

长牡蛎繁殖期糖原含量与类胰岛素基因表达的相关性分析

苗校静^{1,2}, 邓少熙^{1,2}, 李良恒¹, 许飞^{1,3}

(1. 中国科学院海洋研究所实验海洋生物学重点实验室, 山东 青岛 266071; 2. 中国科学院大学, 北京 100049; 3. 中国科学院海洋大科学研究中心, 山东 青岛 266071)

摘要: 为探究长牡蛎在繁殖期间的糖原含量与类胰岛素基因(*cgMIP123*、*cgMIP4*、*cgMILP7*、*cgILP*)和相关转录调控因子(*cgPdx*)相对表达量的相关性, 自2020年5月至2020年10月采集了山东胶南养殖海区的长牡蛎, 测定了血糖含量、糖原含量、条件指数、类胰岛素基因相对表达量及环境因子(酸度、温度、盐度)等数据, 采用多元统计方法对数据进行分析。logistic回归分析结果显示, 长牡蛎配子排放前后血糖含量和内脏糖原含量具有显著差异。构建的回归模型可以通过血糖含量和内脏糖原含量准确判断配子是否已经排放, 区分度 C-index 为 0.903, Hosmer-Lemeshow 拟合优度检验 χ^2 值为 9.06, $P > 0.05$, 验证结果显示该模型可靠。多元线性回归分析结果显示: 条件指数与 *cgMIP123* 和 *cgMIP4* 基因的相对表达量具有相关性, R^2 为 0.91, P 值为 0.007 6, 极显著相关; 内脏糖原和唇瓣组织的糖原含量与 *ILPs* 和 *cgPdx* 相对表达量具有一定的相关性, 其中 *cgMILP7* 的相对表达量与内脏糖原和唇瓣组织的糖原含量呈负相关, *cgPdx* 相对表达量与唇瓣组织的糖原含量呈负相关, *cgMIP4* 和 *cgILP* 的相对表达量与糖原含量呈正相关。

关键词: 长牡蛎; 血糖; 类胰岛素; 胰十二指肠同源异型盒基因; 多元统计

中图分类号: S917.4 文献标识码: A 文章编号: 1000-3096(2022)10-0001-12

DOI: 10.11759/hyxx20210304002

类胰岛素超家族是动物中一个古老的基因家族^[1-2]。在脊椎动物中主要包括胰岛素(insulin, INS)、类胰岛素生长因子(insulin-like growth factor, IGF)、松弛素(relaxin, RLN)和脊椎动物类胰岛素(insulin-like peptide, INSL)。其中 INS 和 IGF 属于 insulin/IGF 家族, 主要参与生长发育、葡萄糖稳态的调节, 这些蛋白质的受体为酪氨酸激酶受体。RLN 和 INSL 属于类松弛素家族, 主要参与生殖和神经内分泌的调节, 受体为 G 蛋白偶联受体。无脊椎动物类胰岛素多具有门类特异性, 比如家蚕素(bombyxin)^[3], 贝类类胰岛素多肽(molluscan insulin-related peptides, MIP)^[4], 线虫类胰岛素多肽(caenorhabditis elegans insulin-like peptides, INS)^[5]等, 由于无脊椎动物门类多样性丰富, 基因在物种形成过程中多发生复制和功能分化, 相关功能尚缺乏系统研究, 推测这些基因在无脊椎动物中也主要参与生长发育、生殖和碳水化合物稳态等过程^[6-8]。

在长牡蛎基因组中共鉴定到 4 个类胰岛素基因: *cgMIP123*、*cgMIP4*、*cgMILP7* 和 *cgILP*^[9]。系统进化分析显示 *cgMIP123* 和 *cgMIP4* 属于贝类类胰岛素多肽 MIP; *cgMILP7* 与果蝇的 *DILP7* 同源性高, 推测可能具

有类似 relaxin 的功能; *cgILP* 与人类的胰岛素和 IGF 同源性较高。氨基酸序列分析将 *cgMIP123*、*cgMIP4*、*cgMILP7* 划分为 β 型, 冠轮动物的大多数 *ILPs* 均属于该类型。而 *cgILP* 则为 γ 型, 该类型的类胰岛素基因在脊索动物、扁形动物、辐射对称动物、海绵等广泛分布^[9]。表达模式分析表明 *cgMIP123* 在肝胰腺和内脏神经节中表达; *cgMIP4* 主要在幼虫发育时期表达, 在成体中的表达量很低; *cgILP* 在肝胰腺中高度表达。2002 年 Hamano 等报道的长牡蛎 *oIRP* 即为 *cgMIP123*, 其基因的季节性表达峰值与长牡蛎软体部生长和配子发育期相对应(3—5 月), 但与糖原积累期(10—11 月)

收稿日期: 2021-03-04; 修回日期: 2021-05-14

基金项目: 国家自然科学基金(41776152); 中国科学院海洋大科学研究中心重点部署项目(COMS2019Q11); 青岛海洋科学与技术试点国家实验室开放基金(QNLM2016ORP0408)

[Foundation: National Natural Science Foundation of China, No. 41776152; Key Deployment Project of Centre for Ocean Mega-Research of Science, Chinese Academy of Science, No. COMS2019Q11; Grant from the Pilot National Laboratory for Marine Science and Technology (Qingdao), No. QNLM2016ORP0408]

作者简介: 苗校静(1996—), 女, 陕西渭南人, 硕士研究生, 主要从事贝类遗传育种研究, E-mail: xiaojing_miao@163.com; 许飞(1982—), 通信作者, 主要从事贝类遗传育种研究, E-mail: xufei@qdio.ac.cn

不重叠^[10-11]。该基因的一些 SNP 位点与长牡蛎体尺性状和软体部重量等具有显著关联性^[12]。目前关于 cgMIP4、cgMILP7 和 cgILP 的功能分析尚欠缺。

本研究对自然海区生长的长牡蛎进行调查, 获得血糖、糖原含量、条件指数、类胰岛素基因相对表达量及环境因子(酸度、温度、盐度、叶绿素 *a* 含量)等数据。并通过 logistic 回归分析、多元线性回归分析等多元统计方法探究其相关性, 以期为解析长牡蛎类胰岛素基因功能及糖原物质积累的调控机制提供数据支持。

1 材料与方法

1.1 实验动物

所用实验动物为来自同一个养殖群体的二龄牡蛎, 处理为单体后集中挂养于山东省青岛市胶南山东头养殖海区。2020 年 5 月至 2020 年 10 月每月随机取回 100 个左右, 现场测定海水的温度、盐度、酸度(以 pH 计量), 同时收集养殖海域水样用于叶绿素 *a* 含量的测定。实际采样时间见表 1。

表 1 2020 年采样时间表

Tab. 1 Sampling time in 2020

取样序号	取样时间
1	5 月 20 日
2	6 月 12 日
3	6 月 30 日
4	7 月 27 日
5	8 月 17 日
6	9 月 14 日
7	10 月 20 日

表 2 引物序列

Tab. 2 Primer sequences

引物名称	基因 ID	引物序列
cgMIP123-qF	LOC105319754	5'-GCTTAGACTTGTTTCCGTAAT-3'
cgMIP123-qR		5'-CGCACGTTTCAACACATCATAG-3'
cgMIP4-qF	LOC105319755	5'-TTGCCGAAATTGGTGCAATAG-3'
cgMIP4-qR		5'-CTGACTGCACAGAGGTGAAA-3'
cgMILP7-qF	LOC105333726	5'-AACGGCAGCTATCCATTACC-3'
cgMILP7-qR		5'-GTCGACGTGTTGATCGAGTT-3'
cgILP-qF	LOC105347687	5'-GATCGATGCTCCAATGGTTAATG-3'
cgILP-qR		5'-TTCTGAGCCACTGCAAAC-3'
cgPdx-qF	LOC105327390	5'-CACAGGCAATTAAGTGAAGATGG-3'
cgPdx-qR		5'-CCAGGTACAGAGGCAGTTATG-3'
cgEF-qF	LOC105338957	5'-TCCACTGGCCATCTCATTTAC-3'
cgEF-qR		5'-TGAAAGAACCCTTTCCCATCTC-3'

