

DOI: 10.12264/JFSC2024-0040

## 鳊胰蛋白酶基因进化及启动子功能分析

胡子俊<sup>1,2</sup>, 卢宏亮<sup>1,2</sup>, 赵晨雨<sup>1</sup>, 何珊<sup>1,2</sup>

1. 华中农业大学水产学院, 湖北 武汉 430070;

2. 长江经济带大宗水生生物产业绿色发展教育部工程研究中心, 湖北 武汉 430070

**摘要:** 为探究胰蛋白酶与鳊(*Siniperca chuatsi*)开口时期活鱼食性之间的联系及鳊关键胰蛋白酶基因的转录调控机制, 利用生物信息学方法对鳊等 11 种不同食性的鱼类胰蛋白酶基因进行数量、序列特征及系统发育分析, 并结合 qPCR、双荧光素酶报告及胰蛋白酶活性检测分析鳊关键胰蛋白酶基因的转录调控机制, 及其与鳊早期胰蛋白酶活性间的联系。结果显示, 鳊基因组中共存在分布在 4 条染色体的 6 个基因座上的 13 个胰蛋白酶基因拷贝, 多于杂食性及草食性鱼类, 表明肉食性鱼类胰蛋白酶基因的扩增与其食物环境需求的高胰蛋白酶相适应。鳊在与人类关键胰蛋白酶基因 PRSS1 直系同源的基因所处关键基因座 2 (包含 SC7-LG17-21898 及 SC7-LG17-21899)和基因座 3 (包含 SC7-LG17-21428、SC7-LG17-21429-1、SC7-LG17-21429-2、SC7-LG17-21430-1 及 SC7-LG17-21430-2)上出现基因扩增, 且基因座 3 上扩增更为显著, 其中鳊胰蛋白酶基因 SC7-LG17-21430-2 亚型的表达水平显著高于其他亚型, 推测其为鳊关键胰蛋白酶基因亚型。双荧光素酶报告结果显示, 构建的 SC7-LG17-21430-2 亚型启动区域的 3 个逐段缺失片段启动子活性均存在显著差异, 且 -593 bp~+20 bp 区域启动子活性最高, 推测为启动子核心区域。此外, 在该核心区域, 发现 PDX1 这一与早期胰腺发育相关且在鳊仔鱼期几乎不表达的转录因子结合位点, 同时 PDX1 结合位点突变后导致核心区域启动子活性显著下降。胰蛋白酶活性检测显示 3 dph 鳊仔鱼的胰蛋白酶活性显著性低于同时期斑马鱼(*Danio rerio*)仔鱼, 表明转录因子 *pdx1* 的低表达可能导致了鳊胰蛋白酶基因关键亚型 SC7-LG17-21430-2 的低表达以及胰蛋白酶的低酶活, 因而鳊通过摄食活鱼苗来补偿内源性胰蛋白酶低活性, 形成自开口起只食活鱼的奇特食性, 为鳊活鱼食性形成机制的深入解析提供了理论依据。

**关键词:** 鳊; 胰蛋白酶; 食性; 酶活性; 启动子活性

中图分类号: S917

文献标志码: A

文章编号: 1005-8737-(2024)05-0524-13

鱼类的食性与消化腺中的消化酶活性有关, 肠道各种消化酶的活性决定了鱼类能否有效地消化其摄入的食物, 这种适应性使得鱼类在摄食中倾向于易消化的食物, 以更有效地获得所需营养<sup>[1-3]</sup>。肉食性鱼类通常表现出高胰蛋白酶活性, 并且胰蛋白酶活性与肉食性鱼类的摄食行为相互影响<sup>[4-7]</sup>。

胰蛋白酶大约由 245 个氨基酸组成, 在结构上具有一个信号肽序列和一个成熟肽序列, 其三级结构由两个  $\beta$  折叠和 3 个  $\alpha$  螺旋组成, 具有一

个活性位点, 该活性位点包括 3 个氨基酸残基: 组氨酸(His)、丝氨酸(Ser)和天冬氨酸(Asp), 其中 Ser 是胰蛋白酶的主要催化残基<sup>[8]</sup>。已有研究表明, 人类丝氨酸蛋白酶 1 基因(Protease, Serine 1, PRSS1)编码阳离子胰蛋白酶原, 是人类最重要的胰蛋白酶基因<sup>[9]</sup>。鱼类中存在着与人 PRSS1 基因直系同源的关键胰蛋白酶基因, 如斑马鱼(*Danio rerio*)基因组中的 *prss1*、*prss59.1* 和 *prss59.2*<sup>[10]</sup>。但不同食性鱼类中, 胰蛋白酶基因的表达模式存在差异, 在杂食性斑马鱼肠道中 *prss59.1* 和

收稿日期: 2024-01-26; 修订日期: 2024-03-23.

基金项目: 国家自然科学基金项目(32172951); 设施渔业教育部重点实验(大连海洋大学)开放课题项目(202317).

作者简介: 胡子俊(1999-), 男, 硕士研究生, 研究方向为水生动物营养与生理. E-mail: huzijun19990728@163.com

通信作者: 何珊, 副教授, 研究方向为鱼类营养调控与饲料利用遗传改良. E-mail: heshan@mail.hzau.edu.cn

*prss59.2* 的表达水平较为接近, 且均高于 *prss1*<sup>[11]</sup>。肉食性的大西洋鳕鱼(*Gadus morhua*)中, 胰蛋白酶 Y 与胰蛋白酶 I 与斑马鱼 *prss59.1* 和 *prss59.2* 同源, 但胰蛋白酶 Y mRNA 的表达水平不足胰蛋白酶 I 的千分之一<sup>[12]</sup>, 同为肉食性的塞内加尔比目鱼(*Solea senegalensis* Kaup)中的同源胰蛋白酶基因亚型的表达规律与大西洋鳕鱼几乎一致<sup>[13]</sup>。

鱼类早期发育阶段, 消化酶的活性至关重要<sup>[14-15]</sup>, 不仅直接影响着仔鱼的存活和发育<sup>[16]</sup>, 还与鱼类的摄食密切相关<sup>[17]</sup>。鱼类在早期发育阶段依赖于胰腺酶类对食物中的营养素进行消化水解<sup>[18-19]</sup>, 仔鱼的快速生长需要大量的蛋白质, 因此蛋白质消化能力对仔鱼尤为重要<sup>[20]</sup>。胰腺酶类中的胰蛋白酶是胰腺分泌的关键碱性蛋白酶<sup>[21]</sup>, 其不仅是鱼类早期发育过程中水解蛋白质的主要消化酶, 还能激活其他蛋白水解酶<sup>[22]</sup>。

鳊(*Siniperca chuatsi*)是我国名特优淡水鱼类, 其食性奇特, 自开口摄食起就以活鱼为食。鳊作为典型的肉食性鱼类, 其基因组中的胰蛋白酶基因数目及关键亚型的表达模式尚不明确。此外, 鳊开口仅食活鱼的特殊食性可能与其开口时期的胰蛋白酶表达活性有关。然而, 鳊开口时期的活鱼食性与胰蛋白酶转录调控之间的联系有待解析。本研究以鳊为主要研究对象, 通过挖掘鳊、尼罗罗非鱼及草鱼等不同食性的鱼类基因组中胰蛋白酶基因的数目, 判断鱼类食性与胰蛋白酶基因数目的关系, 并通过实时荧光定量 PCR (quantitative real-time PCR, qPCR)及双荧光素酶报告确定关键胰蛋白亚型及其核心调控区域, 以期为鱼类食性形成机制及鳊活鱼食性决定机制的解析奠定理论基础。

## 1 材料与方法

### 1.1 实验用鱼及实验条件

实验所用鳊[体重(721.73±99.23) g; 体长(33.57±5.15) cm]来自华中农业大学鳊鱼研究中心; 斑马鱼[体重(0.56±0.12) g; 体长(2.48±0.15) cm]来自中国科学院水生生物研究所。通过人工繁殖获得野生型鳊及斑马鱼受精卵, 于 27 °C 恒温水

