

花鲈 *Claudins* 基因家族的鉴定、进化分析及对环境盐度的表达响应

李俊杰, 于 朋, 李 昀, 田 源, 张凯强, 王海亮, 齐 鑫, 温海深

(中国海洋大学 海水养殖教育部重点实验室, 山东 青岛 266003)

摘要: Claudin (cldn)是细胞间紧密连接蛋白的重要功能和结构组件, 通过控制细胞旁通路通透性参与渗透压平衡的调节。本研究对花鲈(*Lateolabrax maculatus*) *cldns* 基因家族进行了系统的鉴定分析, 结果表明: 花鲈 *cldns* 基因家族共包含 55 个成员, 分布于 16 条染色体上, 其中 34 个 *cldns* 基因在哺乳动物体内有对应直系同源基因, 其余 21 个 *cldns* 是硬骨鱼所特有。选择压力分析显示预测的花鲈 *cldns* 蛋白的跨膜区域存在 14 个受正选择作用的氨基酸位点, 这些正选择位点可能与家族内基因的多样性和功能分化有关。此外, 基于花鲈鳃转录组的基因表达结果表明, 至少有 24 个 *cldns* 基因在鳃中表达, 其中 *cldn8c*、*cldn10b*、*cldn10c*、*cldn10d*、*cldn27a* 和 *cldn30b* 在不同环境盐度处理后表达差异显著, 表明其可能是花鲈鳃组织中调节渗透压平衡的候选功能基因。本研究成果为研究花鲈及其他广盐性鱼类 *cldns* 基因家族成员的渗透调节机制奠定了理论基础。

关键词: *Claudins* 基因家族; 花鲈; 盐度; 进化分析

中图分类号: S917.4

文献标识码: A

文章编号: 1000-3096(2022)08-0065-14

DOI: 10.11759/hyxx20210203001

紧密连接通过平衡细胞间通路调节细胞旁运输, 在渗透调节中发挥重要作用。紧密连接复合蛋白由封闭蛋白(claudin)、闭合蛋白(ooccludin)、缝隙连接蛋白(ZO-1)以及激动蛋白细胞骨架(actin cytoskeleton)构成^[1]。其中, Claudin 是一个多基因家族, 常简称为 cldn, 是构成细胞间紧密连接必不可少的组成部分, 可调节水和离子在细胞垂直方向跨细胞紧密连接的细胞旁运动^[2], 对于细胞间密封的形成和紧密连接的渗透性的控制都至关重要^[3]。

Claudin 是一种四跨膜蛋白, 其中, 第 1 个与第 2 个跨膜结构域高度保守, 氨基与羧基都分布在细胞内, 羧基末端富含丰富的丝氨酸、苏氨酸及酪氨酸残基^[4]。Claudin 对细胞连接处选择渗透性的调节主要通过蛋白激酶途径实现, cldn 上特异性的氨基酸靶点与蛋白激酶 A 或蛋白激酶 C 结合, 磷酸化后紧密连接功能下降, 表现为氯化物渗透性增强^[5]。不同的 *cldn* 成员存在组织特异性的表达, 并对离子交换的选择性不同, 它们以多种组合方式参与构成紧密连接复合蛋白, 增加了紧密连接的结构和功能的多样性, 形成了不同类型的屏障功能^[6]。目前, 已有大量 *cldns* 基因被鉴定出来, 不同物种间基因数存在较大差异: 无脊椎动物如秀

丽隐杆线虫(*Caenorhabditis elegans*)、黑腹果蝇(*Drosophila melanogaster*)只有 4 到 5 个 *cldns* 相关基因^[7-8], 在哺乳动物中, 有 24 个 *cldns* 家族成员被发现^[3], 而在硬骨鱼中, 由于硬骨鱼特异的全基因组复制(WGD)和串联重复事件, 先后共有 63 个 *cldns* 被发现^[9]。最大的扩增事件发生在 *cldn4* 上: 人类 *cldn4* 与 13 个鱼类 *cldns* 基因显示出同源关系(*cldn27a-d*、*cldn28a-c*、*cldn29a-b*、*cldn30a-d*)^[10]。

鱼类的 *cldn* 基因与渗透压调节和盐度适应存在着密切的关系。鳃是鱼类调节渗透压最重要的器官之一。透射电镜观察结果表明, 当鱼体从淡水转入海水中时, 鳃上皮离子细胞和附属细胞间紧密连接的超微结构由封闭的“栅栏”变为泄漏的“孔状”^[11]。这两种细胞类型间存在着泄漏的紧密连接, 对于鳃上皮分泌多余的 Na^+ 提供了至关重要的细胞旁通路^[12],

收稿日期: 2021-02-03; 修回日期: 2021-03-09

基金项目: 国家自然科学基金课题(32072947); 现代农业产业技术体系专项(CARS-47)

[Foundation: National Natural Science Foundation of China, No. 32072947; China Agriculture Research System, No. CARS-47]

作者简介: 李俊杰(1996—), 男, 江苏苏州人, 硕士研究生, 研究方向为鱼类生理与繁育, E-mail: 1121938079@qq.com; 温海深(1963—), 通信作者, 男, 教授, 研究方向为鱼类生理与繁育, E-mail: wenhaishen@ouc.edu.cn

而 *cldns* 对于调节水和离子在细胞垂直方向跨细胞紧密连接的细胞旁运动至关重要。实验结果显示, 高渗环境诱导大西洋鲑鱼鳃中 *cldn10e* 的表达, 低渗环境则诱导大西洋鲑鱼鳃中 *cldn27a* 和 *cldn30* 的表达^[13]。而在罗非鱼的鳃组织中, 低渗环境诱导 *cldn28a* 和 *cldn30* 的表达^[14]。此外, 在两种河豚(*Tetraodon nigroviridis* 和 *T. biocellatus*)鳃中均发现了 *cldn27* mRNA 的盐度依赖性差异表达^[15-16]。Marshall 等^[17]探究了底鳃 *cldn10* 亚型在不同盐度下的适应变化, 结果表明环境盐度的增加提升了 *cldn10* 不同亚型的表达量。对于瓦氏雅罗鱼(*Leuciscus waleckii*)的 *cldn* 基因家族的选择压力分析显示, *cldn7* 为正选择基因, 表明其在促进离子转运和酸碱调节以适应高盐碱的自然环境中发挥潜在作用^[18]。

花鲈(*Lateolabrax maculatus*)是我国重要的海产经济鱼类, 具有典型的广盐性特征, 可在盐度 0~45 的范围内正常存活、生长, 是研究鱼类盐度耐受性及渗透压调节机制的理想模型。本文对花鲈 *cldn* 基因进行了全基因组的鉴定和注释, 并对其基因结构及进化特征进行分析, 同时开展了它们在不同盐度下鳃组织中的表达特征分析, 以期筛选与盐度适应和渗透压调节相关的 *cldn* 基因, 为进一步研究花鲈对急性盐度胁迫的生理响应以及花鲈的耐盐机制研究奠定基础。

1 材料与方法

1.1 实验用鱼及暂养条件

实验用 1 龄花鲈[(100.00±2.34) g]取自山东省东营市双赢水产苗种有限责任公司, 实验开始前花鲈暂养在盐度为 30 的水泥池(5 m×5 m×1 m)中, 温度为(21.0±0.5) °C, pH 值为 7.98~8.04, 溶解氧含量为 6.90~8.54 mg/L, 亚硝酸盐<0.1 mg/L, 光照周期 14L:10D。

1.2 盐度处理实验

本实验共设计 3 个盐度组, 分别为 0(淡水组)、12(等渗组)、30(海水组), 每个盐度组设置 3 个重复。实验开始后, 将暂养的花鲈转移至不同盐度的水桶中, 实验每个重复养殖 5 尾花鲈, 实验进行 30 d, 每日 8:30 和 16:30 投喂配合饲料, 每 2 d 换一次水, 换水量为 50%。实验结束后, 每个盐度组至少随机选取 3 尾鱼进行鳃组织取样并迅速保存于液氮中, 随后置于-80 °C 超低温冰箱中保存备用。

