

文章编号: 1674 - 5566(2011)01 - 0034 - 10

团头鲂主要组织相容性复合体 I 类基因全长 cDNA 的克隆及组织表达分析

马晓茜¹, 刘至治¹, 李思发², 唐文乔¹, 杨金权¹

(1. 上海海洋大学 水产种质资源发掘与利用教育部重点实验室, 上海 201306; 2. 上海海洋大学 农业部水产种质资源与利用重点开放实验室, 上海 201306)

摘要: 为了解团头鲂 MHC I 类基因的结构, 采用 cDNA 末端快速扩增(Rapid-amplification of cDNA ends, RACE)技术, 首次成功克隆了团头鲂“浦江 I 号”F₀ 基础群体及选育 F₈ 世代的 MHC I 类基因, 共获得 3 条 cDNA 全长序列。序列全长为 2 040 ~ 2 079 bp, 含有 87 ~ 102 bp 的 5'-UTR, 1 035 ~ 1 044 bp 的编码区(包括信号肽, alpha 1, alpha 2 和 alpha 3 三个结构域, 跨膜区和胞质区)及 911 ~ 946 bp 的 3'-UTR 区, 分别编码 347, 344 个氨基酸。F₈ 世代序列的核苷酸/氨基酸的同源性很高, 为 89.0%/93.0%, 而与 F₀ 世代的同源性较低(75.3% ~ 77.0% 和 71.1% ~ 71.4%), 呈现出明显的分子多态性。分析表明, 团头鲂具有经典的 MHC I 类分子的空间结构, 与人类 HLA-A2 的抗原肽结合区的晶体结构相比, 二者的差异主要集中在 5 个(I ~ V)区上。在 Neighbor-joining(NJ)法构建的分子进化树中, 团头鲂与草鱼的亲缘关系最近, 与鲑鳟鱼类、爬行类、鸟类、哺乳类及人类的亲缘关系则渐远。RT-PCR 组织表达分析表明, 团头鲂 MHC I 类基因在所检测的 8 个组织中均表达, 在鳃、头肾和血液中有较强的转录本, 中等强度表达于肝脏和脾脏, 在后肾、肠和肌肉中表达较弱。

研究亮点: (1) 以团头鲂“浦江 I 号”F₀ 基础群体、选育 F₈ 世代的样本为材料, 时间跨度 20 多年, 材料难得。(2) 利用脊椎动物中变异最大的基因 MHC 为手段, 从免疫遗传学角度, 首次克隆、测序团头鲂 MHC I 类分子, 分析其空间结构, 了解其分子多态性, 为深入分析人工选择压力对团头鲂群体遗传变异的影响奠定了基础。

关键词: 团头鲂; 主要组织相容性复合体; cDNA; 克隆; 组织表达

中图分类号: S 917

文献标识码: A

主要组织相容性复合体(Major histocompatibility complex, MHC)是目前已知存在于所有脊椎动物体内的与免疫功能密切相关的高度多态基因群^[1-2], 它编码的糖蛋白广泛参与调控动物机体免疫应答^[3-4]。近 20 多年来, 国内外学者已对斑马鱼(*Brachydanio rerio*)^[5]、鲤(*Cyprinus carpio*)^[6]、线纹狼鲈(*Morone saxatilis*)^[7]、虹鳟(*Oncorhynchus mykiss*)^[3]、牙鲮(*Paralichthys olivaceus*)^[8]、大菱鲆(*Scophthalmus maximus*)^[9]、草鱼(*Ctenopharyngodon idellus*)^[10]、赤点石斑鱼(*Epinephelus akaea*)^[11]、半滑舌鳎(*Cynoglossus semilaevis*)^[12]等鱼类的 MHC 基因进行了分析。

MHC 不仅具有高度多态性的特点, 而且能提供鱼类个体适应性的信息^[13-15], 因此, 在鱼类种群结构的鉴定、种群遗传学等方面的研究中得到广泛应用^[13-14]。

团头鲂(*Megalobrama amblycephala*)是中型草食性鱼类, 为我国主要养殖对象之一^[16]。团头鲂“浦江 I 号”是经 15 年的系统选育产生的优良品种^[17], 至 2009 年底已完成第 9 代选育(F₉)。然而, 最近的研究结果表明, F₉ 的多态位点百分数要比基础群体下降 30.29%^[18]。遗传多样性的下降, 对生活在野外多变环境中的群体来说是一种极大威胁。本研究采用 cDNA 末端快速扩增

收稿日期: 2010-07-21 修回日期: 2010-09-14

基金项目: 国家自然科学基金重点项目(30630051); 上海市重点学科建设项目(S30701)

作者简介: 马晓茜(1986-), 女, 硕士研究生, 专业方向为鱼类分子进化与免疫遗传学。E-mail: happyxiaoxi1986@163.com

通讯作者: 刘至治, Tel: 021-61900425, E-mail: zzliu@shou.edu.cn

(Rapid Amplification of cDNA Ends, RACE) 技术,首次克隆团头鲂“浦江 I 号”基础群体 (F_0) 及人工选育世代 (F_8) 的 MHC I 类基因 cDNA 的全长序列,预测和分析其理化性质及高级结构,为进一步了解长期(20 多年)人工选育对团头鲂群体遗传变异的影响程度奠定基础。

1 材料与方 法

1.1 实验材料

团头鲂基础群体 (F_0)、“浦江 I 号”选育世代 (F_8) 样本各 1 尾,体重 800 ~ 1 200 g,于 2007 年 5 月采自上海海洋大学南汇水产种质资源试验站。取团头鲂肝脏,迅速放入液氮中冻存,转至 $-80\text{ }^\circ\text{C}$ 冰箱中保存备用。

1.2 引物设计与 PCR 扩增条件

根据已报道的人、鼠、鸡、蛙和斑马鱼的 MHC

I 类基因序列,在形成二硫键的两个半胱氨酸保守区设计一对兼并引物 (Meam- α_2 f 和 Meam- α_2 r),用于扩增团头鲂 MHC I 类基因 α_2 结构域的部分片段 (222 bp)^[19]。用 Primer Premier 5.0 (<http://www.bio-soft.net>) 软件设计两组特异性引物 (MSP5 和 MSP3),分别用于扩增 MHC I 类 cDNA 序列的 5' 和 3' 末端 (表 1)。以上引物均由上海生工生物工程有限公司合成。5'-RACE 上游引物 UPM 和 3'-RACE 下游引物 M4 分别来自反转录试剂盒 BD SMARTTM-RACE cDNA Amplification kit (Clontech) 和 RNA PCR kit (AMV) Ver. 3.0 (TaKaRa)。由于 3' 和 5' 分别存在多态性,又根据所得序列重新设计特异性引物一对 (Meam-F 和 Meam-R) (表 1),用来扩增 MHC I 类基因编码区序列。