1.2 样品处理

长牡蛎在水族室暂养一晚后, 即进行组织样品采集。随机挑选 30 个无寄居蟹等状况的健康牡蛎用于测定血糖、糖原含量及基因相对表达量。取样均在冰上进行低温操作。长牡蛎开壳后, 用注射器吸取血淋巴(hemolymph, Hem), 离心后取上清冻存, 用于血糖含量测定; 取唇瓣(labial Palps, L)、内脏团(统一取中肠所在的幽门凸起部位 pyloric process, P)、透明闭壳肌(transparent adductor muscle, T)剪碎后储存于强碱溶液中, 用于糖原含量的测定; 取脏神经节(visceral ganglion, VG)、肝胰腺组织(hepatopancreas, Hep)^[13]于 Trizol 中-20 °C 冻存, 用于后续基因相对表达量的测定, 每 3 个个体混合为一个样品。同时随机挑选 20 个健康长牡蛎, 清理贝壳后记录总重, 将软体部放入 55 °C 烘箱烘干 72 h, 称量其干重。

叶绿素 *a* 样本的获取: 用 20 μm 筛绢过滤海水除去大型浮游动植物, 取 0.5 L 海水经 5 μm 滤膜抽滤, 获取滤膜后吸干多余水分并避光暂存于放有冰袋的样品箱中, 运回实验室转移至-80 °C 冰箱。

1.3 实验方法

1.3.1 实时荧光定量 PCR

肝胰腺、神经组织采用 Trizol 法提取总 RNA, 测定浓度及凝胶电泳检测质量后进行去除基因组 DNA 的反转录(艾科瑞 AG11711)。实时荧光定量 PCR 采用诺唯赞(Q711-02)试剂盒在 QuantStudio 6 Flex Real-Time PCR System(ThermoFisher, USA)上进行。以 *cgEF* 为内参基因, 以 5 月 20 日样品作为对照组, 各基因的相对表达量计算公式参考相关文献^[14]。实验所用的引物信息见表 2。

1.3.2 血糖、糖原含量、条件指数及叶绿素 *a* 含量的测定

血糖、糖原含量的测定分别采用 SIGMA (GAGO20)和南京建成试剂盒(F006)并按照说明书进行操作。条件指数(condition index, C_i)是贝类软体部生长和肥满度的衡量指标,有多种计算方法,本研究以干质量/总质量表示^[15]。叶绿素 *a* 含量用 90%丙酮萃取,采用分光光度法测定。

1.3.3 数据分析

数据统计分析使用 R 语言(4.0.2)。

1.3.3.1 方差分析

采用非参数方差分析及 Wilcoxon 秩和检验(用 Benjamini & Hochberg 方法对 *P* 值进行矫正)对血糖含量、糖原含量、条件指数、各基因相对表达量的各月份间差异进行分析,血糖、糖原含量、条件指数、基因相对表达量每周期的生物学重复依次为 30、30、20、10。

1.3.3.2 logistic 回归分析

数据集的构建:采用 5 至 7 月观测个体的性腺状态(gonad condition, G_C , $n=94$)进行分析,以布尔值表示性腺状态(“0”为未排放,“1”为排放),未排放个体 $n_0=32$,排放个体 $n_1=62$ 。

表达式的构建与验证: G_C 与血糖(G)、闭壳肌(G_T)、内脏团(G_P)及唇瓣(G_L)糖原含量、性别(g)的关系表达式如下:

$$\text{logit}(G_C) \sim G + G_T + G_P + G_L + g, \quad (1)$$

其中, $\text{logit}(G_C)$ 表示对 G_C 进行 logit 变换,即 $\ln[P/(1-P)]$, P 代表配子已排放这一事件发生的概率, $P/(1-P)$ 定义为配子排放的优势比。针对建模组采用逐步回归法筛选出具有显著贡献的自变量构建模型,绘制列线图将回归模型可视化。对模型进行数据内部验证:用 Hosmer-Lemeshow test 和校准度曲线图(calibration curve plot)检验模型的拟合优度,即观测结果与预测结果的一致性;用 C-index 值验证模型的区分度,即预测模型区分两种结果的能力,在数学上 C-index 值与接受者操作特征曲线(ROC)曲线下面积(AUC 值)相等^[16]。

1.3.3.3 多元线性回归分析

数据集的构建:①计算每次取样日期与 5 月 20 日的间隔天数形成数据集 time;②参考荧光定量 PCR 的混样方案,将相应个体的生理指标取平均值,从而形成生理指标与基因表达数据的一一对应关系;③由于条件指数(C_i)与基因检测样本来源于不同个体,因此采用各变量在各月份的平均值进行分析。

表达式的构建:以时间(t)、各基因相对表达量作为自变量对 C_i 、 G 、 G_T 、 G_P 和 G_L 分别建立回归模型,以环境因子[叶绿素 *a* 含量、pH 值、盐度(S)、温度(T)]为自变量对各基因相对表达量、 C_i 、 G 、 G_T 、 G_P 和 G_L 分别建立回归模型,探索变量间的相关性。利用逐步回归法筛选出 AIC 值(the Akaike information criterion, 赤池信息准则)最小的模型,并对其进行正态性、独立性、线性、同方差性、多重共线性检验,计算出各表达式自变量的相对权重^[17]。因变量与自变量的关系表达式如下:

$$C_i/G/G_T/G_P/G_L \sim t + A + B + C + D + E \quad (2)$$

$$E/C_i/G/G_T/G_P/G_L \sim F + H + S + T. \quad (3)$$

其中, A : *cgMIP123* 基因相对表达量; B : *cgMIP4* 基因相对表达量; C : *cgMIP7* 基因相对表达量; D : *cgPdx* 基因相对表达量; E : *cgILP* 基因相对表达量; F : *Chla* 含量; H : pH 值; S : 盐度; t : 时间; T : 温度。

2 结果与分析

2.1 牡蛎繁殖期海区环境因子的变化

采样海区在 2020 年 5 月 20 日至 10 月 20 日期间盐度(31.19~29.11)和 pH(7.98~8.38)变化不大,温度呈季节性变化,8 月 17 日达到最高温 28.14 °C(图 1a),由于 8 月份降雨较多,海区盐度和 pH 出现小幅度变化。海区微型藻类(5~20 μm)叶绿素 *a* 含量在 6 月和 7 月均处于较高水平(图 1b)。通过观察样品性腺状态,推测排放期开始于 6 月 12 日左右。

2.2 血糖含量的变化

对血糖含量进行正态分布和方差齐性检验, P 值均小于 0.05,表明血糖含量数据不符合方差分析的要求,因此进行非参数方差分析。结果表明,各月份血糖含量具有极显著差异($P < 0.01$)。Wilcoxon 秩和检验结果如图 2 所示,长牡蛎血糖含量在 2020 年 5 月 20 日至 7 月 27 日期间呈下降趋势,7 月 27 日至 9 月 14 日期间呈上升趋势,10 月 20 日时出现轻微下降,但与 8 月 17 日和 9 月 14 日的血糖数据仍被划分为一组;7 月 27 日测得的平均血糖含量最低,5 月 20 日测得的平均血糖含量最高。

