

不同倍性鲫鱼高迁移率族蛋白 1 基因的克隆和原核表达*

刘东** 刘臻** 刘少军*** 刘良国 尤翠平 陈琳
钟欢 刘筠

湖南师范大学生命科学学院, 教育部蛋白质化学及鱼类发育生物学重点实验室, 长沙 410081

摘要 用 RACE PCR 技术, 分别获得三倍体鲫鱼、四倍体鲫鱼、红鲫和团头鲂的高迁移率族蛋白 1 基因(*HMG1*) mRNA 全长序列, 其开放阅读框包含 579 nt, 翻译成 193 个氨基酸. 不同倍性鲫鱼 *HMG1* 的 cDNA 和氨基酸序列比较分析表明: 在 cDNA 水平上, 四倍体鲫鱼与母本红鲫的同源性(99%)高于同父本团头鲂的同源性(97%); 三倍体鲫鱼与父母本同源性(95%)低于亲本之间的同源性(98%); 在氨基酸水平上, 四倍体鲫鱼与父母本的同源性(100%)高于三倍体鲫鱼与亲本的同源性(97%). 结果表明: 远源物种间的杂交对两性不育的三倍体鲫鱼 *HMG1* 基因造成了一定的冲击, 分子遗传效应表现为三倍体鲫鱼 *HMG1* 基因位点发生了变异; 四倍体鲫鱼 *HMG1* 氨基酸序列与亲本的完全一致性, 克服了等位基因的杂交不亲和性, 为两性可育的异源四倍体鲫鱼遗传稳定奠定了基础. 生物信息学分析表明: *HMG1* 蛋白二级结构具有 8 个螺旋和 3 个转角, *HMG1* 蛋白三级结构于 N-端具有两个 DNA 结合基序, C-端具有一长为 23 个重复 D 或 E 的氨基酸尾. 这种结构决定了 *HMG1* 能够与核 DNA 发生“蛋白-DNA”的相互作用, 参与多种核内生物学功能的完成. 另外, 以 *HMG1* 氨基酸序列构建了鱼类、两栖类、爬行类、鸟类和哺乳类的进化树, 结果提示 *HMG1* 是一种古老的蛋白质, 并在物种演化中具有保守性; 首次构建了鱼类 *HMG1* 原核表达载体, 外源的鱼类 *HMG1* 基因编码蛋白在原核细胞中得到了表达, 这为下一步 *HMG1* 蛋白的制备和生物学功能、尤其与 DNA 转座子相互作用的研究提供了条件, 将助于了解物种间的杂交形成异源多倍体的机制.

关键词 鱼类 多倍体 高迁移率族蛋白基因 系统进化 原核表达

物种间的远源杂交将对杂交种的基因组产生强烈的冲击, 长期以来, 这种冲击被认为是物种分化的机制^[1]. 冲击结果在基因组水平上表现为杂交种的染色体的倍性变化, 例如虹鳟(*Oncorhynchus mykiss*, $2n=60$)(♀)和红点鲑(*Salvelinus fontinalis*, $2n=84$)(♂)杂交, 后代为三倍体的杂交种($3n=102$)^[2].

我们曾以团头鲂(*Megalobrama amblycephala*, $2n=48$)(♂)(鲤形目 Cypriniformes、鲤科 Cyprinidae、鲮亚科 Abramidinae^[3])和红鲫(*Carassius auratus* red var., $2n=100$)(♀)(鲤科、鲤亚科 Cyprininae^[3])杂交获得了两种染色体类型的杂交种, 即性腺可育的四倍体鲫鱼($4n=148$)和性腺不育的三

2009-03-04 收稿, 2009-04-02 收修稿稿

* 国家杰出青年科学基金(批准号: 30725028)、国家重点基础研究发展计划(批准号: 2007CB109200)和湖南省“芙蓉学者”特聘教授基金资助项目

** 同等贡献作者

*** 通信作者, E-mail: lsj@hunnu.edu.cn

倍体鲫鲂($3n=124$)^[4]. 杂交种的染色体呈整套的增加: 四倍体鲫鲂的染色体中两套来自红鲫, 另两套来源于团头鲂; 三倍体鲫鲂的染色体中两套来自红鲫, 另一套来源于团头鲂. 鲫鲂远源杂交对基因组冲击的分子遗传效应研究表明, 基因组内“沉默”的转座子发生了活化, 产生新的转座子(资料未发表). 转座子为染色体内的独立的 DNA 序列, 是生物基因组的结构特征之一, 在基因组内能够从一个位点跳跃到另一个位点^[5]. 但对真核生物体基因组内的转座子研究表明, 寄宿于现物种的转座子表现为转座失活^[6]. 远源杂交能够激活转座子的活性, 其分子机制之一认为是与一种核蛋白——高迁移率族蛋白(HMG)有关^[7]. HMG 与转座子之间的“蛋白-DNA”的相互作用, 促使转座子 DNA 序列弯曲, 有利于转座酶的剪切^[8]. 此外, HMG 蛋白家族广泛参与多种重要的核内生物学功能的完成, 包括调节 DNA 复制、转录、重组和修复等^[9].

HMG 蛋白由英国科学家 Goodwin 于 1973 年在牛胸腺细胞中首次发现, 因其在聚丙烯酰胺凝胶电泳中的高迁移率得名^[10]. 广义的 HMG 包括了两大类蛋白: HMG-基序蛋白和经典的 HMG 蛋白. 前者蛋白序列中含有一个或数个 HMG 的 DNA 结合结构域, 如 SRY (sex determining region Y protein). 经典的 HMG 蛋白分为 3 类: HMGB (原 HMG-1/-2), HMGA 和 HMGN^[9]. 最近, NCBI 数据库(www.ncbi.nlm.nih.gov)将 HMGB 按最初的 HMG1 和 HMG2 分类. 虽然 HMG 蛋白具有高度的保守性, 因其具有特异性的功能基序, 所以能与组织特异性的调控因子、转录或复制机制成分发生相互作用^[11]. 目前 *HMGI* 基因的相关研究, 主要是二倍体脊椎动物基因组内 *HMGI* 的克隆和表达方面, 如鳖^[12] 和人^[13] 等, 而远缘物种间的杂交是否对 *HMGI* 基因造成冲击尚未见有报道. 本实验利用 3'RACE 和 5'RACE 结合 PCR 扩增、克隆、测序等技术, 克隆出不同倍性鱼 *HMGI* 基因 mRNA 全序列, 基于 *HMGI* 的 cDNA 序列和氨基酸序列, 分析了远缘物种间的杂交对 *HMGI* 基因的冲击效应, 并探讨了 *HMGI* 的进化及其与转座子的关系, 将有助于了解物种间的杂交形成异源多倍体的机制. 本文还构建了 *HMGI* 原核表达系统.

