

文章编号: 1000-0615(2018)03-0322-12

DOI: 10.11964/jfc.20161110620

转录组测序研究三角帆蚌珍珠颜色相关基因

毛媛媛¹, 郑荣泉^{1,2,3*}, 张启鹏¹, 梅祎芸¹, 孔绅绅¹, 刘志芳¹

(1. 浙江师范大学化学与生命科学学院, 浙江 金华 321004;
2. 浙江省野生动物生物技术与保护利用重点实验室, 浙江 金华 321004;
3. 浙江师范大学行知学院, 浙江 金华 321004)

摘要: 为筛选影响三角帆蚌珍珠颜色的候选基因, 采用转录组测序平台 Illumina HiSeq™ 2500对蚌壳珍珠层颜色为紫色和白色的三角帆蚌中央膜组织进行高通量测序。所得序列经质控、组装后比对到NR、Swiss-Prot、COG、KOG、KEGG、GO和Pfam数据库中注释, 并进行差异基因聚类分析。结果显示, 获得干净数据共72 800 581 bp, 拼接获得了89 529个Unigene, 差异表达基因2644个, 其中上调表达基因1323个, 下调表达基因1321个。根据GO功能分类可分为分子功能、细胞组分、生物过程等3大类50分支; 根据KEGG代谢通路分析可以分为187类。对差异表达基因分析发现, 部分基因与色素合成(眼黄素、黑色素、喋啶色素)及金属离子转运相关, 还有6个细胞色素P450, 3个细胞色素b561和1个细胞色素b560。研究表明, 这些基因可能在珍珠颜色形成中发挥作用。

关键词: 三角帆蚌; 珍珠颜色; 差异表达; 转录组测序

中图分类号: Q 785; S 917.4

文献标志码: A

三角帆蚌(*Hyriopsis cumingii*)属软体动物门(Mollusca), 双壳纲(Bivalvia), 珠蚌目(Unionoida), 蚌科(Unionidae), 帆蚌属(*Hyriopsis*), 主要分布于长江中下游地区, 为中国特有的淡水贝类。三角帆蚌是中国主要的淡水育珠蚌种, 它的珍珠层光泽鲜艳且厚, 所产珍珠质量相对较好, 生产的珍珠占世界总淡水珍珠98%以上^[1]。我国淡水珍珠产量较高, 但销售额仅为世界珍珠的8%^[2], 主要因为我国生产的珍珠品质普遍不高。因此, 如何提高珍珠质量成了我国珍珠产业首要解决的问题之一。珍珠的质量与形状、大小、颜色、光洁度、有核珍珠层厚度和光泽有关, 其中颜色是评价珍珠质量的重要指标^[3]。珍珠培育研究中, 影响珍珠呈色的因素一直是研究的热点^[4], 目前, 主要因素有: 有机物、结构、金属离子致色^[5-6]。珍珠是由珍珠质构成的, 研究珍珠的颜色即为研究贝壳珍珠质层的颜色, 贝壳珍珠层颜色主要由遗传决定, 且该

性状属于质量性状, 可稳定遗传^[7-8]。Miyashita等^[9]通过蛋白质组学研究发现酪氨酸酶是导致合浦珠母贝(*Pinctada fucata*)所产珍珠出现蓝色的原因, 但未能定位到具体的基因; Li等^[10]筛选了载脂蛋白基因; 罗红瑞^[11]对紫色内壳色和白色内壳色三角帆蚌外套膜组织和珍珠囊进行比较转录组文库分析, 从中筛选得到*HcGIFLP1*、*HcGIFLP2*、*HcCUBDC*、*HcCA2*与珍珠层颜色形成相关的基因; Chen等^[12]克隆了三角帆蚌的2个酪氨酸酶基因*HcTyr*和*HcTyp-1*, 通过荧光定量表达和原位杂交信号检测发现*HcTyr*与三角帆蚌珍珠层颜色形成相关, *HcTyp-1*与珍珠层和角质层颜色形成相关。珍珠的颜色是由多种因素引起的, 目前发现的这些基因尚不能解释清楚珍珠颜色形成的分子机理。因此, 更多珍珠颜色相关基因的筛选对于阐明珍珠颜色形成分子机制至关重要。