1.3 花鲈 *cldn* 基因的鉴定和序列分析

从 NCBI (<http://www.ncbi.nlm.nih.gov/>)数据库中检索获取人(*Homo sapiens*)、小鼠(*Mus musculus*)、红鳍东方鲀(*T. rubripes*)和斑马鱼(*Danio rerio*)4 种物种的所有 *cldn* 基因的氨基酸序列, 利用 TBLASTN (1e-5)对花鲈参考基因组(NCBI 登录号: PRJNA408177)和转录组数据库(NCBI 登录号: SRR4409341 和 SRR4409397)进行检索, 初步获得花鲈 *cldn* 候选基因序列。利用在线工具 open reading frame (ORF) finder (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>)筛选出完整的开放阅读框, 并预测氨基酸序列, 通过 smartblast 比对 NCBI 非冗余(nr)蛋白序列数据库进行进一步验证。利用 NCBI-CDD 数据库(<https://www.ncbi.nlm.nih.gov/cdd/>)对鉴定得到的花鲈 *cldn* 基因的特征结构域进行验证。使用 ExPASyProt-Param 工具(<https://web.expasy.org/protparam/>)计算每个预测的 *cldn* 蛋白的分子量(MW, kDa)和等电点(pI)。根据其他几种脊椎动物的基因组数据库对不同物种的 *cldns* 基因拷贝数进行比较。

1.4 *Cldns* 的系统发育分析及染色体分布

对花鲈、人、小鼠、斑马鱼以及硬骨鱼红鳍东方鲀的 *cldns* 氨基酸序列进行系统发育分析。利用 ClustalW 进行比对(默认参数)。通过 MEGA 7.0 软件, 采用 Neighbor-Joining (NJ)法构建系统进化树, bootstrap 数值选择 1 000^[19]。利用在线软件 Interactive Tree Of Life (<http://itol.embl.de/>)对进化树进行作图。通过花鲈基因组数据库获取染色体位置与基因密度信息, 通过 MCScanX (默认参数)挖掘共线性信息^[20]。根据每个 *cldn* 基因在基因组上的坐标定位其染色体位置, 并使用 TBtools 软件进行结果可视化^[21]。

1.5 花鲈 *Cldns* 基因家族的进化选择压力分析

使用 MUSCLE 方法对 55 条花鲈 *cldns* 的预测的氨基酸序列进行比对。使用 MEGA 7.0 程序的最大似然法对花鲈 *cldns* 构建系统发育树, 基于物种内部 *cldns* 序列的系统发育关系和后续分析的需要, 本文将 *cldns* 家族成员分为 4 组: Class I、Class II、Class III 和 Class IV。使用 PAL2NAL (<http://www.bork.embl.de/pal2nal/>)获得准确比对的核苷酸序列, 用于获得后续选择压力分析。

基于花鲈 *cldns* 家族物种内的系统发育树, 采

用 PAML 4.9i 的 CODEML 程序^[22], 利用不同的模型(枝模型、位点模型、枝位点模型), 对 4 组主要的进化枝进行选择压力分析。计算非同义替换与同义替换的比值 ω 值, 即 d_N/d_S 。当 $\omega=1$ 时, 为中性进化(Neutral selection), 即不受选择; 当 $\omega>1$ 时, 为正选择(positive selection); 当 $0<\omega<1$ 时, 为负选择(negative selection), 也叫净化选择或纯化选择(purifying selection)。本文利用 EasyCodeML 程序对 3 种模型的各项参数进行计算^[23], 采用似然比检验(LRT)比较模型对的拟合程度, 采用贝叶斯法对正选择位点进行识别^[24]。

1.6 基因表达分析

利用花鲈在不同盐度条件下的鳃组织转录组数据库(NCBI 登录号: PRJNA515986)进行 meta 分析,

对每百万 reads 中来自于某基因每千碱基长度的 reads 数(RPKM)值进行归一化处理后, 基于 RPKM 计算 *cldns* 在每个组之间的表达倍数变化(海水组/等渗组、淡水组/等渗组、淡水组/海水组)。满足总 read 数 ≥ 5 、 $\log_2|(\text{差异表达倍数})| \geq 1$ 、 P 值 ≤ 0.05 的基因被认为是显著差异表达基因。

2 结果

2.1 花鲈 *cldns* 基因家族成员的鉴定

在花鲈全基因组范围内共鉴定获得 55 条 *cldns* 序列(表 1), CDS 长度 387~1 284 bp, 编码蛋白长度为 128~427 个氨基酸。除 *cldn10b* 和 *cldn28b* 外, 分子量均在 80 kDa 以下。此外, 各 *cldns* 等电点和保守结构域的位置在表 1 中列出。

表 1 花鲈 *cldns* 基因家族成员的序列特征

Tab. 1 Summary of characteristics of *cldns* genes in spotted sea bass

基因	CDS 长度/bp	分子量/kDa	等电点	预测氨基酸长度	保守结构域位置	NCBI 登录号
<i>cldn1a</i>	636	52.41	5.14	211	13~546	MW575141
<i>cldn1b</i>	603	49.31	5.13	200	46~369	MW575142
<i>cldn2</i>	822	67.51	5.04	273	67~546	MW575143
<i>cldn3a</i>	648	54.21	5.10	215	10~513	MW575144
<i>cldn3b</i>	657	54.08	5.12	218	13~543	MW575145
<i>cldn3c</i>	651	54.75	5.11	216	7~540	MW575146
<i>cldn3d</i>	672	55.60	5.11	223	13~468	MW575147
<i>cldn5a</i>	648	52.89	5.11	215	13~543	MW575148
<i>cldn5b</i>	645	52.84	5.09	214	16~504	MW575149
<i>cldn5c</i>	555	45.39	5.15	184	13~543	MW575150
<i>cldn6</i>	636	53.09	5.10	211	1~534	MW575151
<i>cldn7a</i>	642	53.62	5.10	213	10~546	MW575152
<i>cldn7b</i>	684	57.03	5.07	227	283~606	MW575153
<i>cldn8b</i>	522	44.14	5.15	173	1~297	MW575154
<i>cldn8c</i>	843	71.51	5.07	280	1~297	MW575155
<i>cldn8d</i>	876	72.90	5.05	291	169~693	MW575156
<i>cldn10a</i>	597	49.81	5.13	198	10~483	MW575157
<i>cldn10b</i>	1 038	86.10	5.05	345	58~414	MW575158
<i>cldn10c</i>	750	63.10	5.14	253	22~549	MW575159
<i>cldn10d</i>	771	64.12	5.08	256	22~540	MW575160
<i>cldn11a</i>	657	55.35	5.09	218	16~525	MW575161
<i>cldn11b</i>	651	54.74	5.06	216	10~522	MW575162
<i>cldn14</i>	741	62.88	5.07	246	64~543	MW575163
<i>cldn15a</i>	669	55.16	5.14	222	13~528	MW575164
<i>cldn15b</i>	675	55.92	5.15	224	13~528	MW575165

