

太湖鲂鲢 F₂ 代 GH 基因结构、系统发育和表达特征*

刘加林^{1,2} 贾永义¹ 刘士力¹ 郑建波¹ 迟美丽¹
程 顺¹ 李 飞¹ 尹绍武² 顾志敏^{1①}

(1. 浙江省淡水水产研究所 农业农村部淡水渔业健康养殖重点实验室/浙江省淡水水产遗传育种重点实验室 湖州 313001;
2. 南京师范大学生命科学学院 南京 210023)

摘要 采用 PCR 及 T-A 克隆技术, 从基因组 DNA 成功克隆三角鲂、太湖鲂鲢、太湖鲂鲢 F₂ 代含完整阅读框的生长激素基因全序列, 经测序拼接获得序列全长分别为 6009、6042 和 6058bp, 转录单元长分别为 2049、2014 和 2045bp。三种群体的 GH 基因序列均包括五个外显子、四个内含子及 5' 侧翼区和 3' 侧翼区; 编码氨基酸序列均为 633bp, 编码 210 个氨基酸。对 3 个目 14 种鱼类进行同源性及系统发育分析发现, 太湖鲂鲢 F₂ 代与其祖父母的遗传距离相比亲本较近。另取 10 尾太湖鲂鲢 F₂ 代, 对其进行 GH 基因编码区 SNPs 筛选, 共得到 7 处 SNPs, 其中有 3 处位点的突变导致其氨基酸序列发生了明显的变化; 对其蛋白质结构分析发现, 突变后的氨基酸残基导致蛋白质二级结构发生改变。我们将该 10 尾个体按体重分为小个体组和大个体组, 经组织表达分析发现, 大个体组中的个体具有较高的转录水平, 进一步分析发现其氨基酸残基有所不同。

关键词 太湖鲂鲢; 属间杂交; 生长激素; 克隆; SNPs; 表达分析

中图分类号 Q789; S965 **doi:** 10.11693/hyhz20200100006

太湖鲂鲢 (*Culter alburnus*♀×*Megalobrama terminalis*♂) 是以经两代群体选育和两代异源雌核发育诱导的翘嘴鲌 (*Culter alburnus*) 为母本, 连续三代群体选育的三角鲂 (*Megalobrama terminalis*) 为父本, 通过远缘杂交获得的国家水产新品种 (GS-02-001-2017), 具有生长快、饲料转化率高、体型适中、抗逆性强等特点, 适宜在全国可控水域推广养殖。与大多杂种优势利用类似, 太湖鲂鲢也是父母本杂交一代的经济性状优势利用, 其自交后代具有显著的性状分离现象 (姜建湖等, 2019), 不适宜推广养殖, 但在自交后代中不乏生长、体型、鳞片等显著优势的个体, 可作为后代品种选育的优良育种素材。

生长激素 (growth hormone, GH) 是能够调节控制动物内分泌机制, 使进入机体内的营养物质按动物生长的需求朝着有益的方向进行再分配功能的各类激素的总称, 可以调控信号传导、肢体的发育以及细

胞的定向分化等 (严美姣等, 2010)。鱼类的生长激素与催乳素 (prolactin, PRL) 和胎盘催乳素 (placental lactogen, PL) 属于同一个基因家族 (杨学明等, 2008), 是影响动物生长发育的主效基因 (Møller *et al.*, 2007)。杨学成等 (1991) 最早从大麻哈鱼 (*Oncorhynchus keta*) 脑垂体 cDNA 文库中克隆得到全长为 1.12kb 的 GH cDNA。随着技术的不断发展, 除了在禽畜物种上对 GH 基因的大量研究外 (宋成义等, 2000; 王文君等, 2003), 对鱼类 GH 基因也有大量报道。已有研究表明, 鲤鱼 (*Cyprinus carpio*) (Hong *et al.*, 1993)、草鱼 (*Ctenopharyngodon idellus*) (Ho *et al.*, 1991)、鲶鱼 (*Silurus asotus*) (Tang *et al.*, 1993) 等大多数鱼类存在 5 个外显子和 4 个内含子, 斜带石斑鱼 (*Epinephelus coioides*) 和翘嘴鲌 (*Siniperca chuatsi*) 等少数鱼类具有 6 个外显子和 5 个内含子 (Almuly *et al.*, 2000; 鲁双庆等, 2008; Ma *et al.*, 2012)。本团队已对翘嘴鲌 GH 基

* 浙江省重点研发计划项目, 2016C02055-1 号。刘加林, 硕士研究生, E-mail: 1961954694@qq.com

通信作者: 顾志敏, 研究员, E-mail: guzhimin2006@163.com

收稿日期: 2020-01-06, 收修改稿日期: 2020-05-24

因(刘士力等, 2019)及两种生长激素受体(GHR)进行了分析(刘士力等, 2020), 均发现了微卫星位点且部分位点与生长性状具有关联性。在 GH 基因的 5'端侧翼区发现 2 个微卫星位点, 且与体长、体质量等生长性状具有关联性; GHR1 和 GHR2 中均发现了微卫星序列, 其中 GHR2 基因所发现的 4 个微卫星位点均具有多态性, 且均与生长性状相关。本研究通过克隆获得三角鲂、太湖鲂鮰、太湖鲂鮰 F₂ 代的 GH 基因全长序列, 分析其序列所具有的差异性; 并单独对太湖鲂鮰 F₂ 代 GH 基因进行 SNPs 的筛选以及组织表达分析, 探究基因结构的改变对基因表达带来的影响。本研究结果将有助于阐明决定快速生长相关性状的分子机制, 也将有助于对未来太湖鲂鮰新品种标记辅助选择育种提供良好的理论基础。

1 材料与方法

1.1 样品采集

实验所用翘嘴鲂、三角鲂、太湖鲂鮰及太湖鲂鮰 F₂ 代均来自于浙江省淡水水产研究所综合试验基地, 均剪取少量尾鳍于无水乙醇中-20°C 保存, 用于后续

DNA 的提取; 从 F₂ 代群体中筛选出体重(BW)、体长(BL)和体高(BH)等性状差异较大的 10 尾鱼, 对每尾鱼进行形态测量后分别剪取尾鳍用于 DNA 的提取, 并取适量的肌肉组织分成 3 份经样品保护液(Sample Protector for DNA/RNA) 4°C 过夜后于-80°C 保存, 用于 RNA 的提取。

1.2 GH 基因的克隆及 SNPs 筛选

本课题组已获得翘嘴鲂GH基因序列全长(刘士力等, 2017), NCBI登录号为KX925976。利用已获得的三角鲂、太湖鲂鮰、太湖鲂鮰F₂代转录组中GH的mRNA序列进行Blast, 发现均与翘嘴鲂和团头鲂(*Megalobrama amblycephala*)(登录号: AF463498)推测的mRNA序列相似度极高, 根据其保守区域和已知序列设计7对特异性引物(表1)。引物由生工生物工程(上海)股份有限公司合成。PCR反应体系及反应程序均按照刘加林等(2019)方法进行实施, PCR产物克隆后送生工生物工程(上海)股份有限公司进行测序。