转录组测序(transcriptome sequencing)是对某

收稿日期: 2016-11-28 修回日期: 2017-04-28

资助项目: 浙江省十二五水产育种专项(2012C12907-9; 2012C12907-5)

通信作者: 郑荣泉, E-mail: zhengrq@zjnu.cn

一物种的mRNA进行的高通量测序。转录组测序能快速全面地获得某一物种特定细胞或组织在某一状态下的基因表达情况, 可用于研究基因功能和结构、新转录本预测和可变剪接等^[13]。RNA-seq具有通量高、成本低、灵敏度高、无需基因组序列信息等优势, 使得其在非模式生物中得到了广泛应用^[14]。本实验利用Illumina HiSeq™ 2500测序平台对蚌壳珍珠层颜色为紫色和白色的三角帆蚌中央膜的基因表达情况进行分析, 找到差异表达的基因, 并对其进行功能注释分析, 以期筛选更多与珍珠颜色形成相关的候选基因。

1 材料与方法

1.1 实验材料

三角帆蚌采自浙江省金华市威旺养殖新技术有限公司, 该公司自2000年开始采用家系和群体结合选育出了紫色系和白色系三角帆蚌^[8]。随机选取同一养殖池塘里健康的三角帆蚌, 挑选贝壳内壳色分别为紫色和白色的体型大小无显著差异的2龄蚌(2龄雌雄蚌的生长并无显著差异^[15])各3只。实验蚌取回后放置实验室养殖一周, 养殖结束后取每只蚌的中央膜组织(三角帆蚌的中央膜主要参与珍珠层形成), 然后将紫色蚌和白色蚌所取的中央膜分别等量混合置于灭菌的贮存管中, 将贮存管迅速放入液氮中速冻, 并存放在-80 °C冰箱中保存待用。

RNA提取试剂盒RNAisoPlus购自杭州皓丰生物技术有限公司, DEPC水购自上海生工, 异丙醇、三氯甲烷、无水乙醇购自金华医药公司。

1.2 文库构建与高通量测序

使用TaKaRa RNAisoPlus试剂盒提取三角帆蚌中央膜的总RNA, 提取方法按试剂盒说明书进行, 再用1%的琼脂糖电泳检测总RNA是否降解, 采用Nanodrop 2000检测RNA样品的纯度、浓度和完整性等, 以保证转录组测序所使用的样品合格。

转录组测序文库的制备及测序由浙江大学完成。对检测合格的总RNA样品构建cDNA文库, 再分别使用Agilent 2100和Qubit 2.0对插入片段大小(insert size)和文库的浓度进行检测, 使用Q-PCR方法对文库的有效浓度进行准确定量, 以保证文库质量。库检合格后, 用HiSeq 2500进行

高通量测序, 测序读长为双端100 bp。

1.3 从头组装与功能注释

HiSeq™2500测序所得的原始序列通过去除低质量序列、去接头等过程得到高质量序列。质控后, 通过Ttinity(v2.2.0)对无参转录组进行组装^[16], 使用TransDecoder软件进行Unigene的编码区序列及其对应氨基酸序列的预测, 使用BLAST^[17]软件将Unigene序列与NR^[18]、Swiss-Prot^[19]、GO^[20]、COG^[21]、KOG^[22]、KEGG^[23]数据库比对, 预测完Unigene的氨基酸序列之后使用HMMER^[24]软件与Pfam^[25]数据库比对, 获得Unigene的注释信息。

1.4 差异表达基因

采用Bowtie^[26]将各样品测序得到的Reads与Unigene库进行比对, 根据比对结果, 结合RSEM^[27]进行表达量水平估计。利用FPKM值表示对应Unigene的表达丰度。FPKM计算公式如下:

$$\text{FPKM} = \frac{\text{cDNA Fragments}}{\text{Mapped Fragments (Millions)} / \text{Transcript Length (kb)}}$$