体中孵化受精卵, 并于孵化后第 3 天(3 dph)对鳊和斑马鱼仔鱼及成鱼全鱼进行取样, 在液氮中快速冷冻, 随后保存于-80 °C 超低温冰箱, 用于后续胰蛋白酶酶活测定及基因表达分析。

### 1.2 胰蛋白酶基因挖掘

从 NCBI 中获取团头鲂(*Megalobrama amblycephala*)、草鱼(*Ctenopharyngodon idella*)、青鳉(*Oryzias latipes*)、斑马鱼、红鳍东方鲀(*Takifugu rubripes*)、斑点叉尾鲷(*Ictalurus punctatus*)、尼罗罗非鱼(*Oreochromis niloticus*)、白斑狗鱼(*Esox lucius*)、梭鲈(*Sander lucioperca*)和尖吻鲈(*Lates calcarifer*)的基因组文件、蛋白质序列文件以及注释文件, 从本实验室所有鳊鱼基因组数据库(<http://genomes.igb-berlin.de/>)下载鳊的基因组文件、蛋白质序列文件以及注释文件。从 Pfam 数据库(<http://pfam.xfam.org/>)中查找获取 *trypsin* 基因的隐马尔可夫模型。利用 *hmmer search* 在各鱼类的基因组文件中筛选质量较高的基因(*E-value*<1×10<sup>-20</sup>), 并使用 TBtools 的 Sequence extract 工具利用获取的候选基因编号并进行氨基酸序列的提取, 从而获得备选的基因家族成员的合集。结合胰蛋白酶的序列特征及 NR (non-redundant)库注释, 完成对备选基因的筛选。

### 1.3 系统发育树构建、Motif 分析、序列差异分析及共线性分析

对团头鲂、草鱼、青鳉、斑马鱼、红鳍东方鲀、斑点叉尾鲷、尼罗罗非鱼、白斑狗鱼、梭鲈、尖吻鲈和鳊的胰蛋白酶基因的氨基酸序列进行分析, 利用 MEGA 5.0 进行多序列比对, 使用最大似然法(maximum likelihood, ML)构建系统发育树, 迭代次数 bootstrap=1000, 其余参数设置为默认值<sup>[23]</sup>。再通过在线软件 MEME (<https://meme-suite.org/meme/tools/meme>)识别其氨基酸序列中的保守结构域<sup>[24]</sup>, 使用 TBtools 使其可视化并联合系统发育树进行分析。同时使用 Megalign 对人 *PRSSI* 基因及鳊胰蛋白酶基因的氨基酸序列进行序列差异系数及同源百分比分析, 最后利用基因组注释文件, 通过 NCBI genome date viewer 功能进行基因组共线性分析, 确定备选基因在染色体上的排列规律。

1.4 鳊胰蛋白酶基因关键亚型筛选

利用 Primer premier 5.0 软件设计鳊关键座位上的胰蛋白酶基因特异性引物(表 1), 对鳊仔鱼、成鱼进行 qPCR, 检测鳊在仔鱼、成鱼期不同胰蛋白酶基因亚型的相对表达量, 以确定鳊关键胰蛋白酶基因亚型。每个样本均用 1 mL TRIzol (TaKaRa, 日本)提取总 RNA。利用 HiScript® II

Reverse Transcriptase (Vazyme, 中国)获得 cDNA。使用 AceQ qPCR SYBR Green Master Mix (Vazyme, 中国)对每个样品进行 3 个技术重复的实时荧光定量 PCR (生物学重复  $n=6$ )。3 步 qPCR 程序包括 95 °C (5 min)酶激活步骤和 95 °C (10 s), 进行 40 个循环, 每个引物组的退火温度(30 s)和 72 °C (30 s), 应用核糖体蛋白 L13A (*rpl13a*)作为内参基因。

表 1 实时荧光定量 PCR 引物信息  
Tab. 1 Primer information of qPCR

| 用途<br>usage                 | 名称<br>name       | 引物序列(5'-3')<br>sequence of primers (5'-3')                                                  | 退火温度/°C<br>annealing<br>temperature | 片段长度/bp<br>size of<br>segment |
|-----------------------------|------------------|---------------------------------------------------------------------------------------------|-------------------------------------|-------------------------------|
| qPCR                        | SC7-LG17-21428   | F: CTCCGTCTGCAGTGTGTTGGA<br>R: AGGGCTACCAGAGTCACCAT                                         | 60                                  | 144                           |
|                             | SC7-LG17-21429-1 | F: TGTCACCGTGAACCAGTACG<br>R: ATGGGGACTTCCACACACTG                                          | 60                                  | 147                           |
|                             | SC7-LG17-21429-2 | F: GAGCACAACATCAAGGCCAC<br>R: ACAGGCTGCACATACTGGTT                                          | 58                                  | 156                           |
|                             | SC7-LG17-21430-1 | F: TGTTCTGCGCTGGATACCTG<br>R: TTTCAGCACATCCGTAGCCC                                          | 58                                  | 120                           |
|                             | SC7-LG17-21430-2 | F: TCCCCATCCTGTCTTTCAGC<br>R: CCAGCCATTTCGTTGAACAGG                                         | 60                                  | 231                           |
|                             | SC7-LG17-21898   | F: TTGTTGGAGGCTACGAGTGC<br>R: TGGGGTGACGGATGACCTTA                                          | 60                                  | 225                           |
|                             | SC7-LG17-21899   | F: CCTATCCTGAGCGACAGCAG<br>R: CCATAACCCCAGGACACCAC                                          | 60                                  | 167                           |
|                             | <i>rpl13a</i>    | F: CACCCTATGACAAGAGGAAGC<br>R: TGTGCCAGACGCC CAAG                                           | 58                                  | 100                           |
|                             | $\beta$ -actin   | F: AGAGGGAAATCGTGCGTGAC<br>R: ATACCGAGGAAGGAAGGCTG                                          | 58                                  | 193                           |
|                             | <i>pdx1</i>      | F: CCTGACATTGCCCCCTACAG<br>R: CGTCCTCGTGCGTTTGT                                             | 60                                  | 245                           |
| 载体构建<br>vector construction | 21430-2-R        | CCGGAATGCCAAGCTTGGCTGTGTGAACCTGTCGAT                                                        | 60                                  |                               |
|                             | A-F              | CAGTACCGGAATGCCAAGCTTGGTGGCTGTGTGAACCTGTCG                                                  | 58                                  | 2189                          |
|                             | B-F              | GATAGGTACCGAGCTCTCACTCTGAGACCAGCAGAAC                                                       | 60                                  | 1866                          |
|                             | C-F              | GATAGGTACCGAGCTCTCCTGTCTCTCTCAGCTGTTT                                                       | 60                                  | 612                           |
| 点突变<br>point mutation       | DTB              | F: CATGTGTGGCCAGTAAGA<br>AACATGTAATAATGTAATA<br>R: GTTCTTACTGGCCACAC<br>ATGACTTGTCACATAGGAT | 60                                  | 4818                          |

1.5 关键胰蛋白酶基因亚型启动子表达载体构建

1.5.1 关键胰蛋白酶基因启动子不同片段克隆  
将关键胰蛋白酶基因起始密码子上游总计 2200 bp 的序列提交至 BDGP ([https://www.fruitfly.org/seq\\_tools/promoter.html](https://www.fruitfly.org/seq_tools/promoter.html))进行转录起始位点预测。依据转录起始位点预测结果, 利用 NCBI primer

blast 对所提交序列分段, 并设计特异性扩增引物 (引物两端加上酶切位点 Hind III 和 Sac I 的同源臂), 引物信息见表 1。利用 Hind III 和 Sac I 对 pGL3-basic 载体进行双酶切, 使用 Hieff Clone® Plus Multi One Step Cloning Kit (Yeasen, 中国), 利用同源重组反应构建重组载体。随后在含氨苄

抗性的 LB 培养基上涂板培养, 挑取单菌落后, 利用 pGL3-basic 载体通用引物 (RVprimer3 和 GLprimer2) 进行菌落阳性鉴定。测序正确的阳性质粒分别命名为 pGL3-A (-2169 bp~+20 bp)、pGL3-B (-1846 bp~+20 bp) 和 pGL3-C (-593 bp~+20 bp)。