从F₂代群体中挑选出10尾体型具有明显差异的个体, 通过克隆及测序获得每尾鱼的GH基因全序列, 对该10条序列进行对比, 筛选出编码区SNPs。

表 1 GH 基因克隆 PCR 引物序列
Tab.1 Primer sequences used for GH gene cloning

引物	序列(5'—3')	产物长度(bp)	用途
GH-1	F: CGGTAGACAGTCATCAGAAT R: CCACAACCATCCAATCAAIT	664	1—604
GH-2	F: GTGGAAGTAAAGTTACAGTCATC R: GTCGCACGGGTATATTTCTA	1397	10—1406
GH-3	F: AATGAAGTCCTTAGCAATGC R: ATCACATCCATACCTCTAGC	1146	1214—2359
GH-4	F: CTTAGTGCCAACAACATCAT R: GCTCTTCTGTGTTTCATCTT	1219	2107—3325
GH-5	F: GTTATCAAGGAGGACAACCT R: ATACAGCAGACACATTGGAT	1588	3203—4790
GH-6	F: CTCCCATAICTAAACCCACTACT R: GCTGAATACACGACTCCTAA	1291	4423—5713
GH-7	F: ACATCCTTCCAGAATCCTTC R: CAACCTACGCTACCATCTAA	868	5160—6058
GH	F: CTGTTGTCGGTGGTGCTGGTTAG R: ACGGATGACTGCGTTGTTGAAGAG	87	定量 PCR
EFa1	F: ACATGAGGCAGACCGTTGCT R: GGCAGCCTTCTGTGCAGATT	220	内参

1.3 GH 基因的生物分析

所获得的 DNA 片段和已知的 mRNA 序列拼接及对比按照刘加林等(2019)等的方法进行, 以确定三角鲂、太湖鲂鮰、太湖鲂鮰 F₂ 代 GH 基因正确性。GH 基因的 mRNA 及编码的氨基酸序列按照刘士力等(2016)等的方法进行分析。微卫星和小卫星的查找分别通过 SSRhunter 1.3 和在线软件 Repfind 进行。利用 MEGA 7.0 软件采用邻接法(neighbor-joining method)

将翘嘴鲂、三角鲂、太湖鲂鮰、太湖鲂鮰 F₂ 代同鲂形目(Perciformes)、鲤形目(Cypriniformes)、鲇形目(Siluriformes)在内的共 14 种鱼类进行系统发育分析。

利用在线软件分别对 10 尾 F₂ 代个体 GH 基因蛋白质结构进行预测; ExPASy 和 NCBI 中 Conserved Domains (CD-Search)程序对有无跨膜结构及保守的蛋白质结构域进行预测; SignalP 和 PredictProtein 对信号肽及蛋白质二级结构进行预测。

1.4 GH 基因的定量表达分析

将 RNA 进行提取与检测以及反转录合成 cDNA, 具体方法按照说明书进行。根据本实验中已经得到的 F₂ 代 GH 基因的开放阅读框 cDNA 序列, 设计合成一对跨内含子的正反向特异荧光定量引物 GH, 并且以内参基因 EFa1 作为对照(表 1)。实时定量操作具体步骤按照刘士力等(2019)的方法进行, 目标基因转录本的相对表达量用 qRT(Roche)软件按照 Livak 等(2001)描述的方法计算, 使用 Excel 绘制直方图。为了保证实验结果的可靠性, 每个样品, 实验组和对照组的反应都进行了三次重复。

2 结果与分析

2.1 太湖鲂鲢与亲本及太湖鲂鲢 F₂ 代 GH 基因序列全长分析

采用长片段 PCR 及 T-A 克隆技术, 从基因组 DNA 成功克隆三角鲂、太湖鲂鲢(翘嘴鲢♀×三角鲢♂)和太

湖鲂鲢 F₂ 代含完整阅读框的生长激素基因全序列。本课题组已成功克隆出翘嘴鲢 GH 基因全长序列, 为 5966bp, 转录单元长 1648bp。经测序拼接获得三角鲂、太湖鲂鲢和太湖鲂鲢 F₂ 代序列全长分别为 6009、6042 和 6058bp; 转录单元长分别为 2049、2014 和 2045bp。四种群体 GH 基因序列均包括五个外显子、四个内含子及 5'侧翼区和 3'侧翼区; 编码氨基酸序列均为 633bp, 编码 210 个氨基酸。其序列组成及长度见表 2。

太湖鲂鲢和亲本及太湖鲂鲢 F₂ 代 GH 全基因序列的各自 A、T、C、G 的碱基含量近乎一致, 其中翘嘴鲢和三角鲢的 AT 总含量均为 65.0%, 太湖鲂鲢和太湖鲂鲢 F₂ 代的 AT 总含量均为 64.8%, 该四种群体均显示出了一致的 AT 偏好性。由此可知, 太湖鲂鲢 F₂ 代的 GH 基因序列全长虽然稍大于其亲本及祖父母本, 但其 AT、CG 的总含量与该三种群体无明显差异, 说明太湖鲂鲢 F₂ 代在碱基组成方面具有稳定遗传性。

表 2 四种群体 GH 基因的序列长度(bp)
Tab.2 Sequence lengths of GH genes in four populations (bp)

区域名称	翘嘴鲢	三角鲢	太湖鲂鲢	F ₂
5'侧翼区	2282	2339	2301	2359
第一外显子非编码区	54	54	54	54
第一外显子编码区	10	10	10	10
第一内含子	271	247	247	247
第二外显子	140	140	140	140
第二内含子	454	457	456	456
第三外显子	117	117	117	117
第三内含子	437	471	437	469
第四外显子	162	162	162	162
第四内含子	136	136	136	135
第五外显子编码区	204	204	204	204
第五外显子非编码区	51	51	51	51
3'侧翼区	1648	1621	1727	1654
全长	5966	6 009	6042	6058

2.2 太湖鲂鲢与亲本及太湖鲂鲢 F₂ 代 GH 基因序列与氨基酸序列对比分析

四种群体中太湖鲂鲢 F₂ 代 GH 基因序列最长, 比翘嘴鲢、三角鲢及太湖鲂鲢分别长 59、49 和 16bp。该差异的产生主要是由于各序列 5'侧翼区、内含子及 3'侧翼区长度存在不同程度的差异, 在所有外显子处序列的长度均无差异。以下图例中翘嘴鲢、三角鲢、

太湖鲂鲢、太湖鲂鲢 F₂ 代分别以 I、II、III、IV 表示。

5'侧翼区, 太湖鲂鲢 F₂ 代与三角鲢在同一位置存在 (AAT)₇ 的微卫星序列, 而太湖鲂鲢与翘嘴鲢一致, 在相同位置均存在 (AAT)₈ 的微卫星序列(图 1); 3'侧翼区, 太湖鲂鲢 F₂ 代与太湖鲂鲢和三角鲢微卫星序列相同, 在相同位置存在 (TTC)₆T(TAA)₆ 的微卫星序列, 而翘嘴鲢存在 (TTC)₅T(TAA)₈ 的微卫星序列(图 1)。