基因表达量可直接用于比较不同样品间的基因表达差异。采用DESeq^[28]进行样品组间的差异表达分析, 获得两种颜色蚌的差异表达基因集。在筛选过程中, 将FDR<0.01作为阈值且差异倍数FC(Fold Change)≥2作为筛选标准, 其中FC表示两样品之间表达量的比值, 并在此基础上统计上下调基因的个数。随后对差异表达基因进行KEGG通路富集分析, GO功能及Pfam结构域注释。

2 结果

2.1 Unigene的组装

将测序获得的原始数据去除低质量序列后, 得到72 800 581 bp的干净数据(clean data), 从头组装后得到89 529条Unigene, 平均长度为814 bp, 其中长度在1 kb以上的Unigene有19 721条, N50的长度为1477(图1)。

2.2 Unigene 功能注释

对获得的89 529个Unigene进行功能注释, 选择BLAST参数E-value不大于10⁻⁵和HMMER参数E-value不大于10⁻¹⁰, 最终获得24 058个有注释

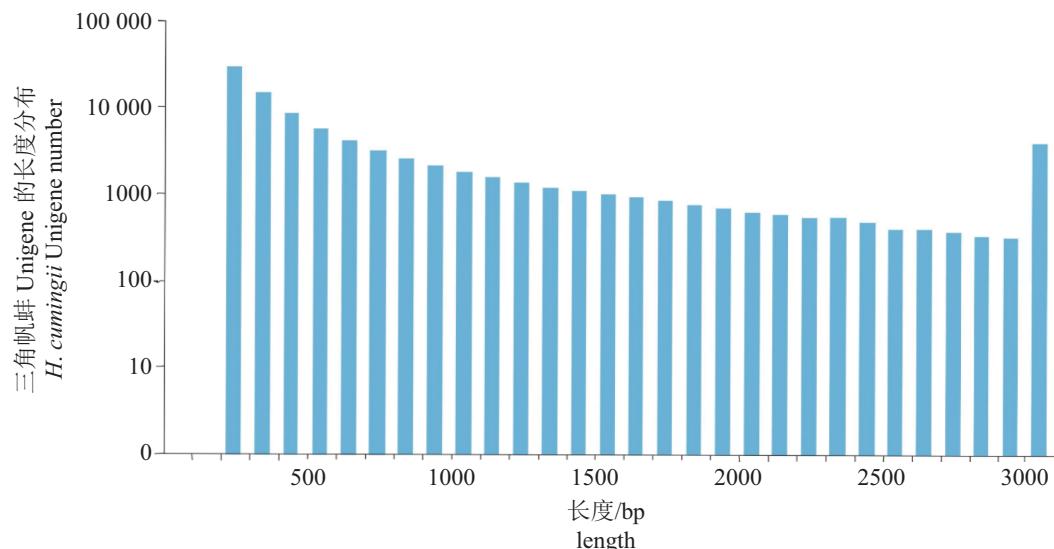


图1 三角帆蚌Unigene的长度分布

Fig. 1 Length distribution of Unigene sequence in *H. cumingii*

信息的Unigene，对注释到NR、Swiss-Prot、COG、KOG、Pfam、GO、KEGG数据库的基因进行统计(表1)。

表1 基因注释数量统计

Tab. 1 Statistics of annotation gene number

数据库类别 gene	注释基因数/个 annotation gene number	注释基因比例/% ratio of annotation gene number
Nr	23 518	80.27
Swiss-Prot	12 912	47.36
COG	5825	24.21
KOG	13 882	57.7
Pfam	14 724	61.2
GO	7173	29.82
KEGG	1654	6.88
总数 total	24 058	100

将拼接所得的Unigene与Nr蛋白数据库进行同源性比对，有23 518条与已知基因同源，相似序列匹配的近缘物种中，相似序列所占比例最高的是牡蛎(*Crassostrea gigas*, 40%)，随后依次是霸王莲花青螺(*Lottia gigantea*, 19%)、海蜗牛(*Aplysia californica*, 11%)、囊舌虫(*Saccoglossus kowalevskii*, 4%)、海蠕虫(*Capitella teleta*, 3%)、文昌鱼(*Branchiostoma floridae*, 3%)、紫海胆(*Strongylocentrotus purpuratus*, 2%)、海葵

