

文章编号: 1674-5566(2011)05-0649-06

草鱼 *MKLNI* 基因可变剪接和微卫星序列的多态性

徐冰¹, 徐义平², 鲍宝龙¹

(1. 上海海洋大学 省部级共建水产种质资源发掘与利用教育部重点实验室, 上海 201306; 2. 浙江省乐清市海洋与渔业局, 浙江 乐清 325600)

摘要: *MKLNI* 基因编码的 muskelin 蛋白主要通过其羧基端的 kelch 重复结构域与其它蛋白形成复合物行使功能。通过分析草鱼鳃组织的基因转录谱, 发现其中的 *MKLNI* 基因存在 (CT)₁₅ 的微卫星序列, 通过检测 3 个不同的草鱼种群, 又发现 (CT)₉、(CT)₁₀、(CT)₁₁ 和 (CT)₁₃ 4 种多态性序列。通过克隆该段 *MKLNI* 基因组序列并与斑马鱼和红鳍东方鲀的 *MKLNI* 基因组序列比较, 发现克隆得到的草鱼基因序列是由一个内含子和两个外显子组成, 共编码 144 个氨基酸。通过进一步调查不同组织该基因的转录本, 发现存在内含子剪接和内含子未被剪接的两个转录本。由于内含子序列未被剪接的转录本在内含子中存在“TGA”终止密码子, 导致 *MKLNI* 基因翻译提前终止, 翻译产物为失去部分 kelch 结构域的 muskelin 蛋白。相对于内含子正常剪接的转录本, 没有被剪接内含子的转录本在各种组织中的表达量要明显低得多。简而言之, 本研究发现了草鱼 *MKLNI* 转录本中的微卫星序列是来自未被剪切的内含子, 而未剪切的内含子导致该转录本编码部分 kelch 结构域缺失的 muskelin 蛋白。

研究亮点: 目前对于可直接锚定功能基因的 I 型微卫星标记研究很少。本文首次发现 *MKLNI* 基因 EST 序列中的二核苷酸的微卫星是来自可变剪接导致的未被剪切掉的内含子, 表明 RNA 可变剪接是造成 ESTs 序列中含微卫星序列的一种途径。

关键词: 草鱼; *MKLNI*; 微卫星多态性; 可变剪接

中图分类号: S 917

文献标志码: A

草鱼 (*Ctenopharyngodon idellus*) 是我国重要的淡水养殖鱼类, 围绕草鱼育种和草鱼抗病的分子水平上的研究, 已越来越受重视^[1-2]。我们前期对草鱼肠道、鳃和肝脏组织的转录谱进行了分析, 得到了 6 269 个表达序列标签 (ESTs), 其中一个 ESTs 为 *MKLNI*, 包含二核苷酸的微卫星序列^[3]。*MKLNI* 所编码 muskelin 的蛋白是一个非常保守的胞内蛋白, 可通过羧基端的 CTLH 基序与 RanBPM (小 GTP 酶 Ran 结合蛋白)、Twa1、p48EMLP 和 p44CTLH 等蛋白一起共同组成复合物行使细胞粘附的功能^[4-5]。此外, 还具有 kelch 重复结构域, 每一个 kelch 结构域大约由 44~56 个氨基酸组成, 主要依靠各个重复结构域相互结合在一起形成一种 β 螺旋结构, 作为多分子复合物的组织者将其它蛋白联系在一起行使

信号转导的功能^[6]。为了确定草鱼 *MKLNI* 的微卫星序列多态性以及这种多态性对该基因编码的影响, 我们克隆并分析了草鱼 *MKLNI* 的基因组序列, 并意外发现该基因还存在 RNA 可变剪接现象。

1 材料与方法

1.1 样品来源

用作测试微卫星多态性的草鱼样品分别来自上海市青浦养殖池塘^[7]、湖北石首老河长江家鱼原种场和匈牙利多瑙河养殖场收集的野生群体^[8]; 长江和匈牙利草鱼样本剪取尾鳍, 置 95% 酒精中保存, 备作基因组 DNA 提取; 抽取总 RNA 样品的草鱼来自上海市临港新城古棕路菜市场。