为进一步研究 HMG1 的生物学功能, 尤其与 DNA 的相互作用奠定了基础.

1 材料与方法

1.1 材料和总 RNA 提取

三倍体鲫鲂、四倍体鲫鲂、团头鲂和红鲫均取自湖南师范大学教育部多倍体鱼繁殖及育种技术工程研究中心, 其中三倍体和四倍体鲫鲂是通过红鲫 ♀ × 团头鲂 ♂ 获得. 1 龄实验鱼的 0.5 g 肝脏用液氮充分研磨后, 按总 Total RNA Kit II (OMEGA) 试剂盒说明书操作提取总 RNA, 然后用 DNase I 酶消化去除总 RNA 中残留的 DNA.

1.2 引物设计与 RT-PCR 扩增

根据已知斑马鱼 *Danio rerio* (BC165307), 鲶鱼 *Ictalurus punctatus* (DQ399438) 和虹鳟 *Oncorhynchus mykiss* (X02666) 等鱼的 *HMGI* 保守 cDNA 序列和氨基酸序列, 利用 Primer premier 5.0 和 Oligo 6 软件设计一对兼并引物, 引物 S 和 A 序列见表 1. 以 RT-PCR 试剂盒 (TIANGEN) 的 Oligo (dT)₁₂₋₁₈ 为引物, 逆转录酶 M-MLV 以总 RNA 为模板, 在 42℃ 经 1.5 h 合成 cDNA, 然后进行 PCR 扩增. 20 μL PCR 反应体系为: 10 × PCR buffer 2.5 μL, MgCl₂ (25 μmol/L) 2.0 μL, 引物 S 和 A (10 μmol/L) 各 0.5 μL, dNTPs (10 μmol/L) 0.5 μL, Taq DNA 聚合物 (1 U) 0.5 μL 和 cDNA 模板 (100 μmol/L) 1 μL. PCR 扩增程序为: 94℃ 预变性 5 min, 然后 94℃ 30 s, 58℃ 30 s, 72℃ 1 min, 共 30 个循环, 最后在 72℃ 延伸 5 min. PCR 扩增产物用胶纯化试剂盒 (上海生工) 回收后, 目的片段插入 TA 载体, 大肠杆菌 DH5α 亚克隆, 用含 X-gal 和 IPTC 的 LB 平板进行蓝白斑筛选, 每个片段选择 3 个阳性克隆送往上海生工测序.

1.3 RACE PCR 扩增、序列拼接和同源性分析

兼并引物 S/A 扩增的已知序列为模板, 利用引物设计软件获得 3' 和 5' 基因特异引物 (GSP) (表 1). 参考 SMART™ RACE cDNA 试剂盒 (CLONTECH) 说明书, 首先合成 3RACE 和 5RACE cDNA, 然后采用巢式 PCR 扩增, 产物经大肠杆菌 DH5α 亚克隆和测序后, 便获得了 3' 和 5' 末端

序列, 与兼并引物 S/A 扩增的序列一起经计算机分析和拼接, 最终得到 *HMG1* 基因 cDNA 全长. 用 Blast 软件于 NCBI 数据库 (www.ncbi.nlm.nih.gov) 中进行序列的同源性搜索, 利用 ClustalW 软件 (www.ebi.ac.uk/clustalw/) 在线进行序列同源性比较. 克隆基因命名后递交到 GenBank 数据库, 得到基因登录号.

表 1 不同倍性鲫鲂 *HMG1* 基因 PCR 扩增引物名称和序列

引物名称	引物序列(5'→3')
兼并	S GAGGA(G/A)CA(C/T)AAGAAGAAACACC
	A CTC(G/A)TA(C/T)TTCTCCTTCAGC(T/C)TGLong: CTAATACGACTCACTATAGGGCAAGCA
UPM	GTGGTATCAAGCAGAGT
	Short: CTAATACGACTCACTATAGGGC
NUP	AAGCAGTGGTA TCAACGCAGAGT
RACE	3'GSP1 CAAGACAAGGTCCGTTACGAGAGG
	3'GSP2 TCCGACCCAAGTGAAGGAAGAG
	5'GSP1 CGTAAACGGACCTTG TCTTGTTTGGC
	5'GSP2 CTTGGCTGACA TAGTCTTCCATCGC

1.4 生物信息学和系统进化分析

红鲫 *HMG1* 基因推测的蛋白氨基酸序列, 用软件 Predictprotein (www.predictprotein.org) 在线分析蛋白二级结构, 预测其“蛋白-DNA”相互作用的结构单位; 将 *HMG1* 氨基酸序列输入 NCBI 数据库中, 进行蛋白三级结构检索 (www.ncbi.nlm.nih.gov/structure/cdd/wrpsb.cgi), 预测其是否具有保守的同源框. 从 GenBank 数据库中下载不同物种的 *HMG* 蛋白氨基酸序列, 用于系统进化分析. 具 *HMG1* 的物种分别是斑马鱼 (AAH45917)、鲑鱼 *Salmo salar* (ACM09402)、虹鳟 (AAA58771)、爪蟾 *Xenopus laevis* (NP-001080836)、猪 *Sus scrofa* (AAW29969)、短尾负鼠 *Monodelphis domestica* (XP-001363686)、小家鼠 *Mus musculus* (CAA56631)、牛 *Bos Taurus* (NP-788785)、马 *Equus caballus* (XP-001488130)、褐家鼠 *Rattus norvegicus* (XP-573095)、珍珠鸟 *Taeniopygia guttata* (XP-002194652)、红原鸡 *Gallus gallus* (NP-990233)、黑猩猩 *pan troglodytes* (XP-509611)、人 *Homo sapiens* (AAP36330) 和中华鳖 *Pelodiscus sinensis*