(*Nematostella vectensis*, 1%)、隆头蛛(*Stegodyphus mimosarum*, 1%)、罗布斯塔泽蛭(*Helobdella robusta*, 1%)和其他物种(15%)。与Swiss-Prot数据库的比对中，有12 912(53.7%)条Unigene注释成功。

COG数据库是对基因产物进行同源分类的数据库，本实验中注释到COG数据库中的Unigene数为5825条，可归属于26个大类，其中数量最多的为一般功能预测类，达到1443条，占24.77%；其次为复制、重组和修复类，占5.84%；信号转导机制类占3%；转录类，占2.52%。此外，有964条单一基因被归属为未知功能类，占16.55%。与Pfam数据库比对发现有14 724条Unigene比对上。

将Unigene在KOG真核直系同源数据库中进行比对，结果有13 882条比对到KOG中，可归属于26类，其中一般功能预测类2431条，占17.51%；信号转导机制1956条，占14.9%；转录后修饰、蛋白质转换、分子伴侣类1177条，占8.48%；未知功能类2819条，占20.31%。

Blast2GO软件分析得到每个Unigene蛋白GO功能注释，GO功能分类分为分子功能、细胞组分和生物学过程3大类50分支，总共有7173条序列富集至GO数据库中。在分子功能类型中：催化活性(catalytic activity GO: 0003824)和蛋白结合(binding GO: 0005488)所占比例最高，蛋白结合转录因子(protein binding transcription factor activity GO: 0000988)比例最低；在细胞组分中：细胞部

分(cell part GO: 0044464)和细胞(cell GO: 0005623)为主要的功能类型, 细胞外基质部分(extracellular matrix part GO: 0044420)所占比例最小; 在生物学过程中: 代谢过程(metabolic process GO: 0008152)、细胞过程(cellular process GO: 0009987)和单生物过程(single-organism process GO: 0044699)被标注成功的比例最高, 节律性过程(rhythmic process GO: 0048511)比例最小。

注释到KEGG数据库中的Unigene有1654条, 被归类至187个代谢通路, 其中基因数量排名靠前的代谢途径有泛素介导的蛋白降解途径(ubiquitin mediated proteolysis, 86条, 45.99%), 核糖体途径(ribosome, 70条, 37.43%), RNA转运途

径(RNA transport, 42条, 22.46%), 嘌呤代谢途径(purine metabolism, 39条, 20.86%), 氧化磷酸化途径(oxidative phosphorylation, 37条, 19.79%), 磷脂酰肌醇信号系统途径(phosphatidylinositol signaling system, 36条, 19.25%)等。在187个代谢通路中, 有6个代谢途径可能与珍珠颜色形成相关, 分别是卟啉和叶绿素代谢途径(porphyrin and chlorophyll metabolism)、外源性物质细胞色素P450代谢(metabolism of xenobiotics by cytochrome P450)、细胞色素P450药物代谢(drug metabolism-cytochrome P450)、ABC转运体(ABC transporters)^[29]、Notch信号通路(notch signaling pathway)^[30]和黑素生成(melanogenesis)(表2)。

表2 颜色相关代谢通路注释统计

Tab. 2 Annotation statistics of color related metabolic pathway

代谢通路 metabolic pathway	基因数量/个 gene number	代谢通路编号 number of metabolic pathway
porphyrin and chlorophyll metabolism	4	ko00860
metabolism of xenobiotics by cytochrome P450	4	ko00980
drug metabolism-cytochrome P450	6	ko00982
ABC transporters	20	ko02010
notch signaling pathway	12	ko04330
melanogenesis	6	ko04916