2.3 各组织糖原含量的变化

各组织糖原含量数据均不符合方差分析要求,因此进行非参数方差分析。结果表明,长牡蛎各组织糖原含量具有极显著差异($P < 0.01$)。其中闭壳肌的糖原含量远低于唇瓣和内脏团。不同月份间的糖原含量也具有极显著差异($P < 0.01$)。在 5 月 20 日至 7 月 27 日期间唇瓣

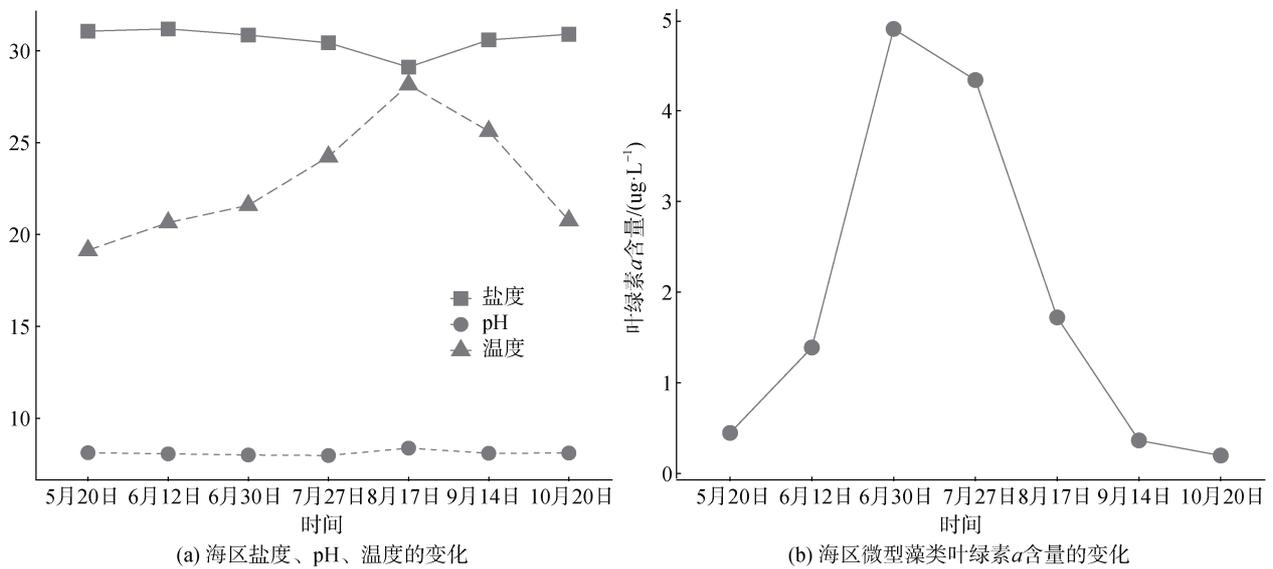


图 1 采样海域各月份环境指标变化
Fig. 1 Variation of environmental parameters around the sampling sea region

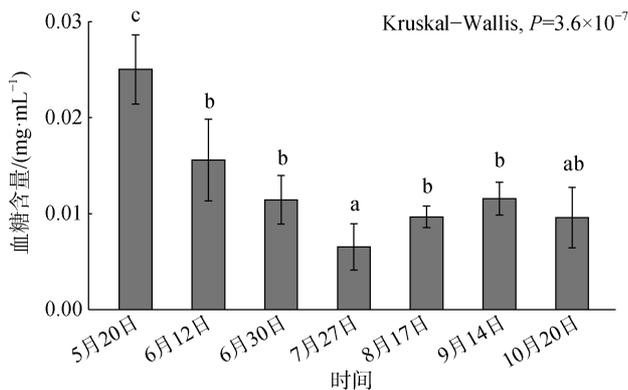


图 2 各月份长牡蛎血糖含量变化
Fig. 2 Seasonal variations of hemolymph glucose content in *Crassostrea gigas*
注: 不同字母表示差异具有显著性($P < 0.05$)

组织的糖原含量显著高于内脏团, 且二者均呈下降趋势。8月17日之后, 唇瓣和内脏团开始逐渐积累糖原, 且二者水平相当。闭壳肌糖原含量变化趋势与唇瓣和内脏团相反, 在7月27日的样品中含量最高(图3)。

2.4 条件指数的变化

长牡蛎条件指数数据不符合方差分析要求, 因此进行非参数方差分析。结果表明, 条件指数的变化趋势与糖原和血糖类似, 在5月20日至7月27日期间快速下降, 之后开始缓慢增长。6月12日的数据分布较为分散, 很可能是由于繁殖期取样个体性腺排放状态不一致造成的。

2.5 基因表达变化

非参数方差分析和 Wilcoxon 秩和检验结果表明, 牡蛎类胰岛素基因各月份间的表达具有显著差异 P 均小于 0.05, *cgMIP123* 在 5 月 20 日的相对表达量最高, 6 月 30 日最低; *cgMIP4* 和 *cgMILP7* 相对表达量无明显季节性变化; *cgILP* 在 2020 年 8 月 17 日的相对表达量最高, 比相对表达量最低的 5 月 20 日高出 4.5 倍, 推测可能与 8 月血糖含量突增有关。转录因子 *cgPdx* 的相对表达量在 6 月 12 日最高, 随后缓慢下降。各基因相对表达量的标准差较大, 表明个体间表达水平具有很大差异。

2.6 logistic 回归分析

配子排放概率的 logistic 回归模型为:

$$P = \frac{e^{3.39 - 0.31G_p - 32.14G + 1.97g_M}}{1 + e^{3.39 - 0.31G_p - 32.14G + 1.97g_M}} \quad (4)$$

将上式进行对数变换可转变为:

$\ln(P/1-P) = 3.39 - 0.31G_p - 32.14G + 1.97g_M$, (5)
其中, P 代表配子排放的概率, G_p 代表内脏团糖原含量, G 代表血糖含量, g_M 代表性别为雄性。表达式的含义为内脏团糖原含量每增加一个单位, 配子排放的优势比($P/1-P$)乘以 0.73 ($e^{-0.31} = 0.73$), 葡萄糖含量每增加一个单位, 配子排放的优势比($P/1-P$)乘以 1.1 ($e^{-32.14} = 1.1$), 若为雄性, 配子排放的优势比($P/1-P$)乘以 7.16 ($e^{1.97} = 7.16$)。用列线图将回归模型可视化如图 6 所示, 含义为: 每个变量测量值(G, G_p, g_M)

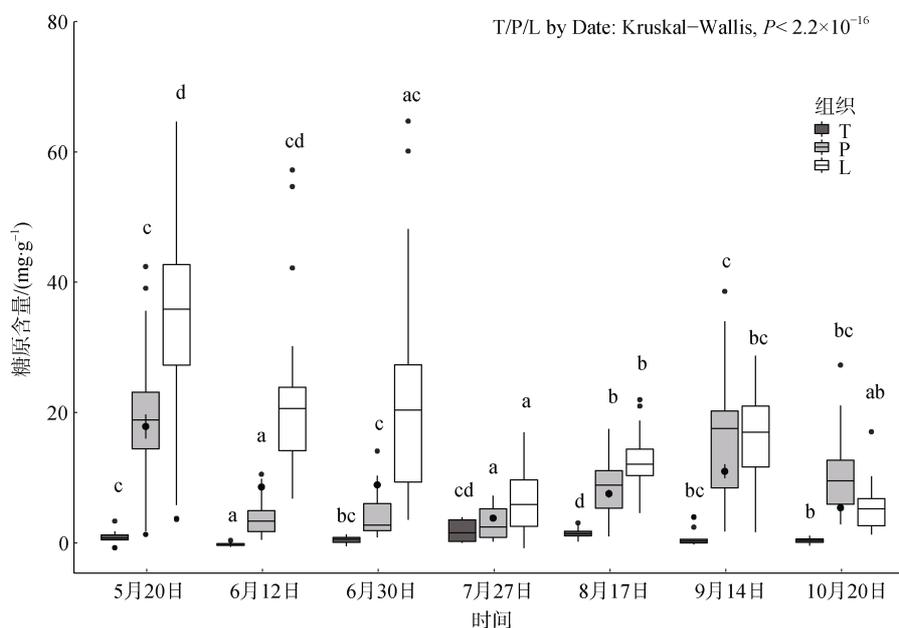


图3 各月份长牡蛎不同组织糖原含量变化

Fig. 3 Seasonal variations in glycogen content in different tissues

注: 不同字母表示同组间差异具有显著性($P < 0.05$)