表 1 团头鲂 MHC I 基因 PCR 引物序列、退火条件及产物

Tab. 1 Primer sequences, PCR product and annealing temperatures of MHC I gene in *Megalobrama amblycephala*

引物名称	引物序列 (5'→3')	退火温度 (°C)	扩增对象	产物大小 (bp)
MSP5-1	GCTGGGTGATGATTCCTTGCTGCACTGGGG	69.7	F_0 样本	约 530
MSP3-1	ATGTA'GGCTGTGAATGGGATG	58		约 100
MSP5-2	CAGTCCAGGTCAGGGTGGCCTTATCC	68	F_8 样本	约 530
MSP3-2	CGGCTGTGAGCTGGATGAAGATGGC	58		约 100
UPM	CTAATACGACTCACTATAGGGC			
M4	GTTTTCCCACTCACCAC			
Meam- α_2 f	CARHNGATGTAYGGNTGT*	51.5	部分 α_2 结构域	222
Meam- α_2 r	YTTNARCCAYTCNATRCA			
Meam-F	ATGCGATCTGTAGTGGTTTTCCTCCTCGG	54	F_0 和 F_8 样本	约 100
Meam-R	TATCGGAAAAAGAAAGGCTTTAAACCTGT			
RT-f	CCCCACTGGAGTAACAA	58	F_9 样本	115
RT-r	ATCGCCGCACTGGTTCTTGAC	58		
β -actin-F	ATCGCCGCACTGGTTCTTGAC	58		334
β -actin-R	CCTGTTGGCTTTGGTTC	58		

注: 兼并引物中 R = A + G, H = A + C, N = A + T + C, Y = T + C。

PCR 反应条件: 5'-RACE, $94\text{ }^\circ\text{C}$ 预变性 3 min, 接着 $94\text{ }^\circ\text{C}$ 30 s、退火 30 s、 $72\text{ }^\circ\text{C}$ 2 min, 30 个循环后 $72\text{ }^\circ\text{C}$ 延伸 10 min, $4\text{ }^\circ\text{C}$ 保温。3'-RACE, $94\text{ }^\circ\text{C}$ 预变性 3 min, 接着 $94\text{ }^\circ\text{C}$ 30 s、退火 30 s、 $72\text{ }^\circ\text{C}$ 1 min, 30 个循环后 $72\text{ }^\circ\text{C}$ 延伸 10 min, $4\text{ }^\circ\text{C}$ 保温。编码区扩增: $94\text{ }^\circ\text{C}$ 预变性 3 min, 接着 $94\text{ }^\circ\text{C}$ 30 s、 $54\text{ }^\circ\text{C}$ 退火 30 s、 $72\text{ }^\circ\text{C}$ 90 s, 35 个循环后 $72\text{ }^\circ\text{C}$ 延伸 10 min, $4\text{ }^\circ\text{C}$ 保温。

1.3 总 RNA 的提取和 fcDNA 的合成

取少量肝脏组织,使用动物组织总 RNA 提取试剂盒 (RNAprep pure Tissue Kit, TIANGEN) 提取总 RNA。利用反转录试剂盒 BD SmartTM-RACE

cDNA Amplification kit (Clontech) 合成 5'-RACE-Ready-cDNA, 利用反转录试剂盒 RNA PCR kit (AMV) Ver. 3.0 (TaKaRa) 反转录得到 3'-RACE-Ready-cDNA。所得 fcDNA 于 $-20\text{ }^\circ\text{C}$ 冰箱保存备用。

1.4 PCR 产物的纯化、克隆及测序

取 $50\text{ }\mu\text{L}$ 的 PCR 产物,经 1% 的琼脂糖凝胶电泳,切下目的片段,用 Biospin Gel Extraction Kit 回收试剂盒,按说明书提供的方法进行纯化。将纯化产物与 T-easy (Promega) 载体进行连接。取连接液 $10\text{ }\mu\text{L}$ 转化于 *Escherichia coli* TOP 10 感受态细胞 (TIANGEN)。根据蓝白斑原理,通过

PCR法(用M13+/-通用引物)验证白色菌斑,挑取阳性克隆进行扩大培养,菌液送至上海英骏(Invitrogen)生物工程有限公司,用ABI 3730全自动测序仪(用M13+/-引物)进行双向测序。每个体测序6个克隆。

1.5 序列分析与蛋白质结构预测

用Clustal X^[20]软件进行序列比对、排序并辅以人工校正后,将获得的3'-RACE、5'-RACE序列用ContigExpress软件进行拼接,得到团头鲂MHC I类基因的cDNA全长序列。利用NCBI的Blast server对全长序列进行核酸和蛋白质的同源性检索。利用SignalP 3.0 Server在线预测可能的信号肽及信号肽剪切位点。用TMHMM Server v. 2.0工具在线分析预测蛋白的跨膜区域。利用PredictProtein工具分析预测蛋白质的一级和二级结构,并利用ProtParam在线分析蛋白质氨基酸组成、分子量及等电点。蛋白质的三级结构通过Swiss-Model程序同源建模获得,用Spdbv 4.0软件进行观看和分析。用DnaSP 4.0^[21]分析核苷酸序列和氨基酸序列的变异位点。系统发育关系树的构建采用Mega 4.0软件^[22]的邻接法(Neighbor-joining, NJ法),系统树的分支置信度采用自展法(Bootstrap analysis, BP)重复检验1 000次。

1.6 MHC I类基因在不同组织中的RT-PCR分析

分别提取团头鲂(“浦江I号”选育世代F₀样本3尾,体重500 g)的脾脏、肝脏、肠道、头肾、肌肉、鳃、后肾及血液共8个组织的总RNA。参照PrimscripTM Reverse transcriptase (TaKaRa)说明书进行逆转录合成cDNA第一链,作为RT-PCR模板。MHC I类基因的检测引物为RT-f和RT-r,内参β-actin的引物为β-actin-F和β-actin-R(表1)。PCR反应条件:94℃预变性3 min,接着94℃45 s,退火30 s,72℃45 s,30个循环后,72℃延伸10 min,4℃保温。

2 结果

2.1 团头鲂MHC I类基因全长cDNA的克隆与分析

利用简并引物(Meam-α₂f和Meam-α₂r),扩

增得到222 bp的α₂结构域的部分cDNA序列,经Blast分析,与草鱼MHC I类基因的同源性最高(92%),确认为团头鲂MHC I类基因序列。根据此序列设计特异性引物(表1),对团头鲂F₀、F₈世代的共2尾个体进行5'-和3'-RACE-PCR,分别得到约530 bp和1 500 bp的特异性条带。经克隆、测序,发现3'端和5'序列确实存在序列多态性。为此,利用重新设计的特异性引物(Meam-F和Meam-R),对MHC I类基因编码区进行扩增、克隆、测序,把所得的多个序列与前面所获的3'及5'序列分别进行比对、拼接后,最终获得F₀世代1条、F₈世代2条cDNA全长序列(对应克隆号为F₀-02-01、F₈-01-01和F₈-01-02)。