收稿日期: 2011-03-11 修回日期: 2011-04-20

基金项目: 上海市教育委员会重点学科建设项目 (S30701 和 J50701)

作者简介: 徐冰 (1985—), 女, 硕士研究生, 研究方向为发育生物学。E-mail: 13633220282-435@163.com

通讯作者: 鲍宝龙, E-mail: blbao@shou.edu.cn

1.2 基因组 DNA 提取

长江和匈牙利草鱼样本的基因组 DNA 提取方法参照文献[8]。上海草鱼样本基因组 DNA 提取来自新鲜草鱼鳍条组织,具体方法参考文献[9]。

1.3 微卫星多态性的检测及序列测定

以实验室构建的草鱼鳃组织 cDNA 文库中 *MKLNI* 基因编码区序列为基础,在微卫星两翼设计上游引物:5'-CAGGATTTCTGCCACATACCACCA-3'和下游引物 5'-TCCATGCACATCTGAGAGAGAGAGA-3'。PCR 反应体系总体积 10 μL , 0.5 μL 模板 DNA, 上游引物 0.5 μL , 下游引物 0.5 μL , 3.5 μL ddH₂O, 2 \times Mix 5 μL 。PCR 反应程序为 94 $^{\circ}\text{C}$ 5 min, 35 个循环(94 $^{\circ}\text{C}$ 30 s, 58 $^{\circ}\text{C}$ 30 s, 72 $^{\circ}\text{C}$ 10 min)。扩增产物用 8% 的聚丙烯酰胺凝胶电泳分离,银染法显色。根据电泳条带显色情况计算多态性基因的等位基因数,然后割胶,把所割各胶分别置于 4 $^{\circ}\text{C}$ 无菌水中过夜。以该无菌水为模板,再用以上相同的 PCR 体系和 PCR 扩增条件进行扩增,琼脂糖电泳后割胶纯化,克隆后测序。

1.4 *MKLNI* 基因组部分序列的克隆和分析

以本实验室构建的草鱼鳃组织 cDNA 文库中 *MKLNI* 基因编码区序列为基础,设计引物克隆 *MKLNI* 部分基因组 DNA(H09ORF:5'-AACGGTCCAGTGCTCGCTC-3', H09ORR:5'-GTGGAAGAGCATGCAGTGCCAATG-3'), PCR 反应体系总体积 50 μL , 1 μL 模板 DNA, 上游引物 1 μL , 下游引物 1 μL , 22 μL ddH₂O, 2 \times Mix 25 μL 。PCR 反应程序为 94 $^{\circ}\text{C}$ 5 min, 35 个循环(94 $^{\circ}\text{C}$ 30 s, 58 $^{\circ}\text{C}$ 30 s, 72 $^{\circ}\text{C}$ 1 min 45 s), 72 $^{\circ}\text{C}$ 10 min, 割胶测序。最后将得到的草鱼 *MKLNI* 基因组部分序列与斑马鱼 (*Danio rerio*)^[10] 和红鳍东方鲀 (*Takifugu rubripes*)^[11] *MKLNI* 基因进行比对,预测其基因组结构 (<http://useast.ensembl.org/index.html>)。

1.5 RT-PCR 及可变剪接转录本的克隆

分别取 100 mg 左右的草鱼鳃、精巢、头肾、皮肤、肝脏脏、肠道、肌肉和肾脏组织用剪刀剪碎置于 1 mL 的 Trizal 中匀浆,总 RNA 具体的抽提方法参考文献[12]。