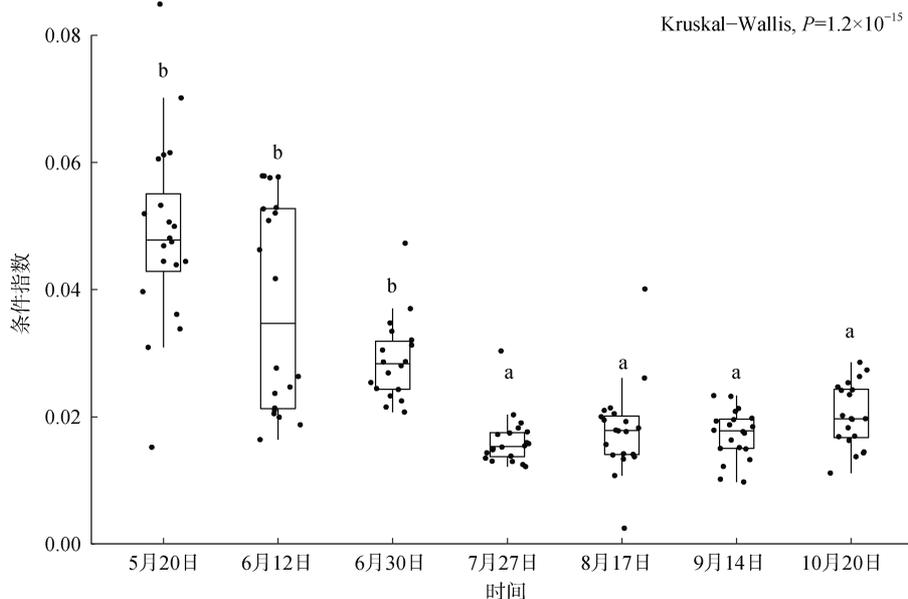


图4 各月份长牡蛎条件指数的变化

Fig. 4 Seasonal variations of the condition index

可以在 points 上可得到该变量的得分, 计算各变量得分和即为 total points, 通过 total points 投射到 Probability of gamete release 轴上, 即可得出对应的配子排放概率 p 。

对模型进行内部验证: C-index 值为 0.904(95% 置信区间: 0.83~0.977, 图 7a), 表明模型区分度极好, 能够准确推算出性腺状态; 由校准度曲线图

(图 7b)可以看出当预测概率在 0.65~0.9 之间时模型高估了配子排放的概率, 在 0.45~0.65 和大于 0.9 时模型低估了配子排放的概率, 但偏离幅度不大, 总体来看基本在 45°线附近, Hosmer-Lemeshow 拟合优度检验 χ^2 值为 9.06, $P = 0.34 > 0.05$, 表明模型预测值与实际值无统计学差异。综上, 模型构建成功。

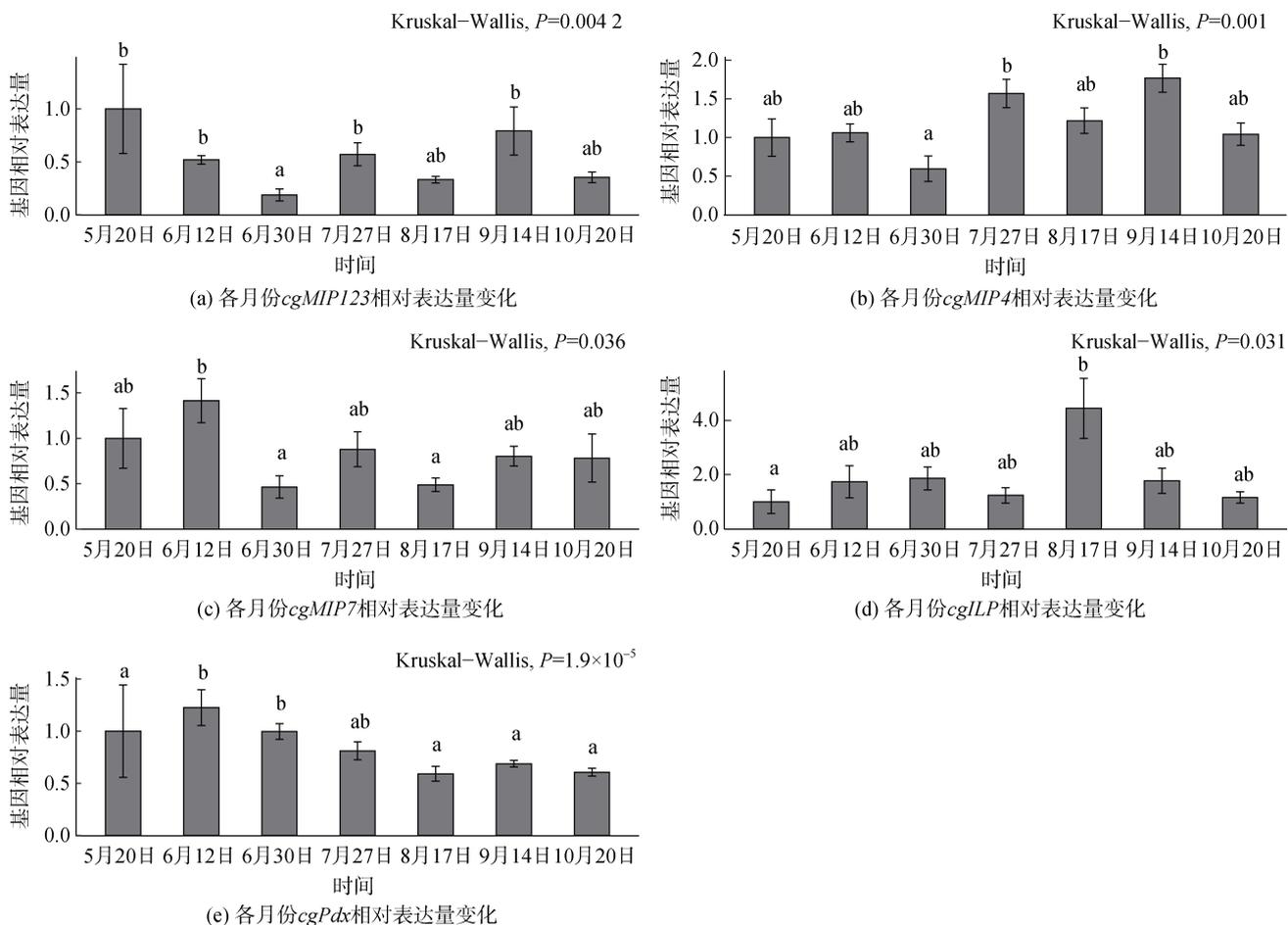


图5 各月份 *cgILPs* 和 *cgPdx* 基因相对表达量变化
Fig.5 Seasonal variations of the relative expression level of *cgILPs* and *cgPdx*