I TCACAATGCAAAAAATAAAATAATAATAATAATAATAATAATGA-----TATTAA----- 5'(920nt-984nt)
 II TCACAATGCAAAAAATAAAATAATAATAATAATAATAATAATTATTAGGCTTTTTTAATGTATAGTTTT 5'(930nt-995nt)
 III TCACAATGCAAAAAATAAAATAATAATAATAATAATAATAATAATGA-----TATTAA----- 5'(920nt-984nt)
 IV TCACAATGCAAAAAATAAAATAATAATAATAATAATAATAATTATTAGGCTTTTTTAATGTATAGTTTT 5'(920nt-984nt)
 I CATTATATTATCAATTTCTTCTTCTTCTTCTTAATAATAATAATAATAATAATAA CCTACTCTTT 3'(1200nt-1265nt)
 II CATTATATTATTACATTTCTTCTTCTTCTTCTTCTTAATAATAATAATAATAATAATAGCCTACTCTTT 3'(1200nt-1265nt)
 III CATTATATTATTACATTTCTTCTTCTTCTTCTTCTTAATAATAATAATAATAATAATAGCCTACTCTTT 3'(1200nt-1265nt)
 IV CATTATATTATTACATTTCTTCTTCTTCTTCTTCTTAATAATAATAATAATAATAATAGCCTACTCTTT 3'(1200nt-1265nt)

图 1 5'及 3'侧翼区微卫星序列比较

Fig.1 Comparison of microsatellite sequences in 5'and 3' flanking regions
注: 阴影加粗部分为微卫星序列, 下划线处表示差异碱基及间隔碱基“T”

第一内含子, 存在8bp以上的差异位点有2处, 8bp以下的差异位点有6处, 太湖鲂 F₂代在该部分与太湖鲂的相似度为100%, 且两者在第一内含子处主要遗传于三角鲂; 第二内含子, 共有16处差异位点, 太湖鲂序列在其中10处只来自于父本三角鲂, 2处来只自于母本翘嘴鲂, 有4处均不来源于亲本, 可能是由于碱基的突变导致。在太湖鲂发生突变的其中1处位点及未发生突变的12处差异位点中, 太湖鲂 F₂代在该13处位点均与其一致(图2); 第三内含子, 太湖鲂 F₂代在全部16处差异位点中发生突变或缺失, 与太湖鲂、翘嘴鲂和三角鲂均不一致(图2); 第四内含子, 太湖鲂 F₂代有3处位点均发生了碱基的突变或缺失, 有2处分别来源于太湖鲂和三角鲂。由此可判断, 在第一、二内含子中, 太湖鲂 F₂代主要遗传于太湖鲂及三角鲂, 遗传较为稳定; 在第三、四内含子中, 太湖鲂 F₂代遗传不稳定, 发现了多个位点碱基的突变或缺失。另外, 在四种群体的内含子中均未发现微卫星序列。

四种群体的GH基因DNA序列均含完整阅读框及起始密码子、终止密码子。根据图3可以看出, 在第4、5外显子中共有3处差异性位点, 自交F₂代在第4外显子区有1处位点完全变异, 该位点位于密码子第3位碱基, 但并未改变其编码的氨基酸; 另外太湖鲂在第5外显子区有2处差异位点, 该两处位点均位于密码子第二位碱基, 导致其编码的氨基酸发生了改变, 太湖鲂 F₂代及其祖父母在该2处编码的氨基酸分别为谷氨酸、天冬氨酸, 而太湖鲂均为甘氨酸。

2.3 4 个群体与其他鱼类 GH 基因同源性分析及系统进化树构建

将4个群体同鲤形目鱼类; 团头鲂(*Megalobrama Amblycephala* AF463498)、鲢(*Hypophthalmichthys molitrix* M94348)、草鱼(*Ctenopharyngodon idella* X60419)、斑马鱼(*Danio rerio* AJ937858); 鲈形目鱼类; 褐石斑鱼(*Epinephelus bruneus* GU138644)、河川沙塘鳢(*Odontobutis potamophila* MH717101)、斜带石斑鱼(*Epinephelus coioides* KR269816)、翘嘴鲈(*Siniperca chuatsi* EF205280); 鲇形目鱼类; 斑点叉尾鲴(*Ictalurus punctatus* AF267989)、黄颡鱼(*Pelteobagrus fulvidraco* KU323395), 通过DNAMAN软件进行GH基因编码区核苷酸序列及氨基酸序列同源性分析, 如表3所示。太湖鲂 F₂代与其亲本的同源性为93.80%, 与祖父母相似度最高(均为99.84%), 其次为鲤形目其他鱼类(89.26%—99.84%), 与鲇形目的同源性(75.51%—76.94%)明显高于鲈形目(52.44%—58.84%), 该结果与氨基酸序列同源性分析结果基本一致。

利用MEGA 7.0软件, 对本实验所获得的太湖鲂与亲本及太湖鲂 F₂代, 从GenBank数据库中获得的包括鲈形目、鲤形目、鲇形目在内的共14种鱼类GH编码区核苷酸序列, 以构建NJ系统进化树。图4表明, 整个进化树共分为3个分支, F₂代与鲤形目鱼类遗传距离相对最近, 其次为鲇形目, 以100%的置信度形成完整的一支, 鲈形目中的褐石斑鱼单独组成一支, 另外三种鱼类以87%的置信度组成另一支, 由此组成鱼类GH基因的系统进化树。该进化

I CCATTA~~AA~~AGTCTATTCAAAAATTGCAGATCTTAACCTGTGATTAGAACAAAAACGATCGATCACACC In2(10nt-77nt)

II CCATTA~~AA~~ATTCTATTCAAAAATTGCAGATCTTAACCTGTGATTAGAACAAAA-----TCGATCACACC In2(10nt-77nt)

III CCATTA~~AA~~ATTCTATTCAAAAATTGCAGATCTTAACCTGTGATTAGAACAAAA-----TCGATCACACC In2(10nt-77nt)

IV CCATTA~~AA~~ATTCTATTCAAAAATTGCAGATCTTAACCTGTGATTAGAACAAAA-----TCGATCACACC In2(10nt-77nt)

I ~~AGT~~-----TTAAAATTCTGTCAACATTATTGC~~CC~~CTCAATTTGTTCCAAACCTGTATGAATTCT In2(100nt-167nt)

II AGGACTCAAATTTAAAATTCTGTCAACATTATTGC~~CC~~CTCAATTTGTTCCAAACCTGTAAAGAATTCT In2(100nt-167nt)

III AGGACTCAAATTTAAAATTCTGTCAACATTACTTGC~~CC~~CTCAATTTGTTCCAAACCTGTAAGAATTCT In2(100nt-167nt)

IV AGGACTCAAATTTAAAATTCTGTCAACATTATTGC~~CC~~CTCAATTTGTTCCAAACCTGTAAGAATTCT In2(100nt-167nt)

I CTCTGCTGAACACAAAGAAGATATTTGAAGAACATGTTCTAGCCAACAGTTGATGGTCCCCAT In2(70nt-236nt)

II CTCTGCTGAACACAAAGAAGATATTTGAAGAACATGGG-TAGCCAACAGTTGATGGTCCCCAT In2(70nt-236nt)

III CTCTGCTGAACACAAAGAAGATATTTGAAGAACATGGG-TAGCCAACAGTTGATGGTCCCCAT In2(70nt-236nt)

IV CTCTGCTGAACACAAAGAAGATATTTGAAGAACATGGG-TAGCCAACAGTTGATGGTCCCCAT In2(70nt-236nt)

I ATACTATGGAAGTCACTGCACATTCTTCAAATATTGCTTTTGTGATCAGCAAAGAAAGAAACAAC In2(260nt-327nt)

II ATACTATGGAAGTCACGCACATTCTTCAAATATTGCTTTTGTGATCAGCAAAGAAAGAAACAAC In2(260nt-327nt)