(AY673604); 酵母 *Saccharomyces cerevisiae* *HMG1* (CAA86503) 作为外群. 真鲷 *Pagrus major* (AAP20177) *HMG* 用以确定 *HMG1* 的进化时序. 用 ClustalX 软件^[14] 对各物种 *HMG* 氨基酸序列排序后, 输入 MEGA 4.0 软件中, 基于邻接法 (NJ) 模型构建系统进化树, 置信度 (Bootstrap) 经 1000 次重复抽样检验^[15].

1.5 重组 *HMG1* 原核表达载体的构建

根据已知的红鲫 *HMG1* 基因 cDNA 全长序列 (GenBank 登录号: FJ785327), 利用软件设计一对引物, 上游 S: 5' CCGGAATTTCATGTCCTCAT-ACGCATACTTT3', 下游 A: 5' CCGCTCGAGT-GTTGTGGTGGTTTATTTCTC3'. 其中 S 5' 端引入 *EcoRI* 酶切位点 (G[↓]AATTC) 和 3 个保护碱基 (CCG), A 5' 端引入 *XhoI* 酶切位点 (C[↓]TCGAG) 和 3 个保护碱基 (CCG). 按“1.2 引物设计与 RT-PCR 扩增”获得红鲫 cDNA. PCR 扩增条件为: 94 °C 5 min, 然后 94 °C 30 s, 61 °C 30 s, 72 °C 1 min, 共 30 个循环, 最后 72 °C 延伸 5 min. PCR 扩增产物经 1.2% 琼脂糖凝胶电泳, 回收 *HMG1* 目的片段, 然后将 *HMG1* 目的片段插入载体 pMD18-T, 重组载体命名为 pMD18-*HMG1*, 将重组载体 pMD18-*HMG1* 转化 DH5 α 感受态细胞, 蓝白斑挑选阳性克隆测序, 序列分析后, 挑取具载体 pMD18-*HMG1* 的菌液在 LB 培养液 (含 0.1 g · L⁻¹ 氨苄青霉素) 中扩大培养 (1:500), 提取质粒, 用 *EcoRI* 和 *XhoI* 对质粒 pMD18-*HMG1* 双酶切, 回收 *HMG1* 目的片段. 将表达载体质粒 pGEX-4T 同样用 *EcoRI* 和 *XhoI* 双酶切, 回收表达载体质粒 pGEX-4T 的线性片段. 在 T4 DNA 连接酶的作用下, 将 *HMG1* 片段和表达载体 pGEX-4T 的线性片段进行连接反应, 即构建了原核表达的重组载体质粒, 命名为 pGEX-*HMG1*.

1.6 原核表达和 SDS-PAGE 检测

将重组表达载体质粒 pGEX-*HMG1* 转化 LB21 感受态细胞, 经氨苄青霉素筛选, 阳性克隆测序, 经序列分析后, 挑取单个经 pGEX-*HMG1* 转化的序列正确的 LB21 菌落继续培养至 A_{600nm} 约为 0.6 时, 加入 IPTG 诱导 6 h. 将诱导菌和非诱导菌离心, 沉淀部分用于 SDS-PAGE 鉴定^[16].

2 结果

2.1 RT-PCR 扩增结果

利用兼并引物 S/A 对三倍体鲫鱼、四倍体鲫鱼、团头鲂和红鲫的总 RNA 进行 *HMG1* 基因的 RT-PCR 扩增, 结果在 4 种鱼中各扩增出 1 条约 400bp 的条带(图 1), 与先前预测的大小相同. 经测序并将该序列在 NCBI 数据库内经 BLASTN 比对, 结果表明克隆序列与斑马鱼 *HMG1* (BC165307) 具有很高的序列同源性 ($ID = 92\%$, $E = 2e - 161$), 说明我们已成功获得了实验鱼的 *HMG1* 的中间部分序列.

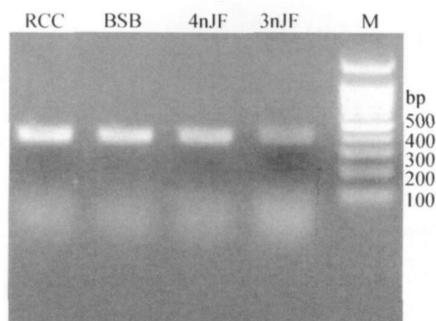


图 1 不同倍性鱼 *HMG1* 基因 RT-PCR 扩增结果

RCC, 红鲫; BSB, 团头鲂; 4nJF, 四倍体鲫鱼; 3nJF, 三倍体鲫鱼; M, DNA 标志 (100bp)

2.2 不同倍性鱼 *HMG1* 基因 cDNA 全长拼接及同源性比较

用引物 GSP1 (3'GSP1 或 5'GSP1) 和 UPM 扩增 RACE cDNA 模板 20 个循环后, 将此 PCR 稀释产物 (1:20) 作模板, 用 GSP2 (3'GSP2 或 5'GSP2) 和 NUP 进行第 2 轮 PCR 扩增后 5' RACE 获得 200bp 的条带, 3' RACE 获得 450bp 的条带. 将目的片段克隆测序, 序列经计算机分析, 去除引物 NUP 和载体序列, RACE 扩增序列与兼并引物扩增序列拼接后, 即获得不同倍性鱼 *HMG1* 基因 cDNA 全长序列(图 2). 该序列包括 1 个 579nt 的开放阅读框, 5' 末端和 3' 末端各 1 个非编码区, 开放阅读框以 ATG 为起始密码子, TAG 为终止密码子, 编码 193 个氨基酸(图 3). 3' 末端非编码区具有两个 mRNA 不稳定信号 (ATTTA) 和 1 个 mRNA 的加尾信号 (AATAAA)(图 2).