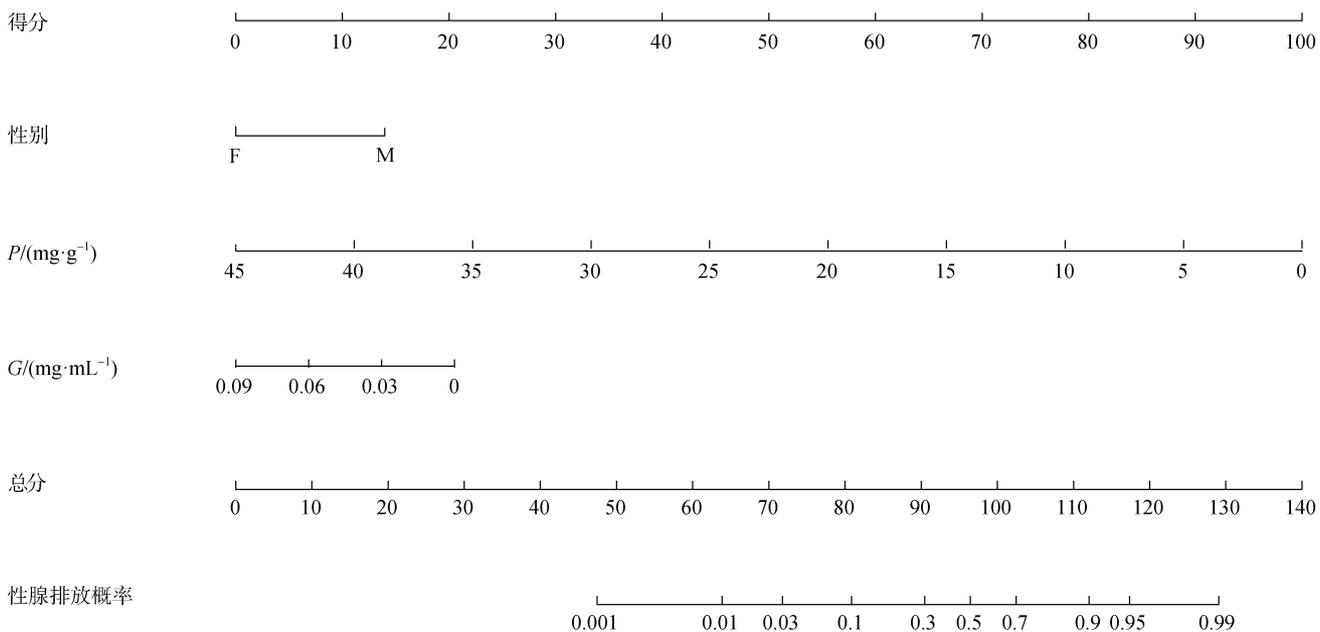


图6 配子排放概率 logistic 模型列线图
Fig. 6 Nomogram of the logistic model of gamete release

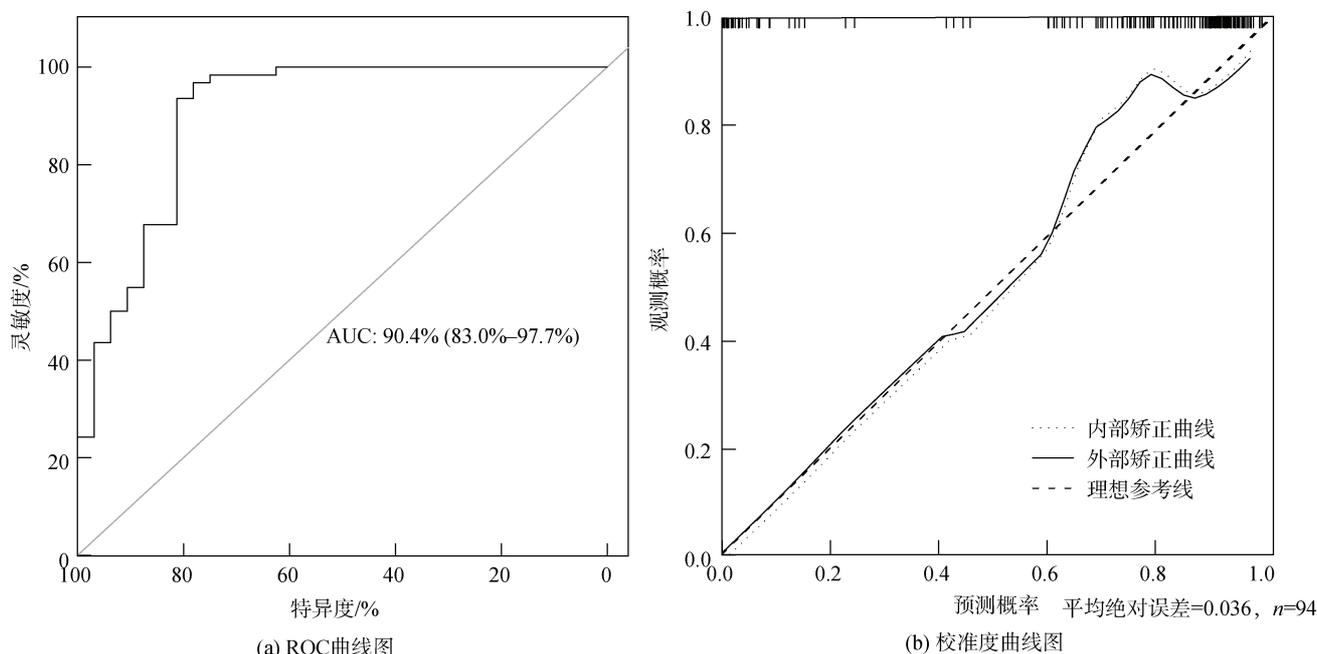


图 7 接受者操作特征曲线(ROC)曲线图及校准度曲线图
Fig. 7 ROC curve for the prediction model and the calibration curve

2.7 多元线性回归模型

采用 5 月 20 日至 10 月 20 日(总体数据)、5 月 20 日至 7 月 27 日(所测生理指标下降趋势阶段)、7 月 27 日至 10 月 20 日(所测生理指标上升趋势阶段)期间所测得的条件指数(C_i)、血糖(G)、各组织糖原含量(G_T 、 G_P 、 G_L)数据与取样时间(t)、各基因相对表达量建立线性回归模型。经逐步回归法最终筛选出 AIC 值最小的模型,并利用相对权重法模拟出各自变量对模型的贡献程度^[17]。

2020 年 5 月 20 日至 10 月 20 日期间,长牡蛎条件指数(C_i)与 $cgMIP123$ 和 $cgMIP4$ 相对表达量建立的回归模型 R^2 达到 0.91,且 P 小于 0.01,具有统计学意义(表 3),其中自变量 C_i 与 $cgMIP123$ 正相关,与

$cgMIP4$ 负相关,所占权重 $cgMIP123$ 比 $cgMIP4$ 略低(表 4);内脏团组织的糖原含量与叶绿素 a 含量的相关性为 0.55,与叶绿素 a 含量呈负相关。 $cgMIP4$ 与 $cgILP$ 与各环境因子具有较高的相关性, R^2 达到 0.88 和 0.98,各环境因子中 pH 和盐度所占权重较高(表 4)。7 月 27 日至 10 月 20 日期间, G_P 和 G_L 与基因表达水平的回归模型 R^2 相对较低,分别为 0.64 和 0.51, P 分别为 9.66×10^{-7} 和 7.68×10^{-4} ,均具有统计学意义(表 3),可以看出 t 与 G_P 和 G_L 呈正相关,与这一时期糖原开始逐渐积累相吻合; $cgMILP7$ 与糖原含量呈负相关,且所占权重较低; $cgMIP4$ 和 $cgILP$ 与 G_P 和 G_L 均呈正相关, $cgPdx$ 与 G_L 呈负相关,其中 $cgMIP4$ 在 G_P 和 G_L 的糖原积累中所占权重最高。