III ATACTATGGAAGTCACAGACATTCTTCAAATATTGCTTTTGTGATCAGCAAAGAAAGAAACAAC In2(260nt-327nt)

IV ATACTATGGAAGTCACAGACATTCTTCAAATATTGCTTTTGTGATCAGCAAAGAAAGAAACAAC In2(260nt-327nt)

I AGGGTGAGTAA In2(330nt-340nt) AAATGTTATTTGAGTGAGTAAATACTATTTTCCTTTT-GTTGTCA In2(416nt-460nt)

II ATGGTGAGTAA In2(330nt-340nt) AAATGTTATTTGAGTGAGTAAATACTATTTTCCTTTTTGTTATCA In2(416nt-460nt)

III ATGGTGAGTAA In2(330nt-340nt) AAATGTTATTTGAGTGAGTAAATACTATTTTCCTTTT-GTTATCA In2(416nt-460nt)

IV ATGGTGAGTAA In2(330nt-340nt) AAATGTTATTTGAGTGAGTAAATACTATTTTCCTTTT-GTTATCA In2(416nt-460nt)

I TTATGCTTTAGTCAACGATATATAGAACTATGAAACAAATAGTAGTATTGTACATGCATGATGCAAGTG In3(40nt-108nt)

II TTATGCTTTAGTCAACGATATATAGAACTATGAAACAAATAGTAGTATTGTACATGCATGATGCAAGTG In3(40nt-108nt)

III TTATGCTTTAGTCAACGATATATAGAACTATGAAACAAATAGTAGTATTGTACATGCATGATGCAAGTG In3(40nt-108nt)

IV TTATGTTTTAGTCAACGATATATAGAACTATGAAACAAATAGTAGTATAGTACATGCATGATGCAAGTG In3(40nt-108nt)

I AATATTGTTGTTGCTTCACAATTGCTCAAACACCAGATTAAGTCCATCATCACTACACTCTAAAA In3(110nt-178nt)

II AATATTTTTGTTGCTTCACAATTGCTCAAACACCAGATTAAGTCCATCATCACTACGCTCTAAAA In3(110nt-178nt)

III AATATTGTTGTTGCTTCACAATTGCTCAAACACCAGATTAAGTCCATCATCACTACACTCTAAAA In3(110nt-178nt)

IV CATATTTTTGTTGCTTCACAATTGCTCAAAAATCAGATTAAGTCCATCATCACTACACTGTA AAA In3(110nt-178nt)

I AATTATTTTCACTATTTATCACATTTTTCTTTTGTCAAATCAACTTAATTAATGTGGTTCAGATAACATA In3(180nt-250nt)

II AATTATTTTCACTATTTATCACATTTTTCTTTTGTCAAATCAACTTAATTAATGTGGTTCAGATAACATA In3(180nt-250nt)

III AATTATTTTCACTATTTATCACATTTTTCTTTTGTCAAATCAACTTAATTAATGTGGTTCAGATAACATA In3(180nt-250nt)

IV ATTATTTTCACTATTTGTCACATTTTTCTTTTGTCAAATCAACTTAATTAACGTGGTTCAGATAACATA In3(180nt-250nt)

I TCACCTT-----CTGAAACTTATTTTTT-AAGTTAAACCA In3(280nt-349nt)

II TCACCTTTGATGGTTTAACTCAAATTTTAAATTCAATGAACTGAAACTTATTTTTT-AAGTTAAACCA In3(280nt-349nt)

III TCACCTT-----CTGAAACTTATTTTTT-AAGTTAAACCA In3(280nt-349nt)

IV TCACCTTTGATGGTTTAACTCAAATTTTAAATTCAATGAACTGAAACTTATTTTTTAAAGTTAAACCA In3(280nt-349nt)

I TAATCTTTTTTCACAGTGATGATTTTCT In3(351nt-380nt)TGAAGCCTCCAGTATTATTTTTCTCT In3(440nt-466nt)

II TAATCTTTTTTCACAGTGATGATTTTCT In3(351nt-380nt)TGAAGCCTCCAGTATTATTTTTCTCT In3(440nt-466nt)

III TAATCTTTTTTCACAGTGATGATTTTCT In3(351nt-380nt)TGAAGCCTCCAGTATTATTTTTCTCT In3(440nt-466nt)

IV CAATCTTTTTTACAGTGATGATGTTTCT In3(351nt-380nt)TGAAGCCTCCAGTATTTTTTTTCTTT In3(440nt-466nt)

图2 内含子差异位点序列比较
 Fig.2 Sequence comparison of intron differential sites
 注: 下划线处阴影加粗部分为差异碱基

表 3 4 个群体与其他鱼类 GH 基因编码区核苷酸序列及氨基酸序列的同源性分析(%)

	Tab.3 Homology analysis of the nucleotide sequence and amino acid sequence in the GH gene coding region of four populations and other fishes (%)												
	<i>C. alburnus</i>	<i>M. terminalis</i>	<i>C. alburnus</i> ♀ × <i>M. terminalis</i> ♂	F ₂	<i>M. amblycephala</i>	<i>H. molitrix</i>	<i>C. idella</i>	<i>D. rerio</i>	<i>O. potamophil</i>	<i>E. coioides</i>	<i>E. bruneus</i>	<i>S. chuatsi</i>	<i>P. fulvidraco</i>
<i>C. alburnus</i>	100.00	99.05	100.00	100.00	98.57	98.10	91.90	43.60	53.30	46.67	53.30	74.76	75.71
<i>M. terminalis</i>	100.00	99.05	100.00	100.00	98.57	98.10	91.90	43.60	53.30	46.67	53.30	74.76	75.71
<i>C. alburnus</i> ♀ × <i>M. terminalis</i> ♂	99.68	99.84	99.05	99.05	98.57	97.14	90.95	43.60	53.30	46.19	53.30	74.76	75.24
F ₂	99.84	99.84	93.80	100.00	98.57	98.10	91.90	43.60	53.30	46.67	53.30	74.76	75.71
<i>M. amblycephala</i>	100.00	99.68	99.84	99.84	98.57	98.10	91.90	43.60	53.30	46.67	53.30	74.76	75.71
<i>H. molitrix</i>	98.58	98.58	98.58	98.42	98.58	98.57	91.43	43.13	53.77	44.76	53.77	74.29	74.76
<i>C. idella</i>	98.26	98.26	97.95	98.10	98.10	98.57	90.95	43.13	53.30	45.71	53.30	74.29	74.76
<i>D. rerio</i>	89.42	89.42	89.10	89.26	88.78	88.78	51.89	43.13	52.36	43.81	52.83	73.33	73.81
<i>O. potamophila</i>	51.97	51.97	52.13	52.44	52.76	50.86	51.89	70.45	69.61	53.73	71.08	44.33	45.32
<i>E. coioides</i>	59.94	59.94	59.62	59.78	59.31	59.78	59.94	59.94	67.16	67.16	98.53	53.88	53.88
<i>E. bruneus</i>	53.55	53.55	52.92	53.55	52.92	53.71	53.24	58.91	70.57	67.16	67.65	46.08	46.57
<i>S. chuatsi</i>	58.84	58.84	59.15	58.84	58.22	58.53	59.62	69.81	91.06	71.87	67.65	53.88	53.40
<i>P. fulvidraco</i>	75.51	75.51	75.20	75.51	75.36	74.88	74.09	55.39	59.58	52.90	58.94	94.50	94.50
<i>I. punctatus</i>	76.94	76.94	76.62	76.94	76.46	76.15	74.88	54.98	59.90	53.57	58.94	92.54	92.54