取以上抽提的总 RNA,利用 18T 引物进行逆转录,逆转录方法参照 M-MLVRT 试剂盒

(Promega 公司)所提供的方法。用克隆 *MKLNI* 基因组序列的引物 H09ORF 和 H09ORR 进行 PCR 扩增。反应体系总体积为 25 μL , 包括 2 μL 模板 cDNA, 1 μL 上游引物, 1 μL 下游引物, 16 μL ddH₂O, 2 \times PCR Buffer Mix 2.5 μL , MgCl₂ (25 mmol/L) 1.5 μL , *Taq* 酶 0.5 μL , dNTP (10 mmol/L) 0.5 μL 。PCR 反应程序为 94 $^{\circ}\text{C}$ 5 min, 35 个循环(94 $^{\circ}\text{C}$ 30 s, 58 $^{\circ}\text{C}$ 30 s, 72 $^{\circ}\text{C}$ 1 min), 72 $^{\circ}\text{C}$ 10 min。琼脂糖电泳后割胶、纯化、克隆和测序。

2 结果

2.1 *MKLNI* cDNA 序列及其所包含的微卫星序列多态性

在草鱼鳃 cDNA 文库中得到的 SFU-GT7-A008-H09 序列长为 629 bp, 此序列已提交至 GenBank(接受号为 GT226477), 通过进一步将该克隆测序,得到 1 122 bp 序列,通过 GenBank 比对确定该序列为草鱼的 *MKLNI* 基因。利用草鱼该基因片段的核苷酸序列进行 Blastn 比对,显示其与斑马鱼的 *MKLNI* 有 99% 相似性;利用该基因片段编码的氨基酸序列进行 Blastn 比对,显示其与斑马鱼及红鳍东方鲀的 *MKLNI* 有 100% 相似性,与近十几个物种有 99% 的相似性,说明该基因在进化过程中高度保守,这可能与编码蛋白质的功能保守性有关。

MKLNI cDNA 序列在 707 bp 处有 15 个 CT 重复的微卫星序列,利用微卫星两翼引物对分别来自长江、匈牙利和上海的草鱼群体进行多态性检测,并进行克隆验证,发现共有 4 条多态性序列,其重复数分别为 (CT)₉、(CT)₁₀、(CT)₁₁ 和 (CT)₁₃(图 1)。每一个体最多只有两个多态位点,表明草鱼 *MKLNI* 为单拷贝基因。根据斑马鱼和红鳍东方鲀全基因组序列,*MKLNI* 在这两种鱼中也是单拷贝基因。出现两个多态位点个体,为杂合子,而只有单个位点的个体,则为纯合子。

2.2 微卫星序列位于 *MKLNI* 第 10 内含子区域

把 1 122 bp 的 *MKLNI* cDNA 序列与斑马鱼和红鳍东方鲀的 *MKLNI* 的基因组序列进行比对分析,发现含 CT 重复的微卫星序列来自 *MKLNI* 基因的第 10 内含子区域。利用引物扩增草鱼 *MKLNI* 基因组第 10 内含子区域,扩增到一个长度为 593 bp 的序列(图 2)。第 10 内含子序列含有典型的“GT-AG”结构,微卫星序列位于草鱼的

第 10 内含子区域。

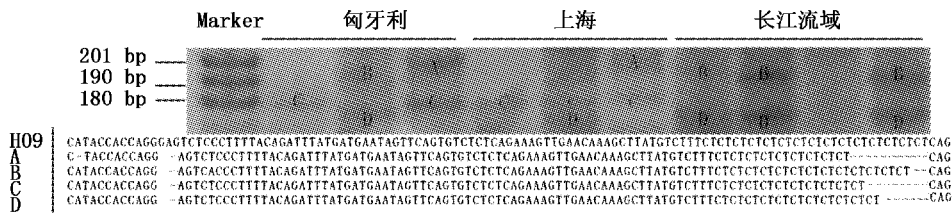


图 1 不同草鱼群体的微卫星多态性及序列差异

Fig. 1 The microsatellite polymorphism and sequence contrast of different grass carp populations

M 为 pBR322 Marker; H09 为草鱼鳃 cDNA 文库中得到的 SFU-GT7-A008-H09 克隆核苷酸序列, 其微卫星序列重复数为 (CT)₁₅; A、B、C、D 分别代表重复数为 (CT)₉、(CT)₁₃、(CT)₁₀、(CT)₁₁ 的多态性序列。