**1.5.2 双荧光素酶报告分析及启动子核心区关键转录因子结合位点预测** 将状态良好的 HEK 293T 细胞接种到 24 孔细胞板中, 每孔  $2 \times 10^5$  细胞, 并加入 1 mL 含 5% FBS 的 DMEM 高糖培养基 (Gibco, 美国)。将 2  $\mu$ g 质粒 DNA、0.4  $\mu$ g pHRL-TK (内参质粒) 和 3~5  $\mu$ L Lipo293 (碧云天, 中国) 转染试剂分别加入 50  $\mu$ L opti-MEM 培养基中 (Gibco, 美国)。将上述溶液混合, 吹打均匀后加入细胞培养板中, 并在 37  $^{\circ}$ C、含 5% CO<sub>2</sub> 的细胞培养箱中培养。24 h 后, 利用 Dual Luciferase Reporter Gene Assay Kit (Yeasen, 中国) 裂解细胞, 并于多功能酶标仪 (PE Enspire) 上测定其荧光活性, 载体相对活性用萤火虫荧光素酶值/海肾荧光素酶值表示。通过在线软件 JASPAR (<http://jaspar.genereg.net>) 预测启动子核心区潜在的关键转录因子结合位点。

**1.5.3 关键转录因子结合位点点突变** 通过在线软件 JASPAR 选取合适的点突变位点, 并利用 Vazyme 点突变引物设计工具设计所需点突变引物 (<https://crm.vazyme.com/cetool/singlepoint.html>),

点突变见表 1。同源重组反应及双荧光素酶报告实验步骤同 1.5.1 及 1.5.2。

**1.5.4 3 dph 时期鳊关键转录因子基因相对表达水平** 实时荧光定量 PCR 操作同 1.4, 选取  $\beta$ -actin 为内参基因, 相关引物见表 1。

**1.6 鳊与斑马鱼仔鱼胰蛋白酶活力对比**

准确称取 3 dph 的鳊及斑马鱼仔鱼, 按重量 (g): 体积 (mL)=1:9 的比例, 加入双蒸水进行匀浆, 利用胰蛋白酶测定试剂盒 (南京建成, 中国) 于多功能酶标仪 (Synergy2) 对鳊及斑马鱼仔鱼的胰蛋白酶活力进行测定。

**2 结果与分析**

**2.1 不同食性鱼类胰蛋白酶基因数目**

隐马尔可夫模型同源比结果表明, 草食性的团头鲂和草鱼基因组中胰蛋白酶基因数目分别为 11 个和 9 个; 杂食性的青鳉、斑马鱼、红鳍东方鲀、斑点叉尾鮰及尼罗罗非鱼分别为 9 个、10 个、7 个、8 个和 8 个; 肉食性的白斑狗鱼、梭鲈、尖吻鲈和鳊分别为 14 个、13 个、13 个和 13 个 (图 1)。肉食性鱼类的胰蛋白酶基因数量高于杂食性和草食性鱼类。其中, 鳊的胰蛋白酶基因为 SC7-LG17-21428、SC7-LG17-21429-1、SC7-LG17-21429-2、SC7-LG17-21430-1、SC7-LG17-21430-2、SC7-LG17-21898、SC7-LG17-21899、SC7-LG17-

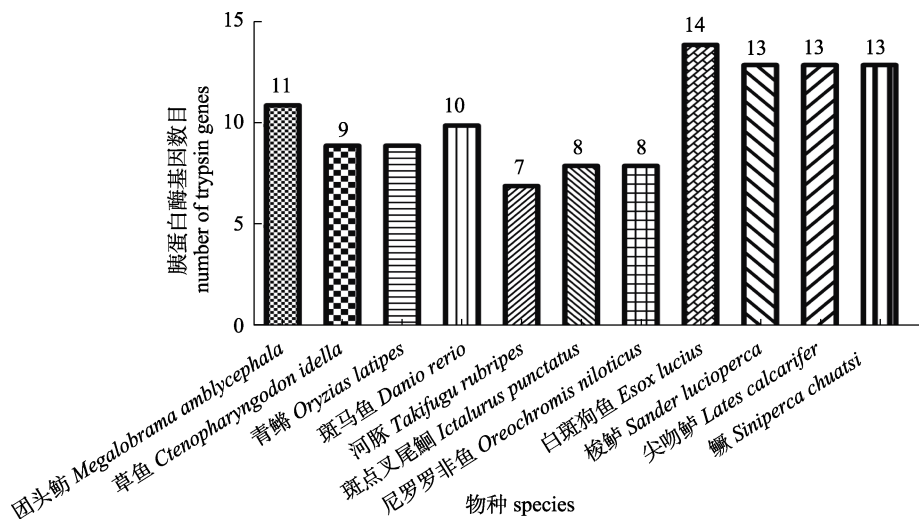


图 1 鳊等 11 种鱼类胰蛋白酶基因数目  
Fig. 1 Number of trypsin genes in 11 kinds of fish species including *Siniperca chuatsi*

21883、SC7-LG01-1496、SC7-LG04-05416、SC7-LG04-05415、SC7-LG10-13115 及 SC7-LG10-13116, 这 13 个基因分布在鳊的 1 号、4 号、10 号以及 17 号染色体的 6 个基因座位上。

## 2.2 系统发育树构建、Motif 分析、序列差异分析及共线性分析结果

根据不同食性鱼类的胰蛋白酶基因系统发育树聚类结果可以看出, 胰蛋白酶基因家族可以划分成七大分支, 每一分支都存在肉食性鱼类和杂食性鱼类, 而草食性鱼类在一号和四号分支中未出现。肉食性鱼类在五号、六号分支(斑马鱼 *prss1*、*prss59.1* 及 *prss59.2* 基因所在分支)及二号、四号分支中均占据数量优势, 但在七号分支中, 草食性和杂食性鱼类占据优势, 在一号及三号分支中不同食性鱼类的比例无较大差异(图 2)。利用 MEME 和 TBtools 对胰蛋白酶基因序列进行 Motif 和进化联合分析, 由图 3 可知, 鳊 13 个胰蛋白酶基因同有 5 个 Motif, 分别为 Motif 4、Motif 5、Motif 6、Motif 7 和 Motif 10。由基因组共线性分析发现, 在图 4 所示基因座 1 上, 肉食性的鳊相较于杂食性的斑马鱼和草食性的草鱼, 在 *has1* 与 *fpr1* 基因之间进化出了 SC7-LG17-21883 这一胰蛋白酶基因亚型。此外, 基因座 2 上, *Im: 7147486* 和 *si:ch211-261d7.3* 之间, 斑马鱼和草鱼仅有一个胰蛋白酶基因亚型(*prss1*), 而鳊存在两个胰蛋白酶基因亚型(SC-LG17-21898 和 SC-LG17-21899)。在基因座 3 上, 草鱼存在一个胰蛋白酶基因亚型(*prss59.1*), 斑马鱼存在两个胰蛋白酶基因亚型(*prss59.1* 和 *prss59.2*), 而鳊存在 5 个胰蛋白酶基因亚型(SC7-LG17-21428、SC7-LG17-21429-1、SC7-LG17-21429-2、SC7-LG17-21430-1 和 SC7-LG17-21430-2)。同时发现在鱼类中, 胰蛋白酶基因所处关键基因座(基因座 2 和基因座 3)上, 肉食性鱼类的胰蛋白酶基因数目相较于杂食性和草食性鱼类出现了扩增的现象, 并且这种扩增现象在基因座 3 上更为明显(图 4)。多序列比对结果显示, 胰蛋白酶基因在进化过程中高度保守, 但鳊胰蛋白酶基因亚型 SC7-LG17-21898 在结构上缺失了关键的功能保守区域(图 5a), 且鳊 SC7-LG17-