注: 对角线上方为氨基酸序列的同源性, 下方为编码区核苷酸序列的同源性

I AGCCTGACCGT Ex4(85nt-95nt) GAGAGAGCTTT Ex5(95nt-105nt) CCTGGATTCAAACCTGCACC Ex5(180nt-198nt)
 II AGCCTGACCGT Ex4(85nt-95nt) GAGAGAGCTTT Ex5(95nt-105nt) CCTGGATTCAAACCTGCACC Ex5(180nt-198nt)
 III AGCCTGACCGT Ex4(85nt-95nt) GAGGAGCTTT Ex5(95nt-105nt) CCTGGTTCAAACCTGCACC Ex5(180nt-198nt)
 IV AGCCTAACCGT Ex4(85nt-95nt) GAGAGAGCTTT Ex5(95nt-105nt) CCTGGATTCAAACCTGCACC Ex5(180nt-198nt)
 I VLIKGLDQGPNMDENDSLPLPFEDFYLTMGESSLRE^SFRLLACFKKDMHKVETYL^RVANCRRSL^DSN Aa(140nt-207nt)
 II VLIKGLDQGPNMDENDSLPLPFEDFYLTMGESSLRE^SFRLLACFKKDMHKVETYL^RVANCRRSL^DSN Aa(140nt-207nt)
 III VLIKGLDQGPNMDENDSLPLPFEDFYLTMGESSLRE^SFRLLACFKKDMHKVETYL^RVANCRRSL^GSN Aa(140nt-207nt)
 IV VLIKGLDQGPNMDENDSLPLPFEDFYLTMGESSLRE^SFRLLACFKKDMHKVETYL^RVANCRRSL^DSN Aa(140nt-207nt)

图3 外显子及推导氨基酸序列分析比较

Fig.3 Comparison of exons and derived amino acid sequences
 注: 下划线处阴影加粗部分为差异碱基

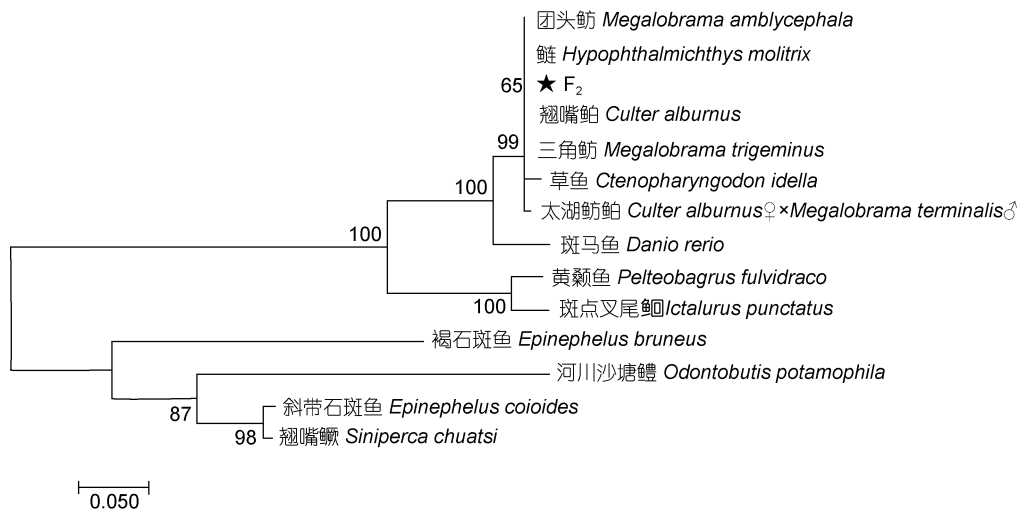


图4 4个群体与其他鱼类系统进化树

Fig.4 Phylogenetic trees of four species and other fishes

树与同源性分析结果具有最大的一致性, 证明了对F₂代进行深度遗传挖掘的重要性。

2.4 太湖鲂鮰 F₂代 GH 基因 SNPs 的筛选及基因结构

经太湖鲂鮰自交后的F₂代个体之间具有较为明显的性状分离现象(图5), 翘嘴鲂与三角鲂的各种形态特征组合在F₂代中得到了较为全面的遗传, BW、BL和BH等性状在F₂代群体中的差异较为显著。对10尾太湖鲂鮰F₂代个体(编号分别为S1—S10)的GH基因外显子序列对比发现, 第一外显子中并未发现突变位点; 在第二外显子中发现两个突变位点, 即在86nt以1:9的比例存在G/A转换、135nt以1:9的比例存在A/G转换, 值得注意的是, 135nt突变碱基改变了氨基酸序列, 由天门冬氨酸(D)转变为天门冬酰胺(N); 第三外显子存在一个突变位点, 在104nt以1:9的比例存在T/C转换, 致使苏氨酸(T)转变为异亮氨酸(I); 第

四外显子在90nt发现一个突变位点, 以3:7的比例存在A/G转换, 未改变其氨基酸序列; 第五外显子突变位点最多, 为3个, 分别在19nt以1:9的比例存在T/C转换, 在238nt以4:6的比例存在A/C转换, 在247nt以4:6的比例存在T/G转换, 其中19nt突变碱基改变了氨基酸序列, 即脯氨酸(P)转变为丝氨酸(S)。差异氨基酸序列见图6。

根据序列长度及位点的变化情况, 制作了太湖鲂鮰F₂代基因结构的变异图(图7)。由图可知, 经自交后的F₂代个体在GH基因编码区序列长度和核苷酸发生了显著的变化, 从序列长度来看, 只有3个个体与其亲本相近, 其余7个个体序列长度均遗传于祖父母本, 这与2.3所描述结果具有一致性; 从编码区核苷酸角度出发, F₂代发生突变的7处位点与其祖父母本在同一区域同一位置均发生变异, 因此太湖鲂鮰