2.3 差异表达基因

不同内壳色的两个样品所获得的转录本利用DESeq软件进行比较, 最终获得差异表达基因2644个, 其中上调表达的基因1323个, 下调表达的基因1321个。差异表达基因的GO注释如图2所示, 分别有170、225和443个Unigene注释到GO数据库的生物学过程、分子功能和细胞组分中。其中, 部分差异表达基因注释至色素及金属离子结合的分子功能中, 包括3个色素沉积、1个血红素结合及1个金属离子结合。将基因根据参与的KEGG代谢通路分为6个分支: 人类疾病(human diseases), 遗传信息处理(genetic information processing), 代谢(metabolism), 细胞过程(cellular processes), 环境信息处理(environmental information processing), 有机系统(organismal systems)(图3)。差异表达基因中有775条Unigene注释到Pfam数据库中, 在这些差异基因中有33条Unigene与色素和金属离子代谢相关, 4个包含酪氨酸酶的结构域, 2条具有多铜氧化酶结构域, 10条注释到动

物血红素超氧化酶结构域, 含有细胞色素P450、细胞色素b561和细胞色素b560结构域的Unigene分别有6条、3条和1条, 具有血红素转运蛋白D、锰离子转运、锌离子转运及ABC转运蛋白各1条, 2条具有色氨酸加氧酶结构域, 还有1条Unigene产物注释至GTP环化水解酶 I (表3)。

3 讨论

3.1 三角帆蚌转录组测序分析

近年来, 转录组技术已广泛应用于软体动物研究中, 包括文蛤(*Meretrix meretrix*)^[30]、黑蝶贝(*Pinctada margaritifera*)^[31]、长牡蛎(*Crassostrea gigas*)^[32]、长肋日月贝(*Amusium pleuronectes*)^[33]、欧洲鲍螺(*Haliotis tuberculata*)^[34]和织锦巴非蛤(*Paphia textile*)^[35]等。Bai等^[36]利用454测序技术对内壳色为紫色和白色的三角帆蚌外套膜和珍珠囊进行转录组测序及比较分析, 获得386个差异表达基因; Zhang等^[37]对三角帆蚌的性腺、肝胰

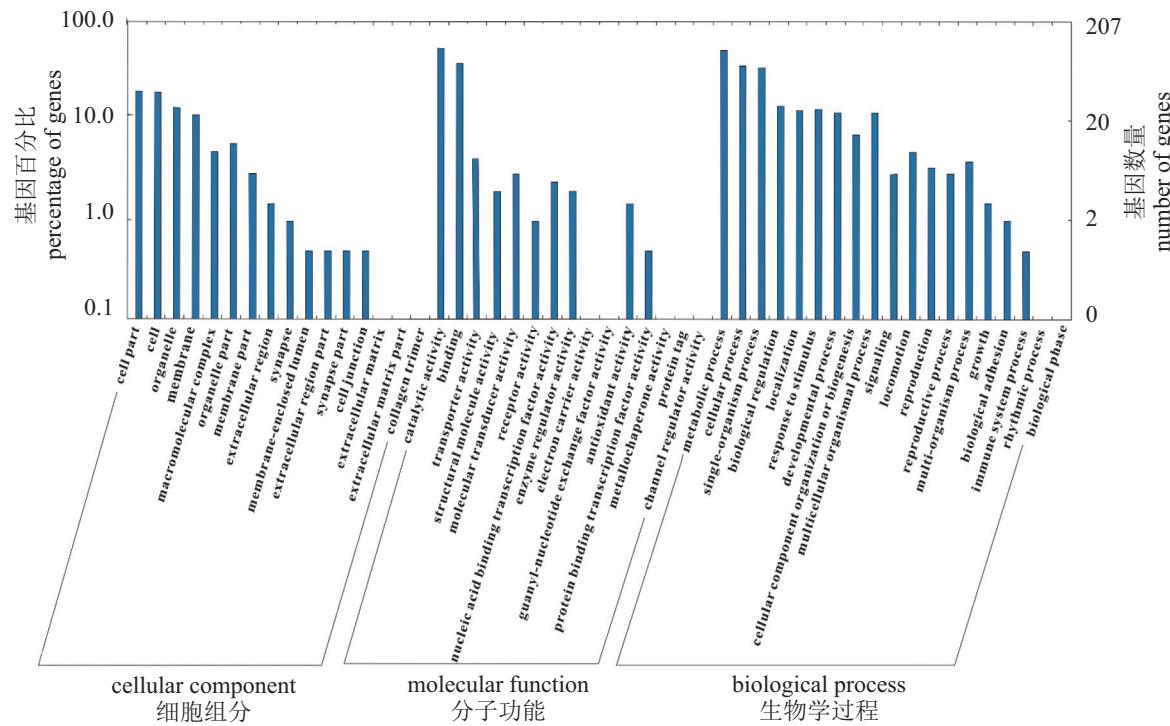


图2 三角帆蚌差异表达基因的GO功能分类

Fig. 2 Gene ontology classification of putative functions of differentially expressed genes of *H. cumingii*