续表

基因	CDS 长度/bp	分子量/kDa	等电点	预测氨基酸长度	保守结构域位置	NCBI 登录号
<i>cldn15c</i>	660	54.45	5.11	219	7~537	MW575166
<i>cldn15d</i>	675	56.11	5.09	224	10~537	MW575167
<i>cldn18a</i>	780	64.27	5.10	259	10~594	MW575168
<i>cldn18b</i>	675	54.74	5.14	224	55~363	MW575169
<i>cldn18c</i>	660	53.52	5.14	219	133~441	MW575170
<i>cldn19</i>	669	55.46	5.10	222	10~546	MW575171
<i>cldn20a</i>	729	60.92	5.08	242	10~543	MW575172
<i>cldn20b</i>	837	69.75	5.04	238	64~543	MW575173
<i>cldn23a</i>	840	68.76	5.02	279	52~552	MW575174
<i>cldn23b</i>	759	62.32	5.09	252	22~618	MW575175
<i>cldn27a</i>	390	32.64	5.22	129	7~270	MW575176
<i>cldn27b</i>	747	62.08	5.07	248	34~543	MW575177
<i>cldn27c</i>	390	32.11	5.22	129	7~258	MW575178
<i>cldn27d</i>	621	52.34	5.11	206	10~504	MW575179
<i>cldn28a1</i>	639	53.39	5.11	212	10~504	MW575180
<i>cldn28a2</i>	639	53.39	5.11	212	10~504	MW575180
<i>cldn28b</i>	1 284	107.49	4.98	427	358~789/844~1 155	MW575181
<i>cldn29a</i>	633	52.10	5.14	210	13~504	MW575182
<i>cldn29b</i>	387	32.22	5.23	128	7~264	MW575183
<i>cldn30a</i>	624	52.14	5.12	207	13~504	MW575184
<i>cldn30b</i>	666	54.74	5.13	221	10~543	MW575185
<i>cldn30c</i>	645	53.31	5.12	214	13~543	MW575186
<i>cldn30d</i>	756	61.52	5.10	251	22~324	MW575187
<i>cldn31a</i>	648	54.53	5.10	215	10~543	MW575188
<i>cldn31b</i>	669	54.60	5.11	222	10~549	MW575189
<i>cldn32a</i>	636	53.03	5.10	211	13~510	MW575190
<i>cldn32b</i>	690	57.68	5.08	229	10~546	MW575191
<i>cldn33a</i>	699	57.98	5.07	232	10~542	MW575192
<i>cldn33b</i>	624	51.40	5.13	207	19~516	MW575193
<i>cldn34</i>	756	62.51	5.07	251	55~588	MW575194

2.2 花鲈 *cldns* 基因拷贝数分析及染色体定位

对 *cldns* 各基因在花鲈、人、小鼠、鸡(*Gallus gallus*)、斑点叉尾鲷(*Ictalurus punctatus*)、斑马鱼及红鳍东方鲀 7 种脊椎动物中的拷贝数进行了总结分析。如表 2 所示,硬骨鱼 *cldns* 基因数目显著高于高等脊椎动物的 *cldns* 数目。花鲈、红鳍东方鲀和斑点叉尾鲷的 *cldns* 数目是人类的 2 倍以上,表明 *cldns* 基因家族在硬骨鱼中产生了显著扩增。*cldn4*、*cldn9*、

cldn16、*cldn17*、*cldn22* 及 *cldn24* 为高等脊椎动物所特有,而其直系同源物在硬骨鱼中缺失;根据斑马鱼和红鳍东方鲀的命名规则:硬骨鱼与哺乳动物(人)有直系同源关系的基因参照哺乳动物命名,而与人无直系同源关系的基因从 *cldn25* 开始命名^[10],因此,*cldn25* 之后的基因为硬骨鱼所特有。花鲈 *cldn1* 比其他硬骨鱼多了一个基因拷贝,而 *cldn12* 在花鲈基因组中缺失,*cldn18* 拥有 3 个拷贝,多于所对比的任何一种硬骨鱼(表 2)。

表 2 *cldns* 各基因拷贝数在几种典型脊椎动物中的比较

Tab. 2 Comparison of gene copy numbers of *cldn* genes among selected vertebrate genomes

名称	人	鼠	鸡	斑点叉尾鮰	斑马鱼	红鳍东方鲀	花鲈
<i>Cldn1</i>	1	1	1	1	1	1	2
<i>Cldn2</i>	1	1	1	1	1	1	1
<i>Cldn3</i>	1	1	1	4	4	4	4
<i>Cldn4</i>	1	1	0	0	0	0	0
<i>Cldn5</i>	1	1	1	2	2	3	3
<i>Cldn6</i>	1	1	0	0	0	1	1
<i>Cldn7</i>	1	1	0	2	2	2	2
<i>Cldn8</i>	1	1	1	3	3	4	3
<i>Cldn9</i>	1	2	0	0	0	0	0
<i>Cldn10</i>	1	1	1	5	4	5	4
<i>Cldn11</i>	1	1	1	3	2	2	2
<i>Cldn12</i>	1	1	1	1	1	1	0
<i>Cldn13</i>	0	1	0	0	0	1	0
<i>Cldn14</i>	1	1	1	3	0	2	1
<i>Cldn15</i>	1	1	0	2	3	2	3
<i>Cldn16</i>	1	1	1	0	0	0	0
<i>Cldn17</i>	1	1	0	0	0	0	0
<i>Cldn18</i>	1	1	1	2	1	1	3
<i>Cldn19</i>	1	1	1	1	1	1	1
<i>Cldn20</i>	1	1	1	1	0	2	2
<i>Cldn21</i>	0	0	0	0	0	0	0
<i>Cldn22</i>	1	1	0	0	0	0	0
<i>Cldn23</i>	1	1	0	2	2	2	2
<i>Cldn24</i>	1	1	0	0	0	0	0
<i>Cldn25</i>	1	1	1	1	0	1	0
<i>Cldn26</i>	0	0	0	1	0	1	0
<i>Cldn27</i>	0	0	0	2	0	4	4
<i>Cldn28</i>	0	0	0	2	3	3	3
<i>Cldn29</i>	0	0	0	2	3	2	2
<i>Cldn30</i>	0	0	0	4	3	4	4
<i>Cldn31</i>	0	0	0	2	3	1	2
<i>Cldn32</i>	0	0	0	2	2	2	2
<i>Cldn33</i>	0	0	0	3	1	3	2
<i>Cldn34</i>	0	0	0	0	1	0	1
<i>Total</i>	24	26	14	52	43	56	55

花鲈的 55 个 *cldns* 基因分布于 16 条染色体上, 染色体 5 和染色体 12 分布着多个 *cldns* 串联重复基因(图 1)。如图 2 所示, 两基因簇分别与人类的 *cldn3* 和 *cldn4* 具有较高的同源性。此外, 3 号、8 号、16 号、18 号和 20 号

染色体分别包含 1 个 *cldn* 基因, 15 号、17 号和 23 号染色体分别包含 2 个 *cldn* 基因, 10、11、19 号染色体各含有 3 个 *cldn* 基因, 而 2 号染色体包含 4 个 *cldn* 基因。花鲈的 *cldn18c* 定位于 scaffold524 上, 未挂载到染色体上(图 1)。

