较少, 目前只在斑马鱼 (*Danio rerio*)、条纹鲈 (*Morone saxatilis*)、花鲈 (*Lateolabrax japonicus*)、褐牙鲈 (*Paralichthys olivaceus*)

收稿日期: 2009-10-13; 修订日期: 2009-12-17.

基金项目: 国家自然科学基金资助项目 (30271022); 国家公益性行业 (农业) 科研专项 (200903045); 国家重点基础研究发展计划项目 (2004CB117401); 广东省科技计划 (2005B20301007); 广东省科技厅农业攻关子项目 (2005A20105001); 国家科技基础条件平台建设子项目 (2006DKA30470-008)。

作者简介: 钟茂春 (1984-), 男, 硕士研究生. Tel: 020-81616509; E-mail: 030901.student@sina.com

通讯作者: 郑光明, 研究员. Tel: 020-81616509; E-mail: zgmyl@tom.com

中克隆出基因组序列;在红鳍东方鲀(*Fugu rubripes*)、大西洋鲑(*Salmo salar*)、虹鳟(*Oncorhynchus mykiss*)、鲤(*Cyprinus carpio*)、大口黑鲈(*Micropterus salmoides*)中得到cDNA序列。

目前,对鲮的研究主要集中于分子标记与遗传多样性方面^[7-9],而关于相关功能基因的研究则很少,因此本研究选择珠江水系的东江、北江、西江3个江段种群的野生鲮作为实验对象,先克隆到*Myf5*基因的cDNA序列并进一步获得基因组序列。通过生物信息学方法,分析基因结构,蛋白质结构,以及预测基因生物学功能,再用SSCP(single strand conformation polymorphism, SSCP)的方法对鲮*Myf5*基因进行SNPs(Single Nucleotide Polymorphisms)检测^[10-11],以获得与生长性状相关的SNPs位点,并对该位点的基因型在各江段种群中的遗传分布及群体遗传结构进行研究,为鲮分子标记辅助育种奠定基础。

1 材料与方法

1.1 材料

实验材料鲮采样于珠江水系3个江段:东江(惠州)、北江(清远)、西江(梧州)。

1.2 总RNA的提取

肌肉总RNA的提取,取鲮背部肌肉30 mg,按照Promega公司SV Total RNA Isolation System试剂盒的方法进行,用1.2%的琼脂糖电泳检测提取的RNA。

1.3 DNA的提取

珠江水系各江段鲮各56尾,取肝脏。按萨姆布鲁克等的方法提取DNA^[12],用1.2%的琼脂糖电泳检测。

1.4 引物合成与PCR扩增

根据在GenBank上登录的斑马鱼和鲤的*Myf5*基因的cDNA序列,在其保守区设计引物。引物由上海生工生物工程技术服务有限公司合成。PCR体系(buffer, Mg^{2+} , dNTP, *Taq* 酶)试剂购于上海生工生物工程技术服务有限公司。

RT-PCR 扩增cDNA核心片段:将15 μ L RNA与2 μ L随机引物在70 $^{\circ}$ C混合5 min,再将混合试剂(5 μ L Buffer,0.7 μ L RNA-inhibitor,1.3 μ L dNTP,1 μ L

M-MLVRT)加入到上一步混合液中,42 $^{\circ}$ C,2 h,70 $^{\circ}$ C,5 min,冰浴2 min。稀释10倍作为扩增模板。

PCR 扩增总体积为20 μ L,PCR扩增体系的组成为10 $\times$ PCR reaction buffer 2 μ L、 Mg^{2+} 2 μ L ddH₂O 13.3 μ L、引物(10 pmol/ μ L)各0.5 μ L、模板DNA(50 ng/ μ L)1 μ L。PCR扩增程序:预变性94 $^{\circ}$ C 5 min 1个循环;94 $^{\circ}$ C 30 s,52~60 $^{\circ}$ C 30 s,72 $^{\circ}$ C 1 min,30个循环;最后72 $^{\circ}$ C继续延伸7 min PCR产物用1.2%琼脂糖凝胶电泳检测,结束后用UVP凝胶成像系统分析检测扩增结果。