3条序列的全长在2 040~2 079 bp间,编码区长度为1 044 bp、1 035 bp和1 035 bp,分别编码347、344和344个氨基酸;此外还分别含有87~102 bp的5'非阅读区(5'-Untranslated Region, 5'-UTR)和911~946 bp的3'-UTR区。与F₀世代相比,F₈世代的2个序列在编码区的alpha 2、alpha 3结构域分别存在6个(ACT、ACA)和3个(CCA)碱基——即对应T、N、P这3个氨基酸的缺失变异(表2和图2)。F₈世代内的2条序列内,核苷酸、氨基酸的同源性都比较高,分别为89.0%和93.0%;但在世代之间,序列间的核苷酸、氨基酸同源性较低,分别为75.3%~77.0%和71.1%~71.4%。根据KLEIN等^[23]的命名规则,它们分别被命名为Meam-UAA * F0-0201、Meam-UAB * F8-0101和Meam-UAB * F8-0102(序列号:HM151018 - HM151020)。以Meam-UAA * F0-0201为参照进一步分析发现,在编码区1 044个核苷酸中共存在变异位点215个,多态位点比例为20.6%,其中alpha 1、alpha 2、alpha 3区域存在42、123(含6个缺失位点)、45(含3个缺失位点)个变异位点,分别占变异位点总数的19.53%、57.21%、20.93%,而信号肽区、跨膜区和胞质区(Transmembrane and cytoplasmic region, TM/CY)都只有7个变异位点(仅占3.25%),显然alpha 2区是变异发生的主要区域。

TC ACCAGA CAT CTGAAACGT CTA CAGCAA ATA TCT
 CAT TCT CTA GAACGA TTT TTA CAGACAAAACT GTT TCT CAGGCAACAGCAATG CGA TCT GTA ATG 15
 ←5'-UTR M R S V M 5