21898 基因与人 *PRSSI* 基因序列差异系数最低, 同源百分比最高(图 5b)。

## 2.3 鳊关键胰蛋白酶基因亚型

qPCR 结果如图 6 所示, 在 3 dph 仔鱼及成鱼中, 鳊胰蛋白酶基因亚型 SC7-LG17-21430-2 的相对表达量均显著性高于其余亚型, 推测其为鳊关键胰蛋白酶基因亚型。并且鳊处于基因座 2 的胰蛋白酶基因(SC7-LG17-21428、SC7-LG17-21429-1、SC7-LG17-21429-2、SC7-LG17-21430-1 及 SC7-LG17-21430-2 表达水平整体上低于处于基因座 3 的胰蛋白酶基因(SC7-LG17-21898 及 SC7-LG17-21899)的表达水平。

## 2.4 关键基因亚型 SC7-LG17-21430-2 启动子活性检测

对构建的鳊胰蛋白酶基因关键亚型 SC7-LG17-21430-2 启动子区域重组质粒 pGL3-A (-2169 bp~+20 bp)、pGL3-B (-1846 bp~+20 bp)和 pGL3-C (-593 bp~+20 bp)进行琼脂糖凝胶电泳, 发现 PCR 产物条带与预期条带位置一致, 经测序与启动子序列一致, 表明启动子缺失片段重组质粒构建成功, 可用于启动子活性分析。结果显示, 重组质粒 pGL3-C 的荧光活性显著高于 pGL3-A 及 pGL3-B ( $P<0.05$ ), 说明 SC7-LG17-21430-2 基因的-593 bp~+20 bp 区域具有较强的启动子活性, 为启动子核心区(图 7a)。利用在线软件 JASPAR 在该启动子核心区域预测到胰腺发育相关转录因子 PDX1 的结合位点(图 7b)。将所预测到的 PDX1 结合位点 1 的第一个 T 碱基替换为 G 碱基, 突变示意图见图 7c, 且在发生碱基替换后, 突变型 pGL3-C 载体的相对荧光素酶活性显著性低于未突变 pGL3-C 载体(图 7d)。

## 2.5 3 dph 时期鳊 *pdx1* 相对表达水平

qPCR 结果如图 8 所示, 在 3 dph 时期, 鳊 *pdx1* 基因的相对表达水平极低, 仅为 0.000 4。

## 2.6 鳊与斑马鱼仔鱼开口时期胰蛋白酶活性

利用紫外分光光度计法测得鳊和斑马鱼胰蛋白酶活性结果如图 9 所示, 在 3 dph, 鳊仔鱼的胰蛋白酶活性显著性低于斑马鱼仔鱼的胰蛋白酶活性。

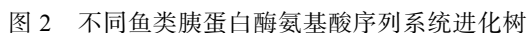


Fig. 2 Phylogenetic tree of trypsin amino acid sequence in different fish species

食性及摄食行为具有重要影响。已有研究表明,肉食性鱼类的肠道蛋白酶活性显著高于草食性鱼<sup>[27]</sup>,且鳊成鱼的胰蛋白酶活性已被证实显著高于草食性的草鱼及杂食性的昆明裂腹鱼(*Schizot*



*horax grahami*)<sup>[6]</sup>。“自适应调节假说”(adaptive modulation hypothesis, AMH)认为, 消化酶活性会随着基因拷贝数的增加而增加, 且与日粮中对应的底物数量有关<sup>[28-29]</sup>。在鱼类中, 有研究表明肉食性的紫灰翼鲷(*Anoplarchus purpureus*)基因组中胰蛋白酶基因的拷贝数多于草食性的猴面线鲷(*Cebidichthys violaceus*)<sup>[30]</sup>。本研究中同样发现以鳊为代表的肉食性鱼类胰蛋白酶基因数目显著性高于以尼罗罗非鱼为代表的杂食性鱼类及以草鱼为代表的草食性鱼类, 而杂食性鱼类与草食性鱼类之间胰蛋白酶基因拷贝数无较大差异。且相较于杂食性及草食性鱼类, 肉食性的鳊在基因座 1、基因座 2 及基因座 3 上均出现了胰蛋白酶基因扩增, 在演化过程中适应性地扩增了胰蛋白酶基因。基因扩增往往导致其所行使的功能增强以适应环境, 如鱼类基因组中嗅觉受体基因 *or* 的扩增能帮助鱼类识别水环境中更为复杂的气味分子, 促进鱼类向不同水生环境过渡<sup>[31]</sup>; 草鱼基因组中甜味受体基因 *t1r2* 发生了扩增以适应其日粮中的

高碳水化合物<sup>[32]</sup>, 这与“自适应调节假说”不谋而合, 因此, 肉食性鱼类胰蛋白酶基因的扩增与其食物环境需求的高胰蛋白酶活性息息相关。

经由多序列比对, 发现鳊位于基因座 2 的胰蛋白酶基因亚型 SC7-LG17-21898 在结构上缺失了关键的功能区域, 但 SC7-LG17-21898 基因在系统发育上却与人关键胰蛋白酶基因 *PRSSI* 同源性最高。由系统发育分析发现鳊 SC7-LG17-21898 和人 *PRSSI* 基因由共同祖先的同一个基因演化而来, 在鱼类食性演化的过程中, 肉食性鱼类的胰蛋白酶基因演化重心从基因座 2 转移到基因座 3, 位于基因座 2 的鳊胰蛋白酶基因 SC7-LG17-21898 和 SC7-LG17-21899 功能会被逐渐弱化, 针对鳊胰蛋白酶基因的 qPCR 结果验证了这一观点。SC7-LG17-21898 基因的功能结构的缺失可能源自于其功能的弱化。虽然 SC7-LG17-21898 基因其他区段的氨基酸序列和人 *PRSSI* 基因的氨基酸序列相似度更高, 同源性最强, 但处于关键位点的 SC7-LG17-21898 保守功能结构的缺失仍可能

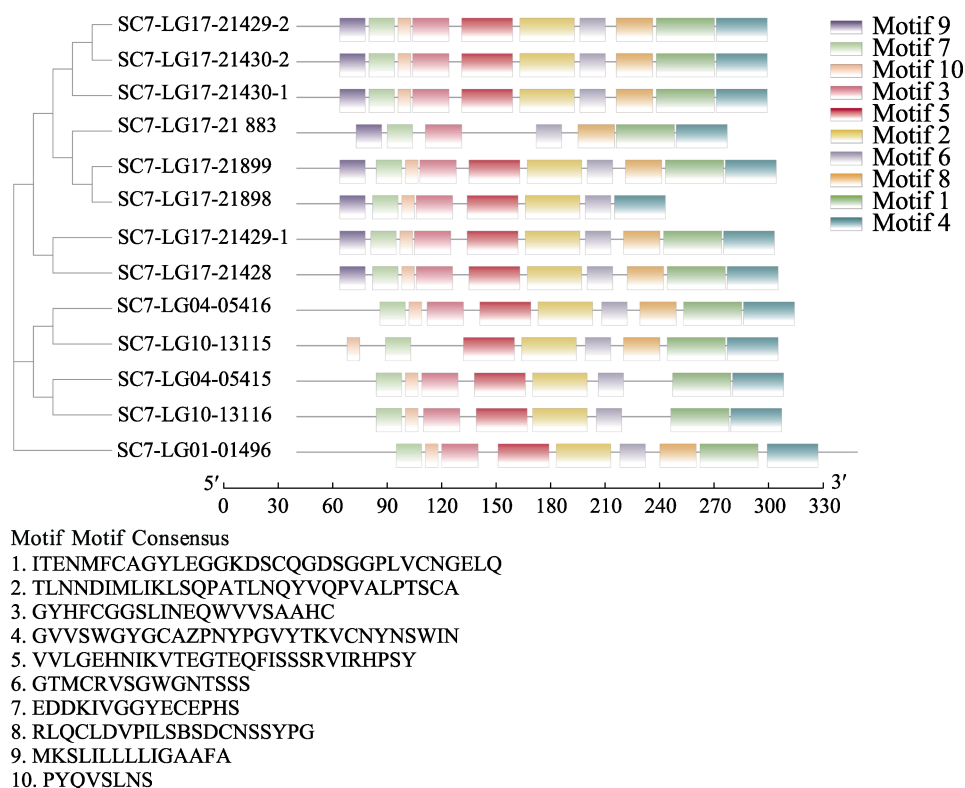


图 3 鳊胰蛋白酶基因氨基酸序列 Motif 分析

Fig. 3 Motif analysis in the amino acid sequence of trypsin genes in *Siniperca chuatsi*