1.5 PCR扩增产物的克隆和测序

扩增产物经1.2%低熔点琼脂糖凝胶电泳并回收纯化,与pMD18-T载体连接,转化感受态大肠杆菌DH5a,涂平板,挑取单菌落,接种到含100 mg/L Amp⁺ LB培养基,37 $^{\circ}$ C过夜。用载体通用引物和扩增引物作菌落PCR检测并将阳性转化子送上海生工生物工程技术服务有限公司测序。

1.6 PCR-SSCP分析

筛选SNPs位点引物合成:基因上设计在ORF(Open reading frame)内DNA序列特异性引物,如表1所示。

PCR程序为:预变性94 $^{\circ}$ C,5 min,1个循环;94 $^{\circ}$ C,30 s;50~58 $^{\circ}$ C,30 s,72 $^{\circ}$ C,30 s,共30个循环;72 $^{\circ}$ C 7 min。反应结束。产物2 μ L与8 μ L变性buffer(甲酰胺9.8 mL,0.5 mol/L EDTA 0.2 mL,加少量溴酚蓝和二甲苯菁;定容至10 mL)混合,98 $^{\circ}$ C,10 min,冰浴5 min。用12%聚丙烯酰胺凝胶15 V/CM,6 h,银染法染色^[11],制成干胶保存。将纯合基因型送上海生工生物工程技术服务有限公司测序,分析突变位置。

1.7 统计分析

1.7.1 基因型分布分析 根据电泳结果直接判断出每一个体的基因型,计算各等位基因在各江段鲮中的基因频率并运用卡方检验和最大似然率法分别进行符合Hardy-Weinberg平衡的检验。

1.7.2 群体遗传结构分析 采用PopGene32(Population Genetic Analysis, Version1.31)软件进行*Myf5*基因在各江段种群内的遗传变异、群体杂合性及群体遗传结构等分析。分析内容涉及:*Myf5*基

表1 PCR-SSCP特异性引物
Tab.1 The GSP of PCR-SSCP

名称 Name	碱基组成 Primer sequence	位置 Location	长度 /bp Length	扩增区域 Amplification area	扩增长度 /bp Amplification length	TM 值 TM value	作用 Usage
P(1-1)	TCCAACCAATCATGGACGTATTCTC	exon	25	1		64.2	筛选 SNP 位点
P(1-2)	TCTCCACCTTGGGAAGTCGCTGA	exon	23	330	330	67.8	筛选 SNP 位点
P(2-1)	CCTCGGCGAACCCAGTCAGC	exon	21	292		70.6	筛选 SNP 位点
P(2-2)	TTCAGTATCAATCTAAACCCCTC	exon	23	520	229	55.2	筛选 SNP 位点
P(3-1)	CTCTTCAAGTCAAACCTGAGGGGT	exon	23	482		59.4	筛选 SNP 位点
P(3-2)	GATGGATAGAATGCCCTGTGTGC	intron	23	694	213	63.3	筛选 SNP 位点
P(4-1)	CTCTCTCTCCAACCTGGAAGTGATT	intron	24	2056		58.8	筛选 SNP 位点
P(4-2)	CTTATTGGCATCTGAAAGCATTAT	exon	24	2271	216	58.1	筛选 SNP 位点

因位点在各品系内的遗传变异基因的观察等位基因数(N_a)、有效等位基因数(N_e)、Shannon' S信息指数(I); *Myf5* 基因位点的群体杂合性: 群体纯合度(Homozygosity, H_{om})、群体杂合度(Heterozygosity, H_{et})、Wright' s固定指数(F_{is})等^[12-13]。参照Botstein等^[13]的方法计算多态信息含量(Polymorphism Information Content, PIC):

$$PIC=1-\sum_{i=1}^n p_i^2-\sum_{i=1}^{n-1} \sum_{j=i+1}^n 2p_i^2 p_j^2$$

式中, P_i 和 P_j 分别为群体*i*和*j*的等位基因频率。

2 结果与分析

2.1 鲮 *Myf5* 基因序列及多肽序列分析

将测序结果除去载体序列并在NCBI上用BLAST搜索其他物种同源序列,使用Vector 8.0将测序所得的*Myf5* 基因序列进行拼接,克隆获得鲮 *Myf5* 基因序列(GenBank注册号: GU289508),基因序列为