CTT TTG CTC CTCGGAGCT CAT CTT GCC TAT GCT GGAACACAC TCT CTGAAA TAC TTC TAC ACT GCT 81
 alpha1 []
 L L L L G A H L A Y A . G T H S L K Y F Y T A 27
 GAGTCT GGAGAC ATT GAT TTC CCA GAG TTC ACT GCAGTT GGT CTG GTA GAT GAAGGGCAGTTT ATG 147
 E S G D I D F P E F T A V G L V D E G Q F M 49
 TAC TTC GAC AGCAAC ATA ATG AAA GCT GTT CCA AAGACAGAGTGG ATC AAACAG AAT GTGGGAGAA 213
 Y F D S N I M K A V P K T E W I K Q N V G E 71
 GAT TAC TGG GAC AGAGAGACT CAG ACT GAA TTT GCC CAG CAT CAG TCG TTC AAAACAAC ATC CAG 279
 D Y W D R E T Q T E F A Q H Q S F K N N I Q 93
 ATC GTA AAGGAA CGG TTT AACCAA TCA AAGGGT GTG CAC ACA GTC CAGGAA ATG TAT GGT TGT GAA 345
 alpha2 []
 I V K E R F N Q S K G V H T V Q E M Y G C E 115
 TGG GAT GAT CAG ACT GGAGAAACA AAT GGG TTC ATC CAGGAC GGT TAT GAT GGAGAGGAC TTT CTG 411
 W D D Q T G E T N G F I Q D G Y D G E D F L 137
 TCT CTG GAT TTG AAGGAG ATGAGATGG ATT TGG CCA GTG CAC CAAGGA ATA AAG ACT GTA CAGAAG 477
 S L D L K E M R W I S P V H Q G I K T V Q K 159
 TGGAAC AAT GACAGAGCC TTT ATT GAGAGT GACAAA CAT TAC CTC AAC ACT GAGTGC GTT GAT TGG 543
 W N N D R A F I E S D K H Y L N T E C V D W 181
 CTGAAGAAG TAT CTG CAG TAT GGAAAGAGCAGC GTGAAGAAAACA GTC TCT CCT CAGGTG TCT CTG 609
 alpha3 []
 L K K Y L Q Y G K S S L K K T V S P Q V S L 203
 TTG CAGAAGACT CCC TCT TCT CCA GTA ACGTGT CAT GCT ACAGGA TTT TAC CCCAGTGGAGTAACA 675
 L Q K T P S S P V T C H A T G F Y P S G V T 225
 ATC TCC TGG ATG AAG AAT GGACAA GAC CAT GAT GAG GAT GTG GAT CTT GGAGAA CTT CTT CCC AAT 741
 I S W M K N G Q D H D E D V D L G E L L P N 247
 GAGGACGGAACC TAC CAGAAGACGAGCAGC CTC AAA GTT AAA CCT GAT GAGTGG AAGAACAACAAA 807
 E D G T Y Q K T S S L K V K P D E W K N N K 269
 TAC AGC TGT GTG GTGGAA CAT CAGGGCAAGAGT GTGACAAAG GAT GAG ATC AGGACA ATT GGGGAA 873
 Y S C V V E H Q G K S V T K D E I R T I G E 291
 ~~~~~  
 CCAGGA GAT GTT CCCATTGGCAT CATTGTTGGAGCTCTTGCTGCTGTCGTCCTGTTGATTCTCATT 939  
 TM/CY [ ]  
 P G D V P I G I I V G A L A A V V L L I L I 313  
 GGT GTT GCT GGG TAT GTGGTGT ATAAGAAGAAACAAGGCTT TAAACCTGTCAGTGGCTCCGATGAT 1005  
 G V A G Y V V Y K K K Q G F K P V S G S D D 335  
 GGT TCA AACAGCTCAGCTCAAGCAGTTCACACAAGCATAAAGATGAGACCAGACGTATGGGAGAGAT 1044  
 G S N S S A Q A V P Q A \* 3'-UTR 347  
 GAAAAT TAT GCT TGTGGTGTGGCTAGCTTATAAAAAAATGCACATACACACTCACACTCAGG  
 CACACACACACACATAGACACAGACACAGATTCTCATAAATTTCTGGGATTGATCTGTTTATG  
 GTTCTG AAT TAA TTA CATCTAAGAATGAATGTTT TGT TTC TACACAGTCTGTT CATCTATGG  
 TTA ACT TTT TAA TTT TAACATGTAGAAAAACAACACATGTTTCACAGAACCTGAAATATGTGATA  
 TTT CTC TTT AGT GACAGAAGAAAAATAGAAATTATTAATTGGCAGGATTAATTTGTAATTAAGA  
 AAGATT GAT TTG TTATTTGTGATGTCAGTACTTAACATCAAAAATAAATATAGTAACCTGGGAAAC  
 ATC TCT CGA TCA GTCTT CAT GTCTAT GTTGAATATGAAAAATCAGGGAAAGACCGCTACAACCT  
 GAAAAA TTG TTT TCT ACTGCAACATT TATATAACTTAAATCAACTTTGAAATGGTATAGAT TGA  
 TAG TTT GTGAAA TTGGTCTTTTCCATCAATTAATTAACCAATATGTAAATTTCTATTTTATTCA  
 TTG TAT GGA ATT TTT GGGCTTTTCTCTTGTATATTTCTTCTTTTATTTCTTGTATTTTA  
 ATG ATT CAC ATA CCTGTATACAGAATAATAGTTGATTCTTGTTAAATAATGTCAGATGCTGTACTA  
 TTAAC ATT TGT TTTTACATATATCTGAACAGTACAGTAAAATATAGATATTTTCTTTGAGGTTA  
 CTA AAT CTA AGT TAAT TATGTGAATTTGTTT TTTTGTATATCAGTTGCATTCAAAAAGTTGT  
 AAT GTT TTC ATT CTCTTAAAAATAAATAAACCCTCACAAAAAGGAAAAAAAAAAAAAAAAAAAA

图1 团头鲂 MHC I cDNA 全长序列以及由此推测的氨基酸序列 (Meam-UAA \* F0-0201)

Fig. 1 The full-length cDNA sequence and putative amino acid sequence of  
*Megalobrama amblycephala* (Meam-UAA \* F0-0201)

-----表示信号肽区域; \_表示糖基化位点; Y S C V V E H 示免疫球蛋白和主要相容性复合体蛋白信号; \* 表示终止密码子。

表 2 3 条团头鲂 MHC I cDNA 序列的结构组成

Tab. 2 The composition of three MHC I cDNA sequences in *Megalobrama amblycephala*

| 名称                              | 5'-UTR | 编码区   |         |         |         |        | 总长       | 3'-UTR | 编码的氨基酸数目 |