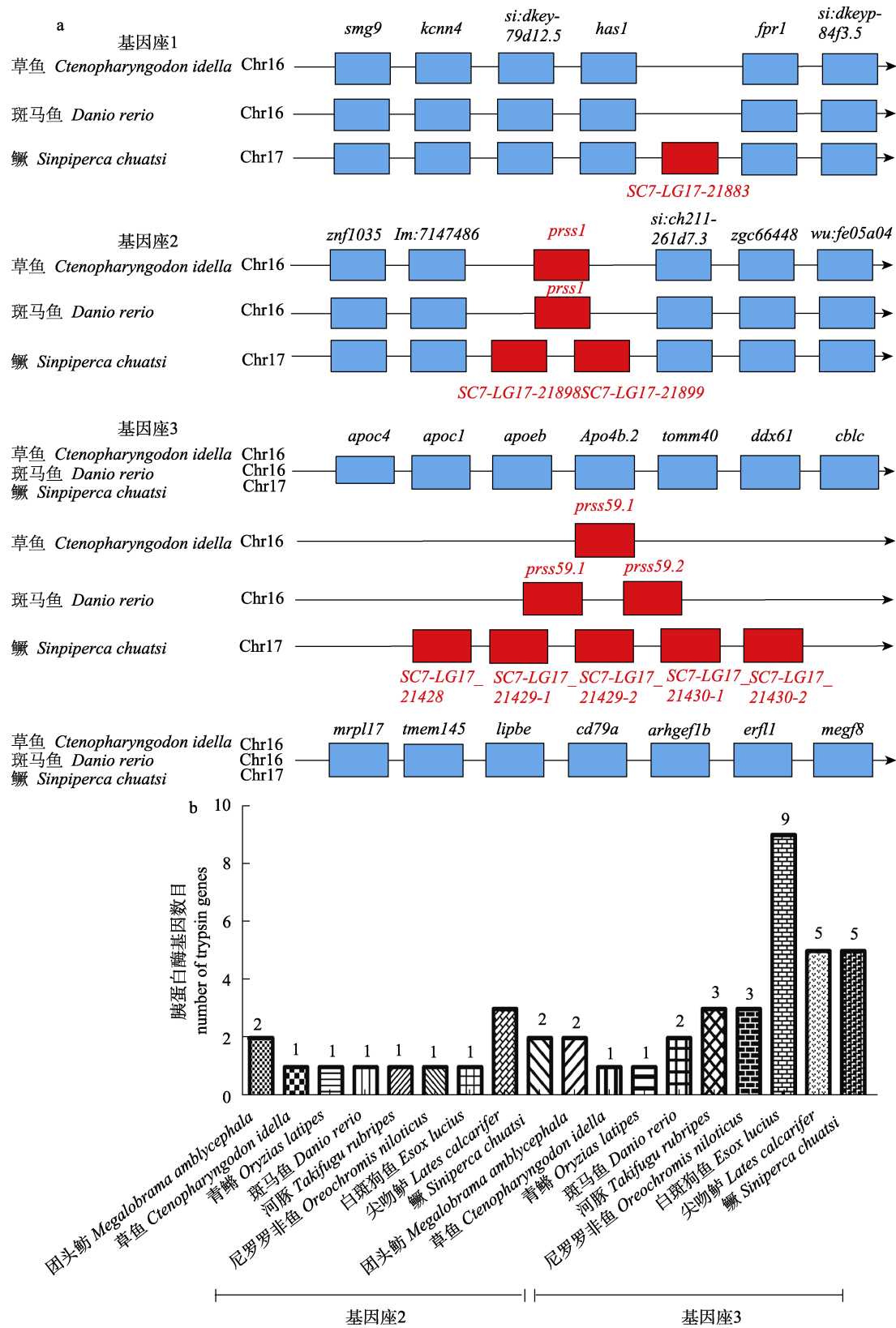


图 4 胰蛋白酶基因共线性分析结果(a)及部分鱼类基因座 2 和基因座 3 上胰蛋白酶基因数目(b)

Fig. 4 Covariance analysis of trypsin genes (a) and number of trypsin genes on locus 2 and locus 3 in selected fish species (b)



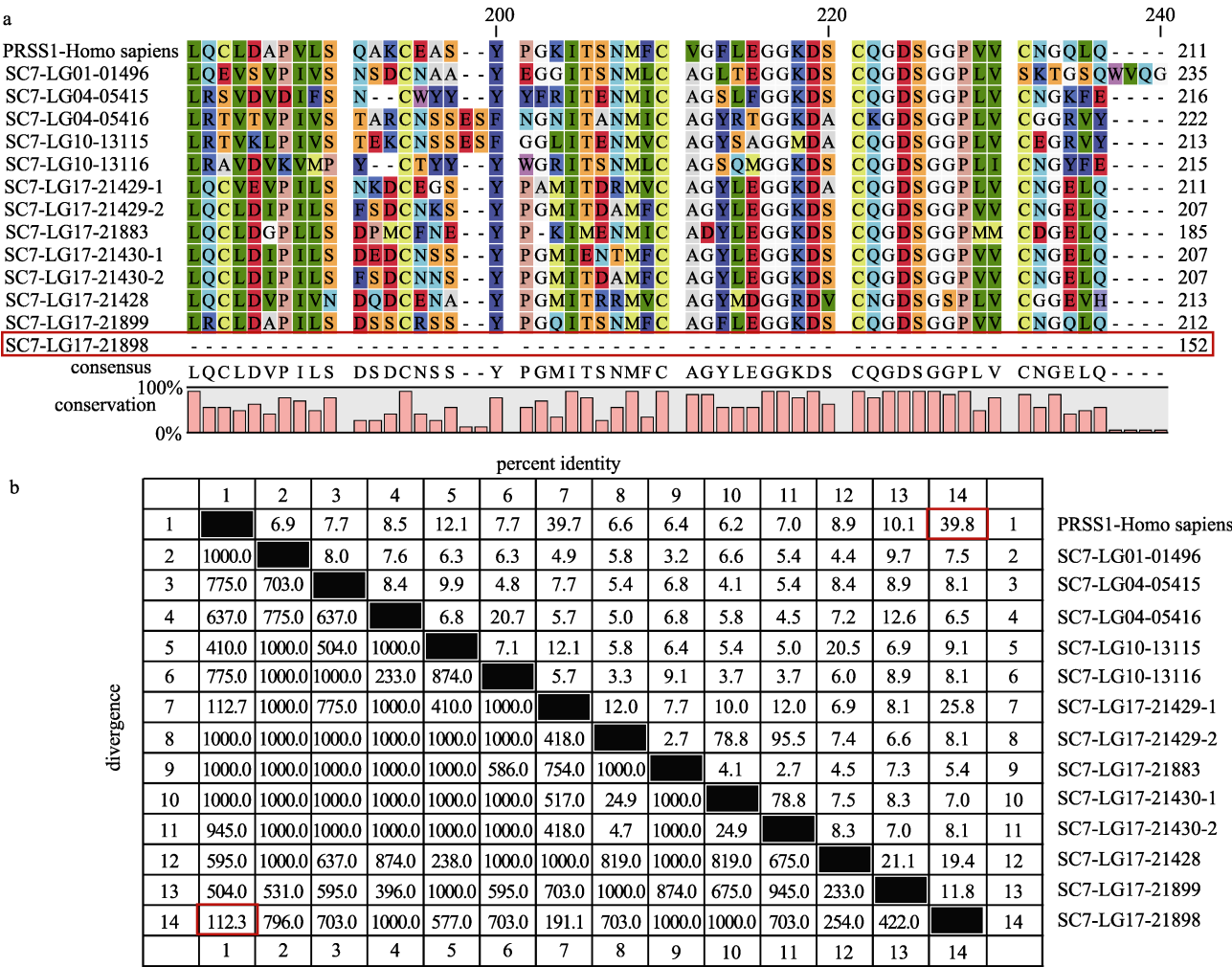


图 5 鳊胰蛋白酶基因氨基酸序列多序列比对(a)及其与人 PRSS1 基因编码的氨基酸序列同源百分比(b)