表 3 线性回归模型表达式

Tab. 3 Regression model expression

时间段	公式	R^2	P
5 月 20 日到 10 月 20 日	1. $C_i=0.042+0.039A-0.030B$	0.91	0.007 6
	2. $G_P=0.74-0.055F$	0.55	0.055
	3. $B=90.78-0.13F-36.45S-53.28H$	0.88	0.070
	4. $E=-193.35+0.22F+66.42S+121.52H+6.40T$	0.98	0.032
7 月 27 日到 10 月 20 日	5. $G_P=-0.31+0.0054t+0.33B-0.25C$	0.64	9.66×10^{-7}
	6. $G_L=0.047+0.0016t+0.16B-0.062C+0.026E-0.14D$	0.51	7.68×10^{-4}

注: A : $cgMIP123$ 基因相对表达量; B : $cgMIP4$ 基因相对表达量; C : $cgMIP7$ 基因相对表达量; D : $cgPdx$ 基因相对表达量; E : $cgILP$ 基因相对表达量; F : Chla 含量; H : pH 值; S : 盐度; t : 时间; T : 温度

表 4 回归模型相关参数

Tab. 4 Parameters of the regression model

公式编号	Y_i	X_i	权重/%	与 Y_i 的相关性
1	C_i	B	55.01	负
		A	44.99	正
2	G_P	F	100.00	负
		F	27.84	负
3	B	S	42.36	负
		H	29.80	负
		F	8.38	正
4	E	S	25.94	正
		H	42.36	正
		T	23.32	正
		t	53.74	正
5	G_P	B	28.37	正
		C	17.90	负
		B	37.69	正
6	G_L	E	28.38	正
		t	21.67	正
		C	6.59	正
		D	5.66	负

注: A : *cgMIP123* 基因相对表达量; B : *cgMIP4* 基因相对表达量; C : *cgMIP7* 基因相对表达量; D : *cgPdx* 基因相对表达量; E : *cgILP* 基因相对表达量; F : *Chl a* 含量; H : pH 值; S : 盐度; t : 时间; T : 温度

3 讨论

长牡蛎糖原含量的高低是评价其肥美程度的一个重要指标,糖原的积累也是目前的一个研究热点。牡蛎糖原储存在囊状结缔组织细胞(vesicular connective tissue cell, VCT)中,主要分布于唇瓣、外套膜、性腺区^[18],参与糖原合成和分解的糖原合酶(glycogen synthase, GS)与磷酸化酶(glycogen phosphorylase, GP)也主要在唇瓣和性腺中表达,两者表达水平最高的时期分别对应于糖原积累和配子发育期^[19]。糖原的储存和动员是牡蛎新陈代谢的重要组成部分,在牡蛎能量储存和配子发育中具有重要作用。糖原代谢可分为两个水平: 1) 糖原细胞的长期储存与动员周期。2) 血糖浓度的动态平衡调节。糖原细胞的长期储存与动员周期和配子发育以及浮游植物生物量和其它环境因子密切相关。血糖含量也与浮游植物生物量密切相关,另外血糖含量可以影响 GS 活性形式的比例从而间接地作用于糖原合成代谢^[20],最新研究表明胰岛素信号可以通过 PPP1R3B 和 PPP1R3G 与糖原合成相联系^[21]。因此本文聚焦于糖原、血糖、

类胰岛素这三者之间的关系。

基于近年来对长牡蛎类胰岛素基因的系统发生学分析及功能预测^[9],本研究筛选了 5 个相关基因,并采集了青岛胶南地区长牡蛎养殖群体自然繁殖期间的血糖、不同组织糖原含量、条件指数等指标,通过多元统计方法与相关基因的表达水平进行联合分析,得到以下主要结果:长牡蛎血糖含量和内脏团糖原含量在配子排放前后具有显著差异,且与性腺排放状态紧密相关;条件指数与 *cgMIP123* 和 *cgMIP4* 的相对表达量具有相关性;内脏团和唇瓣糖原含量与 *ILPs* 和 *cgPdx* 的相对表达量均具有一定的相关性,*cgMILP7* 的相对表达量与糖原含量成负相关,*cgPdx* 的相对表达量与唇瓣组织糖原含量呈负相关,*cgMIP4* 和 *cgILP* 的相对表达量与糖原含量成正相关。相关结果为进一步深入研究牡蛎生长繁殖和糖原积累等重要生物过程的调控机制提供数据支持。

3.1 环境因子与长牡蛎生长性状、各基因表达量的相关性

双壳类动物的繁殖周期和储能周期与水温、盐度和浮游植物等环境因子密切相关^[22-23]。海水温度通过有效积温或浮游植物丰度等直接或间接影响牡蛎性腺发育和生长繁殖。本研究观察到长牡蛎于 6 月中上旬进入性腺排放期,此时取样海区的水温为 20~22 °C。随着夏季水温的升高,该海域浮游植物丰度增加^[24],既有利于幼虫发育,也有利于亲本繁殖后体内营养物质的快速恢复。多元线性回归分析表明内脏团组织的糖原含量与叶绿素 *a* 含量具有一定的相关性 R^2 为 0.55,与叶绿素 *a* 含量呈负相关,也表明繁殖期为长牡蛎能量代谢高峰期。本研究测定了各月份叶绿素 *a* 含量,结果显示养殖海域 6 月和 7 月微型浮游植物生物量均处于较高水平。在夏季繁殖期间,海区浮游植物丰富度较高,但长牡蛎血糖含量逐渐下降,由于繁殖期消耗大量能量供应所致;7 月至 10 月血糖含量逐渐升高可能是由于繁殖期能量消耗减少且浮游植物生物量仍较丰富所致。*cgILP* 在 8 月的相对表达量显著升高,可能是由于血糖含量的增加引起。另外,血糖含量是极易受到个体状态和环境影响的生理指标,由于条件限制,无法从海上获取样品后实时进行分析,因此本研究的血糖含量数据具有一定的滞后性,但考虑到各批次取样流程基本一致,所得的变化趋势应该是准确的。

环境因子对 *cgMIP4* 和 *cgILP* 所构建的表达式表

明这两个基因更易受到环境因素的影响。各环境因子中 pH 和盐度所占权重较大, 温度和叶绿素 *a* 含量次之, 表明当变化相同单位时 pH 和盐度对基因表达量产生的影响最大, 但考虑到实际情况中 pH 和盐度处于稳定的状态, 温度对基因表达的影响会更高。

3.2 生长性状间的相关性

糖原的动员和储存分别与牡蛎的繁殖期(3 月至 10 月)、休眠期(11 月至 2 月)相对应^[22, 25-26], 糖原合成可以看成是对食物供应和高血糖的反应, 而糖原动员与配子形成的开始有关^[18]。本研究结果显示牡蛎繁殖行为造成体内糖原的大量流失, 同时血糖和条件指数(牡蛎肥满度的指征)也大幅下降。内脏团糖原含量在 6 月 12 日至 7 月 27 日一直处于较低水平, 与性腺排放期重合, 推测可能原因有: 第一, 内脏团糖原优先被用于配子的形成; 第二, 配子本身含有糖原, 随着配子被排放, 内脏团糖原储备被清空, 因此含量随之降低。