|---------------------------------|--------|-------|---------|---------|---------|--------|----------|--------|----------|
|                                 |        | SP    | alpha 1 | alpha 2 | alpha 3 | TM/CY  |          |        |          |
| Meam-UAA * F <sub>0</sub> -0201 | 89 bp  | 48 bp | 261 bp  | 279 bp  | 291 bp  | 165 bp | 1 044 bp | 946 bp | 347      |
| Meam-UAB * F <sub>8</sub> -0101 | 87 bp  | 48 bp | 261 bp  | 273 bp  | 288 bp  | 165 bp | 1 035 bp | 918 bp | 344      |
| Meam-UAB * F <sub>8</sub> -0102 | 102 bp | 48 bp | 261 bp  | 273 bp  | 288 bp  | 165 bp | 1 035 bp | 911 bp | 344      |

## 2.2 氨基酸序列分析

团头鲂 MHC I 类分子包含信号肽、3 个结构域、跨膜区及胞质区等结构。以 Meam-UAA \* F<sub>0</sub>-0201 为例(图 1), 它的开放阅读框编码一个含有 347 个氨基酸的前体蛋白, 推测分子量为 38.7 Ku、等电点为 5.42。用 SignalP 在线分析可知, 信号肽与 alpha 1 区的分界点在 16Ala 与 17Gly 位之间, 前 16 个氨基酸为信号肽。酶切后的成熟蛋白含有 331 个氨基酸, 推测分子量为 37.0 Ku、等电点为 5.25。用 TMHMM server 2.0 在线分析可知, 1Met ~ 298Gly 为该蛋白的胞外区, 跨膜区位于 299Ile 与 321Tyr 之间, 322Lys ~ 347Ala 为胞内区。经 ProtParam 预测, 该蛋白在第 86 ~ 88 位上存在着一个 N-糖基化位点(N-X-S/T), 在第 4 位、第 73 位、第 155 位、178 位和 247 位共含有 5 个蛋白激酶 C-磷酸化位点(S/T-X-R/K), 在第 64 位、第 94 位、第 268 位、第 274 位及第 318 位上共含有 5 个酪蛋白激酶 II 磷酸化位点(S/T-X-X-D/E)。但在上述位点中, F<sub>8</sub> 世代的 2 条氨基酸序列有 3 ~ 4 处不存在 N-糖基化位点或蛋白激酶 C-磷酸化位点(图 2)。在 alpha 2 的第 101 位、第 164 位, alpha 3 的第 203 位、第 260 位分别含有一个半胱氨酸残基(C), 形成 2 对二硫键, 这些位点从团头鲂到人都保守(图 2)。另外在第 258 位上存在 1 个组成免疫球蛋白和主要组织相容性复合体蛋白信号区(F/Y-X-C-X-V/A-X-H), 这一区域只在鲑鳟鱼类稍有不同, 而其他脊椎动物都存在(图 2), 表明该位点在抗原递呈并激发特异性免疫反应上起重要作用。

把团头鲂 MHC I 类分子的氨基酸序列与鲤<sup>[6]</sup>、草鱼<sup>[19]</sup>、斑马鱼<sup>[5]</sup>、虹鳟<sup>[3]</sup>、大西洋鲑(*Salmo salar*)<sup>[24]</sup>、红鲑(*Oncorhynchus gorbuscha*)<sup>[25]</sup>、唇鲨(*Triakis scyllium*)<sup>[26]</sup>、非洲爪蟾(*Xenopus laevis*)<sup>[27]</sup>、红原鸡(*Gallus gallus*)<sup>[28]</sup>、鼠<sup>[29]</sup>及人类(HLA-A2)<sup>[30]</sup>的同源序列进行比较分析表明:(1)团头鲂与其他脊椎动物一样, 大部分与抗原

多肽结合的 10 个关键性氨基酸(如 M5、Y7、F22、Y59、Y84、T143、K146、W147、Y159 和 Y171)是保守的, 但也存在变异。如 M5 在团头鲂中变为 L, 斑马鱼中变为 W, 但在鲑鳟类、鼠及人类中都是保守的; F22 仅在鼠中变为 Y, 而 Y84 在团头鲂、草鱼变为 R, K146、W147 在团头鲂中变为 R、L, W147 在草鱼中也变为 L。(2)与  $\beta_2 m$  结合的氨基酸, 如 F8、T10、V25、Q96、Q115、D119、G120、D122、E232、D239、Q242 等, 也是相对比较保守的。(3) 86 ~ 88 位的糖基化位点在所有种属中都存在。(4) 除鲤和大西洋鲑外, 所有种属中均存在第 4 位的蛋白激酶 C 磷酸化位点。(5) 在酪蛋白激酶 II 磷酸化位点上, 团头鲂有 5 个位点(第 64 位、第 94 位、第 268 位、第 274 及第 318 位), 而大西洋鲑、虹鳟也有 1 个位点(第 318 位)。可见, 团头鲂 MHC I 类分子存在与其它许多脊椎动物结构或功能上的相似性, 但亦有明显的种的特征(图 2)。

## 2.3 蛋白质高级结构预测与分析

采用 PredictProtein 在线预测二级结构表明, 团头鲂 MHC I 类分子的组成与人类的 HLA 分子的结构十分类似<sup>[31]</sup>, 主要由  $\alpha$  螺旋、 $\beta$  折叠组成, 兼有少量无规则卷曲。其中  $\alpha$  螺旋(H)占 24.21%,  $\beta$  片层(E)占 32.56%, 无规则卷曲及  $\beta$  转角占 43.23%, 属于混合型蛋白。

从 MHC I 类分子成熟蛋白三级结构预测结果(图 3)看出, 靠近细胞膜的 alpha 3 区由 4 条  $\beta$  片层组成, 位于分子的底部。alpha 1 区和 alpha 2 区位于顶部, 共同组成一个抗原肽结合槽, 槽的两端封闭, 一般只能容纳由 8 ~ 10 个氨基酸残基组成的小肽。alpha 1 区和 alpha 2 区均是由 4 条反向的  $\beta$  片层及  $\alpha$  螺旋组成(图 3A 和 B), 略有不同的是 alpha 1 区存在一长一短两个  $\alpha$  螺旋、中间由无规则卷曲连接, 而 alpha 2 区则只有一个长螺旋。

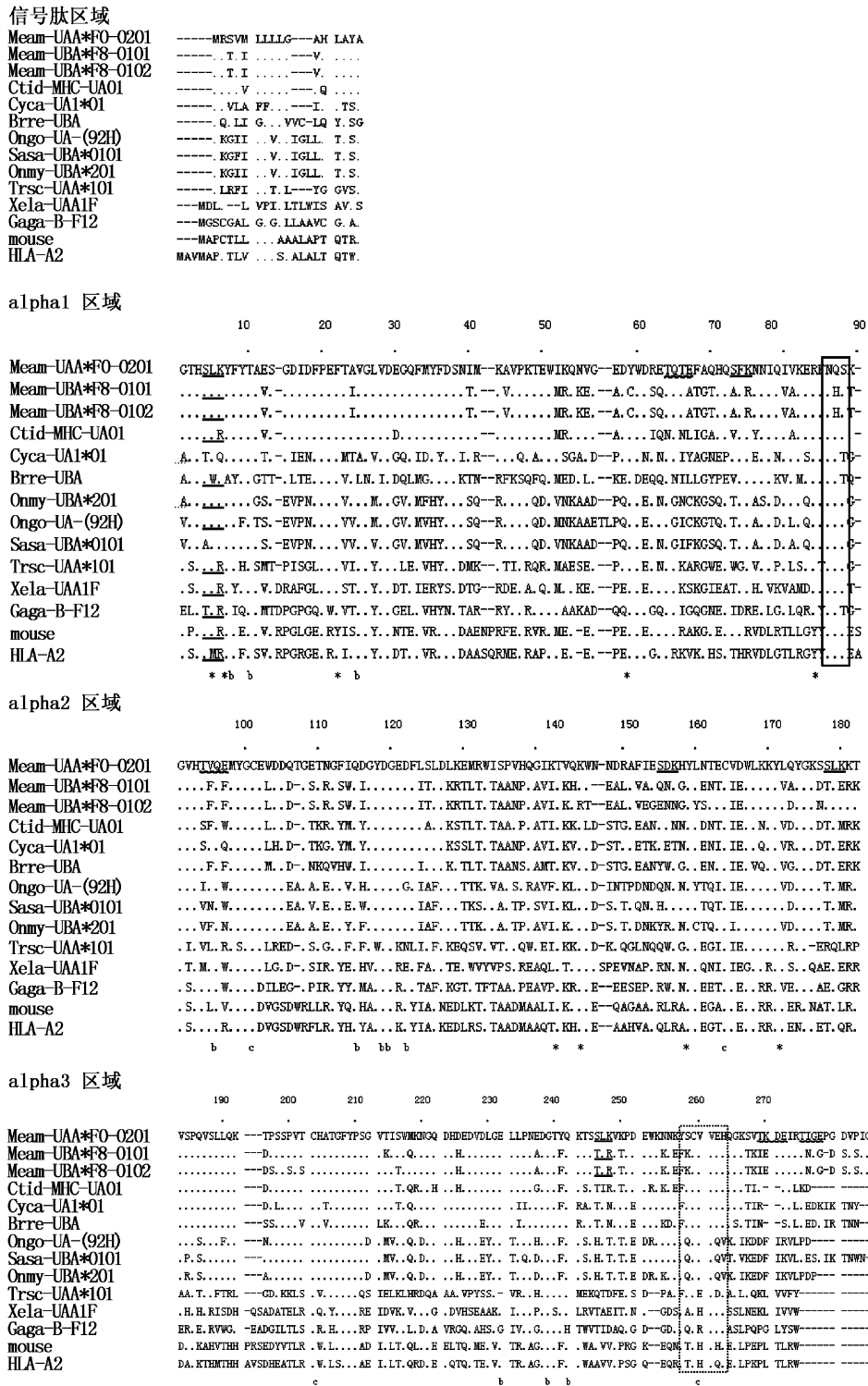


图 2 团头鲂 MHC I 分子氨基酸序列与其他脊椎动物的比较

Fig. 2 The comparison of amino acid sequences of MHC class I molecules in *Megalobrama amblycephala* and other vertebrates

\* 表示人 (HLA-A2) 与抗原多肽结合的关键性氨基酸; b 表示人 (HLA-A2) 与  $\beta_2m$  结合的氨基酸; . 表示与 Meam-UAA \* F0-0201 一致的氨基酸序列; c 表示半胱氨酸位点; - 表示缺失; 实线方框表示糖基化位点。虚线方框表示主要组织相容性复合体信号; 短横线表示蛋白激酶 C 磷酸化位点; 波浪线表示酪蛋白激酶 II 磷酸化位点。比中所用到的各同源序列的序列号分别为鲤 (Cyca, X91015); 草鱼 (Ctid, AB109779); 斑马鱼 (Brre, BC074095); 红鲢 (Ongo, D58386); 大西洋鲑 (Sasa, AF504019); 虹鳟 (Onmy, AF287485); 唇鲨 (Trsc, AF034316); 非洲爪蛙 (Xela, L20733); 红原鸡 (Gaga, AY234768); 鼠 (Mouse, M58156); 人类 (HLA-A2, K02883)。

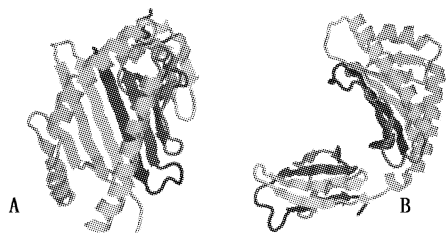


图3 团头鲂 MHC I 类分子成熟蛋白三级结构预测