Fig. 5 Multiple sequence comparison in the amino acid sequence of trypsin genes in *Siniperca chuatsi* (a) and divergence of amino acid sequences between *trypsin* genes of *Siniperca chuatsi* and *PRSS1* gene of human (b)

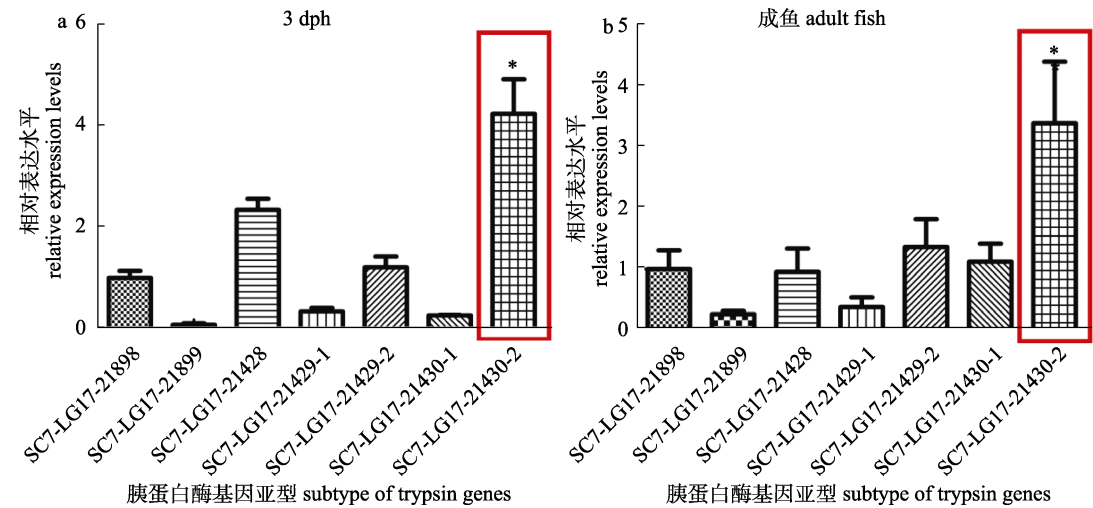


图 6 3 dph 鳊(a)及鳊成鱼(b)胰蛋白酶基因相对表达水平

\*表示该组与其他组之间存在显著性差异( $P<0.05$ ).

Fig. 6 Relative expression levels of *trypsin* genes in *Siniperca chuatsi* at 3 dph (a) and adult *S. chuatsi* (b)

\* indicates a significant difference between this group and the other groups ( $P<0.05$ ).

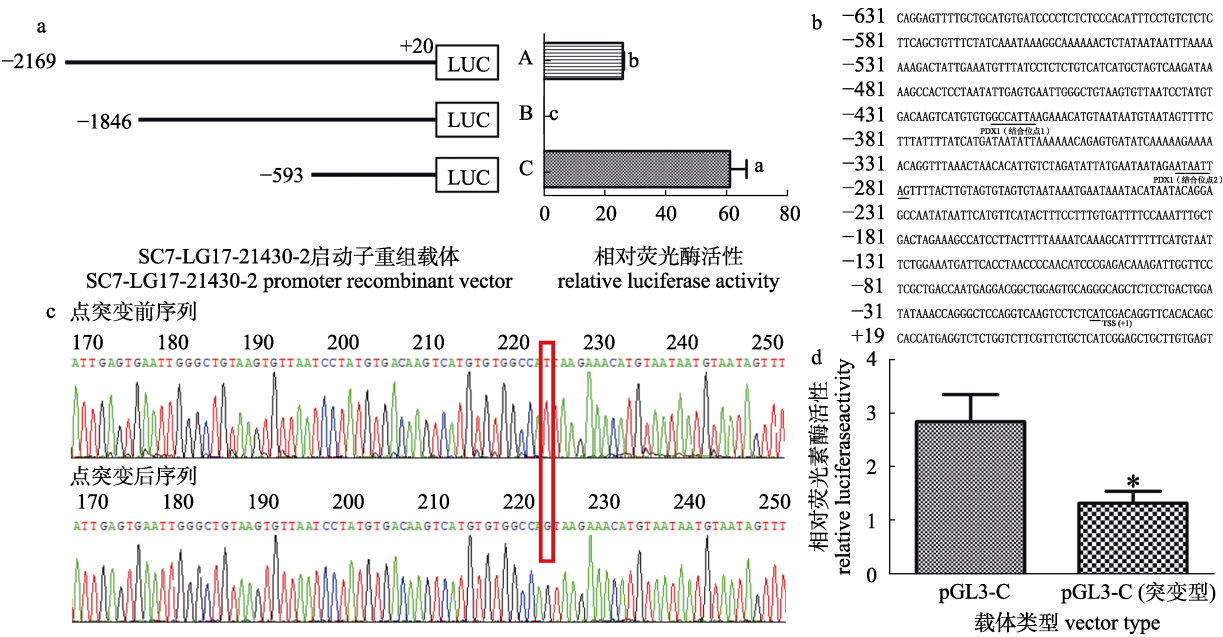


图 7 PDX1 对鳊胰蛋白酶基因 SC7-LG17-21430-2 亚型启动子核心区启动活性的影响

a. 鳊胰蛋白酶基因 SC7-LG17-21430-2 亚型启动子缺失片段相对荧光素酶活性; b. 鳊胰蛋白酶基因 SC7-LG17-21430-2 亚型启动子区域的转录因子结合位点预测; c. PDX1 结合位点突变示意图; d. 突变型载体相对荧光素酶活性。

\*及不同字母代表组间差异显著,  $P<0.05$ 。

Fig. 7 The impact of PDX1 on the initiation activity of the core promoter region of SC7-LG17-21430-2 trypsin subtype in *Siniperca chuatsi*

a. Relative luciferase activity of three constructed deleted fragments from SC7-LG17-21430-2 in *S. chuatsi*; b. Prediction of transcription factor binding sites in the promoter region of SC7-LG17-21430-2 in *S. chuatsi*; c. Schematic diagram of point mutation at the binding site of the PDX1; d. Relative luciferase activity of the point-mutated vector. \* and different letters represent significant difference among species,  $P<0.05$ .

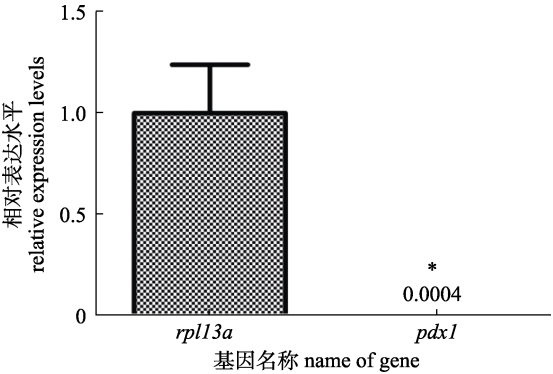


图 8 3 dph 时期鳊仔鱼 *pdx1* 基因相对表达水平

\*表示该组与其他组之间存在显著性差异( $P<0.05$ )。

Fig. 8 Relative expression levels of *pdx1* in *Siniperca chuatsi* at 3 dph

\* indicates significant difference between this group and the other groups ( $P<0.05$ ).

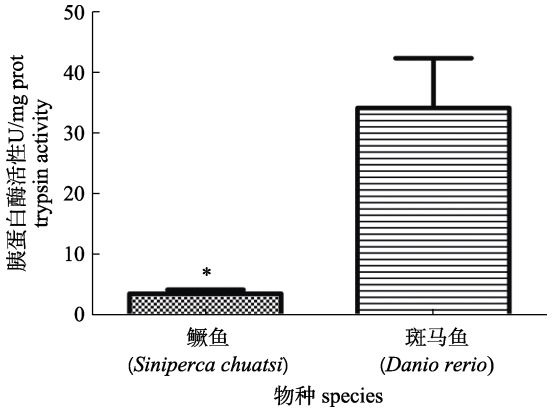


图 9 鳊与斑马鱼仔鱼 3 dph 时胰蛋白酶活性比较

\*表示该组与其他组之间存在显著性差异( $P<0.05$ )。

Fig. 9 Comparison of trypsin activity between *Siniperca chuatsi* and *Danio rerio* at 3 dph

\* indicates significant difference between this group and the other groups ( $P<0.05$ ).