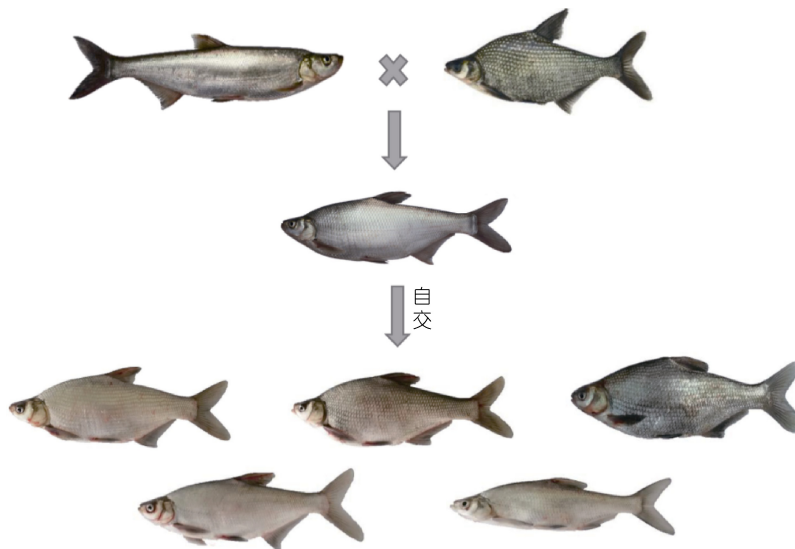


图 5 F₂ 代群体性状分离图

Fig.5 Segregation diagram of F₂ generation population traits

- S1 KMINDFEDNLLPEERRQLSKIFPLSFCNSDSIEAPT**GKDE**TQKSSMLKLLRIS Aa(45nt-97nt)
- S2 KMINDFFEDNLLPEERRQLSKIFPLSFCNSDSIEAPT**GKDE**TQKSSMLKLLRIS Aa(45nt-97nt)
- S3 KMINDFFEDNLLPEERRQLSKIFPLSFCNSDSIEAPT**GKDE**IQKSSMLKLLRIS Aa(45nt-97nt)
- S4 KMINDFFEDNLLPEERRQLSKIFPLSFCNSDSIEAPT**GKDE**TQKSSMLKLLRIS Aa(45nt-97nt)
- S5 KMINDFFEDNLLPEERRQLSKIFPLSFCNSDSIEAPT**GKDE**TQKSSMLKLLRIS Aa(45nt-97nt)
- S6 KMINDFFEDNLLPEERRQLSKIFPLSFCNSDSIEAPT**GKDE**TQKSSMLKLLRIS Aa(45nt-97nt)
- S7 KMINDFFEDNLLPEERRQLSKIFPLSFCNSDSIEAPT**GKDE**TQKSSMLKLLRIS Aa(45nt-97nt)
- S8 KMINDFFEDNLLPEERRQLSKIFPLSFCNSDSIEAPT**GKDE**TQKSSMLKLLRIS Aa(45nt-97nt)
- S9 KMINDFFEDNLLPEERRQLSKIFPLSFCNSDSIEAPT**GKDE**TQKSSMLKLLRIS Aa(45nt-97nt)
- S10KMINDFFEDNLLPEERRQLSKIFPLSFCNSDSIEAPT**GKDE**TQKSSMLKLLRIS Aa(45nt-97nt)
- S1 FRLIESWEPFSQTL**S**GAISNSLTVGNPNQITEKLADLKVGISVLIKGCLDG**Q**P Aa(98nt-150nt)
- S2 FRLIESWEPFSQTL**S**GAISNSLTVGNPNQITEKLADLKVGISVLIKGCLDG**Q**P Aa(98nt-150nt)
- S3 FRLIESWEPFSQTL**S**GAISNSLTVGNPNQITEKLADLKVGISVLIKGCLDG**Q**P Aa(98nt-150nt)
- S4 FRLIESWEPFSQTL**S**GAISNSLTVGNPNQITEKLADLKVGISVLIKGCLDG**Q**P Aa(98nt-150nt)
- S5 FRLIESWEPFSQTL**S**GAISNSLTVGNPNQITEKLADLKVGISVLIKGCLDG**Q**P Aa(98nt-150nt)
- S6 FRLIESWEPFSQTL**S**GAISNSLTVGNPNQITEKLADLKVGISVLIKGCLDG**Q**P Aa(98nt-150nt)
- S7 FRLIESWEPFSQTL**S**GAISNSLTVGNPNQITEKLADLKVGISVLIKGCLDG**Q**P Aa(98nt-150nt)
- S8 FRLIESWEPFSQTL**S**GAISNSLTVGNPNQITEKLADLKVGISVLIKGCLDG**Q**S Aa(98nt-150nt)
- S9 FRLIESWEPFSQTL**S**GAISNSLTVGNPNQITEKLADLKVGISVLIKGCLDG**Q**P Aa(98nt-150nt)
- S10FRLIESWEPFSQTL**S**GAISNSLTVGNPNQITEKLADLKVGISVLIKGCLDG**Q**P Aa(98nt-150nt)