2 516 bp, 包括部分5'侧翼区、3个外显子、2个内含子和部分3'侧翼区序列。外显子和内含子数与人、家鼠、鸡、斑马鱼、鲤等物种相同,其ORF框内外显子大小分别为453 bp、76 bp、194 bp; 外显子1包括起始密码子ATG,外显子3包括终止密码子TGA; 内含子1和内含子2大小分别为1 311 bp和90 bp。

用Primer 5.0和Vector 8.0以及在线(www.expasy.org)翻译将鲮 *Myf5* 基因的ORF译成氨基酸序列(使用标准密码子),共编码240个氨基酸,在线计算得该多肽等电点和分子量为: 6.08, 26.288 31 kD多肽序列通过expasy proteomic server 用SWISS-MODEL预测并用RasMol进行分析,该基因编码的氨基酸序列具有一典型的碱性螺旋—环—螺旋(basic helix-loop-helix, bHLH)结构域,位于第60~124氨基酸处用Vector8.0等生物信息软件分析现已知的10种物种编码区内核苷酸序列同源性,并通过clustalx1.81和MEGA4.0构建分子进化树,如图1所示。

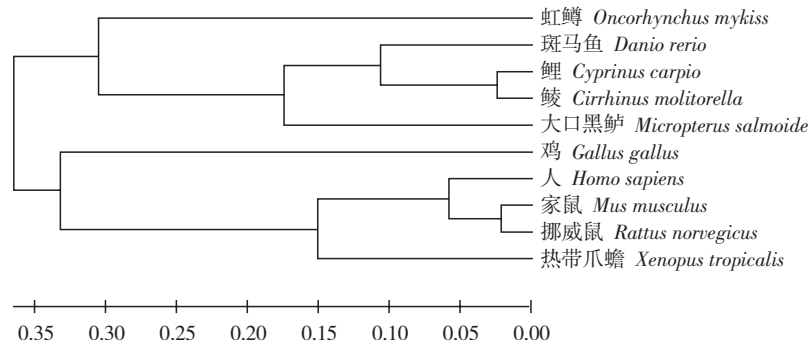


图1 依据不同动物 *Myf5* 核苷酸序列构建的进化树
Fig.1 NJ-tree according to the nucleotide sequence of *myf5*

鲮 *Myf5* 基因编码序列与鲤亲缘关系最近,序列同源性为93.6%;其次与斑马鱼相比序列同源性为85.9%。

2.2 PCR-SSCP结果分析

用表1中所示4对引物的PCR产物进行SSCP分析,其中第2对引物的PCR产物进行SSCP分析表

现出3种基因型,分别命名为AA为野生型, BB为突变型, AB为杂合型(图2)。测序发现AA野生型在412位点处发生C→T突变而形成BB型突变型(图3所示突变碱基位置),突变点好位于密码子的第一个位碱基位,将丝氨酸(Ser, S)变成了脯氨酸(Pro, P),改变了氨基酸序列。

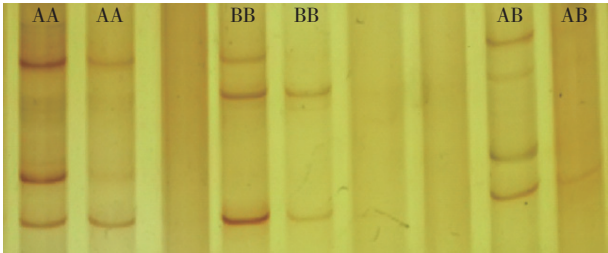


图2 鲮 *Myf5* 基因不同基因型个体的SSCP分析
AA: 野生型, BB: 突变型, AB: 杂合型.
Fig. 2 SSCP analysis on PCR amplification in different individuals of *Myf5* gene
AA: wild, BB: mutation, AB: Heterozygote.

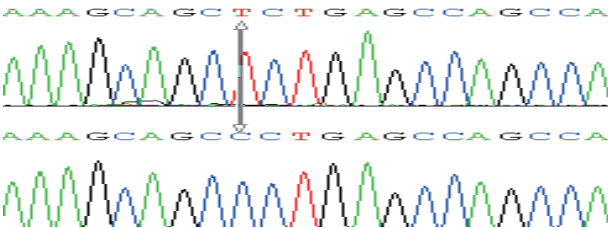


图3 分离自*Myf5*的基因SNP序列分析
箭头表示替换碱基位置.
Fig. 3 Sequence analysis on the isolated genotype of *Myf5* gene
Arrow denotes substitution.