腺、斧足、外套膜、鳃和闭壳肌6个组织进行了转录组测序，并从转录组文库中筛选到8个性别决定基因。本研究对内壳色为紫色和内壳色为白色的三角帆蚌中央膜进行无参转录组测序，得到了非冗余Unigene序列89 529条，其中有11 204条Unigene长度大于1000 bp，有8517条Unigene长度大于2000 bp，平均长度为814 bp，比Bai等^[36]获得的平均长度296 bp长。组装的转录本与组装后的转录本比对上的序列占79.35%，均一性分析结果良好，符合质量要求。

3.2 转录组文库注释分析

对三角帆蚌Unigene进行生物学功能注释，其中有24 058条Unigene被注释，占总Unigene的26.87%，其余还有73.13%未被注释到已知基因，出现这种注释率较低的情况可能是因为相近物种的基因注释信息在目前的国际公共基因数据库中收录较少。海洋动物中也有类似报道，例如虾夷扇贝(*Patinopecten yessoensis*)^[38]转录组分析中只有32.62%被注释，而南极帽贝(*Laternula elliptica*)^[39]只有16.93%的注释率。

序列的长度通常影响比对的结果，而转录组测序获得的短序列很少能与已知基因匹配^[38]。本研究获得的三角帆蚌转录组测序序列中，去

掉了小于300 bp的短序列，以保证基因匹配更有意义。在所有被注释的序列中，序列长度小于1000 bp的序列有44.15%被注释，注释效率与序列长度成正相关。

与Nr数据库比对结果显示：与在公共数据库中已注册的三角帆蚌序列匹配的Unigene仅有0.46%，这可能是因为目前公共数据库中仍然缺乏三角帆蚌的注册信息。Unigene在KOG真核直系同源数据库中比对注释至无机离子转运和代谢(248, 1.79%)、脂类转运和代谢(327, 2.36%)以及信号转导机制(1956, 14.9%)等的基因可能与贝壳珍珠层颜色有关。

在富集至GO数据库生物学过程的基因中，部分基因与色素的沉积有关，例如包含几丁质的成人表皮色素沉积(adult chitin-containing cuticle pigmentation, GO: 0048085)，复眼色素沉积(compound eye pigmentation, GO: 0048072)，表皮色素沉积(cuticle pigmentation, GO: 0048067)，发展色素沉积(developmental pigmentation, GO: 0048066)，眼色素沉积(eye pigmentation, GO: 0048069)，表皮色素沉积的正调控(positive regulation of cuticle pigmentation, GO: 0048081)和包含几丁质的成人表皮色素沉积调控(regulation of adult chitin-contai-