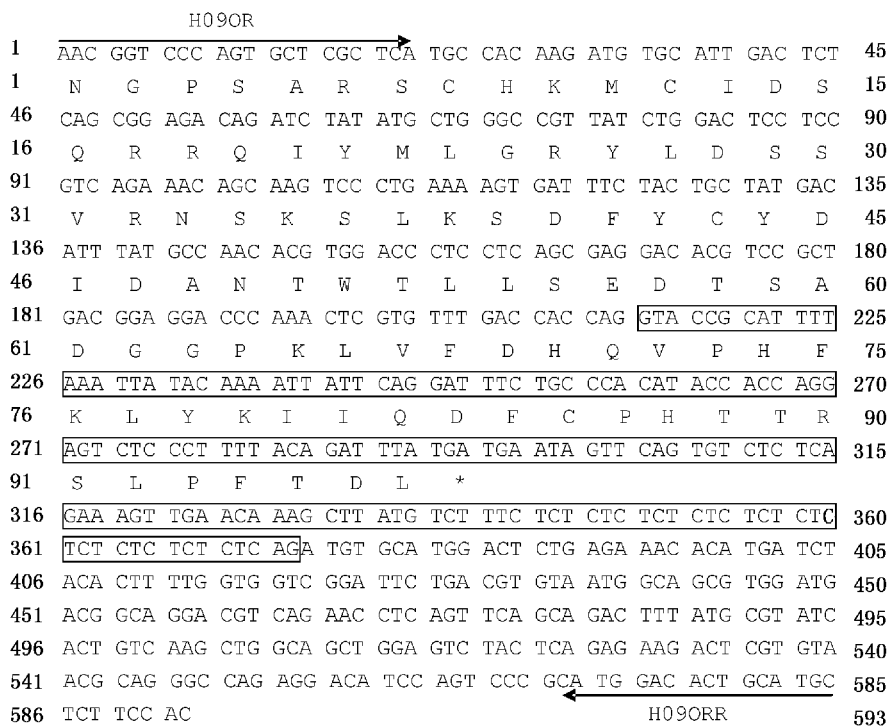


图 2 含第 10 内含子的草鱼 *MKLNI* 基因组部分序列

Fig. 2 The *MKLNI* genome sequence of *Ctenopharyngodon idellus* including intron 10

黑色箭头分别代表上下游引物序列 H09ORF 和 H09ORR; 黑色方框代表内含子 10; 其中黑色下划线代表微卫星序列。

2.3 *MKLNI* 的可变剪接及其组织特异性

PCR 扩增草鱼 *MKLNI* 基因外显子 10、11 区域的 cDNA 序列, 在鳃组织中扩增到一个 593 bp 的长片段和一个 432 bp 的短片段(图 3), 长片段比短片段多出的 161 bp 为第 10 内含子, 表明草鱼 *MKLNI* 基因存在可变剪接现象, 可变剪接模式如图 4 所示, 微卫星序列就位于存在可变剪接的这个内含子区域。

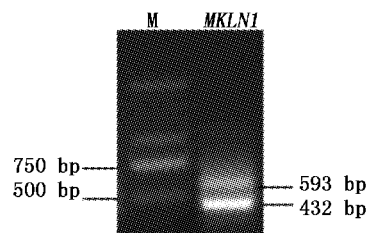


图 3 *MKLNI* 基因在草鱼鳃组织中的表达情况
Fig. 3 The expression of *MKLNI* in the gill of *Ctenopharyngodon idellus*