Fig. 3 Tertiary structure prediction for MHC I maturation protein of *Megalobrama amblycephala*

A 为成熟蛋白三级结构顶面观; B 为成熟蛋白三级结构侧面观; 绿色区、蓝色区、黄色区分别代表 alpha 1 区、alpha 2 区、alpha 3 区域。

比较团头鲂 MHC I 类分子成熟蛋白与人类 HLA-A2 晶体结构(图 4), 二者的结构差异主要分布在如下 5 个区域: I. Ser13 ~ Pro20, II. Ala40 ~ Arg44, III. Glu89 ~ Ser92, IV. Gly104 ~ Phe109, V. Glu128 ~ Ser132。参照 BJORKMAN<sup>[31]</sup> 等对 HLA-A2 的二级结构的预测结果, 发现这些变异的区域都存在于  $\alpha$  螺旋内侧或连接  $\alpha$  螺旋与  $\beta$

折叠的无规则线团中, 而  $\alpha$  螺旋的外侧及  $\beta$  折叠区域都是十分保守的。

## 2.4 系统发生分析

比较不同种属间的 MHC I 类分子的氨基酸序列同源性, 发现团头鲂与草鱼的相似性最高(63.4%), 与鲤的相似性较高(60.1%), 而与斑马鱼、大西洋鲑、红鲑、虹鳟及唇鲨相似性则较小(分别为 51.2%, 51.1%、45.5%、49.7% 和 34.7%), 与非洲爪蛙、红原鸡、老鼠和人类的相似性最小(分别为 32.5%、27.6%、30.0% 和 32.6%)。在 NJ 法构建的分子进化树中, 团头鲂  $F_8$  世代的 2 个 MHC 分子首先聚类, 再与  $F_0$  世代的一个分子相连, 然后与草鱼、鲤、斑马鱼聚类为分支 A; 鲑鳟鱼类、其他脊椎动物和人类则分别聚类为分支 B、C, 而唇鲨(Trsc-UAA \* 101) 却位于硬骨鱼类与其他脊椎动物之间。可见, 团头鲂与鲤科鱼类中的草鱼关系最近, 与斑马鱼、鲤的关系稍远, 而与海水的鲑鳟鱼类关系较远, 与蛙、鸡、鼠等其他脊椎动物和人类的关系则渐远(图 5)。

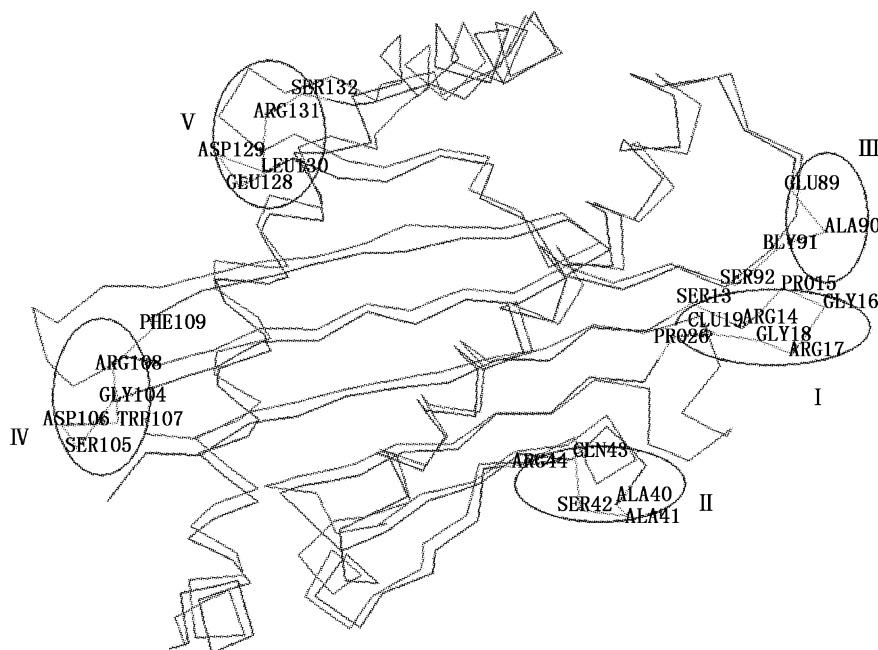


图4 团头鲂 MHC I 类分子与人类 HLA-A2 抗原结合区(PBR)三级结构比较

Fig. 4 Tertiary structure comparison of PBR between *Megalobrama amblycephala* MHC I molecule and human HLA-A2

蓝色线条代表人类 HLA-A2 的抗原结合区(PBR)结合区; 红色线条代表团头鲂 MHC I 类分子 PBR 区; 绿色线条代表人类 HLA-A2 与团头鲂 MHC I 类分子的差异区域。

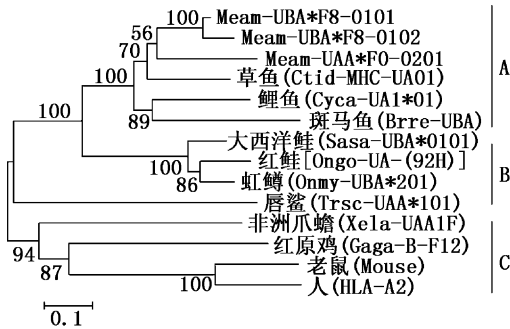


图 5 由 NJ 法构建的团头鲂 MHC I 类基因与其他物种的系统发生树

Fig. 5 Phylogenetic tree of the MHC Class I gene based on amino acid sequences of *Megalobrama amblycephala* and other vertebrates by Neighbour-joining (NJ) method

## 2.5 MHC I 类基因在不同组织中的表达分析

用 RT-PCR 半定量检测方法分析团头鲂 MHC I 类基因在不同组织中的表达情况表明, MHC I 类基因在不同组织中均扩增出约 115 bp 的片段,其中较强表达于鳃、头肾和血液中,中等强度表达于脾脏、肝脏,而在后肾、肠和肌肉中表达较弱(图 6)。

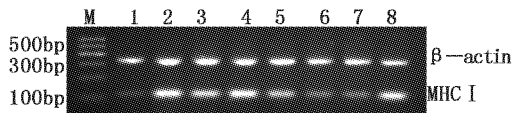


图 6 团头鲂 MHC I 类基因组织表达特异性的 RT-PCR 分析

Fig. 6 RT-PCR analysis of MHC class I gene expression in various tissues from *Megalobrama amblycephala*

M. Marker I; 1. 肌肉; 2. 鳃; 3. 脾脏; 4. 头肾; 5. 肝脏; 6. 后肾; 7. 肠; 8. 血液。

## 3 讨论

### 3.1 团头鲂 MHC I 类基因的序列特征及表达多态性

高度多态性及广泛的组织表达性是哺乳动物经典 MHC I 类分子的主要特征。许多研究表明,鱼类的 MHC 分子也具有相似的特征<sup>[2-4,10,14]</sup>。本研究采用 RACE 技术获得的团头鲂“浦江 I 号”的基础群体( $F_0$ )、选育世代( $F_8$ )共 2 尾个体的 3 条 cDNA 全长序列中,尽管  $F_8$  世代内的 2 条序列的核苷酸/氨基酸的同源性都较高

(0.89/0.93),但在世代之间,核苷酸/氨基酸的同源性却较低(0.75/0.71 左右),显示出明显的分子多态性。编码区变异最大的部分集中在抗原肽结合(PBR)区( $\alpha 1$  和  $\alpha 2$  区占总变异位点的 76.74% 左右),表明在平衡性选择机制(balancing selection)下, MHC 分子的这两个结构域受到的选择压力较大。有趣的是,团头鲂的  $\alpha 3$  区的多态位点(20.93%)也比较高,类似的现象在虹鳟中也有发现<sup>[3]</sup>。本研究中,我们发现选育  $F_8$  世代的 2 条序列之间的变异较大,这是否由于长期的人工选择压力对 MHC 分子造成的影响,有待在以后的群体遗传变异的比较分析中予以证实。

分析表明,团头鲂 MHC I 类分子,一方面具有与其它许多脊椎动物结构或功能上的相似性,另一方面亦有明显的种的特征(图 2)。团头鲂 MHC I 类分子的 PBR 与人类 HLA-A2 分子的晶体结构特征是基本一致<sup>[31]</sup>,草鱼的 MHC I 类分子也具有类似的结构<sup>[19]</sup>。可见,不同种属间的 MHC I 类分子,虽然在氨基酸序列上差异很大,但在进化过程中保留了与抗原多肽结合的关键性氨基酸以及功能性的立体结构。

研究表明,硬骨鱼类的 MHC I 类分子具有广泛的组织表达性,只是不同组织的表达强弱不同。ANTA0<sup>[32]</sup> 等发现斑点叉尾鲷 (*Lctalurus punctatus*) 的 MHC I 基因在鳃、头肾和血液表达最强,在心脏、肝脏、皮肤、脾脏中等强度表达,肠道中表达较弱,而脑、卵和肌肉中几乎检测不到表达。YANG 等<sup>[10]</sup> 发现,草鱼的 MHC I 基因在肾、鳃、肠、心、脾、肝和脑等组织中呈现出不同的表达水平,其中肾的表达量最高。我们在团头鲂的 8 个组织中均可检测到 MHC I 基因的表达产物,其中在鳃、头肾和血液中表达程度最强,中等强度表达于脾脏和肝脏,而在后肾、肠和肌肉中表达较弱,这与 ANTA0 等<sup>[32]</sup> 的研究结果相似。显然,不同鱼类的 MHC I 基因在免疫相关器官如头肾、脾等组织表达较强,而在其他组织表达较弱或检测不到表达量,呈现出明显的表达程度上的组织特异性。

### 3.2 团头鲂 MHC I 基因的系统进化关系

在 NJ 法构建的系统发生树中(图 5),团头鲂 MHC I 类分子与所有硬骨鱼类聚为一大支,这当中,所有鲤科鱼类聚为分支 A,而鲑鳟类聚为分支



B且位于基部,表明二者具有单系性。值得注意的是,软骨鱼(唇鲨, Trsc-UAA \* 01)在系统树中位于硬骨鱼类与其他脊椎动物(分支C)之间,单独为一小支,说明它与其他高等脊椎动物的亲缘关系较硬骨鱼类更近,显示MHC I类基因的分化是晚于软骨鱼类与硬骨鱼类及其他高等脊椎动物分化的。但张玉喜等<sup>[9]</sup>、丁少雄等<sup>[11]</sup>等在构建大菱鲂MHC II A分子、赤点石斑鱼MHC II B分子的系统发生树时,发现软骨鱼类不是与硬骨鱼类聚类却先与其他高等脊椎动物聚在一起,故认为MHC II类基因的分化是早于软骨鱼类与硬骨鱼类及其他高等脊椎动物的分化的<sup>[11]</sup>。我们的研究结果与此不一致。显然,在鱼类中,MHC I类、MHC II类基因由于分属于不同的连锁群而存在不同步进化的现象。

在研究过程中,蔡完其教授给予了很多帮助,唐首杰博士也多次协助采样,在此表示由衷的感谢。