图 6 差异氨基酸序列对比

Fig.6 Comparison of different amino acid sequences

注: 下划线处阴影加粗部分为差异氨基酸

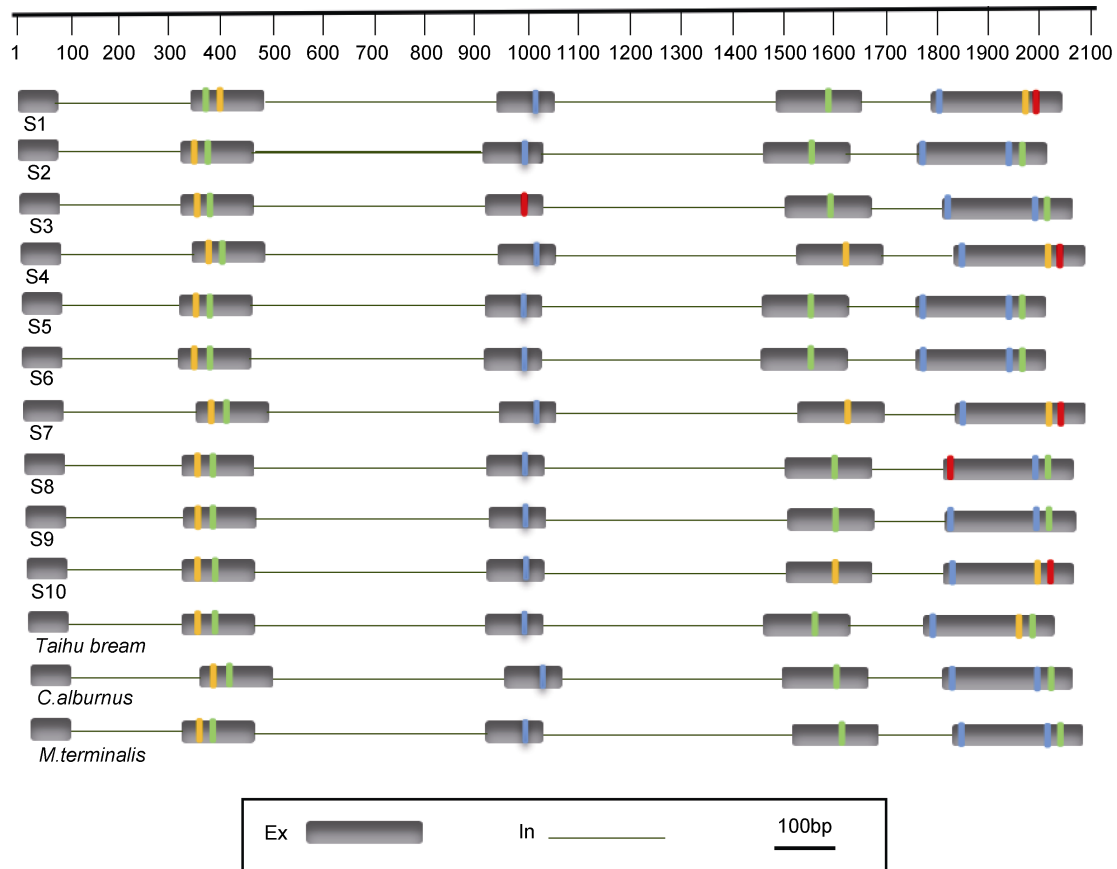


图7 GH 基因结构示意图

Fig.7 GH gene structure diagram

注: 黄色、红色、绿色、蓝色分别代表碱基 A、G、T、C

通过自交 GH 基因结构产生了显著性变化, 下一步对其进行蛋白质结构分析及组织表达分析, 探究基因结构的变异对蛋白质结构及组织表达所带来的影响。

2.5 蛋白质结构预测分析

经蛋白质结构预测发现, 10 尾 F_2 代 GH 基因均无跨膜结构, 且在成熟肽序列中均包含一个信号肽, 全部由 22 个氨基酸组成; 用 NCBI 的 Conserved Domains (CD-Search) 程序分析蛋白质结构域表明, 均具有 1 个保守的蛋白结构域, 且都位于第 28—209 位氨基酸处。由于 S1、S3 和 S8 个体分别在第 49、85 和 150 位氨基酸处发生了突变, 且突变位置均位于蛋白结构域, 可信度(E-value)较氨基酸序列未突变个体($6.54e^{-72}$)发生了变化, 其中 S3($7.65e^{-71}$)的 E-value 增加, 而 S1 和 S8 降低(均为 $4.48e^{-71}$)。蛋白质二级结构预测表明, 突变个体 S1、S3 和 S8 二级结构分别含 52.86%、53.81%和 52.38%的螺旋(helix), 43.81%、42.38%和 44.76%的环(loop), 3.33%、3.81%和 2.86%的链(strand), 而未突变个体二级结构含 52.86%的螺

旋、43.33%的环(loop)和 3.81%的链。由于氨基酸残基的变化导致蛋白质二级结构发生了改变。

2.6 太湖鲂鮈 F_2 代 GH 基因定量表达分析

实时定量分析结果表明, GH mRNA 在所检测的 10 个个体中均有表达, 在 S1 中的转录水平最高, 按表达量大小其次为 S2、S3、S10、S9, 在 S4—S8 中的表达量较低。我们将 10 个个体按体重分为两组(小个体组: S3, S4, S7, S9, S10; 大个体组: S1, S2, S5, S6, S8), 小个体组体重平均值 512.16g, 表达量 0.238; 大个体组平均值 825.08, 表达量 1.087, 结果表明体大的表达量高(图 8)。

3 讨论

GH 基因是鱼类脑垂体中分泌的促进生长的单一亚基的蛋白激素(杨学明等, 2008), 它参与鱼的生长代谢, 能够加速蛋白质合成和脂类降解(Sakamoto *et al*, 2006; Sciara *et al*, 2006), 温海深等(2002)报道了鲤生长激素(cGH)RIA 对鲤科鱼类的测定具有良好的

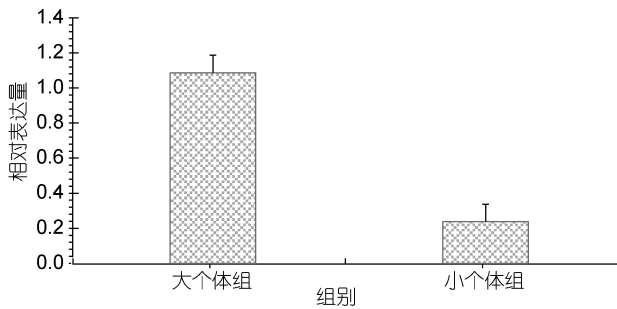


图 8 F₂ 代大、小个体组中 GH mRNA 的表达

Fig.8 Expression of GH mRNA in F₂ generation of large and small individual groups

效果。本实验中,太湖鲂鮰及其亲本和太湖鲂鮰 F₂ 代 GH 基因序列均包括五个外显子、四个内含子及 5' 侧翼区和 3' 侧翼区,其中外显子长度均具有一致性, F₂ 代在第四外显子中发生了 G/A 突变,但未改变其编码的氨基酸序列。相较于外显子,内含子存在相对较大的差异。F₂ 代内含子长度分别与翘嘴鲂、三角鲂和太湖鲂鮰具有 9、4 和 31bp 的差异,且通过序列对比发现, F₂ 代的第一、二内含子遗传较为稳定,主要来源于太湖鲂鮰及三角鲂,第三、四内含子具有较多的碱基突变或缺失,具有不稳定遗传。值得注意的是,在 5' 侧翼区,太湖鲂鮰 F₂ 代与祖父本三角鲂在同一位置存在(AAT)₇ 的微卫星序列,而太湖鲂鮰与母本翘嘴鲂在相同位置均存在(AAT)₈ 的微卫星序列;在 3' 侧翼区,太湖鲂鮰 F₂ 代、太湖鲂鮰与父本三角鲂一致,在相同位置存在(TTC)₆T(TAA)₆ 的微卫星序列,而母本翘嘴鲂存在(TTC)₅T(TAA)₈ 的微卫星序列。预测太湖鲂鮰 F₂ 代 GH 基因位于两端的微卫星位点具有多态性,与生长相关性状之间的关联性较强。

本研究对 3 个目 14 种鱼类进行同源性分析及系统发育分析发现,两者结果具有最高的一致性,符合常规生物学上物种亲缘关系的远近程度, F₂ 代与祖父母的遗传距离稍近于亲本。由于第五外显子出现了两处碱基突变, F₂ 代在保守的蛋白结构域与亲本具有两处氨基酸残基的差异(分别处于 176nt 和 205nt), 同时也在氨基酸序列 176nt 发现一处与鲢鱼之间的差异位点,且仅有一处; F₂ 代在信号肽和保守蛋白区域均发现一处与草鱼具有差异性的氨基酸残基,发生了由缬氨酸-甘氨酸、精氨酸-甘氨酸的变化;在与鲈形目和鲇形目的比较中, F₂ 代与其氨基酸序列的差异主要集中在保守蛋白区域 109nt 至 184nt 之间,相似性较低,特别是与河川沙塘鳢仅有 43.60% 的相似度。该结果对 F₂ 代进行深度遗传挖掘具有重要意义。

陈雪峰(2010)等采用了无需特殊仪器的 Tetra-primers ARMS 与 PCR-RFLP 法对吉富罗非鱼 IGF2 基因进行了 SNPs 的检测,共得到 11 处 SNP 位点, 8 处位于内含子中, 3 处位于外显子中, 且得到内含子 SNPs 的突变率高于外显子的结论。本文通过筛选得到体型具有显著差异的太湖鲂鮰 F₂ 代共 10 尾, 对其 GH 基因外显子序列进行分析。结果显示, 在第二、三、四和五外显子中共得到 7 处 SNP 位点, 其中位于第二、三和五外显子中的 3 处突变位点使氨基酸残基发生变异, 导致 S1、S3 和 S8 个体发生了序列变化。根据其变化情况, 我们对 GH 基因蛋白质结构进行预测, 结果发现变异后的氨基酸残基改变了蛋白质二级结构, 使突变个体及未突变个体间产生不同的 helix、loop 和 strand 的组成占比。我们将 10 个 F₂ 代个体按体重分为小个体组和大个体组, 分析发现, GH mRNA 在大个体组中的转录水平较高, 主要是由于在大个体组中 S1 发生了氨基酸残基的变化(天冬氨酸—天冬酰胺), 导致其产生了最高的转录水平, 使大个体组整体转录水平升高, 该结果可为以后的分子辅助标记育种研究提供理论基础。本研究所取鱼类均为同一时期, 且所取组织相同, 所得结论具有一定得参考意义和理论价值。