M 为 Marker DL2000。

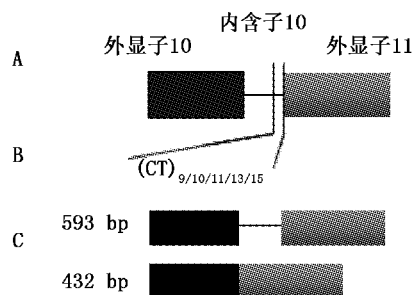


图4 *MKLN1* 基因的基因组结构及其可变剪接模式图

Fig. 4 The gene structure and the alternative splicing of *MKLN1*

A 为 *MKLN1* 基因的结构(黑色细线表示内含子 10, 长方形表示外显子); B 为微卫星序列在内含子中的位置及已克隆得到的重复种类; C 为克隆得到的 *MKLN1* 基因的可变剪接模式。

通过进一步对各可变剪接转录本蛋白质功

能分析发现, 161 bp 内含子被剪切掉的 432 bp 的短片段编码 144 个氨基酸, 而 593 bp 的长片段只编码 97 个氨基酸。如果草鱼 *MKLN1* 基因的外显子 10 之前读码框未被破坏的话, 由于 593 bp 中 161 bp 的内含子 10 未被剪切将会导致该段外显子编码到第 97 个氨基酸时遇到内含子 10 中的“TGA”终止密码子使蛋白质翻译终止(图 2), 这样将会产生失去第二个 kelch 结构域的蛋白, 产生失去功能的蛋白。

通过调查草鱼鳃、精巢、头肾、皮肤、肝胰脏、肠道、肌肉和肾脏组织的 *MKLN1* 基因表达, 发现肠道、皮肤、肝胰脏、肌肉、精巢、肾脏、头肾和鳃中都得到了没有剪切掉内含子 10 的长为 593 bp 的片段和剪切掉内含子 10 的 432 bp 短片段, 各组织中 593 bp 的转录本的表达强度明显要低于 432 bp 转录本的表达强度(图 5)。



图5 *MKLN1* 基因在草鱼各组织中的可变剪接分析

Fig. 5 Different expression of *MKLN1* in the tissues of *Ctenopharyngodon idellus*

3 讨论

微卫星指的是由 1~6 bp 核苷酸为重复单元组成的重复序列^[13], 并且与许多生命活动密切相关, 比如影响染色质的分配, 控制基因活性、基因重组、DNA 复制、细胞分裂和错配修复系统等^[14]。在所有已知的基因组中几乎都有微卫星分布, 并且广泛分布于编码区和非编码区, 例如在人类连锁图谱中发现了 5 264 个 (AC/TG)_n 微卫星多态性位点^[15]。微卫星多态性表现于同一微卫星位点在不同个体之间以及同一个体的正常组织与某些异常组织之间微卫星位点重复单位的数目不同, 本研究在调查微卫星多态性过程中发现该微卫星位点在不同个体间重复数不同, 说明该基因在草鱼中确实具有多态性。

可变剪接指的是 mRNA 前体的不同剪接方式, 其中的某个或某些外显子被代替、添加或删除, 从而加工成不同的成熟 mRNA, 翻译得到多种蛋白质^[16]。可变剪接现象普遍存在于真核生物中并且高频率出现在不同基因中, 比如, 在对

6 400 个人类基因调查中就发现 63% 的基因存在可变剪接现象^[17], 真核生物中 mRNA 前体的可变剪接是单个基因产生不同可变剪接转录本的重要机制, 在调控基因功能中起了重要的作用。本研究发现的草鱼 *MKLN1* 基因微卫星序列位于内含子 10 中, 该内含子未被剪切掉的转录本, 由于在该微卫星序列 39 bp 前(图 2) 出现“TGA”终止密码子导致翻译终止, 因此该微卫星序列的存在及其重复数的改变并不影响读码框的编码。但内含子 10 中的微卫星 (CT) 重复数目的不同, 是否影响该内含子的剪接, 还需要今后进一步的研究。

有较多例子表明可变剪接的存在会导致所编码的蛋白丧失某些特定的功能, 如在草鱼 CC 趋化因子 *CCL24* 和 *CCL26* 基因中, 由于其中一个转录本的内含子没有被剪切掉, 导致翻译后内含子中的终止密码子“TGA”提前终止了完整基因的翻译, 最终导致编码的蛋白功能缺失, 而且各转录本在不同组织中的表达量不尽相同^[12, 18]。通过分析斑马鱼和红鳍东方鲀 *MKLN1* 基因组序