用 ClustalW 在线软件 (www.ebi.ac.uk/clustalw/) 分别对不同倍性鱼 *HMG1* cDNA 序列进行同源性比对(表 2). 结果发现: 红鲫与团头鲂的序列同源性为 98%; 四倍体鲫鱼与其母本红鲫同源性高达 99%, 与其父本团头鲂达 97%, 而与三倍体鲫鱼同源性为 96%; 三倍体鲫鱼与其父母本同源性各为 95%. 用 ClustalW 在线软件分别对多倍体鲫鱼及其亲本 *HMG1* 推测的蛋白氨基酸进行同源性比较(表 2), 发现四倍体鲫鱼与其父母本同源性各达 100%, 而三倍体鲫鱼与其亲本同源性各为 97%.

表 2 多倍体鲫鱼与其原始父母本 *HMG1* cDNA(对角线下) 和蛋白氨基酸(对角线上) 序列同源性比较(单位: %)

实验鱼	RCC	BSB	3nJF	4nJF
RCC		100	97	100
BSB	98		97	100
3nJF	95	95		97
4nJF	99	97	96	

2.3 *HMG1* 生物信息学和系统进化分析

将推测的红鲫的 *HMG1* 193 个蛋白氨基酸, 经 Predictprotein 软件在线预测二级结构, 结果发现氨基酸序列 N-端具有 8 个螺旋(helix)单位, 其位点分别为 3—16, 26—36, 42—44, 48—62, 88—103, 110—122, 126—128, 和 132—146; 3 个转角 (switch) 单位, 其位点分别为 1—6, 63—64, 88—92. C-端具有一长为 23 个以 D 或 E 重复的氨基酸尾. 在 NCBI 上用软件进行蛋白三级结构预测, 结果表明氨基酸序列具有两个 DNA 结合基序. 在 NCBI 数据库, 将红鲫 *HMG1* 蛋白氨基酸序列经 BLASTP 软件比对, 结果表示红鲫 *HMG1* 蛋白氨基酸序列与斑马鱼 *HMG1* (AAH45917) 的相似性为 94%, 与鲑鱼 (ACM09402)、虹鳟 *HMG1* (AAA58771) 和真鲷 *HMG1* (AAP20177) 各为 87%, 与爪蟾 *HMG1* 为 83% (NP-001080836), 与猪 *HMG1* (AAW29969) 和短尾负鼠 *HMG1* (XP-001363686) 各为 85%, 与小家鼠 *HMG1* (CAA56631)、牛 *HMG1* (NP-788785) 和马 *HMG1* (XP-001488130) 各为 86%, 与褐家鼠 *HMG1* (XP-573095) 为 83%, 与珍珠鸟 *HMG1* (XP-002194652) 和红原鸡 *HMG1* (NP-990233) 各为 85%, 与黑猩猩 *HMG1* (XP-509611)

```

RCC -----AAGGATCCAACAAAACCAAGAGGCAAAATGCCTCATACGATACTTTGTCC 60
1nJF ACGCGG-AAGGATCCAACAAAACCAAGAGGCAAAATGCCTCTTACGATACTTTGTCC 60
BSB -ACGCGG-GAGGATCCAACAAAACCAAGAGGCAAAATGCCTCATACGATACTTTGTCC 60
3nJF -ACGCGGAAGGATCCAACAAAACCAAGAGGCAAAATGCCTCTTACGATACTTTGTCC 60
*****

RCC AGACCTGCCGTGAGGAGCATAAGAAGAAACACCCTGAGGCAACGGTCAACTTCTGTGAGT 120
1nJF AGACCTGCCGAGAGGAGCATAAGAAGAAACACCCTGAGGCAACGGTCAACTTCTGTGAGT 120
BSB AGACCTGCCGTGAGGAGCATAAGAAGAAACACCCTGAGGCAACGGTCAACTTCTGTGAGT 120
3nJF AGACCTGCCGAGAGGAGCATAAGAAGAAACACCCTGAGGCAACGGTCAACTTCTGTGAGT 120
*****

RCC TTTCCAAAAGTGCTCAGAGCGATGGAAGACTATGTCAGCCAAGGAAAAGGGAAGTTTG 180
1nJF TTTCCAAAAGTGCTCAGAGCGATGGAAGACTATGTCAGCCAAGGAAAAGGGAAGTTTG 180
BSB TTTCCAAAAGTGCTCTGAGCGATGGAAGACTATGTCAGCCAAGGAAAAGGGAAGTTTG 180
3nJF TTTCCAAAAGTGCTCAGAGCGATGGAAGACTATGTCAGCCAAGGAAAAGGGAAGTTTG 180
*****

RCC AAGATATGGCCAAACAAGACAAGGTCCGTTACGAGAGGGAGATGAAAACACTACATTCCAC 240
1nJF AAGATATGGCCAAACAAGACAAGGTCCGTTACGAGAGGGAGATGAAAACACTACATTCCAC 240
BSB AAGATATGGCCAAACAAGACAAGGTCCGTTACGAGAGGGAGATGAAAACACTACATTCCAC 240
3nJF AGGATATGGCCAAACTTGACAAGGCACGTTACGAGAGGGAGATGAGAACACTACATTCCAC 240
* *****

RCC CCAAAGCGGAGAAGAAAAGAGGTTTAAGGACCCCAATGCTCCAAGAGACCCCGTCTG 300
1nJF CCAAAGCGGAGAAGAAAAGAGGTTTAAGGACCCCAATGCTCCAAGAGACCCCGTCTG 300
BSB CCAAAGCGGAGAAGAAAAGAGGTTTAAGGACCCCAATGCTCCAAGAGACCCCGTCTG 300
3nJF CCAAAGCGGAGAAGAAAAGAGGTTTAAGGACCCCAACGCACCCAAGAGACCCCGTCTG 300
*****

RCC CCTTCTTCATTTCTGCTCTGAGTTCGACCCAAGGTGAAGGAAGAGACCCCTGGTCTCT 360
1nJF CCTTCTTCATTTCTGCTCTGAGTTCGACCCAAGGTGAAGGAAGAGACCCCTGGTCTCT 360
BSB CCTTCTTCATTTCTGCTCTGAGTTCGACCCAAGGTGAAGGAAGAGACCCCTGGTCTCT 360
3nJF CATTCTTCATTTCTGCGCCAATTCCGACCCAAGGTAAAAGAAGAACCCCTGGTCTGT 360
* *****

RCC CTATTGGAGATGTGGCTAAGAACTGGGTGAGATGTGGAACAAAACATCAGCTGAGGAGA 420
1nJF CTATTGGAGATGTGGCTAAGAACTGGGTGAGATGTGGAACAAAACATCAGCTGAGGAGA 420
BSB CCATTGGAGATGTGGCTAAGAACTGGGTGAGATGTGGAACAAAACATCAGCTGAGGAGA 420
3nJF CTATTGGAGATGTGGCCAAGAACTGGGTGAGATGTGGAACAAAACATCATCTGAGGAGA 420
* *****

RCC AGCAGCCATTGAGAAGAAGGCAGCCAGGCTGAAGGAGAAGTACGAGAAGGACATCACCG 480
1nJF AGCAGCCATTGAGAAGAAGGCAGCCAGGCTGAAGGAGAATACGAGAAGGACATCACCG 480
BSB AGCAGCCATTGAGAAGAAGGCAGCCAGGCTGAAGGAGAATACGAAAAGGACATCACCG 480
3nJF AGCAGCCATACGAGAAGAAGGCTGCCAGGCTGAAGGAGAATACGAGAAGGACATCACCG 480
*****
    
```

```

RCC      CCTATCGTCCAAGGCAAAGTGGGAGTCCAGCCAAGGCCCCATCCAAGCCAGACAAGG 540
4nJF     CCTATCGTCCAAGGCAAAGTGGGAGTCCAGCCAAGGCCCCATCCAAGCCAGACAAGG 540
BSB      CCTATCGTCCAAGGCAAAGTGGGAGTCCAGCCAAGGCCCCATCCAAGCCAGACAAGG 540
3nJF     CCTATCGTCCAAGGCAAAGTGGGAGTCCAGCCAAGGCCCCATCCAAGCCAGACAAGG 540
*****

RCC      CTAATGATGATGATGATGAAGACGACGATGATGATGATGATGAGGAAGAAGATGAGGAAGAAG 600
4nJF     CTAATGATGATGATGATGAAGACGACGATGATGATGATGATGAGGAAGAAGATGAGGAAGAAG 600
BSB      CTAATGATGATGATGACGAAGACGACGATGATGATGATGATGAGGAAGAAGATGAGGAAGAAG 600
3nJF     CTAATGATGATGATGACGAAGACGACGATGATGATGATGATGAGGAAGAAGATGAGGAAGAAG 600
*****

RCC      ATGACGATGACGAGTAGATAAATAACATTTAGAGATTAGCTTGTGTCTATAAAGCATT 660
4nJF     ATGACGATGACGAGTAGATAAATAACATTTAGAGATTAGCTTGTGTCTATAAAGCATT 660
BSB      ATGACGATGACGAGTAGATAAATAACATTTAGAGACTAGCTTGTGTCTATAAAGCATT 660
3nJF     ATGACGATGACGAGTAGATAAATAACATTTAGAGACTAGCTTGTGTCTATAAAGCATT 660
*****

RCC      AACCCCTGTACTTTACTGAGGCTATAAAATGGGTAAGAGAAATAAACACCACAAC 720
4nJF     AACCCCTGTACTTTACTGAGGCTATAAAATGGGTAAGAGAAATAAACACCACAAC 720
BSB      AACCCCTGTACTTTACTGAGGCTATAAAATGGGTAAGAGAAATAAACACCACAAC 720
3nJF     AACCCCTGTACTTTACTGAGGCTATAAAATGGGTAAGAGAAATAAACACCACAAC 720
*****

RCC      ATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 754
4nJF     ATTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA----- 754
BSB      ATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA----- 754
3nJF     ATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA----- 754
** *****