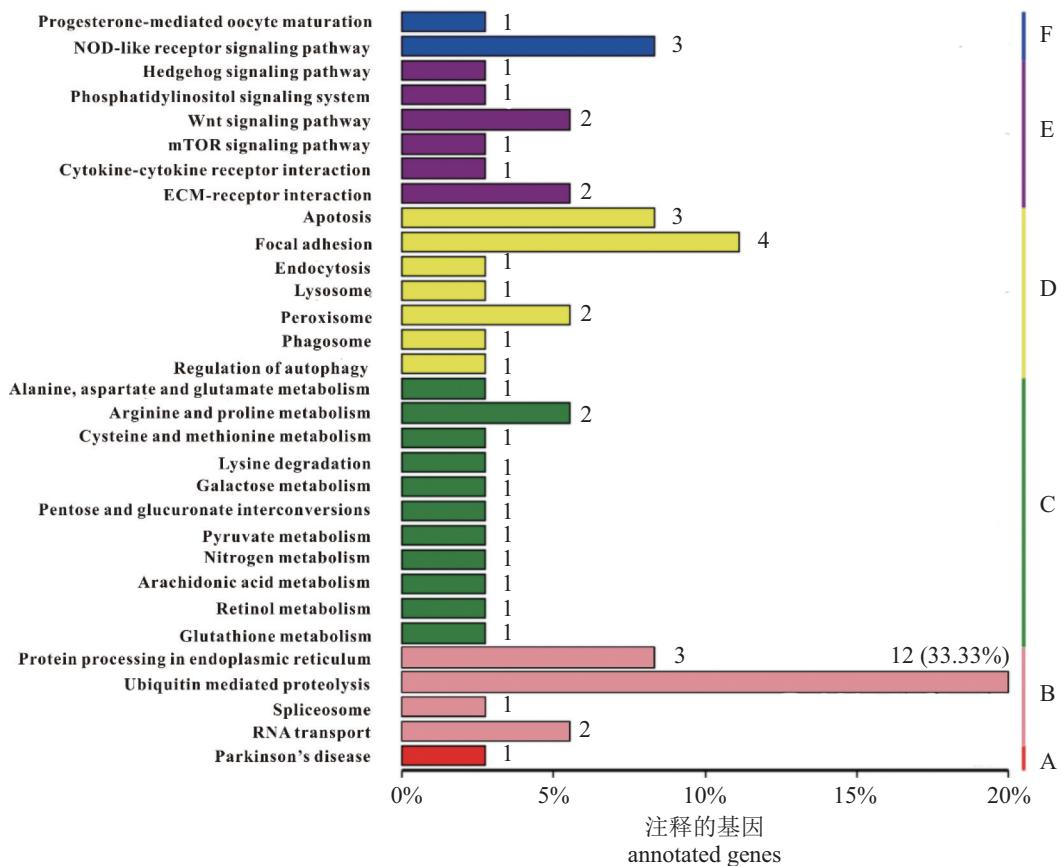


图3 三角帆蚌差异表达基因的KEGG富集分类

纵坐标为KEGG代谢类别名称, 横坐标为注释到该类别下的基因个数及其个数占被注释上的基因总数的比例。将基因根据参与的KEGG代谢通路分为6个分支, A(human diseases, 人类疾病); B(genetic information processing, 遗传信息处理); C(metabolism, 代谢); D(cellular processes, 细胞过程); E(environmental information processing, 环境信息处理); F(organismal systems, 有机系统)。

Fig. 3 KEGG classification of putative functions of differentially expressed genes of the *H. cumingii*

Vertical coordinate means the name of metabolic category of KEGG, abscissa means number and proportion of annotated gene. The gene was divided into six branches according to the participating KEGG metabolic pathway: A (human disease); B (genetic information processing); C (metabolism); D (cellular processes); E (environmental information processing); F (organismal systems).

ning cuticle pigmentation, GO: 0048082), 这些基因可能参与贝壳内壳色形成。KEGG代谢通路注释分析中, 一些代谢通路如卟啉和叶绿素代谢, 外源性细胞色素P450代谢, ABC蛋白转运体^[29]、Notch信号通路^[30]和黑素生成, 都与生物色素的合成和转运代谢相关, 因此, 我们推断这些代谢通路可能参与贝壳珍珠层颜色形成。

3.3 不同内壳色三角帆蚌中央膜差异表达基因的分析

在获得的差异表达基因中, 只有1148条Unigenes被注释, 还有1496条未被注释, 未注释的基因中有些基因表达量较高。

差异表达基因的GO注释 GO统计结果显示: 差异表达基因在生物学过程数据库中注释

的有170个, 其中富集在细胞的占该功能类别的23.53%, 细胞组分占24.12%; 注释富集至分子功能数据库的有225个差异表达基因, 催化活性占该功能类别的34.67%, 结合占30%; 富集至细胞组分的有443个, 代谢过程占24.15%, 细胞过程占16.93%, 单组织过程占16.03%。在生物学过程分类中, 有3个基因与色素沉积有关, 可能参与贝壳珍珠层颜色形成。在分子功能类型中, 有3个基因具有血红素结合(heme binding GO: 0020037)功能, 血红素是高等动物血液和肌肉中的红色色素, 可能参与珍珠颜色形成。部分基因富集至金属离子结合的类别中, 目前, 珍珠颜色与金属离子的相关性研究较多, 例如杨明月等^[4]发现不同颜色珍珠中微量元素存在差别, Zn, Mg, V, Ag和Co使珍珠呈现紫色, Ti和V使