2.3 突变位点在不同江段鲮的基因型和基因频率统计以及遗传变异结果

不同江段鲮 *Myf5* 基因的基因型频率和基因频率统计结果见表2。在各个江段种群内, AA基因型和A等位基因频率明显地高于BB基因型和B等位基因。分别用卡方检验和最大似然率法对各江段种群内鲮 *Myf5* 基因的基因型分布进行Hardy-Weinberg平衡的检验。结果表明,3个江段种群中只有北江江段种群符合Hardy-Weinberg平衡($P>0.05$), 西江江段种群显著的偏离了Hardy-Weinberg平衡($0.01<P<0.05$), 而东江江段种群均极显著的偏离

了Hardy-Weinberg平衡($P<0.01$)。多态信息含量及遗传杂合度等指标是度量群体遗传变异的重要参数,多态信息含量以及遗传杂合度越低,表明遗传一致性越高。由表3得出,各江段种群的多态信息含量均在0.25与0.5之间,该位点属于中度多态性。就群体杂合度而言,3个江段种群的杂合度较高,均高于期望值,通常当杂合子观察频率大于随机交配情况下的预期频率时,固定指数为负值,若存在近交效应时,杂合子频率的观察值减小,则固定指数为正。本实验中3个江段种群的固定指数均为负数,表明杂合子过剩。

表2 *Myf5* 基因在不同江段鲮间遗传分布

Tab. 2 Genetic distribution of *Myf5* gene among different parts of Pearl River

品系 Line	检测个数 Number	基因型分布 Genotype distribution			等位基因频率 Allele frequency		适合性检验法 Chi-square		似然率检验法 Likelihood ratio test	
		AA	AB	BB	A	B	χ^2	P	G^2	P
东江 Dongjiang	56	9/0.1607	43/0.7679	4/0.0714	0.5446	0.4554	16.2780	0.000055	16.2780	0.000055
北江 Beijiang	46	11/0.2391	29/0.6304	6/0.1522	0.5543	0.4674	3.2314	0.072239	3.2969	0.069411
西江 Xijiang	44	10/0.2273	30/0.7273	4/0.0455	0.5909	0.4901	6.2926	0.012124	6.5805	0.010310
总计 Total	146	30/0.2055	102/0.7123	14/0.0890	0.5548	0.4452	16.2780	0.000055	17.4759	0.000029

表3 鲮*Myf5* 基因等位基因位点的群体杂合性及中性选择分析

Tab. 3 Heterozygosity analysis and the Ewens-Watterson test for neutrality of *Myf5* gene loci in different parts of Pearl River

品系 Line	基因个数 Number	观察值 Observed value		Levene's 期望值 Expected value		Nei's 期望杂合度 H_{et}	多态信息含量 PIC	有效等位基因数 Ne Effective number of alleles	固定指数 Wright's fixation index F_{is}	信息指数 Shannon's Information index	期望值 Expected value	
		群体纯合度 H_{om}	群体杂合度 H_{et}	群体纯合度 H_{om}	群体杂合度 H_{et}						观察值 F	SE
东江 Dongjiang	112	0.2321	0.7679	0.4995	0.5005	0.4960	0.3730	1.9842	-0.5481	0.6892	0.5040	0.8163 0.0273
北江 Beijiang	92	0.3696	0.6304	0.5005	0.4995	0.4941	0.3400	1.9766	-0.2759	0.6872	0.5059	0.8123 0.0262
西江 Xijiang	88	0.3182	0.6818	0.5037	0.4963	0.4907	0.2529	1.9635	-0.3895	0.6838	0.5093	0.8028 0.0276
总计 Total	292	0.3014	0.6986	0.5043	0.4957	0.4940	0.3720	1.9763	-0.4142	0.6871	0.5060	0.8480 0.0262