```

图 2 多倍体鲫鲂及其原始亲本 HMG1 cDNA 序列同源性比对

“*”表示相同的核苷酸位点，“-”表示缺失位点，下画线“ATTTA”表示 mRNA 的不稳定信号，下画线“AATAAA”表示 mRNA 的加尾信号

和人 HMG1 (AAP36330) 各为 86%，与中华鳖为 63%。因此，红鲫 HMG1 基因被命名为 *rcHMG1*，团头鲂为 *bsHMG1*，三倍体鲫鲂为 *3njfHMG1*，四倍体鲫鲂为 *4njfHMG1*，GenBank 登录号分别为 FJ785327, FJ785328, FJ785329 和 FJ785330。

以酵母 HMG1 (CAA86503) 的氨基酸序列为外群，将上述物种的 HMG 蛋白氨基酸序列经 CLUSTALX 软件比对后，用 MEGA4.0 软件构建系统进化树(图 4)。

从图 4 可见，不同物种的 HMG1 和真鲷的 HMG 为平行演化关系，在 HMG1 的一支聚类群

中，中华鳖处于分支树的根部，然后分为 2 分支，1 分支包括鱼类(鲑鱼、虹鳟、三倍体鲫鲂、斑马鱼、团头鲂、红鲫和四倍体鲫鲂)，另 1 分支包括两栖类(爪蟾)、鸟类(珍珠鸟和红原鸡)和哺乳类。结果表明，在动物还没有分化之前，HMG1 可能就存在于其体内，在随后已分化的物种中，由于各自的生物进化速率不同，HMG1 出现了序列变异。哺乳类中的鼠、马、人、猪、黑猩猩和牛等聚为一个类群，红鲫和团头鲂及其杂交后代的四倍体鲫鲂聚为另一个类群，表明 HMG1 在已分化的物种中又具有保守性。