通过构建 logistic 回归模型, 可以根据内脏团糖原含量、血糖水平和性别三个指标准确的反映性腺是否已经排放, 表明长牡蛎性腺排放状态与这三个指标的综合效应具有很强的相关性。虽然牡蛎性腺排放状态可以通过肉眼直接观察, 性腺未排放时, 内脏团组织呈丰满的乳白色, 随着生殖细胞的排空, 内脏团逐渐透明(图 8), 但是该模型的目的并非为了推测性腺是否排放, 而是分析配子排放期间相关联的主要生理变化过程。本研究为回顾性分析, 由于长牡蛎所在海域、气候等众多影响因素存在差异, 该回归模型是否适用于其他时段或海域的长牡蛎尚待验证。



图 8 配子排放前(左, 2020 年 5 月 20 日样品)和排放后(右, 2020 年 6 月 12 日样品)长牡蛎的表型对比

Fig. 8 Morphological differences of *Crassostrea gigas* before (left, sampled on May 20, 2020) and after (right, sampled on May 20, 2020) spawning

3.3 类胰岛素基因与长牡蛎生长性状间的相关性

贝类 MIP 的基因功能已有较多研究基础, 它们多在神经组织表达, 可能参与了贝类的生长发育过程。将静水椎实螺(*Lymnaea stagnalis*)的神经破坏后, MIP 基因表达缺失, 会导致其身体和壳的生长停滞, 当重新植入神经簇后, 生长再次恢复^[11]。*cgMIP123* 和 *cgMIP4* 均属于贝类特有的类胰岛素基因家族 MIP, 且二者在基因组上定位于相邻位点, 推测是通过串联重复的方式发生基因复制而产生。其中 *cgMIP123* 有较多研究, 但均未发现其与糖原积累有关, 2012 年 Cong 等发现该基因中有 2 个 SNP 位点与长牡蛎生长性状(壳高、壳长、壳宽、总重和软体部重)呈极显著相关($P < 0.01$)^[12]。2005 年 Hamano 等报道了 *cgMIP123* 在 11 月和次年 1、3、5、7 月的相对表达量, 其中 11 月、次年 1 月和 3 月相对表达量依次升高, 3 月、5 月和 7 月相对表达量逐渐下降^[10-11]。本研究补充了 5 月至 10 月之间的观察数据, 结果显示 *cgMIP123* 在 7 月的相对表达量低于 5 月, 且该基因在糖原含量显著变化时期(5 月至 10 月)无明显的变化趋势, 与 Hamano 等人的研究结论一致。同时, 从多元统计的公式也可以得出 *cgMIP123* 的相对表达量与 P、L 的糖原积累期无关, 而与条件指数呈显著正相关。对于 *cgMIP4*, 在 P、L 糖原回归模型中所占权重均较大, 且呈正相关, 表明与长牡蛎糖原积累具有一定相关性。*cgMIP123* 与 *cgMIP4* 在生长和糖原方面的差异可能与两个基因的功能分化有关。*cgILP* 与胰岛素家族具有较高的同源性, 且在肝胰腺中特异性表达, 可能在血糖调节方面具有一定作用, 但是血糖含量未建立出回归模型, 因此未能确定 *cgILP* 与血糖之间的关系。公式 3 显示 *cgILP* 的相对表达量与唇瓣糖原积累成正相关, 但所占权重较低, 表明 *cgILP* 可能在糖原积累的过程中发挥了些许作用。胰岛素参与的血糖调控过程涉及到糖原合成通路, 但是长牡蛎血糖快速调节和糖原的季节性积累可能具有不同的调控机制, 因此通过糖原积累水平很难筛选到与血糖调节相关的基因。

cgMILP7 具有特殊的氨基酸序列模式, 系统进化分析发现其与果蝇的 *DILP7* 具有高度同源性^[9], 可能都属于类松弛素家族, 在生殖方面具有一定作用。长牡蛎的糖原含量在配子成熟与排放期显著下降, 在繁殖期之后又逐渐上升, 即繁殖过程与糖原

含量呈负相关,同时回归分析发现 *cgMILP7* 的表达水平与糖原含量呈负相关关系,推测 *cgMILP7* 很可能与繁殖过程的调控相关。

本研究所采用的有参线性回归模型对数据有很多的约束性条件,如一些基本的假设条件:因变量值或残差呈正态性、误差的独立性、线性、同方差性、多重共线性等,导致最终只构建成功了3个回归模型。其中糖原积累期的公式3中 R^2 较低,表明所构建的回归模型涉及的自变量有限,还有其他基因、代谢物或物理化学因子等参与了长牡蛎糖原积累的调控。后续研究可以考虑组学和非参数回归分析相结合的方法,进一步筛选血糖和糖原积累相关的调控因子。另外,本研究所用的实验动物来自同一个养殖群体,个体间可能存在亲缘关系,需要在更多的群体中验证相关结果。

4 结论

基于近年来对长牡蛎类胰岛素基因的系统发生学分析及功能预测,本文筛选了4个相关基因,并收集了青岛胶南地区长牡蛎养殖群繁殖期间的血糖、不同组织糖原含量、条件指数等指标,利用多元统计方法进行了数据分析。logistic 回归分析表明长牡蛎性腺是否排放与血糖含量、内脏团糖原含量和性别之间具有一定相关性。多元回归分析显示,繁殖期(2020年5月20日至10月20日)长牡蛎条件指数与 *cgMIP123* 和 *cgMIP4* 的相对表达量具有强相关性, R^2 达到0.91($P=0.0076$); *cgMIP4* 与 *cgILP* 与各环境因子具有较高的相关性, R^2 分别达到0.88和0.98;7月27日至10月20日期间,内脏团糖原含量与 *cgMIP4* 和 *cgMIP7* 基因表达水平具有一定相关性, R^2 达到0.64($P=9.66\times 10^{-7}$),唇瓣糖原含量与 *cgMIP4*, *cgMIP7*, *cgPdx* 和 *cgILP* 的相对表达量具有一定相关性 R^2 达到0.51($P=7.68\times 10^{-4}$)。建立的 logistic 回归模型和多元线性回归模型为研究性状与基因相对表达量相关性提供了一种分析方法,同时为进一步探索牡蛎生长繁殖和糖原积累等重要生物过程的调控机制提供数据支撑。

参考文献:

[1] LECROISEY C, LE PÉTILLON Y, ESCRIVA H, et al. Identification, evolution and expression of an insulin-like peptide in the Cephalochordate Branchiostoma lanceolatum[J]. PLoS One, 2015, 10(3): e0119461.

[2] CHAN S J, STEINER D F. Insulin through the ages: Phylogeny of a growth promoting and metabolic regulatory hormone[J]. American Zoologist, 2000, 40(2): 213-222.

[3] NAGASAWA H, KATAOKA H, ISOGAI A, et al. Amino acid sequence of a prothoracicotrophic hormone of the silkworm Bombyx mori[J]. Proceedings of the National Academy of Sciences of the United States of America, 1986, 83(16): 5840-5843.

[4] SMIT A B, VREUGDENHIL E, EBBERINK R H M, et al. Growth-controlling molluscan neurons produce the precursor of an insulin-related peptide[J]. Nature, 1988, 331(6156): 535-538.

[5] ZHENG S, CHIU H, BOUDREAU J, et al. A functional study of all 40 *Caenorhabditis elegans* insulin-like peptides[J]. Journal of Biological Chemistry, 2018, 293(43): 16912-16922.