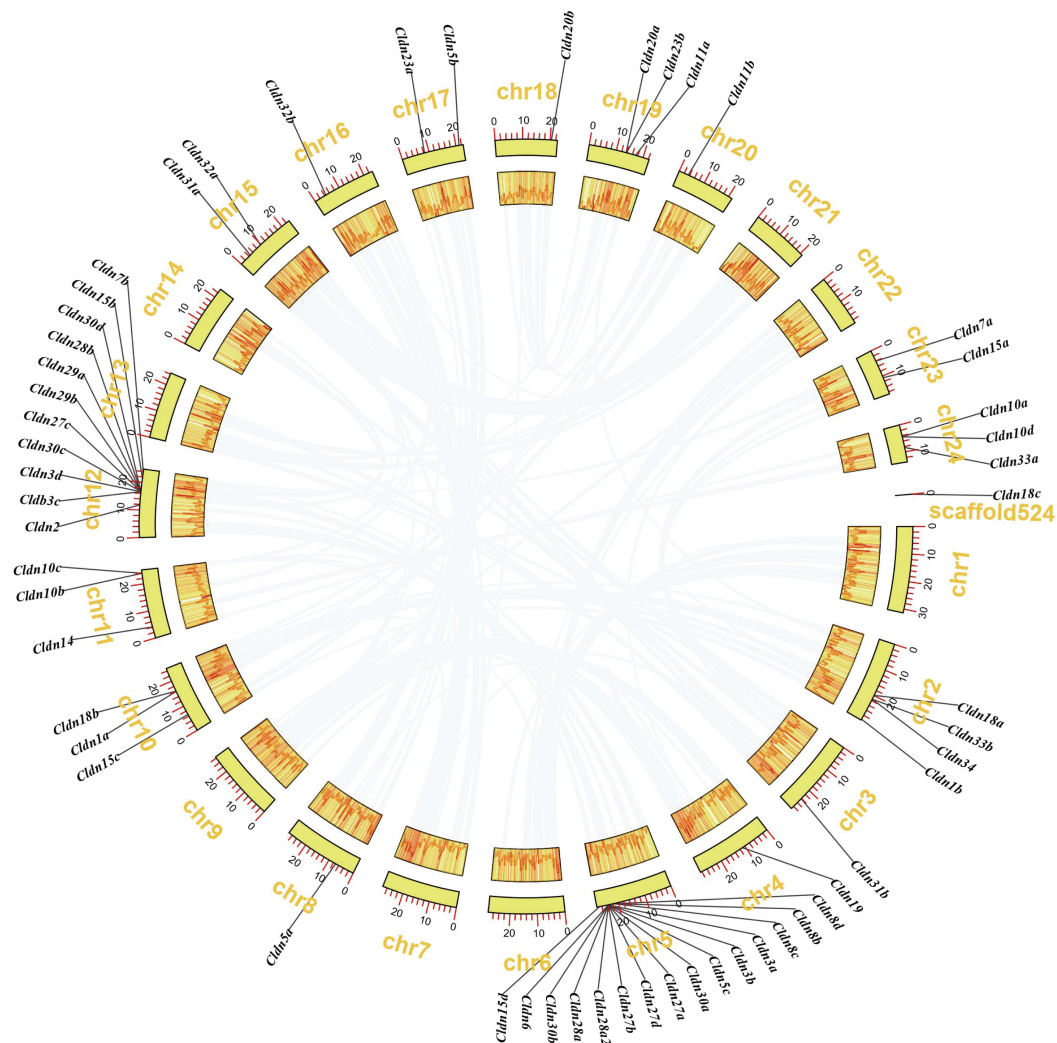


图 1 花鲈 *cldns* 基因的染色体分布情况

Fig. 1 Chromosomal locations of *cldn* genes of spotted sea bass

注：黑色线条代表每个 *cldn* 基因在染色体上的位置，基因组所有基因的共线性关系由灰色线条表示。橙色外环代表每个染色体上的基因密度

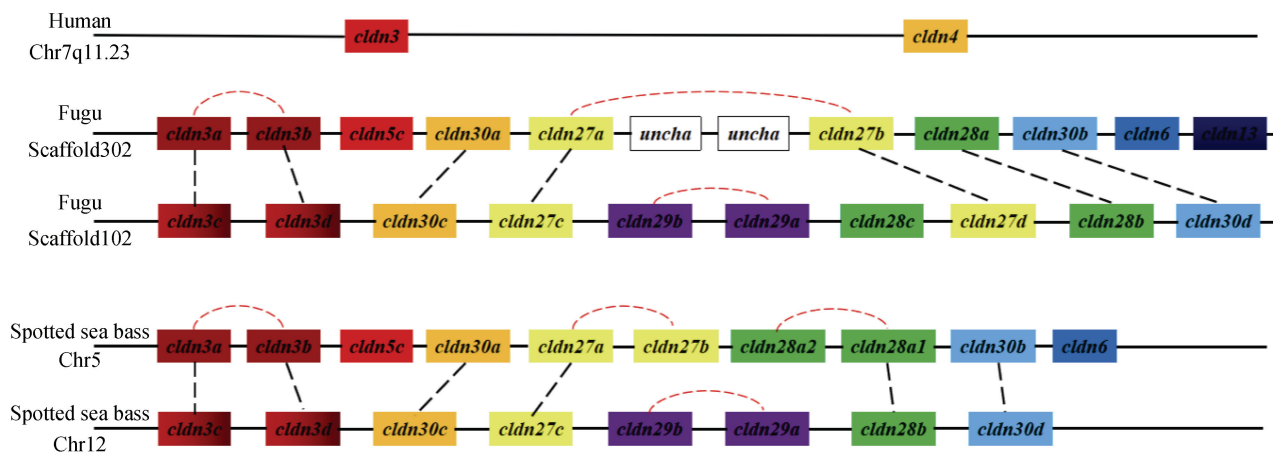


图 2 人 *cldn3*、*cldn4* 及红鳍东方鲀和花鲈 *cldn3*-like、*cldn4*-like 基因簇在染色体上的定位和位置关系

Fig. 2 The order of duplication of the claudin 3- and 4-like genes in fugu and spotted sea bass

注：红色虚线表示染色体内的基因复制，黑色虚线表示染色体片段间的基因复制

2.3 系统发育分析

选取人、鼠、红鳍东方鲀、斑马鱼以及花鲈共 202 条 *cldns* 氨基酸序列构建系统进化树。如图 3 所示, 这些基因共聚为 7 组进化枝。总的来说, 大多数花鲈 *cldns* 基因与已知其他物种的同类基因聚在一起。共有 34 个花鲈 *cldns* 基因与哺乳动物 *cldns* 基因同源, 其余 21 个 *cldns* 是硬骨鱼所特有^[10]。其中, 花鲈 *cldn1* 具有两个拷贝(*cldn1a* 和 *cldn1b*), 由图 3 可

知, 二者的系统发育关系较远。因此, 为进一步证实花鲈 *cldn1* 基因注释的准确性, 我们进行了共线性分析(图 4), 结果显示花鲈 *cldn1a* 与红鳍东方鲀的 *cldn1* 具有高度保守的相邻基因, 而由于 *cldn1* 在红鳍东方鲀为单拷贝, 我们选取双色雀鲷(*Eupomacentrus partitus*)与花鲈的 *cldn1b* 进行共线性分析, 结果证明花鲈 *cldn1b* 与其 *cldn1-like* 具有相似的邻近基因, 证实了花鲈 *cldn1* 基因的命名的准确性。