BSB	MSSYAYFVQTCREEHKKKHPEATVNFSEFSKCKSERWKTMSAKEKGFEDMAKQDKVRYE 60
4nJF	MSSYAYFVQTCREEHKKKHPEATVNFSEFSKCKSERWKTMSAKEKGFEDMAKQDKVRYE 60
RCC	MSSYAYFVQTCREEHKKKHPEATVNFSEFSKCKSERWKTMSAKEKGFEDMAKQDKVRYE 60
3nJF	MSSYAYFVQTCREEHKKKHPEATVNFSEFSKCKSERWKTMSAKEKGFEDMAKLDKARYE 60
	***** **.***
BSB	REMKNYIPPKGEKKRKFDPNAPKRPPSAFFIFCSEFRPKVKEETPGLSIGDVAKKGEM 120
4nJF	REMKNYIPPKGEKKRKFDPNAPKRPPSAFFIFCSEFRPKVKEETPGLSIGDVAKKGEM 120
RCC	REMKNYIPPKGEKKRKFDPNAPKRPPSAFFIFCSEFRPKVKEETPGLSIGDVAKKGEM 120
3nJF	REMKNYIPPKGEKKRKFDPNAPKRPPSAFFIFCAEFRPKVKEETPGLSIGDVAKKGEM 120
	*****;*****
BSB	WNKTSAEKQPFEKKAARLKEYEKDITAYRSKGKVGPPAKAPSKPDKANDDDDEDDDD 180
4nJF	WNKTSAEKQPFEKKAARLKEYEKDITAYRSKGKVGPPAKAPSKPDKANDDDDEDDDD 180
RCC	WNKTSAEKQPFEKKAARLKEYEKDITAYRSKGKVGPPAKAPSKPDKANDDDDEDDDD 180
3nJF	WNKTSSEKQPYEKKAARLKEYEKDITAYRSKGKVGPPAKAPSKPDKANDDDDEDDDD 180
	*****;*****
BSB	DEEEDEEEDDDE 193
4nJF	DEEEDEEEDDDE 193
RCC	DEEEDEEEDDDE 193
3nJF	DEEEDEEEDDDE 193

图 3 多倍体鲫鲂及其原始亲本 HMG1 蛋白氨基酸序列同源性比对
“*”表示相同的核苷酸位点

2.4 原核表达及 SDS-PAGE 检测

用 *Xho* I 和 *Eco*R I 双酶切重组载体质粒 pMD18-HMG1, 获得了约 700 bp 的 *HMG1* 基因片段(图 5), 并构建成重组质粒 pGEX-HMG1. 重组质粒 pGEX-HMG1 转化 BL21 细胞并在 IPTG 的诱导下, 表达了目的蛋白(图 6), 在 SDS-PAGE 胶上 50 ku 处呈现为一条较浓的蛋白带, 其大小与预期的相当; 而不作诱导处理的转化菌, 没有出现这条蛋白带. 说明我们已成功地对 *HMG1* 基因进行了原核表达, 从而为进一步研究 HMG1 蛋白的生物学作用奠定了基础.

3 讨论

红鲫、团头鲂及其杂交后代三倍体鲫鲂和四倍体鲫鲂组织总 RNA 中, 经兼并引物 S/A PCR 扩增后, 都能扩增 1 个条带(图 1), 说明物种基因组虽然遭受种间的远源杂交的冲击^[1], 但 *HMG1* 基因能够稳定地从父母本遗传给其杂交

种. 染色体组型分析表明^[4], 三倍体鲫鲂有两套染色体来自母本红鲫, 1 套染色体来自父本团头鲂; 四倍体鲫鲂有两套染色体来自母本, 另两套染色体来自父本. 在不同倍性鱼中扩增出 *HMG1*, 说明 *HMG1* 基因在鱼类的转录调控不受倍性效应调节^[17]. 事实上, *HMG1* 基因编码蛋白参与了生物体内诸如调节基因转录、核小体重建和细胞凋亡等生物发育的基本过程^[9], 这类功能性基因真实地从亲本遗传给子代, 对杂交种的存活是必要的.

对从不同倍性鱼中克隆出的 *HMG1* cDNA 全长(表 2)分析后发现红鲫和团头鲂之间的序列相似性为 98%, 四倍体鲫鲂和母本红鲫的序列相似性为 99%, 四倍体鲫鲂和父本团头鲂的序列相似性为 97%, 表明两性可育的异源四倍体鲫鲂基因组中, *HMG1* 基因与母本具有较高的同源性. 三倍体鲫鲂 *HMG1* 序列与父母本的相似性各为 95%, 低于父母本之间的序列相似性(98%), 表明 *HMG1* 基因

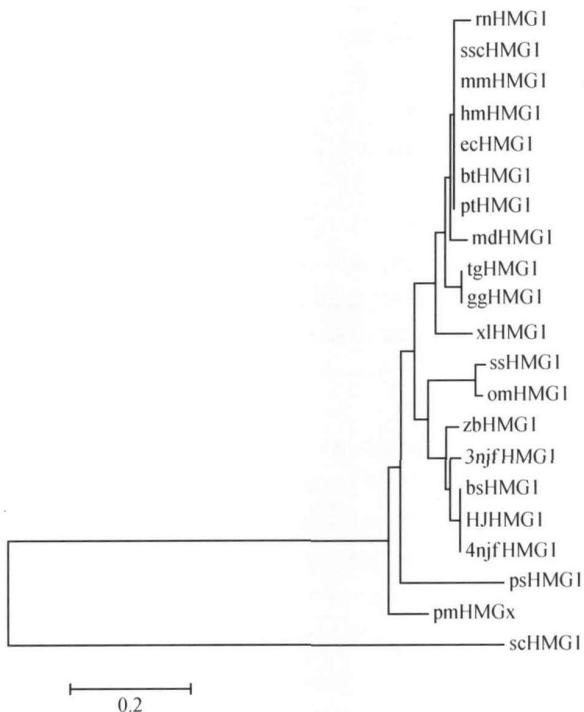


图 4 NJ 法构建的不同物种 HMG1 氨基酸序列系统进化树
 rnHMG1 褐家鼠; ecHMG1 马; hmHMG1 人; sscHMG1 猪;
 ptHMG1 黑猩猩; btHMG1 牛; mmHMG1 小家鼠; mdHMG1 短尾负鼠; tgHMG1 珍珠鸟; ggHMG1 红原鸡; xlHMG1 爪蟾;
 ssHMG1 鲑鱼; omHMG1 虹鳟; 3njfHMG1 三倍体鲫; zbHMG1 斑马鱼; bsHMG1 团头鲂; hJHMG1 红鲫; 4njfHMG1 四倍体鲫; psHMG1 中华鳖; pmHMGx 真鲷; scHMG1 酵母