[6] SATAKE S, MASUMURA M, ISHIZAKI H, et al. Bombyxin, an insulin-related peptide of insects, reduces the major storage carbohydrates in the silkworm Bombyx mori[J]. Comparative Biochemistry and Physiology - Part B: Biochemistry & Molecular Biology, 1997, 118(2): 349-357.

[7] ROBINSON S D, SAFAVI-HEMAMI H. Insulin as a weapon[J]. Toxicon, 2016, 123(2016): 56-61.

[8] RULIFSON E J, KIM S K, Nusse R. Ablation of insulin-producing neurons in flies: Growth and diabetic phenotypes[J]. Science, 2002, 296(5570): 1118-1120.

[9] CHERIF-FEILDEL M, HEUDE BERTHELIN C, ADELIN B, et al. Molecular evolution and functional characterisation of insulin related peptides in molluscs: Contributions of *Crassostrea gigas* genomic and transcriptomic-wide screening[J]. General and Comparative Endocrinology, 2019, 271(22): 15-29.

[10] HAMANO K, AWAJI M, USUKI H. cDNA structure of an insulin-related peptide in the Pacific oyster and seasonal changes in the gene expression[J]. Journal of Endocrinology, 2005, 187(1): 55-67.

[11] HAMANO K, AWAJI M. cDNA cloning and expression analysis of insulin-related peptide gene and prohormone convertase gene in the Pacific oyster, *Crassostrea gigas*[J]. Fisheries Science, 2002, 68: 769-771.

[12] CONG R, LI Q, KONG L. Polymorphism in the insulin-related peptide gene and its association with growth traits in the Pacific oyster[J]. Biochemical Systematics and Ecology, 2012, 46(2013): 36-43.

[13] FOX R. Invertebrate Zoology: A functional evolutionary approach[M]. California: Society of Systematic Biologists, 2007

[14] PFAFFL M W. A new mathematical model for relative quantification in real-time RT-PCR[J]. Nucleic Acids Research, 2001, 29(9): 2002-2007.

- [15] CANO J, ROSIQUE M J, ROCAMORA J. Influence of environmental parameters on reproduction of the European flat oyster (*Ostrea edulis* L.) In a coastal lagoon (Mar Menor, Southeastern Spain)[J]. *Journal of Molluscan Studies*, 1997, 63(2): 245-273.
- [16] HAN K, SONG K, CHOI B W. How to develop, validate, and compare clinical prediction models involving radiological parameters: study design and statistical methods[J]. *Korean Journal of Radiology*, 2016, 17(3): 339-350.
- [17] KABACOFF R I. R in action: Data analysis and graphics with R[M]. New York: Manning Publications Co., 2011.
- [18] BERTHELIN C, KELLNER K, MATHIEU M. Histological characterization and glucose incorporation into glycogen of the Pacific oyster *Crassostrea gigas* storage cells[J]. *Marine Biotechnology*, 2000, 2(2): 136-145.
- [19] BACCA H, HUVET A, FABIoux C, et al. Molecular cloning and seasonal expression of oyster glycogen phosphorylase and glycogen synthase genes[J]. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, 2005, 140(4): 635-646.
- [20] GABBOTT, WHITTLE. Glycogen-Synthetase in the sea Mussel *Mytilus-Edulis* L.2. Seasonal-Changes in glycogen-content and glycogen-synthetase activity in the mantle tissue[J]. *Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology*, 1986, 83(1): 197-207.
- [21] LI Q Q, ZHAO Q Y, ZHANG J Y, et al. The protein phosphatase 1 complex is a direct target of AKT that links insulin signaling to hepatic glycogen deposition[J]. *Cell Reports*, 2019, 28(13): 3406-3422.
- [22] LI Qi, LIU W G, SHIRAS Kunio, et al. Reproductive cycle and biochemical composition of the Zhe oyster *Crassostrea plicatula* Gmelin in an eastern coastal bay of China[J]. *Aquaculture*, 2006, 261(2): 752-759.
- [23] YU Z, LIU C, WANG F, et al. Diversity and annual variation of phytoplankton community in Yesso scallop (*Patinopecten yessoensis*) farming waters of North Yellow Sea of China[J]. *Aquaculture*, 2019, 511: 734266.
- [24] 曲静, 宫相忠, 庄会富, 等. 胶南近海及邻近海域夏、冬季浮游植物群落结构[J]. *海洋湖沼通报*, 2009(3): 143-154.
- QU Jing, GONG Xiangzhong, ZHUANG Huifu, et al. Phytoplankton community structure in coastal water of Jiaonan and adjacent water in summer and winter[J]. *Transactions of Oceanology and Limnology*, 2009(3): 143-154.
- [25] 种金豆, 李琪, 徐成勋, 等. 长牡蛎“海大3号”生长繁殖与营养成分的周年变化[J]. *水产学报*, 2020, 44(3): 411-418.
- CHONG Jindou, LI Qi, XU Chengxun, et al. Seasonal variation in growth and gonadal development of the Pacific oyster “Haida No. 3” in relation to the nutritive components[J]. *Journal of fishery of China*, 2020, 44(3): 411-418.
- [26] 邓传敏, 孔令锋, 于瑞海, 等. 长牡蛎壳金选育群体性腺发育与营养成分的周年变化[J]. *中国水产科学*, 2017, 24(1): 40-49.
- DENG Chuanmin, KONG Lingfeng, YU Ruihai, et al. Seasonal variation in gonadal development and nutritive components in the golden shell colored strain of Pacific oyster (*Crassostrea gigas*)[J]. *Journal of Fishery Sciences of China*, 2017, 24(1): 40-49.

Glycogen content during reproduction seasons of Pacific oyster and the correlation analysis with relative expression of insulin-like genes

MIAO Xiao-jing^{1, 2}, DENG Shao-xi^{1, 2}, LI Liang-heng¹, XU Fei^{1, 3}

(1. Institute of Oceanology, Chinese Academy of Sciences, Qingdao 266071, China; 2. University of Chinese Academy of Sciences, Beijing 100049, China; 3. Center for Ocean Mega-Science, Chinese Academy of Sciences, Qingdao 266071, China)

Received: Mar. 4, 2021

Key words: *Crassostrea gigas*; Pacific oyster; blood glucose content; insulin-like genes; *Pdx*; multivariate statistics

Abstract: In the present study, glycogen content variations during the reproduction period of *Crassostrea gigas* (Pacific oyster) and their correlations with the expression patterns of insulin-like genes (*cgMIP123*, *cgMIP4*, *cgMILP7*, *cgILP*) and the related transcription factor (*cgPdx*) were investigated. The blood glucose and glycogen content, condition index, and relative gene expression level were assayed on oysters cultured on a farm in Jiaonan, Shandong. Furthermore, marine environmental factors such as acidity, temperature, and salinity were collected from May to October 2020. Multivariate logistic regression analysis indicated a high correlation between the spawning and variations of blood glucose and glycogen content around the pyloric process. A regression model was constructed to estimate the gonad discharging status by the two parameters. The concordance index was at 0.903, while the χ^2 value was at 9.06 ($P > 0.05$) for the Hosmer–Lemeshow test, indicating a good fit. Application of the model to verification samples also showed high reliability. The multiple linear regression analysis exhibited correlations of growth or physiological values with the expression pattern of some genes: the condition index was correlated with *cgMIP123* and *cgMIP4* (R^2 : 0.91, P : 0.0076); the glycogen content around the pyloric process and labial palps showed correlation with *cgILPs* and *cgPdx*; the relative expression of *cgMILP7* was negatively correlated with the glycogen content of all the assayed tissues; *cgPdx* was negatively correlated with the glycogen content of labial palps; *cgMIP4* and *cgILP* were positively correlated with glycogen content.

(本文编辑: 杨 悦)