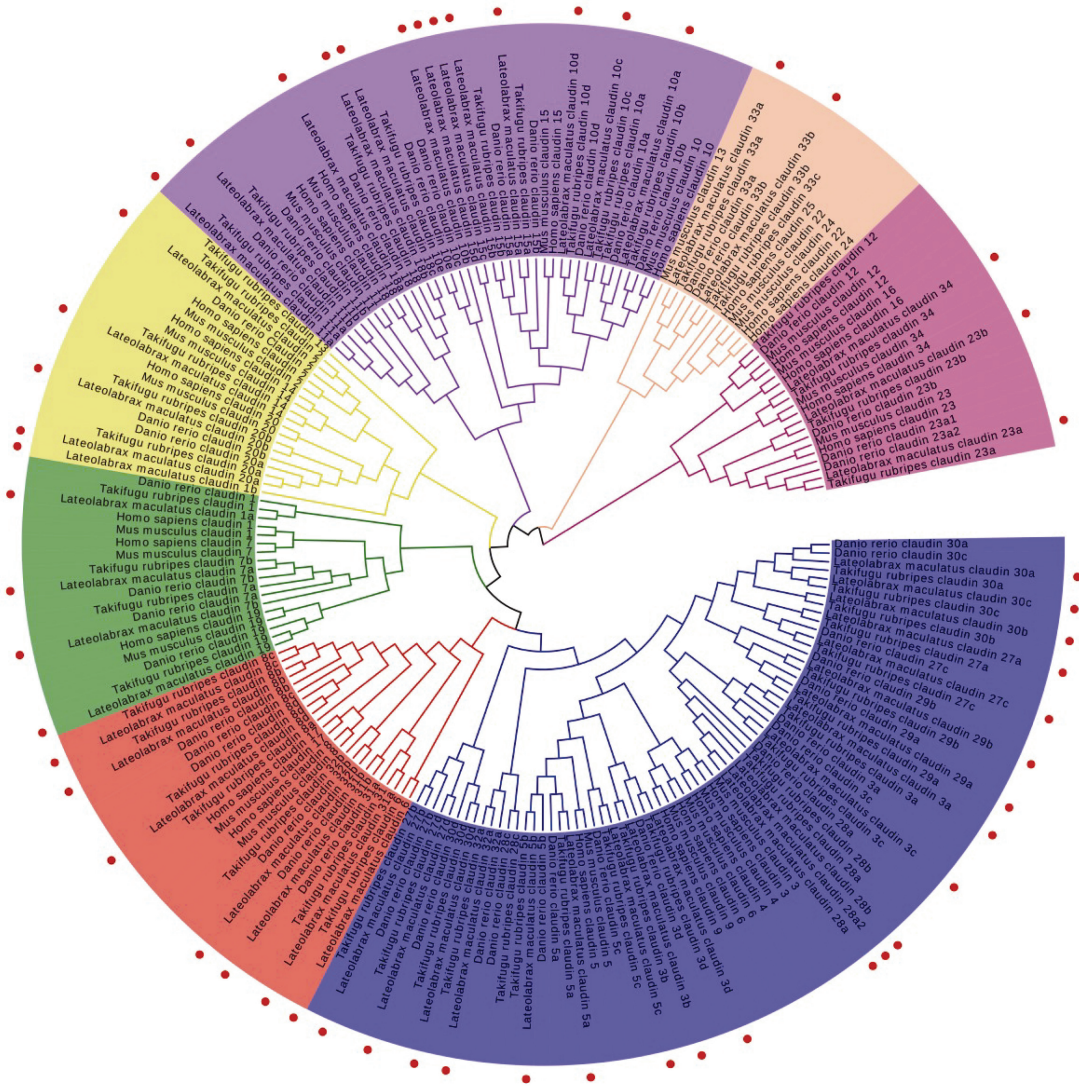


图 3 *cldns* 基因家族各成员系统进化树

Fig. 3 Phylogenetic tree of members of *cldn* genes family

注: 红点表示鉴定出的花鲈 *cldns* 基因, 不同颜色表示不同的进化枝

2.4 花鲈 *cldns* 基因的选择压力分析

将花鲈 55 条 *cldns* 氨基酸序列构建物种内的进化树, 依据聚类关系将 55 条 *cldns* 共分为 4 个大组: Class I 包含 *cldn3-like* 和 *cldn4-like* 等多个旁系同源

基因, 多为硬骨鱼较高等脊椎动物所特有的基因; Class II 包括 *cldn1*、*cldn7*、*cldn10*、*cldn15* 和 *cldn19*; Class III 包括 *cldn1*、*cldn2*、*cldn14*、*cldn20*; Class IV 包括 *cldn11*、*cldn18*、*cldn23*、*cldn33*、*cldn34*(图 5)。

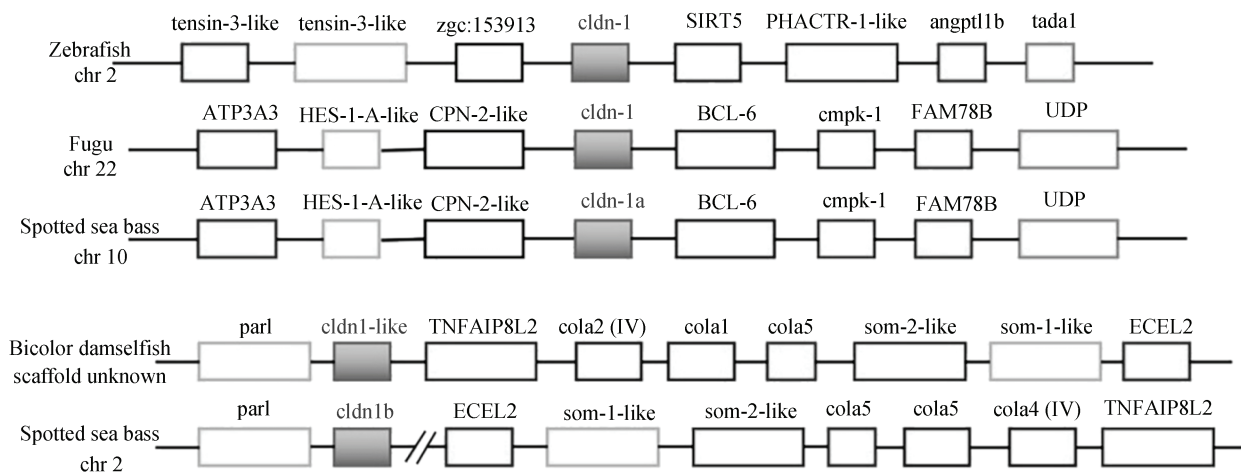


图4 *cldn1* 基因共线性分析
Fig. 4 Syntenic analysis of *cldn1*

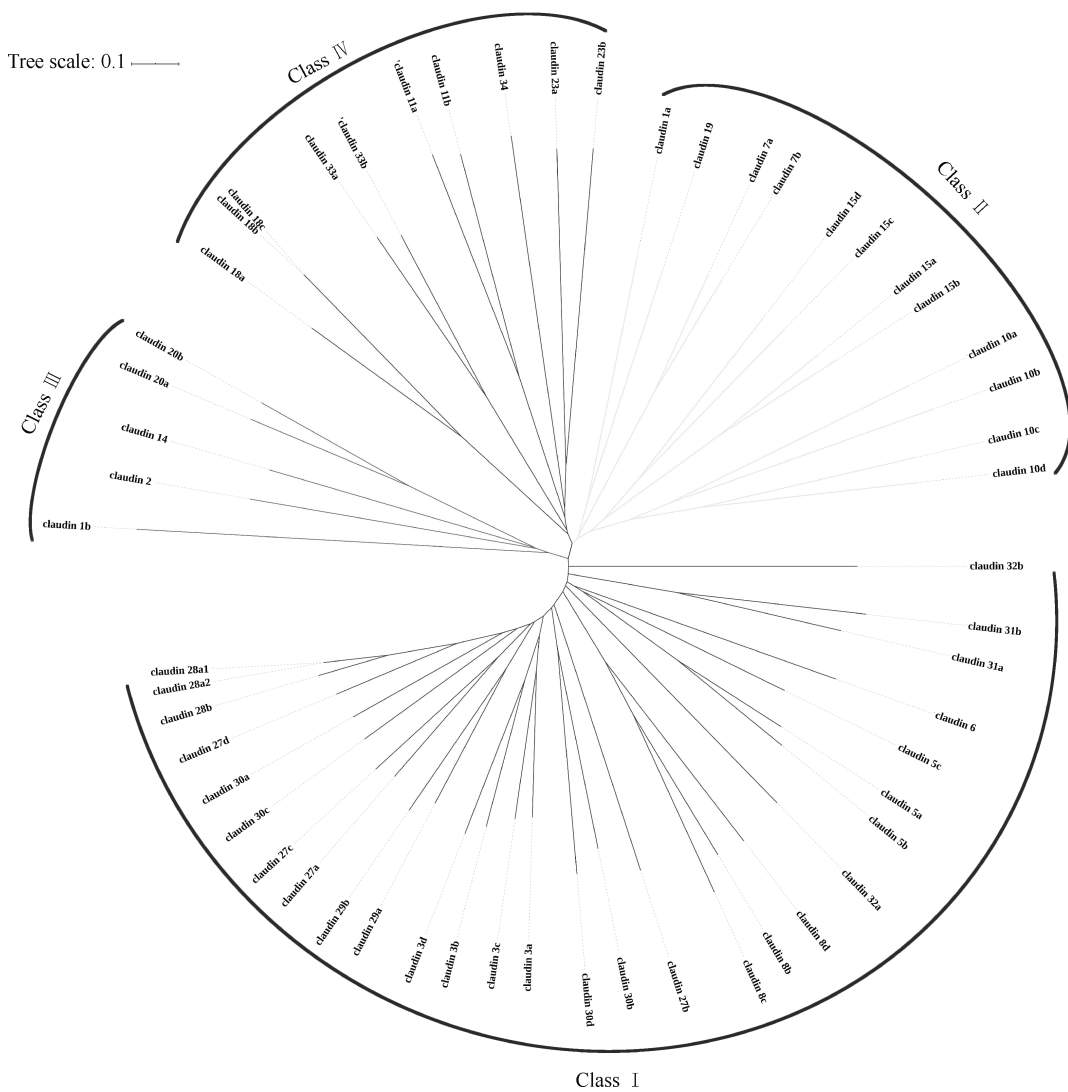


图 5 花鲈 cldns 系统进化树

Fig. 5 Phylogenetic tree of members of cldns in spotted sea bass

采用枝位点模型检测不同分支上不同位点受到的选择压力。结果如表 3 所示：将 Class I、Class II、Class III 和 Class IV 四大枝分别作为前景枝，每次计算时将其他分支标记为背景枝，结果表明，除 Class II 以外均检测到受正选择作用的氨基酸位点，在 Class I 分枝中，有 6.38% 的位点受到正选择压力，139Y 为正选择位点 ($P>0.95$)；在 Class IV 的分支上，