在两性不育的异源三倍体鲫基因组中发生了位点变异，可能是物种间的杂交冲击基因的结果，并提示三倍体鲫不是染色体简单叠加后的多倍体。在植物多倍体的研究中，倍性效应与远源杂交都能引起 DNA 甲基化的变异从而影响基因表达^[18]。三倍体鲫 HMG1 序列与父母本相似性降低，可能是远源杂交的结果，随之基因功能发生改变，造成三倍体鲫与四倍体鲫性腺可育性的差异。HMG1 氨基酸序列分析发现(表 2)，红鲫、团头鲂和四倍体鲫之间的氨基酸序列相似性均为 100%，三倍体鲫与父母本间的相似性为 97%，三倍体鲫在氨基酸序列水平上发生的变化，进一步说明了远源杂交对三倍体鲫 HMG1 基因造成了一定的影响，三倍体鲫 HMG1 蛋白的功能是否发生了改变，有待于深入研究。四倍体鲫 HMG1 氨基酸序列与亲本的 100% 同源性，克服了等位基因的杂交不

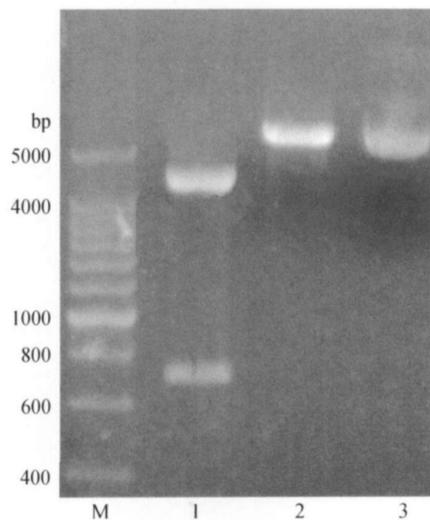


图 5 质粒 pMD18-HMG1 和 pGEX-4T 酶切分析
 1, pMD18-HMG1(EcoRI + XhoI); 2, pGEX-4T(EcoRI + XhoI); 3, pGEX-4T; M, DNA 标志 (200 bp)

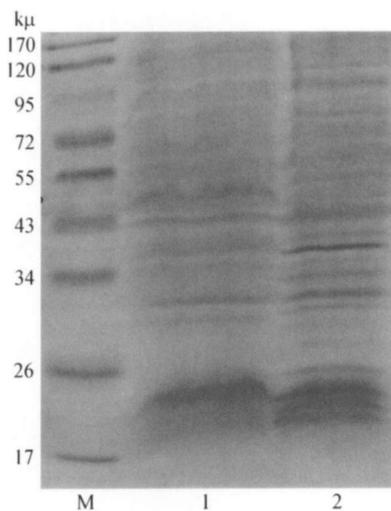


图 6 HMG1 原核表达产物的 SDS-PAGE 分析
 1, 诱导 6h 后表达 50 ku 的融合蛋白 GST-HMG1; 2, 不作诱导处理的阴性对照; M, 蛋白分子质量标志

亲和性，为两性可育的异源四倍体鲫成为自然界中新型多倍体物种奠定了基础。

红鲫 HMG1 蛋白二级结构经 Predictprotein 软件预测，发现具有 8 个螺旋单位，3 个转角单位^[19]。螺旋单位和转角单位能够构成“螺旋-转角-螺旋”结构，这种复合结构可以伸入 DNA 双螺旋的大沟或小沟，每个复合结构采取非特异性的方式与 DNA 结合^[20]。HMG1 蛋白经 NCBI 三级结构预

测,发现具有两个DNA结合区,1个包括了N-端前4个螺旋单位,另1个包括了N-端后4个螺旋单位^[21].HMG蛋白C-端具有一个氨基酸尾^[22],一般包括C尾端连续的约30个氨基酸^[23].进一步研究表明,HMG1这种长的氨基酸尾,能与核小体的核组蛋白H3发生特异性的连接^[24].HMG1蛋白的这种结构,决定了其能够与DNA片段的两端连接,从而使DNA发生弯曲,改变DNA的结构^[25].我们曾研究了红鲫、团头鲂、三倍体鲫鲂和四倍体鲫鲂基因组中的DNA转座子,红鲫未检测出类TC1的转座子,团头鲂和三倍体鲫鲂具有相似的类TC1转座子,而四倍体鲫鲂出现了一种不同于父本的转座子(资料未发表).本文研究结果初步表明一种推测:异源四倍体鲫鲂基因组中存在功能性的HMG1蛋白,该蛋白与母性来源的转座子发生了“蛋白-DNA”的相互作用,导致转座子发生弯曲继而剪切;异源三倍体的HMG1氨基酸序列与亲本相似性降低,这种变化可能削减了HMG1与转座子的结合能力,三倍体鲫鲂基因组中的转座子因此没有发生变化.