#### 参考文献:

- [1] EDWARDS S V, HEDRICK P W. Evolution and ecology of MHC molecules: from genomics to sexual selection [J]. Trends Ecol Evol, 1998, 13(8): 305-311.
- [2] KLEIN J, FIGUEROA F, KLEIN D, et al. Major histocompatibility complex genes in the study of fish phylogeny [M]. San Diego: Academic Press, 1997: 271-281.
- [3] AOYAGI K, DIJKSTRA J M, XIA C, et al. Classical MHC class I genes composed of highly divergent sequence lineages share a single locus in rainbow trout (*Oncorhynchus mykiss*) [J]. J Immunol, 2002, 168(1): 260-273.
- [4] XU S X, REN W H, LI S Z, et al. Sequence Polymorphism and Evolution of Three Cetacean MHC Genes [J]. J Mol Evol, 2009, 69(3): 260-275.
- [5] STRAUSBERG R L, FEINGOLD E A, GROUSE L H, et al. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences [J]. P Natl Acad Sci USA, 2002, 99(26): 16899-16903.
- [6] van ERP S H M, DIXON B, FIGUEROA F, et al. Identification and characterization of a new major histocompatibility complex class I gene in carp (*Cyprinus carpio* L.) [J]. Immunogenetics, 1996, 44(1): 49-61.
- [7] HARDEE J J, GODWIN U, BENEDETTO R, et al. Major histocompatibility complex class II A gene polymorphism in the striped bass [J]. Immunogenetics, 1995, 41(4): 229-238.
- [8] 张玉喜,陈松林. 牙鲆MHC class II B基因多态性及其与鱼体抗病力关系的分析[J]. 水产学报, 2006, 30(5): 633-639.
- [9] 张玉喜,陈松林. 大菱鲂MHC II B基因全长cDNA的克隆与组织表达分析[J]. 高技术通讯, 2006, 16(8): 859-863.
- [10] YANG T Y, HAO H F, JIA Z H, et al. Characterisation of grass carp (*Ctenopharyngodon idellus*) MHC class I domain lineages [J]. Fish Shellfish Immun, 2006, 21(5): 583-591.
- [11] 丁少雄,张之文,杜佳莹,等. 赤点石斑鱼(*Epinephelus akaara*) MHC IIB基因的克隆与表达多态性分析[J]. 海洋学报, 2009, 31(2): 129-138.
- [12] XU T J, CHEN S L, JI X, et al. Molecular cloning, genomic structure, polymorphism and expression analysis of major histocompatibility complex class IIA and IIB genes of half-smooth tongue sole (*Cynoglossus semilaevis*) [J]. Fish Shellfish Immun, 2009, 27(2): 192-201.
- [13] KIM T J, PARKER K M, HEDRICK P W. Major histocompatibility complex differentiation in Sacramento River chinook salmon [J]. Genetics, 1999, 151(3): 1115-1122.
- [14] XIA C, KIRYU I, DIJKSTRA J M, et al. Differences in MHC class I genes between strains of rainbow trout (*Oncorhynchus mykiss*) [J]. Fish Shellfish Immun, 2002, 12(4): 287-301.
- [15] MILLER K M, KAUKINEN K H, BEACHAM T D, et al. Geographic heterogeneity in natural selection on an MHC locus in sockeye salmon [J]. Genetica, 2001, 111(1): 237-257.
- [16] 李思发. 中国淡水主要养殖鱼类种质研究 [M]. 上海: 上海科学技术出版社, 1998.
- [17] 李思发. 鱼类良种介绍团头鲂浦江1号 [J]. 中国水产, 2001(11): 52-52.
- [18] 赵岩,李思发,唐首杰. 团头鲂“浦江1号”选育后期世代群体同野生群体间遗传变异的ISSR分析 [J]. 水产学报, 2009, 33(6): 893-900.
- [19] XIA C, XU G Y, LIN C Y, et al. Molecular cloning and polymorphism of major histocompatibility complex class I genes from grass carp (*Ctenopharyngodon idellus*) [J]. Prog Nat Sci, 2004, 14(5): 403-410.
- [20] THORMPSON J D, GIBSON T J, PLEWNIAC F, et al. The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools [J]. Nucleic Acids Res, 1997, 25(24): 4876-4882.
- [21] ROZAS J, SANCHEZ-DELBARRIO J C, MESSEGUER X, et al. DnaSP, DNA polymorphism analyses by the coalescent and other methods [J]. Bioinformatics, 2003, 19(18): 2496-2497.
- [22] TAMURA K, DUDLEY J, NEI M, et al. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0 [J]. Mol Biol Evol, 2007, 24(8): 1596-1599.
- [23] KLEIN J, BONTROP R E, DAWKINS R L, et al. Nomenclature for the major histocompatibility complexes of different species: a proposal [J]. Immunogenetics, 1990, 31(4): 217-219.
- [24] GRIMHOLT U, HORDVIK I, FOSSE V M, et al. Molecular cloning of major histocompatibility complex class I cDNAs from Atlantic salmon (*Salmo salar*) [J]. Immunogenetics,