115D 在 $P>0.95$ 这个置信水平上被认为是受正选择的氨基酸位点；在 Class III 的分支上，在 $P>0.95$ 的置信水平上在蛋白的跨膜区域有 14 个位点被检测出来：92A、93S、100G、102K、105K、115D、117I、135A、136V、137S、139Y、140A、143V、174A。但就整体而言，四大枝的 ω 值仍远小于 1，处于负选择。

表 3 支位点模型的参数估计与似然比检验

Tab. 3 Parameter estimation and likelihood ratio test of Branch-site model

进化枝	模型比较	参数估计	似然比检验	正选择位点
Class I	Model A (np=111) vs Model A null (np=110)	$P_0=0.748\ 80$	0.000 157 094	139Y ($P>0.95$)
		$P_1=0.187\ 40$		
		$(P_{2a}+P_{2b}=0.063\ 8)$		
		$\omega_0=0.207\ 89$		
Class II	Model A (np=111) vs Model A null (np=110)	$(\omega_1=1.000\ 00)$	0.379 410 220	None
		$\omega_2=1.000\ 00$		
		$P_0=0.799\ 87$		
		$P_1=0.200\ 13$		
Class III	Model A (np=111) vs Model A null (np=110)	$(P_{2a}+P_{2b}=0.000\ 00)$	0.161 845 675	92A 93S 100G 102K 105K 115D 117I 135A 136V 137S 139Y 140A 143V 174A ($P>0.95$)
		$\omega_0=0.204\ 97$		
		$(\omega_1=1.000\ 00)$		
		$\omega_2=1.000\ 00$		
Class IV	Model A (np=111) vs Model A null (np=110)	$P_0=0.385\ 44$	0.003 374 406	115D ($P>0.95$)
		$P_1=0.096\ 42$		
		$(P_{2a}+P_{2b}=0.518\ 14)$		
		$\omega_0=0.209\ 02$		
		$(\omega_1=1.000\ 00)$		
		$\omega_2=1.000\ 00$		
		$P_0=0.658\ 92$		
		$P_1=0.164\ 92$		
		$(P_{2a}+P_{2b}=0.176\ 15)$		
		$\omega_0=0.206\ 93$		
		$(\omega_1=1.000\ 00)$		
		$\omega_2=1.000\ 00$		

2.5 花鲈鳃组织中 *cldns* 基因在不同盐度下的表达模式分析

对不同盐度条件下的花鲈鳃组织的转录组数据进行 Meta 分析，结果如表 4 所示，*cldns* 基因家族中至少有 24 个成员在鳃组织中表达，共有 6 个 *cldns* 基因在不同的盐度实验组中呈现出显著差异表达：*cldn8c*、*cldn10b*、*cldn10c*、*cldn10d*、*cldn27a* 和 *cldn30b*。其中，*cldn8c* 和 *cldn27a* 在淡水中的表达量显著高于等渗水、海水中的表达量；*cldn10b* 和 *cldn30d* 的表达量关系为：海水>等渗水>淡水；*cldn10c*、*cldn10d* 与 *cldn10b* 在淡水中的表达量显著高于海水，其在淡水和等渗水中也存在显著差异。研究结果表明：这 6 个 *cldns* 基因在不同的盐度环境下差异表达，可能参与花鲈鳃组织渗透调节，具备离子屏障等功能。

3 讨论

Cldns 是细胞间紧密连接蛋白的重要功能和结构组件，在调节细胞旁路通透性方面起着至关重要的作用^[8]。本研究在花鲈全基因组范围内共鉴定出 55 个 *cldns* 基因，并对其进行了系统发育分析、基因拷贝数分析、染色体定位、进化分析以及鳃组织 *cldns* 在不同盐度环境下的表达特征分析。与多数硬骨鱼的研究相似^[10, 13, 14, 25-27]，相较于高等脊椎动物，本研究中 *cldns* 家族在花鲈中也出现显著扩增，为理解硬骨鱼的基因进化和全基因组复制提供了一个理想模型。*cldns* 在硬骨鱼中的注释较为复杂，如在斑马鱼中，与人类有直系同源关系的基因以数字后缀命名(*cldn-1*、*cldn-5* 等)，而斑马鱼相比于高等脊椎动物所特有的基因则以字母后缀命名(*cldn-d*、

表 4 花鲈鳃中 *cldns* 基因在不同盐度下的差异表达

Tab. 4 Differentially expression levels of *cldn* genes in spotted sea bass following experimentally challenged with different salinities

基因	盐度		
	淡水组/等渗组	海水组/等渗组	淡水组/海水组
<i>cldn1a</i>	0.361	-0.185	0.546
<i>cldn1b</i>	0.036	-0.378	0.414
<i>cldn2</i>	—	—	—
<i>cldn3a</i>	—	—	—
<i>cldn3b</i>	—	—	—
<i>cldn3c</i>	—	—	—
<i>cldn3d</i>	—	—	—
<i>cldn5a</i>	0.315	-0.339	0.654
<i>cldn5b</i>	0.326	-0.399	0.725
<i>cldn5c</i>	—	—	—
<i>cldn6</i>	—	—	—
<i>cldn7a</i>	—	—	—
<i>cldn7b</i>	0.311	0.139	0.172
<i>cldn8b</i>	-0.512	0.158	-0.670
<i>cldn8c</i>	1.372*	0.626	0.746
<i>cldn8d</i>	0.149	0.048	0.101
<i>cldn10a</i>	—	—	—
<i>cldn10b</i>	0.847	-0.578	1.425*
<i>cldn10c</i>	1.061*	-0.036	1.097*
<i>cldn10d</i>	1.330*	0.228	1.103*
<i>cldn11a</i>	0.007	-0.297	0.304
<i>cldn11b</i>	—	—	—
<i>cldn14</i>	—	—	—
<i>cldn15a</i>	—	—	—
<i>cldn15b</i>	—	—	—
<i>cldn15c</i>	—	—	—
<i>cldn15d</i>	—	—	—
<i>cldn18a</i>	—	—	—
<i>cldn18b</i>	—	—	—
<i>cldn18c</i>	—	—	—
<i>cldn19</i>	-0.190	-0.393	0.204
<i>cldn20a</i>	—	—	—
<i>cldn20b</i>	—	—	—
<i>cldn23a</i>	0.021	0.304	-0.283
<i>cldn23b</i>	—	—	—
<i>cldn27a</i>	1.318*	0.577	0.741
<i>cldn27b</i>	-0.003	-0.110	0.107
<i>cldn27c</i>	—	—	—

续表

基因	盐度		
	淡水组/等渗组	海水组/等渗组	淡水组/海水组
<i>cldn27d</i>	—	—	—
<i>cldn28a1</i>	—	—	—
<i>cldn28a2</i>	0.415	-0.505	0.920
<i>cldn28b</i>	—	—	—
<i>cldn29a</i>	—	—	—
<i>cldn29b</i>	—	—	—
<i>cldn30a</i>	0.417	-0.150	0.567
<i>cldn30b</i>	0.789	-0.452	1.241*
<i>cldn30c</i>	0.264	-0.118	0.382
<i>cldn30d</i>	-0.682	-0.187	-0.495
<i>cldn31a</i>	0.017	-0.117	0.134
<i>cldn31b</i>	—	—	—
<i>cldn32a</i>	0.342	0.102	0.240
<i>cldn32b</i>	—	—	—
<i>cldn33a</i>	—	—	—
<i>cldn33b</i>	0.388	-0.141	0.530
<i>cldn34</i>	—	—	—