Popgene32 计算得 3 个江段鲮种群的基因流为 661.3523, 说明 3 个江段的鲮种群有很频繁的基因交

流。Popgene32 根据 3 个江段鲮种群基因型分布构建遗传距离图。图 4 表明, 东江与北江亲缘关系更近。

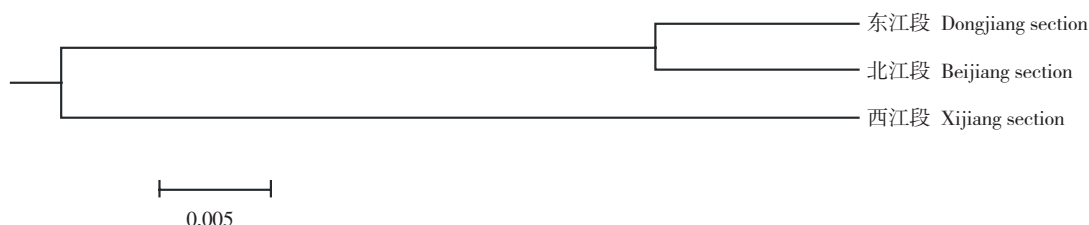


图4 珠江3个江段鲮遗传距离

Fig. 4 NJ-tree of *C. molitorella* from three sections of Pearl River according to genetic distance

3 讨论

对 *Myf5* 的 mRNA 序列推导发现, 起始密码子 AUG 上游的第 3 个核苷酸为腺嘌呤, 而紧跟 AUG 后面的核苷酸为鸟嘌呤, 这与“Kozak 序列”特征一致, 符合大多数真核生物起始密码子所具有的结构特征。

用 Popgene32 计算得 3 个江段鲮的基因流很大, 基因交流频繁, 这符合鲮在该地区的繁殖特点: 鲮卵为浮性卵, 在繁殖季节东江、北江、西江鲮卵随水流而下, 汇集于珠三角, 而后成鱼产卵时候随机逆流到 3 个不同江段, 形成 3 个江段鲮种群交流。

各个江段种群内, AA 基因型和 A 等位基因频率明显地高于 BB 基因型和 B 等位基因, 并都存在杂合子过剩, 3 个江段种群总体上不符合 Hardy-Weinberg 平衡, 说明群体中 A、B 等位基因在传递过程中可能受到过突变、选择或者是迁移、以及遗传漂变等因素影响。

分子进化的中性理论认为, 自然群体中分子水平上的遗传变异在很大程度上是中性的, 各种变异的存在与否完全取决于自然选择和遗传漂变作用。对实验中遗传变异情况的进化形式进行中性选择分析。先计算各群体在该基因座位上杂合度的观察值, 再根据中性理论进行样本含量为 1 000 的模拟群体的遗传变异预测。通过检验表明两者之间存在显著水平的差异 ($P < 0.05$), 即不遵循中性选择模式。

多态信息含量及遗传杂合度等指标是度量群体遗传变异的重要参数, 多态信息含量以及遗传杂合

度越低, 表明遗传一致性越高。3 个江段种群都有中等多态信息含量及杂合度, 说明该位点遗传变异程度高。信息指数高表明各实验群体的遗传变异性较高, 该检测结果与群体杂合度的分析基本一致。较高的 Shannon's 指数也表明, 实验群体的遗传变异性较高, 该检测结果与群体杂合度的分析基本一致。

所以, 在总体都不符合 Hardy-Weinberg 平衡的情况下, AA 基因型个体数大于 BB 基因型个体数, 由群体杂合性及中性分析表明, 该位点突变群体具体良好的多态性, 变异程度较高且不遵循中性选择模式, 则是 BB 基因型个体受到自然选择, 即 A 等位基因突变成 B 等位基因后, 纯合子 BB 基因型个体生殖力下降。杂合子 AB 基因型个体生殖力则不受影响。因此, 突变位点可以作为在选育过程度中的一个分子标记, 在选育过程中, 避开纯合基因型 BB 型个体。

从分子水平上推测, 真核细胞内蛋白质磷酸化位点主要发生在丝氨酸 (Ser)、苏氨酸 (Thr) 和酪氨酸 (Tyr) 残基侧链的羟基上, 不同的蛋白激酶可识别和修饰不同蛋白质的不同位点, 形成磷酸化作出了调控开关^[17]。通过分析基因序列, 翻译在珠江水系 3 个不同江段种群中都有 T412C 的存在, 此突变位点正好位于密码子的第一个碱基位点, 改变氨基酸组成, 将氨基序列中的 138 位丝氨酸 (Ser) 突变成脯氨酸 (Pro)。在多肽的 C-端 132 ~ 148 内富含丝氨酸 (Ser), 已克隆的鱼类 *Myf5* 基因中, 都存有此富含丝氨酸 (Ser) 的区域, 并都以 LPXESSSEPXS (P) SSXS