列,发现两物种的 *MKLNI* 基因中 kelch 结构域都为两个重复。虽然有研究表明 muskelin 中单个 kelch 基序也能行使相应功能,比如在 COS7 或 C2C12 细胞中过表达 muskelin,氨基端的第一个 kelch 基序足够用于增加细胞和 hTSP-1 羧基端球状域的结合能力^[19],但是研究表明 β 螺旋结构的完整性、kelch 结构域本身发生缺失甚至是点突变都会对蛋白质功能产生影响,例如 HCF-1 中 β 螺旋结构的部分缺失及 RAG-2 中的点突变都会导致编码的蛋白失去蛋白结合功能,这些对于 kelch 结构及突变的研究都强调了完整的 β 螺旋结构在行使生物学功能中的重要性^[20-21]。因此本研究在草鱼中发现的一个 *MKLNI* 基因转录本,由于内含子 10 未被剪切导致该段外显子编码到第 97 个氨基酸时遇到内含子 10 中的“TGA”终止密码子使蛋白质翻译终止,这样将会产生失去一个 kelch 结构域的蛋白,从而可能导致编码的蛋白失去原有功能。

参考文献:

- [1] 苏建明,章怀云,肖调义. 草鱼抗病育种研究进展[J]. 内陆水产, 2002, 27(1): 43-45.
- [2] LIU F, WANG D, FU J, et al. Identification of immune-relevant genes by expressed sequence tag analysis of head kidney from grass carp (*Ctenopharyngodon idella*) [J]. Comparative Biochemistry and Physiology, Part D, Genomics & proteomics, 2010, 5(2): 116-23.
- [3] XU B, WANG S, JIANG Y, et al. Generation and analysis of ESTs from the grass carp, *Ctenopharyngodon idellus* [J]. Animal Biotechnology, 2010, 21(4): 217-225.
- [4] UMEDA M, NISHITANI H, NISHIMOTO T. A novel nuclear protein, Twa1, and Muskelin comprise a complex with RanBPM [J]. Gene, 2003, 303: 47-54.
- [5] KOBAYASHI N, YANG J, UEDA A. RanBPM, Muskelin, p48EMLP, p44CTLH, and the armadillo-repeat proteins ARMC8alpha and ARMC8beta are components of the CTLH complex [J]. Gene, 2007, 396(2): 236-247.
- [6] ADAMS J, KELSO R, COOLEY L. The kelch repeat superfamily of proteins: propellers of cell function [J]. Trends in Cell Biology, 2000, 10(1): 17-24.
- [7] 蒋燕,谢彩霞,杨桂梅,等. 16S DNA 文库随机测序法分析草鱼肝胰脏细菌群落组成[J]. 上海海洋大学学报, 2010, 19(4): 440-446.
- [8] 宋晓,李思发,王成辉,等. 草鱼中国土著群体与欧美日移居群体遗传差异的线粒体序列[J]. 水生生物学报, 2009, 33(4): 709-716.
- [9] 杨琳琳,蒋燕,杨桂梅,等. 银鲟皮肤和肌肉组织的基因表达谱分析[J]. 上海水产大学学报, 2008, 17(4): 390-395.
- [10] SONG H D, SUN X J, DENG M, et al. Hematopoietic gene expression profile in zebrafish kidney marrow [J]. Proceedings of the National Academy of Sciences, 2004, 101(46): 16240-16245.
- [11] APARICIO S, CHAPMAN J, STUPKA E, et al. Whole-genome shotgun assembly and analysis of the genome of *Fugu rubripes* [J]. Science, 2002, 297(5585): 1301-1310.
- [12] 杨琳琳,谢彩霞,龚小玲,等. 草鱼 CC 趋化因子基因 CCL24 及其可变剪接[J]. 上海海洋大学学报, 2010, 19(3): 296-301.
- [13] RAMEL C. Mini- and microsatellites [J]. Environmental Health Perspectives, 1997, 105(4): 781-789.
- [14] LI Y C, KOROL A B, FAHIMA T, et al. Microsatellites within genes: structure, function, and evolution [J]. Molecular Biology and Evolution, 2004, 21(6): 991-1007.
- [15] DIB C, FAURE S, FIZAMES C, et al. A comprehensive genetic map of the human genome based on 5,264 microsatellites [J]. Nature, 1996, 380(6570): 152-154.
- [16] LOPEZ A J. Alternative splicing of pre-mRNA: developmental consequences and mechanisms of regulation [J]. Annual Review of Genetics, 1998, 32: 279-305.
- [17] KAN Z, STATES D, GISH W. Selecting for functional alternative splices in ESTs [J]. Genome Research, 2002, 12(12): 1837-1845.
- [18] 鲍宝龙,李家乐,汪桂玲,等. 斑点叉尾鲷趋化因子 SCYA126 基因及其可变剪接[J]. 中国水产科学, 2007, 14(1): 1-7.
- [19] ADAMS J C, SEED B, LAWLER J. Muskelin, a novel intracellular mediator of cell adhesive and cytoskeletal responses to thrombospondin-1 [J]. The EMBO Journal, 1998, 17(17): 4964-4974.
- [20] CALLEBAUT I, MORNON J P. The V(D)J recombination activating protein RAG2 consists of a six-bladed propeller and a PHD fingerlike domain, as revealed by sequence analysis [J]. Cellular and Molecular Life Sciences, 1998, 54(8): 880-891.
- [21] WILSON A C, FREIMAN R N, GOTO H, et al. VP16 targets an amino-terminal domain of HCF involved in cell cycle progression [J]. Molecular and Cellular Biology, 1997, 17(10): 6139-6146.