- 1993,37(6):469-473.
- [25] KATAGIRI T, HIRONO I, AOKI T, *et al.* Isolation of major histocompatibility complex class I cDNA from pink salmon (*Oncorhynchus gorbusha*) [J]. *Dev Comp Immunol*, 1996, 20(3):217-228.
- [26] OKAMURA K, OTOTAKE M, NAKANISHI T, *et al.* The most primitive vertebrates with jaws possess highly polymorphic MHC class I genes comparable to those of humans [J]. *Immunity*, 1997, 7(6):777-790.
- [27] SHUM B P, AVILA D, DU PASQUIER L, *et al.* Isolation of a classical MHC class I cDNA from an amphibian. Evidence for only one class I locus in the *Xenopus* MHC [J]. *J Immunol*, 1993, 151(10):5376-5386.
- [28] HUNT H D, FULTON J E. Analysis of polymorphisms in the major expressed class I locus (BF-IV) of the chicken [J]. *Immunogenetics*, 1998, 47(6):456-467.
- [29] HORTON R M, HILDEBRAND W H, MARTINKO J M, *et al.* Structural analysis of H-2Kf and H-2Kfm1 by using H-2K locus-specific sequences [J]. *J Immunol*, 1990, 145(6):1782-1787.
- [30] KOLLER B H, ORR H T. Cloning and complete sequence of an HLA-A2 gene: analysis of two HLA-A alleles at the nucleotide level [J]. *J Immunol*, 1985, 134(4):2727-2733.
- [31] BJORKMAN P J, SAPER M A, SAMRAOUI B, *et al.* Structure of the human class I histocompatibility antigen, HLA-A2 [J]. *Nature*, 1987, 329(6139):506-512.
- [32] ANTAO A, WILSON M, WANG J, *et al.* Genomic organization and differential expression of channel catfish MHC class I genes [J]. *Dev Comp Immunol*, 2001, 25(7):579-595.

## Full length cDNA cloning and tissue expression of major histocompatibility complex (MHC) class I from blunt snout bream (*Megalobrama amblycephala*)

MA Xiao-qian<sup>1</sup>, LIU Zhi-zhi<sup>1</sup>, LI Si-fa<sup>2</sup>, TANG Wen-qiao<sup>1</sup>, YANG Jin-quan<sup>1</sup>

(1. Key Laboratory of Exploration and Utilization of Aquatic Genetic Resources, Ministry of Education, Shanghai Ocean University, Shanghai 201306, China; 2. Key Laboratory of Aquatic Genetic Resources and Utilization, Ministry of Agriculture, Shanghai Ocean University, Shanghai 201306, China)

**Abstract:** RACE (Rapid Amplification of cDNA Ends) technique was used to clone "Pujiang No. 1" blunt snout bream (*Megalobrama amblycephala*) MHC class I gene (Meam-MHC I), and total three full-length cDNA sequences in the F<sub>0</sub> and F<sub>8</sub> generations were obtained by cloning and sequencing method, respectively. The sequences were 2 040 - 2 079 bp in length, including an 87 - 102 bp 5'-UTR, a 1 035 - 1 044 bp coding region, containing signal region, three ( $\alpha$ 1,  $\alpha$ 2,  $\alpha$ 3) domains, a transmembrane and cytoplasmic region (TM/CY), and 911 - 946 bp 3'-UTR. The coding region had 344 - 347 amino acids. Within F<sub>8</sub> generation, the nucleotide/amino acid sequences similarities were as high as 89.0%/93.0%. But the values were relatively as low as 75% - 77% /71% between generations, indicating obviously sequence polymorphism. The Meam-MHC I exhibited a classical 3D molecular structure, and compared to the PBR region of HLA-A2, the major differences were mainly in the five regions (I - V). It was revealed in the neighbor-joining tree that the blunt snout bream had the closest relationship with grass carp, far from salmon and trout, and still farther from, chicken and human. Semi-quantitative RT-PCR analysis demonstrated that Meam-MHC I transcript was detected at high level in gill, head kidney and blood, moderate in liver and spleen, and low in trunk kidney, gut and muscle.

**Key words:** *Megalobrama amblycephala*; major histocompatibility complex (MHC); cDNA; clone; tissue expression