注: 数值代表 $\log_2$ (差异表达倍数); *: 显著差异表达基因

cldn-e 等)^[25]。在红鳍东方鲀的相关研究中, 所有的 *cldns* 基因均以数字后缀命名^[9]。Baltzegar 等人在 2013 年对斑马鱼 *cldns* 基因的注释根据系统发育关系进行了修订^[2]。本文通过与斑马鱼和红鳍东方鲀的同源关系对花鲈的 *cldns* 进行注释, 并通过系统发育关系进行验证。对于产生扩增且系统发育关系聚类不好的基因拷贝(如 *cldn1*)通过共线性分析对基因的注释进行进一步验证。此外, 位于 5 号染色体上的 *cldn28a* 通过串联重复产生了一个旁系同源物, 二者所编码的蛋白及理化性质完全一致, 我们参考斑马鱼的注释原则, 在本文中将其注释为 *cldn28a1* 和 *cldn28a2*。

在高等哺乳动物的基因组中, 共发现 27 种 *cldns*^[28], 而在硬骨鱼中, 已经有 63 个 *cldns* 在 16 种硬骨鱼中被发现。在本次研究中, 共获得 55 个花鲈的 *cldns*。值得注意的是, 仅在硬骨鱼中发现了 *cldns* 直系同源基因的多拷贝现象, 表明硬骨鱼 *cldns* 基因相较于高等脊椎动物的显著扩增与硬骨鱼的全基因组复制(WGD)有密切的关系。据报道, 这次硬骨鱼特异的全基因组复制事件是导致硬骨鱼物种多样化的重要原因^[29]。而全基因组复制事件后部分基因丢失

而硬骨鱼的 *cldns* 基因数量却是高等脊椎动物的两倍还多, 这表明单是 WGD 不足以解释硬骨鱼中 *cldns* 基因的扩增, 例如本研究中 5 号和 12 号染色体上出现的多个 *cldns* 基因形成的基因簇解释了该现象。在基因复制事件后, 产生的新基因存在 3 种情况: 1) 丢失(例如成为假基因或沉默基因); 2) 关键通路的效能需增加一倍; 3) 有足够的环境选择压力来推动新功能的发展, 新基因保存下来并获得新的功能^[30-31]。为此, 我们对花鲈的 *cldns* 基因进行了选择压力分析, 枝模型和位点模型的计算结果表明, *cldns* 基因家族整体主要受负选择。枝位点模型的计算结果表明, 在不同的主要分枝上, *cldns* 蛋白的某些氨基酸位点受正选择作用。所筛选到的正选择位点主要分布在 *cldns* 蛋白跨膜区域, 这些位点上的正选择压力可能是不同 *cldns* 基因的组织表达特异性和功能多样性的重要原因。

由于鳃组织与周围水环境直接接触, 因此其对维持硬骨鱼体内的稳态至关重要。几十年来, 人们普遍认为鳃上皮的不同的旁细胞特性对鳃组织的功能至关重要^[12], 但是鳃组织紧密连接复合蛋白的分子生理学直到近几年才成为人们关注的焦点^[1]。*cldns* 基

因在鳃组织中的重要性已经在多个硬骨鱼中进行探讨^[10, 32], 并在多中硬骨鱼的鳃组织中找到至少 44 个 *cldns*^[1]。一些研究表明, 当青斑河豚和双斑鲳由海水转入淡水的低渗环境后, *cldn3* 和 *cldn8* 亚型的 mRNA 表达量上调, 表明该 *cldns* 的功能可能是形成细胞间屏障以限制体内离子的流失^[16, 33]。本研究中, 我们同样发现花鲈 *cldn8c* 在淡水中的表达量显著高于等渗水中, 且高于其在海水中的表达量; 此外, *cldn30d* 在淡水中的表达量显著高于其在海水中的表达量, 且表达量在海水、等渗水和淡水中逐级上调, 表明 *cldn30d* 在低渗环境下同样起到细胞间“屏障”功能, Engelund 等人对大西洋鲑 *cldn30* 的研究结果与上述结果一致, 其将大西洋鲑的 *cldn30* 转染至哺乳动物的肾脏细胞系后降低了上皮细胞的电导率和一价阳离子的细胞旁通路渗透性^[34]。对斑马鱼和大西洋鲑的研究表明, 在高渗环境下, *cldn10* 的 mRNA 水平上调, 作为离子细胞(MRCs)和附属细胞(AC)间的“泄漏孔洞”使体内多余的 Na⁺排出体外^[13-14, 25], 而在本研究中却相反, *cldn10b*、*cldn10c* 和 *cldn10d* 在高渗环境中的表达较低, 在淡水中的表达显著高于海水中的表达, 且未在鳃组织中发现高渗环境显著上调的 *cldns* 基因。推测由于各物种应对环境盐度变化的响应机制具有物种特异性, 以及实验处理的条件不同、实验用鱼的规格不同所导致。关于花鲈的 *cldns* 基因在应对环境盐度变化的具体功能, 还需进一步深入研究。

4 结论

本研究在花鲈全基因组范围内共鉴定获得 55 个 *cldns* 基因, 通过系统发育和共线性分析确保了基因注释的准确性; 基因拷贝数和染色体定位分析证实硬骨鱼 *cldns* 基因家族扩增为全基因组复制和串联重复事件共同作用的结果; 选择压力分析表明花鲈 *cldns* 基因家族整体处于负选择, 但在蛋白的跨膜区域存在 14 个受环境正选择作用的氨基酸位点, 这些位点上的正选择压力可能是 *cldns* 基因多样性和功能分化的重要依据; 此外, 在不同的盐度环境下, 鳃组织中共有 6 个 *cldns* 基因呈现显著的差异表达, 表明它们在应对盐度环境变化时发挥潜在作用。

参考文献:

[1] CHASIOTIS H, KOLOSOV D, BUI P, et al. Tight junc-

- tions, tight junction proteins and paracellular permeability across the gill epithelium of fishes: a review[J]. *Respiratory Physiology & Neurobiology*, 2012, 184(3): 269-281.
- [2] BALTZEGAR D A, READING B J, BRUNE E S, et al. Phylogenetic revision of the claudin gene family[J]. *Marine Genomics*, 2013, 11: 17-26.
- [3] VAN ITALLIE C M, ANDERSON J M. Claudins and epithelial paracellular transport[J]. *Annual Review of Physiology*, 2006, 68: 403-429.
- [4] MORIN P J. Claudin proteins in human cancer: promising new targets for diagnosis and therapy[J]. *Cancer Research*, 2005, 65(21): 9603-9606.
- [5] 雷军, 夏祥国. 紧密连接蛋白 claudin 的研究进展[J]. *泸州医学院学报*, 2009(6): 655-657.
LEI Jun, XIA Xiangguo. Research progress of tight junction protein Claudin[J]. *Journal of Luzhou Medical College*, 2009(6): 655-657.
- [6] FURUSE M, SASAKI H, TSUKITA S. Manner of interaction of heterogeneous claudin species within and between tight junction strands[J]. *The Journal of Cell Biology*, 1999, 147(4): 891-903.
- [7] ASANO A, ASANO K, SASAKI H, et al. Claudins in *Caenorhabditis elegans*: their distribution and barrier function in the epithelium[J]. *Current Biology*, 2003, 13(12): 1042-1046.
- [8] BEHR M, RIEDEL D, SCHUH R. The claudin-like megatrachea is essential in septate junctions for the epithelial barrier function in *Drosophila*[J]. *Developmental Cell*, 2003, 5(4): 611-620.
- [9] KOLOSOV D, BUI P, CHASIOTIS H, et al. Claudins in teleost fishes[J]. *Tissue Barriers*, 2013, 1(3): e25391.
- [10] LOH Y H, CHRISTOFFELS A, BRENNER S, et al. Extensive expansion of the claudin gene family in the teleost fish, *Fugu rubripes*[J]. *Genome Research*, 2004, 14(7): 1248-1257.
- [11] HWANG P P, LEE T H, LIN L Y. Ion regulation in fish gills: recent progress in the cellular and molecular mechanisms[J]. *American Journal of Physiology. Regulatory, Integrative and Comparative Physiology*, 2011, 301(1): 28-47.
- [12] SARDET C. The surface epithelium of teleostean fish gills. Cellular and junctional adaptations of the chloride cell in relation to salt adaptation[J]. *The Journal of Cell Biology*, 1979, 80(1): 96-117.
- [13] TIPSMARK C K, KIILLERICH P, NILSEN T O, et al. Branchial expression patterns of claudin isoforms in Atlantic salmon during seawater acclimation and smoltification[J]. *American Journal of Physiology Regulatory Integrative & Comparative Physiology*, 2008, 294(5): 1563-1574.