Polymorphic microsatellite in the alternative splicing transcription of *MKLNI* in *Ctenopharyngodon idellus*

XU Bing¹, XU Yi-ping², BAO Bao-long¹

(1. Key Laboratory of Exploration and Utilization of Aquatic Genetic Resources, Shanghai Ocean University, Ministry of Education, Shanghai 201306, China; 2. Bureau of Marine and Fisheries, Yueqing 325600, Zhejiang, China)

Abstract: Muskelin coded by *MKLNI* gene participates in signal transduction, cell-matrix adhesion and cellular component movement as an intracellular protein. In previous study, *MKLNI* containing 15 repeats of dinucleotide (CT) was identified from a cDNA library of grass carp, *Ctenopharyngodon idellus*. To further investigate the polymorphism of the dinucleotide, we examined ten individuals from three different grass carp population in this research and identified other different repeats of dinucleotide as (CT)₉, (CT)₁₀, (CT)₁₁ and (CT)₁₃. Each individual owning at most two loci indicates only one copy *MKLNI* in the genome of grass carp. Compared with its orthologue in zebrafish (*Danio rerio*) or fugu (*Takifugu rubripes*) genome, the dinucleotide microsatellite was found not in an exon, but in one intron corresponding 10th intron of *MKLNI* in *Danio rerio* or *Takifugu rubripes*. Further investigation mRNA fragment spanning from 10th to 11th exons showed that there existed mRNA alternative splicing in *MKLNI* and the microsatellite was indeed located in unspliced 10th intron in grass carp. Terminal code "TGA" locating in unspliced intron leads the transcript to translate into truncated muskelin. The truncated muskelin might lose partial function because of kelch domain broken. Moreover, the abundance of unspliced transcript was higher than that of spliced transcript in gill, intestine, head kidney, skin, hepatopancreas, muscle and kidney. Taken together, the microsatellite we found in one EST of *MKLNI* is from an unspliced intron, and the unspliced intron leads the transcript to encode truncate muskelin.

Key words: *Ctenopharyngodon idellus*; *MKLNI*; polymorphic microsatellite; alternative splicing transcription