形式存在(X为任意氨基酸),此位点可能是可磷酸化位置,可以起磷酸化功能开关作用。氨基酸序列中丝氨酸(Ser)突变成脯氨酸(Pro)导致位点消失,无法形成磷酸化调控。蛋白质功能由其高级结构决定,而一级结构决定了多肽的高级结构。突变型纯合子BB总体上数量较少,是否因为突变后部分功能丢失,导致个体生活力下降,这还需作进一步实验证明。

参考文献:

- [1] 潘炯华. 广东淡水鱼类志[M]. 广东科技出版社,1991: 174-177.
- [2] Braun T, Arnold H H. The four human muscle regulatory helix-loop-helix proteins Myf3-Myf6 exhibit similar hetero-dimerization and DNA binding properties [J]. Nucleic Acids Res. 1991, 19: 5645-5651.
- [3] Murre C, Mc Caw P S, Baltimore D. A new binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myoG proteins [J]. Cell, 1989, 56: 777-783.
- [4] 赵高峰,陈宏,雷初朝,等. 秦川牛 GHR 基因 SNP 及其与生长性状关系的研究[J]. 遗传, 2007, 29 (3): 60-65.
- [5] Langlands K, Yin X, Anandi G, et al. Differential interactions of Id proteins with basic helix loop helix transcription factors [J]. J Biol Chem. 1997, 272: 19785-19793.
- [6] Jaynes J B, Johnson J E, Buskin J N, et al. The muscle creatine kinase gene is regulated by multiple upstream elements, including a muscle-specific enhancer [J]. Mol Cell Biol, 1988, 8: 62-70.
- [7] Zheng GM, Zhu X P, Zhang Y. Genetic Differential Research on the Mud Carp (*Cirrhina molitorella*) in different parts of Pearl River Valley [J]. J Agr Biotechnol, 2001, 9 (2): 178-182.
- [8] 郑光明,朱新平,张跃,等. RAPD 技术鉴定 3 种鲮的研究[J]. 华中农业大学学报, 1999, 18 (4): 371-374.
- [9] 郑光明,朱新平,张跃,等. 鲮、麦鲮和野鲮之 RAPD 的遗传标记[J]. 上海水产大学学报, 1999, 8.3: 216-220.
- [10] 张海军,陈宏,房兴堂,等. 山羊生长激素释放激素 (GHRH) 基因多态性研究中用 PCR-SSCP 和测序技术研究波尔山羊 (*Capra hircus*) 和徐淮白山羊 (*C. hircus*) [J]. 遗传, 2007, 29 (3): 112-116.
- [11] 孙文浩,朱庆,蒋小松,等. 鸡 *Myf6* 基因遗传多态性及其遗传效应研究[J]. 遗传, 2008, 1. 30 (1): 71-76.
- [12] 萨姆布鲁克 J, 拉塞尔 D W. 分子克隆实验指南[M]. 科学出版社, 2002: 463-470.
- [13] 许绍斌,陶玉芬,杨昭庆,等. 褚嘉祐简单快速的 DNA 银染和胶保存方法[J]. 遗传, 2002, 24 (3): 335-336.
- [14] Nei M. Kumar S. Molecular Evolution and Phylogenetics [M]. Oxford: Oxford University Press, 2000.
- [15] Avise J C. Molecular markers, natural history and evolution. New York: Chapman & Hall, 2000.
- [16] Botstein D, White R L, Skolnick M, et al. Construction of a genetic linkage map in man using restriction fragment length polymorphisms [J]. Am J Hum Gen, 1980, 32 (3): 314-331.
- [17] 赵凌志,刘颖,覃征. Weighted SVM 在蛋白质磷酸化位点预测中的应用[J]. 计算机工程与应用, 2006, 03, 155-167.
- [18] 姜运良,李宁. 肌肉生成的分子生物学研究进展[J]. 农业生物技术学报, 1999, 7 (2): 201-204.