- [14] TIPSMARK C K, BALTZEGAR D A, OZDEN O, et al. Salinity regulates claudin mRNA and protein expression in the teleost gill[J]. *American Journal of Physiology Regulatory Integrative & Comparative Physiology*, 2008, 294(3): 1004-1014.
- [15] BAGHERIE-LACHIDAN M, WRIGHT S I, KELLY S P. Claudin-8 and -27 tight junction proteins in puffer fish *Tetraodon nigroviridis* acclimated to freshwater and seawater[J]. *Journal of Comparative Physiology B*, 2009, 179(4): 419-431.
- [16] DUFFY N M, BUI P, BAGHERIE-LACHIDAN M, et al. Epithelial remodeling and claudin mRNA abundance in the gill and kidney of puffer fish (*Tetraodon biocellatus*) acclimated to altered environmental ion levels[J]. *Journal of Comparative Physiology B Biochemical Systemic & Environmental Physiology*, 2011, 181(2): 219.
- [17] MARSHALL W S, BREVES J P, DOOHAN E M, et al. claudin-10 isoform expression and cation selectivity change with salinity in salt-secreting epithelia of *Fundulus heteroclitus*[J]. *The Journal of Experimental Biology*, 2018, 221(1): jeb168906.
- [18] XU J, LI J T, JIANG Y, et al. Genomic basis of adaptive evolution: The survival of Amur ide (*Leuciscus waleckii*) in an extremely alkaline environment[J]. *Molecular Biology & Evolution*, 2017, 34(1): 145-159.
- [19] SUDHIR K, GLEN S, KOICHIRO T. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets[J]. *Molecular Biology & Evolution*, 2016, 33(7): 1870-1874.
- [20] WANG Y P, TANG H B, DEBARRY J D, et al. MCS-canX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity[J]. *Nucleic Acids Research*, 2012, 40(7): e49.
- [21] CHEN C, CHEN H, ZHANG Y, et al. TBtools: An Integrative toolkit developed for interactive analyses of big biological data[J]. *Molecular Plant*, 2020, 13(8): 1194-1202.
- [22] YANG Z H. PAML 4: phylogenetic analysis by maximum likelihood[J]. *Molecular Biology and Evolution*, 2007, 24(8): 1586-1591.
- [23] GAO F, CHEN C, ARAB D A, et al. EasyCodeML: A visual tool for analysis of selection using CodeML[J]. *Ecology and Evolution*, 2019, 9(7): 3891-3898.
- [24] BIELAWSKI J P, YANG Z. Maximum likelihood methods for detecting adaptive evolution after gene duplication[J]. *Journal of Structural and Functional Genomics*, 2003, 3(1): 201-212.
- [25] KOLLMAR R, NAKAMURA S K, KAPPLER J A, et al. Expression and phylogeny of claudins in vertebrate primordia[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2001, 98(18): 10196-10201.
- [26] MUKENDI C, DEAN N, LALA R, et al. Evolution of the vertebrate claudin gene family: insights from a basal vertebrate, the sea lamprey[J]. *The International Journal of Developmental Biology*, 2016, 60(1/3): 39-51.
- [27] SUN L Y, LIU S K, BAO L S et al. Claudin multigene family in channel catfish and their expression profiles in response to bacterial infection and hypoxia as revealed by meta-analysis of RNA-Seq datasets[J]. *Comparative Biochemistry & Physiology Part D Genomics & Proteomics*, 2015, 13: 60-69.
- [28] MINETA K, YAMAMOTO Y, YAMAZAKI Y, et al. Predicted expansion of the claudin multigene family[J]. *FEBS Lett*, 2011, 585(4): 606-612.
- [29] HURLEY I A, MUELLER R L, DUNN K A, et al. A new time-scale for ray-finned fish evolution[J]. *Proceedings of the Royal Society B: Biological Sciences*, 2007, 274(1609): 489-498.
- [30] PASQUIER J, CABAU C, NGUYEN T, et al. Gene evolution and gene expression after whole genome duplication in fish: the PhyloFish database[J]. *BMC Genomics*, 2016, 17(1): 368.
- [31] TAYLOR J S, PEER Y V D, MEYER A. Genome duplication, divergent resolution and speciation[J]. *Trends in Genetics*, 2001, 17(6): 299-301.
- [32] BUI P, KELLY S P. Claudins in a primary cultured puffer fish (*Tetraodon nigroviridis*) gill epithelium[J]. *Methods in Molecular Biology*, 2011, 762: 179-194.
- [33] MAZDAK B L, WRIGHT S I, KELLY S P. Claudin-3 tight junction proteins in *Tetraodon nigroviridis*: cloning, tissue-specific expression, and a role in hydromineral balance[J]. *American Journal of Physiology Regulatory Integrative & Comparative Physiology*, 2008, 294(5): 1638-1647.
- [34] ENGELUND M B, YU A S L, LI J, et al. Functional characterization and localization of a gill-specific claudin isoform in Atlantic salmon[J]. *American Journal of Physiology Regulatory Integrative & Comparative Physiology*, 2012, 302(2): 300-311.

Identification and evolutionary analysis of Claudins gene family in spotted sea bass and its expression characteristics in gill tissue under different salinity

LI Jun-jie, YU Peng, LI Yun, TIAN Yuan, ZHANG Kai-qiang, WANG Hai-liang, QI Xin, WEN Hai-shen

(The Key Laboratory of Mariculture, Ministry of Education, Ocean University of China, Qingdao 266003, China)

Received: Feb. 3, 2021

Key words: *Claudins* gene family; *Lateolabrax maculatus*; salinity; evolutionary analysis

Abstract: Claudin (cldn) is an important functional and structural component of the intercellular tight junction protein, which is involved in the regulation of osmotic pressure balance by controlling the permeability of cellular bypass pathways. In the present study, *cldn* gene family was systematically identified and characterized in spotted sea bass (*Lateolabrax maculatus*). The results showed that there were 55 members in *cldns* gene family of spotted sea bass, which were distributed on 16 chromosomes. Among them, 34 members have their direct homologous genes in mammals, and the remaining 21 *cldns* genes seemed to be teleosts-specific. Selection pressure analysis showed that 14 amino acid sites were positively selected in the predicted transmembrane region of cldns protein, which may be associated with the gene diversity and functional differentiation within the family. Additionally, the expression levels of at least 24 *cldns* genes were detected in the transcriptome of gills in spotted sea bass and *cldn8c*、*cldn10b*、*cldn10c*、*cldn10d*、*cldn27a* and *cldn30b* showed significant expression differences between different salinity environment, suggesting their key roles in regulating osmotic pressure balance in gills of spotted sea bass. This study provides a theoretical basis for studying the osmotic regulation mechanism of *cldns* gene family in spotted sea bass and other teleosts.

(本文编辑: 杨 悦)