邻接法(NJ)构建的系统进化树(图4)分析可见,真鲷的HMG与其他脊椎动物的HMG1是一种平行演化关系^[26],两种基因在进化历程中,可能因功能分化而采取不同的进化方式.从HMG1的聚类树可见,爬行动物的中华鳖处于聚类树的根部,随后分为2分支,1分支包括鱼类,另1分支包括两栖类、鸟类和哺乳类.郑济芳等^[12]用HMG1的两个DNA结合区序列构建了鱼类、爬行类和哺乳类等物种的进化树,中华鳖HMG1也是处于进化树的根部.HMG1的聚类研究结果证明了一个观点,即HMG1被认为是进化上较为古老的一种蛋白,起源可能早于5亿年前,生物还未分化之前就已存在^[22].此外,黑猩猩、人、牛、马、猪和鼠等不同的哺乳动物聚为一群,表明HMG1虽然历经了漫长的演化,可能因为在生物发育方面有重要的调控作用,其氨基酸序列又具有保守性.由此断定,物种的分化与HMG1的演化可能速率不一致,由此在新物种形成时,HMG1能正常行使其生物学功能,从而保证了新物种的正常发育、存活和以至于繁殖后代.通过红鲫和团头鲂杂交所获得的可育四倍体鲫鲂便是一个很好的例证.

我们用红鲫HMG1 cDNA开放阅读框序列构建了原核表达载体.实验表明,外源的红鲫HMG1基因在原核表达载体中能够表达融合蛋白.鱼HMG1基因编码蛋白在原核细胞中得到了表达,为下一步该蛋白的制备和生物学功能研究,尤其是与DNA转座子相互作用的研究奠定了基础.

致谢 该实验得到本研究室姚占洲,罗凯坤,陶敏老师,曾小明和邹拓迷研究生的帮助,在此一并致谢.

参 考 文 献

- 1 Evgen'ev M. Mobile elements and genome evolution. *Mol Biol*, 2007, 41(2): 203-213
- 2 Takayoshi U, Yoshio O, Ryozo S, et al. Triploid hybrids between female rainbow trout and male brook trout. *Bull Jap Soc Sci Fish*, 1984, 50(8): 1331-1336
- 3 余先觉. 中国淡水鱼类染色体. 北京: 科学出版社, 1989, 44-75
- 4 Liu SJ, Qin QB, Xiao J, et al. The formation of the polyploidy hybrids from different subfamily fish crossings and its evolutionary significance. *Genetics*, 2007, 176: 1023-1034
- 5 Finnegan DJ. Transposable elements in eukaryotes. *Int Rev Cytol*, 1985, 93: 281-326
- 6 Miskey C, Izsvak Z, Kawakami K, et al. DNA transposons in vertebrate functional genomics. *CMLS, Cell Mol Life Sci*, 2005, 62: 629-641
- 7 Hatem Z, Izsvak Z, Khare D, et al. The DNA-bending protein HMG1 is a cellular cofactor of *Sleeping Beauty* transposition. *Nucleic Acids Research*, 2003, 31(9): 2313-2322
- 8 Dik CG, Hiom K, Tanya TP, et al. Stimulation of V(D)J cleavage by high mobility group proteins. *EMBO J*, 1997, 16(10): 2665-2670
- 9 徐佳, 刘志锋, 姜勇. 高迁移率族蛋白与真核基因表达调控. *生物化学与生物物理进展*, 2005, 32(5): 397-402
- 10 Goodwin GH, Sander C, Johns EW. A new group of chromatin associated proteins with a high content of acidic and basic amino acids. *Eur J Biochem*, 1973, 38(1): 14-19
- 11 Bustin M. Regulation of DNA-dependent activities by the functional motifs of the high-mobility-group chromosomal proteins. *Mol Cell Biol*, 1999, 19(8): 5237-5246
- 12 郑济芳, 胡弼, 吴端生. 中华鳖HMG1基因的克隆与序列分析. *水生生物学报*, 2006, 30(2): 192-197
- 13 别良峰, 于文彬, 程晓东, 等. 人高迁移率族蛋白1基因的克隆和表达. *第四军医大学学报*, 2003, 24(3): 232-234
- 14 Thompson JD, Higgins DG, Gibson TJ. Clustal W: Improving

- the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*, 1994, 22: 4673—4680
- 15 Koichiro T, Joel D, Masatoshi N, et al. MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol*, 2007, 24(8): 1596—1599
- 16 张维铭(主编). 现代分子生物学实验手册. 北京: 科学出版社, 2003, 138—144
- 17 Chen ZJ, Ni Z. Mechanisms of genomic rearrangents and gene expression changes in plant polyploidys. *Bioessays*, 2006, 28(3): 240—252
- 18 彭海, 张静, 吴先军. 植物基因表达中的倍性效应: 研究进展、问题与展望. *中国科学 C*, 2008, 38(1): 1—7
- 19 Malin Y, Kent K, KEN AD, et al. Predicting conformational switches in proteins. *Protein Sci*, 1999, 8: 1752—1764
- 20 Saito K, Kikuchi T, Yoshida M. The mechanism of sequence non-specific DNA binding of HMG1/2-boxB in HMG1 with DNA. *Protein Eng*, 1999, 12: 235—242
- 21 Aron MB, Anderson JB, Derbyshire MK, et al. CDD: A conserved domain database for interactive domain family analysis. *Nucleic Acids Res*, 2007, 35: 237—240
- 22 Lotze MT, Tracey KJ. High-mobility group box 1 protein (HMGB1): Nuclear weapon in the immunarsenal. *Nat Rev*, 2005, 5: 331—342
- 23 Tsuda K, Kikuchi M, Mori K, et al. Primary structure of non-histone protein HMG1 revealed by the nucleotide sequence. *Biochemistry*, 1988, 27: 6159—6163
- 24 Toshifumi K, Koichi S, Tetsuya U, et al. Distinct domains in HMGB1 are involved in specific intramolecular and nucleosomal interactions. *Biochemistry*, 2008, 47: 13991—13996
- 25 Gaillard C, Strauss F. High affinity binding of proteins HMG1 and HMG2 to semicatenated DNA loops. *BMC Mol Biol*, 2000, 1: 1
- 26 黄大卫. 支序系统学概论. 北京: 农业出版社, 1996, 23—55