导致鳊丧失了一个发挥功能的胰蛋白酶基因。这可能与鳊开口时期胰蛋白酶活性低有关, 因而鳊通过摄食活鱼苗来补偿内源性胰蛋白酶低活性, 形成自开口起只食活鱼的奇特食性。

有研究对鱼类部分胰蛋白酶基因亚型的相对表达水平进行了比较<sup>[12-13]</sup>, 但现有研究仅涵盖部分胰蛋白酶基因亚型, 仍无法确定鱼类关键的胰

蛋白酶基因亚型。本研究选择含有与人 *PRSSI* 基因直系同源基因的座位 2 和与斑马鱼 *prss59.1* 和 *prss59.2* 基因直系同源基因的座位 3, 利用 qPCR 对以上两个基因座中的鳊胰蛋白酶基因进行组织表达分析, 从结果推测鳊关键胰蛋白酶基因亚型为 SC7-LG17-21430-2, 且 SC7-LG17-21430-2 亚型启动区域的 3 个逐段缺失片段启动子活性均存在显著差异, 其中-593 bp~+20 bp 区域启动子活性最高, 推测为启动子核心区域。此外, 在该核心区域, 发现 PDX1 这一与早期胰腺发育相关的转录因子的结合位点, PDX1 是调控胰腺发育及胰腺酶类生物活性的主要正调控转录因子, 对胰腺干细胞的形成和维持中起着重要的调控作用, 是胰腺早期发育的重要转录因子<sup>[33-34]</sup>。研究表明缺乏 Pdx1 的小鼠腺泡细胞会加速衰老, 导致胰蛋白酶分泌缺陷<sup>[35]</sup>。本研究中, 利用点突变手段将 Pdx1 的结合位点进行点突变, 结果导致核心区域启动子活性显著下降, 且 3 dph 时期鳊 RNA-seq 数据及 qPCR 结果表明, *pdx1* 基因在鳊发育早期几乎不表达, 可能导致鳊开口时期胰腺发育迟缓, 胰蛋白酶活性显著低于 3 dph 时期的斑马鱼仔鱼。鱼类孵化初期消化系统发育并不完善, 消化能力受限<sup>[36]</sup>, 部分鱼类在开口期需利用来自猎物的消化酶在肠道中协助激活酶原, 改善消化<sup>[37-38]</sup>。开口时期低胰蛋白酶活性的鳊仔鱼能消化利用活饵料鱼, 可能是得益于活饵料鱼体内的胰蛋白酶, 而卤虫等浮游动物胰蛋白酶含量较低<sup>[39]</sup>, 无法给予鳊充足的外源性胰蛋白酶, 因此鳊拒食低胰蛋白酶含量的浮游动物, 转而选择摄食高胰蛋白酶含量的活鱼苗, 以补偿低活性的内源性胰蛋白酶, 这可能是鳊开口时期活鱼食性形成的重要因素。

综上所述, 本研究从胰蛋白酶基因的进化、胰蛋白酶活性、关键胰蛋白酶基因的转录调控机制入手, 解析了鳊胰蛋白酶基因数目、结构及表达模式与其开口时期活鱼食性间的联系, 为进一步解析鳊活鱼食性的形成机制提供了理论依据。

#### 参考文献:

[1] Pusey B J, Read M G, Arthington A H. The feeding ecology of freshwater fishes in two rivers of the Australian wet

- tropics[J]. *Environmental Biology of Fishes*, 1995, 43(1): 85-103.
- [2] Garrison L P, Link J S. Dietary guild structure of the fish community in the Northeast United States continental shelf ecosystem[J]. *Marine Ecology Progress Series*, 2000, 202: 231-240.
- [3] Chaudhuri A, Mukherjee S, Homechaudhuri S. Diet composition and digestive enzymes activity in carnivorous fishes inhabiting mudflats of Indian sundarban estuaries[J]. *Turkish Journal of Fisheries and Aquatic Sciences*, 2012, 12(2): 265-275.
- [4] Drewe K E, Horn M H, Dickson K A, et al. Insectivore to frugivore: Ontogenetic changes in gut morphology and digestive enzyme activity in the characid fish *Brycon guatemalensis* from Costa Rican rain forest streams[J]. *Journal of Fish Biology*, 2004, 64(4): 890-902.
- [5] Gioda C R, Pretto A, de Souza Freitas C, et al. Different feeding habits influence the activity of digestive enzymes in freshwater fish[J]. *Ciência Rural*, 2017, 47(3): e20160113.
- [6] Jiao F, Zhang L, Limbu S M, et al. A comparison of digestive strategies for fishes with different feeding habits: Digestive enzyme activities, intestinal morphology, and gut microbiota[J]. *Ecology and Evolution*, 2023, 13(9): e10499.
- [7] Colchen T, Dias A, Gisbert E, et al. The onset of piscivory in a freshwater fish species: Analysis of behavioural and physiological traits[J]. *Journal of Fish Biology*, 2020, 96(6): 1463-1474.
- [8] Muhlia-Almazán A, Sánchez-Paz A, García-Carreño F L. Invertebrate trypsins: A review[J]. *Journal of Comparative Physiology B*, 2008, 178(6): 655-672.
- [9] Scheele G, Bartelt D, Bieger W. Characterization of human exocrine pancreatic proteins by two-dimensional isoelectric focusing/sodium dodecyl sulfate gel electrophoresis[J]. *Gastroenterology*, 1981, 80(3): 461-473.
- [10] Rana R, Hossain F, Ali H, et al. Biochemical characterization of zebrafish *Prss59.1*[J]. *Biochemical and Biophysical Research Communications*, 2021, 560: 32-36.
- [11] Chen W B, Li W G, Zhang Z, et al. Molecular cloning, sequence analysis and tissue expression of three forms trypsinogens in zebrafish, *Danio rerio*[J]. *Acta Scientiarum Naturalium Universitatis Sunyatseni*, 2013, 52(1): 111-117. [陈文波, 李卫国, 张珍, 等. 斑马鱼 3 种胰蛋白酶原基因的克隆, 序列分析及组织表达[J]. 中山大学学报(自然科学版), 2013, 52(1): 111-117.]
- [12] Pálsdóttir H M, Gudmundsdóttir Á. Development of a qRT-PCR assay to determine the relative mRNA expression of two different trypsins in Atlantic cod (*Gadus morhua*)[J].