Cloning and SNPs analysis of *Myf5* gene from *Cirrhinus molitorella*

ZHONG Maochun^{1,2}, ZHENG Guangming¹, ZHAO Jian¹, ZHU Xinping¹, MA Lisa¹, PAN Debo¹, CHEN Kunci¹, XIE Wenping¹, SHI Yan¹

(1. Pearl River Fisheries Research Institute, Chinese Academy of Fishery Sciences, Guangzhou 510380, China; 2. College of Aqua-life Science & Technology, Shanghai Ocean University, Shanghai 200090, China)

Abstract: Several molecular markers were used in mud carp genetic diversity research. There is no choiceness varietal mud carp (*Cirrhinus molitorella*) in present culture. Three part of Pearl-river drainage wild mud carp (*Cirrhinus molitorella*) were sampled as experimentation fish, and PCR-SSCP (single strand conformation polymorphism, SSCP) was used to screen *Myf5* ORF (open reading frame) DNA sequence, expecting to find SNPs relative to growth and trait. Mud carp (*Cirrhinus molitorella*) belongs to Cypriniformes, Cyprinidae, distributing in Pearl River, Ming River, Han River, Hainan, Taiwan and southeast of Asian. It is an important fishery production in Pearl-river drainage area. *Myf5* gene (myogenic factor 5) is a member of *MRFs* gene family, which include *Myf5*, *MyoG*, *MyoD* and *MRF4*. The primary function of the gene family is advance myoblast multiplication and differentiation. The function of *MyoD* and *Myf5* genes is determined to initiate the differentiation process of that muscular satellite cells to be muscle derived stem cell. And that of *MyoG* and *MRF4* genes was thought to be involved in regulating the differentiation process of muscle derived stem cell to be myotube and muscle fibers. All of those functions were implementing by the typical trait b-HLH structure of MRF family, which form dimeric by combining with E-protein. Then Key gene is activated and pressed by the dimeric recognize upstream E-box. Many species of *Myf5* gene or cDNA have been cloned. Such as *Mus musculus*, *Bos Taurus*, *Sus scrofa*, *Ovis ammon*, *Gallus gallus*, *Danio rerio*, *Cyprinus carpio*, *Micropterus salmoide* and *Oncorhynchus mykiss*.

Myf5 gene of mud carp (*Cirrhinus molitorella*) has been cloned, sequenced and analyzed by PCR-SSCP. Mud carp Genome DNA and total RNA was extracted from flesh muscle of mud carp. *Myf5* gene was cloned including three exons and two introns of which the length is 453 bp, 76 bp, 194 bp and 1 311 bp and 90 bp (GeneBank accession: GU289508). The cDNA sequence of *Myf5* gene was obtained by RT-PCR. The length of the cDNA is 1 104 bp, and 723 bp ORF (open reading frame) was identified, which shared 93.6% homology with that of *Cyprinus carpio*, 85.9% homology with *Danio rerio*. It encodes a peptide of 240 amino acids whose calculated molecular weight is 26.288 31 kD and theoretical isoelectric is 6.08. AA sequence analysis showed that the *Myf5* protein have the typical trait b-HLH structure of MRF family. The population genetic information about genetic distribution, variation and heterozygosity of *Myf5* gene was analyzed by using PCR-SSCP in three part of the Pearl River. The only one mutation C/T at base position 412 of the whole CDS was found among individuals in whole group. Through further analysis of the position, it could be inferred that the genotypes distribution of *Myf5* in three parts of the Pearl River were not in accordance with the Hardy-Weinberg equilibrium except the north of the Pearl River subgroup, and not in accordance with the Hardy-Weinberg equilibrium in whole group. All of the three groups have high heterozygosity and moderate PIC, which indicate that the mutation is high degree in the groups and fine genetic diversity. Furthermore, there is obvious differential between observed values and expected values that indicate the evolution is discord with theory of neutral selection. In the whole of the Pearl River, A and B alleles are that not in accordance with the Hardy-Weinberg equilibrium, high degree mutation, fine genetic diversity and discord with theory of neutral selection. The number of AA type is more than that of BB type evidently, and the conclusion can be made that BB type individual has been selected. Consequently, the SNP can be used in the breeding process, avoiding the BB type as parents

From the molecular level, to speculate, phosphorylation switch site changed or disappeared because of the mutation, resulting in decline in individual life force. This is in the need of further experiments. [Journal of Fishery Sciences of China, 2010, 17(4): 681–688]

Key words: mud carp; *Myf5* gene; SSCP; genetic diversity

Corresponding author: ZHENG Guangming. Tel: 020-81616509; E-mail: zgmzy1@tom.com