4 结论

本研究利用分子生物学技术, 克隆获得三角鲂、太湖鲂鮰、太湖鲂鮰 F₂ 代的 GH 基因全长序列, 从分子角度分析其遗传差异性; 单独对太湖鲂鮰 F₂ 代 GH 基因进行 SNPs 的筛选以及对 GH 基因蛋白质结构进行了预测, 并进行了组织表达分析, 结果表明太湖鲂鮰自交后引起的氨基酸残基的改变导致蛋白质二级结构发生改变及基因表达产生显著变化。

参 考 文 献

- 王文君, 陈克飞, 任 军等, 2003. 中外不同猪品种生长激素基因遗传多态性检测. 农业生物技术学报, 11(1): 103—104
- 刘士力, 张爱菊, 练青平等, 2016. 泥鳅和大鳞副泥鳅细胞色素 P450c17- I (CYP17- I) 基因的克隆及组织表达分析. 浙江农业学报, 28(11): 1853—1861
- 刘士力, 贾永义, 刘加林等, 2020. 翘嘴鲂两种生长激素受体基因结构及微卫星多态性与生长性状的相关性. 水产学报, 44(6): 894—906
- 刘士力, 贾永义, 蒋文枰等, 2017. 翘嘴鲂生长激素(GH)基因与侧翼区的克隆及分析. 浙江农业学报, 29(8): 1281—1289
- 刘士力, 蒋文枰, 程 顺等, 2019. 翘嘴鲂生长激素基因侧翼

- 2 个微卫星位点与生长性状的关联分析. 浙江农业学报, 31(1): 62—68
- 刘加林, 刘士力, 蒋文桦等, 2019. 河川沙塘鳢 *GH* 基因及侧翼的克隆与生物信息学分析. 浙江农业学报, 31(9): 1461—1470
- 严美姣, 吴旭, 李钟杰, 2010. 暗纹东方鲀生长激素基因克隆与同源性分析. 江西农业大学学报, 32(2): 236—242
- 杨学成, 杨志兴, 刘伟民等, 1991. 大麻哈鱼基因文库的构建及其生长激素基因的克隆. 生物技术, 1(1): 13—16
- 杨学明, 张立, 黄光华等, 2008. 革胡子鲶生长激素基因克隆与序列同源分析. 西南农业学报, 21(2): 483—486
- 宋成义, 经荣斌, 2000. 畜禽 GH 基因多态性与生产性能相关研究进展. 国外畜牧科技, 27(2): 25—27
- 陈雪峰, 杨国梁, 俞菊华等, 2010. 吉富罗非鱼 *IGF2* 基因分离及其单核苷酸多态性与体型、增重相关性. 动物学杂志, 45(2): 107—114
- 姜建湖, 沈斌乾, 陈建明等, 2019. “太湖鲢鲈”及其亲本肌肉营养成分的分析与评价. 水生生物学报, 43(2): 388—394
- 鲁双庆, 刘峰, 刘臻等, 2008. 三种鳊鱼 (*Siniperca*) 生长激素基因的克隆及序列比较. 海洋与湖沼, 39(4): 354—361
- 温海深, 林浩然, 肖东等, 2002. 野生鲢鱼生长激素分泌的季节变化及其神经内分泌调控. 动物学报, 48(2): 213—220
- Almuly R, Cavari B, Ferstman H *et al*, 2000. Genomic structure and sequence of the gilthead seabream (*Sparus aurata*) growth hormone-encoding gene: Identification of minisatellite polymorphism in intron I. *Genome*, 43(5): 836—845
- Ho K K, Wong M W, Chan A P Y, 1991. Cloning and sequencing of the grass carp (*Ctenopharyngodon idellus*) growth hormone gene. *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression*, 1090(2): 245—248
- Hong Y H, Scharf M, 1993. Sequence of the growth hormone (GH) gene from the silver carp (*Hypophthalmichthys molitrix*) and evolution of GH genes in vertebrates. *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression*, 1174(3): 285—288
- Livak K J, Schmittgen T D, 2001. Analysis of relative gene expression data using Real-Time Quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods*, 25(4): 402—408
- Ma Q, Liu S F, Zhuang Z M *et al*, 2012. Genomic structure, polymorphism and expression analysis of the growth hormone (GH) gene in female and male half-smooth tongue sole (*Cynoglossus semilaevis*). *Gene*, 493(1): 92—104
- Møller N, Copeland K C, Nair K S, 2007. Growth hormone effects on protein metabolism. *Endocrinology and Metabolism Clinics of North America*, 36(1): 89—100
- Sakamoto T, McCormick S D, 2006. Prolactin and growth hormone in fish osmoregulation. *General and Comparative Endocrinology*, 147(1): 24—30
- Sciara A A, Rubiolo J A, Somoza G M *et al*, 2006. Molecular cloning, expression and immunological characterization of pejerrey (*Odontesthes bonariensis*) growth hormone. *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology*, 142(3—4): 284—292
- Tang Y, Lin C M, Chen T T *et al*, 1993. Structure of the channel catfish (*Ictalurus punctatus*) growth hormone gene and its evolutionary implications. *Molecular Marine Biology and Biotechnology*, 2(4): 198—206

STRUCTURE, PHYLOGENY AND TISSUE DISTRIBUTION OF GH IN *CULTER ALBURNUS*♀×*MEGALOBrama TERMINALIS*♂

LIU Jia-Lin^{1,2}, JIA Yong-Yi¹, LIU Shi-Li¹, ZHENG Jian-Bo¹, CHI Mei-Li¹, CHENG Shun¹, LI Fei¹,
YIN Shao-Wu², GU Zhi-Min¹

(1. Key Laboratory of Healthy Freshwater Aquaculture of Ministry of Agriculture and Rural Development / Key Laboratory of Freshwater Aquatic Animal Genetic and Breeding of Zhejiang Province, Zhejiang Institute of Freshwater Fisheries, Huzhou 313001, China; 2. School of Life Science, Nanjing Normal University, Nanjing 210023, China)

Abstract Using polymerase chain reaction (PCR) and T-A cloning technology, we successfully obtained complete sequences of growth hormone (GH) gene with complete open reading frames from complete genomic DNA from three fish populations *Megalobrama terminalis*, *Culter alburnus*♀ × *Megalobrama terminalis*♂ (*Ca*×*Mt*), and the F₂ of *Ca*×*Mt*. The total length of the sequence was 6009, 6042, and 6058bp, respectively, and the length of the transcription unit was 2049, 2014, and 2045bp, respectively. The GH gene sequence of the three populations includes five exons, four introns, 5'UTR (untranslated region) and 3'UTR. All of them encoded 633bp amino acid sequence and 210 amino acids. The homology and phylogeny analysis on 14 fish species from 3 orders showed that the genetic distance between the F₂ and its grandparents was closer than that of their parents. Seven SNPs (single nucleotide polymorphism) were obtained from 10 individuals of the F₂, from which three mutations were observed causing significant changes in amino acid sequence. The analysis of protein structure showed that the mutated amino acid residue resulted in the change of protein secondary structure. We divided the 10 individuals into small and large body groups according to their body weight. Through tissue expression analysis, we found that the individuals in the large group had higher transcription level, and further analysis found that their amino acid residues were different.

Key words *Culter alburnus*♀×*Megalobrama terminalis*♂; intergeneric cross; growth hormone; cloning; SNPs; expression analysis