- Comparative Biochemistry and Physiology B: Biochemistry & Molecular Biology, 2007, 146(1): 26-34.
- [13] Machado M, Infante C, Asensio E, et al. Molecular characterization and gene expression of six trypsinogens in the flatfish Senegalese sole (*Solea senegalensis* Kaup) during larval development and in tissues[J]. Comparative Biochemistry and Physiology B: Biochemistry & Molecular Biology, 2008, 149(2): 334-344.
- [14] Rønnestad I, Yúfera M, Ueberschär B, et al. Feeding behaviour and digestive physiology in larval fish: Current knowledge, and gaps and bottlenecks in research[J]. Reviews in Aquaculture, 2013, 5(S1): S59-S98, S270.
- [15] Ghasemi N, Imani A, Noori F, et al. Ontogeny of digestive tract of stellate sturgeon (*Acipenser stellatus*) from hatching to juvenile stage: Digestive enzymes activity, stomach and proximal intestine[J]. Aquaculture, 2020, 519: 734751.
- [16] Anderson K, Kuo C Y, Lu M W, et al. A transcriptomic investigation of digestive processes in orange-spotted grouper, *Epinephelus coioides*, before, during, and after metamorphic development[J]. Gene, 2018, 661: 95-108.
- [17] Odedeyi D O, Fagbenro O A. Feeding habits and digestive enzymes in the gut of *Mormyrus rume* (Valenciennes 1846) (Osteichthyes Mormyridae)[J]. Tropical Zoology, 2010, 23(1): 75-89.
- [18] Chi L, Liu Q H, Xu S H, et al. Maternally derived trypsin may have multiple functions in the early development of turbot (*Scophthalmus maximus*)[J]. Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology, 2015, 188: 148-155.
- [19] Whitcomb D C, Lowe M E. Human pancreatic digestive enzymes[J]. Digestive Diseases and Sciences, 2007, 52(1): 1-17.
- [20] Rønnestad I, Tonheim S, Fyhn H, et al. The supply of amino acids during early feeding stages of marine fish larvae: A review of recent findings[J]. Aquaculture, 2003, 227(1): 147-164.
- [21] Mir I N, Ahmad Bhat I, Ahmad Dar S, et al. Expression of alpha-amylase and growth-related genes during early larval developmental stages of *Clarias magur*[J]. Aquaculture, 2019, 507: 69-74.
- [22] Chen J M, Kukor Z, Le Maréchal C, et al. Evolution of trypsinogen activation peptides[J]. Molecular Biology and Evolution, 2003, 20(11): 1767-1777.
- [23] Tamura K, Nei M, Kumar S. Prospects for inferring very large phylogenies by using the neighbor-joining method[J]. Proceedings of the National Academy of Sciences of the United States of America, 2004, 101(30): 11030-11035.
- [24] Bailey T L, Williams N, Misleh C, et al. MEME: Discovering and analyzing DNA and protein sequence motifs[J]. Nucleic Acids Research, 2006, 34(S2): W369-W373.
- [25] Srichanun M, Tantikitti C, Utarabhand P, et al. Gene expression and activity of digestive enzymes during the larval development of Asian seabass (*Lates calcarifer*)[J]. Comparative Biochemistry and Physiology Part B: Biochemistry & Molecular Biology, 2013, 165(1): 1-9.
- [26] Mata-Sotres J A, Moyano F J, Martínez-Rodríguez G, et al. Daily rhythms of digestive enzyme activity and gene expression in gilthead seabream (*Sparus aurata*) during ontogeny[J]. Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology, 2016, 197: 43-51.
- [27] Day R D, Tibbetts I R, Secor S M. Physiological responses to short-term fasting among herbivorous, omnivorous, and carnivorous fishes[J]. Journal of Comparative Physiology Part B: Biochemical Systems and Environmental Physiology, 2014, 184(4): 497-512.
- [28] Axelsson E, Ratnakumar A, Arendt M L, et al. The genomic signature of dog domestication reveals adaptation to a starch-rich diet[J]. Nature, 2013, 495: 360-364.
- [29] German D P, Foti D M, Heras J, et al. Elevated gene copy number does not always explain elevated amylase activities in fishes[J]. Physiological and Biochemical Zoology, 2016, 89(4): 277-293.
- [30] Le N, Heras J, Herrera M J, et al. The genome of *Anop-larchus purpurescens* (Stichaeidae) reflects its carnivorous diet[J]. Molecular Genetics and Genomics, 2023, 298(6): 1419-1434.
- [31] Liu H, Chen C H, Lv M L, et al. A chromosome-level assembly of blunt snout bream (*Megalobrama amblycephala*) genome reveals an expansion of olfactory receptor genes in freshwater fish[J]. Molecular Biology and Evolution, 2021, 38(10): 4238-4251.
- [32] Yuan X C, Liang X F, Cai W J, et al. Expansion of sweet taste receptor genes in grass carp (*Ctenopharyngodon idellus*) coincided with vegetarian adaptation[J]. BMC Evolutionary Biology, 2020, 20(1): 25.
- [33] Yang X D, Raum J C, Kim J, et al. A PDX1 cistrome and single-cell transcriptome resource of the developing pancreas[J]. Development, 2022, 149(13): dev200432.
- [34] Holland A M, Góñez L J, Naselli G, et al. Conditional expression demonstrates the role of the homeodomain transcription factor Pdx1 in maintenance and regeneration of  $\beta$ -cells in the adult pancreas[J]. Diabetes, 2005, 54(9): 2586-2595.
- [35] Horiguchi M, Yoshida M, Hirata K, et al. Senescence caused

- by inactivation of the homeodomain transcription factor Pdx1 in adult pancreatic acinar cells in mice[J]. *FEBS Letters*, 2019, 593(16): 2226-2234.
- [36] Tao T, Peng J R. Liver development in zebrafish (*Danio rerio*)[J]. *Journal of Genetics and Genomics*, 2009, 36(6): 325-334.
- [37] Lauff M, Hofer R. Proteolytic enzymes in fish development and the importance of dietary enzymes[J]. *Aquaculture*, 1984, 37(4): 335-346.
- [38] Kolkovski S, Tandler A, Kissil G W, et al. The effect of dietary exogenous digestive enzymes on ingestion, assimilation, growth and survival of gilthead seabream (*Sparus aurata*, Sparidae, Linnaeus) larvae[J]. *Fish Physiology and Biochemistry*, 1993, 12(3): 203-209.
- [39] Moraiti-Ioannidou M, Castritsi-Catharios J, Miliou H, et al. Biochemical composition and digestive enzyme activity during naupliar development of *Artemia* spp. from three solar saltworks in Greece[J]. *Aquaculture*, 2008, 286(3): 259-265.

## Analysis of the evolution and promoter function of trypsin genes in mandarin fish (*Siniperca chuatsi*)

HU Zijun<sup>1,2</sup>, LU Hongliang<sup>1,2</sup>, ZHAO Chenyu<sup>1</sup>, HE Shan<sup>1,2</sup>

1. College of Fisheries, Huazhong Agricultural University, Wuhan 430070, China;

2. Engineering Research Center of Green Development for Conventional Aquatic Biological Industry in the Yangtze River Economic Belt, Ministry of Education, Wuhan 430070, China

**Abstract:** This study aimed to investigate the relationship between trypsin and the feeding habits of live prey fish in *Siniperca chuatsi* and to elucidate the transcriptional regulatory mechanism of the key trypsin gene in *S. chuatsi*. In this study, we used bioinformatics methods to conduct quantitative, structural, and phylogenetic analyses of trypsin genes in 11 fish species with different feeding habits, combining qPCR, dual-luciferase reporter, and trypsin activity assays to analyze the transcriptional regulation mechanism of the key trypsin gene subtype in *S. chuatsi* and its correlation with trypsin activity at the first feeding stage. The results showed that a total of 13 trypsin genes distributed at six loci on four chromosomes were identified in *S. chuatsi*, which was more than that in omnivorous and herbivorous fish, indicating that the amplification of trypsin genes in carnivorous fish is highly adapted to the food environment. Carnivorous *S. chuatsi* was amplified at key gene loci 2 (SC7-LG17-21898 and SC7-LG17-21899) and 3 (SC7-LG17-21428, SC7-LG17-21429-1, SC7-LG17-21429-2, SC7-LG17-21430-1, and SC7-LG17-21430-2), which are homologous to the key trypsin gene *PRSS1* in humans. More significant gene amplification occurred at locus 3, where SC7-LG17-21430-2 had a significantly higher expression level than the other subtypes, indicating that it was the key trypsin subtype in *S. chuatsi*. The dual-luciferase reporter assay results showed that the promoter activity of three constructed deleted fragments in the 5' flanking region of the SC7-LG17-21430-2 subtype had significant differences, and the region of -593 bp—+20 bp had the highest promoter activity, suggesting it as the core promoter region. Furthermore, PDX1, a transcription factor associated with early pancreatic development, was barely expressed in *S. chuatsi* at 3 dph and was observed in the core promoter region. Moreover, point mutations in the binding site of PDX1 result in a significant decrease in the promoter activity of the core region. The trypsin activity of *S. chuatsi* at 3 dph was significantly lower than that of *Danio rerio*, indicating that the low expression of *pdx1* might lead to the low expression of SC7-LG17-21430-2 and low trypsin activity, which could be associated with the unique feeding habits of live prey fish in *S. chuatsi*. These results provide the groundwork for analyzing the mechanisms underlying the special feeding habits of *S. chuatsi*.

**Key words:** *Siniperca chuatsi*; itrypsin; feeding habit; enzyme activity; promoter activity

**Corresponding author:** HE Shan. E-mail: heshan@mail.hzau.edu.cn