表3 颜色相关差异表达基因结构域注释统计

Tab. 3 Domain annotation statistics of color related differentially expressed genes

基因编号 gene number	Pfam注释 Pfam annotation	FDR (E值) FDR (E value)	匹配物种 matched species
c49581.graph_c3	animal haem peroxidase	0.007 207 758	<i>Crassostrea gigas</i>
c57055.graph_c0	animal haem peroxidase	6.09E-10	<i>Crassostrea gigas</i>
c59859.graph_c2	animal haem peroxidase	3.34E-09	<i>Crassostrea gigas</i>
c63520.graph_c0	animal haem peroxidase	0	<i>Crassostrea gigas</i>
c42862.graph_c1	animal haem peroxidase	3.53E-06	<i>Crassostrea gigas</i>
c63884.graph_c3	animal haem peroxidase	1.94E-11	<i>Lottia gigantea</i>
c42862.graph_c0	animal haem peroxidase	0.000 582 976	<i>Crassostrea gigas</i>
c43211.graph_c0	animal haem peroxidase	8.18E-05	<i>Crassostrea gigas</i>
c62774.graph_c0	animal haem peroxidase	7.77E-16	<i>Crassostrea gigas</i>
c43211.graph_c3	animal haem peroxidase	5.10E-06	<i>Euprymna scolopes</i>
c28793.graph_c0	common central domain of tyrosinase	2.79E-13	<i>Crassostrea gigas</i>
c61644.graph_c0	common central domain of tyrosinase	1.01E-07	<i>Crassostrea gigas</i>
c37807.graph_c0	common central domain of tyrosinase	0	<i>Crassostrea gigas</i>
c63690.graph_c1	putative tyrosinase-like protein tyr-3	4.50E-10	<i>Crassostrea gigas</i>
c58190.graph_c0	m multicopper oxidase	0.004 325 873	<i>Branchiostoma floridae</i>
c63029.graph_c0	m multicopper oxidase	0.000 685 611	<i>Crassostrea gigas</i>
c46043.graph_c0	cytochrome P450	0.000 119 012	<i>Lottia gigantea</i>
c35868.graph_c0	cytochrome P450	1.54E-05	<i>Crassostrea gigas</i>
c54847.graph_c0	cytochrome P450	0.004 837 581	<i>Solen grandis</i>
c63864.graph_c1	cytochrome P450	4.20E-08	<i>Lottia gigantea</i>
c54827.graph_c0	cytochrome P450	5.20E-08	<i>Crassostrea gigas</i>
c47081.graph_c0	cytochrome P450	0.000 411 941	<i>Crassostrea gigas</i>
c60291.graph_c0	eukaryotic cytochrome b561	9.21E-05	<i>Lottia gigantea</i>
c56731.graph_c2	eukaryotic cytochrome b561	0.000 835 624	<i>Crassostrea gigas</i>
c63813.graph_c0	eukaryotic cytochrome b561	0.006 152 549	<i>Lottia gigantea</i>
c20829.graph_c0	cytochrome b560	1.03E-09	<i>Caligus clemensi</i>
c52513.graph_c0	GTP cyclohydrolase I	6.06E-05	<i>Lottia gigantea</i>
c56361.graph_c2	heme exporter protein D (CcmD)	5.00E-10	
c49422.graph_c0	magnesium transporter NIPA	0.000 141 097	<i>Crassostrea gigas</i>
c59297.graph_c0	tryptophan 2,3-dioxygenase	3.49E-10	<i>Azumapecten farreri</i>
c47339.graph_c0	tryptophan 2,3-dioxygenase	7.95E-11	<i>Azumapecten farreri</i>
c60821.graph_c1	ABC transporter	0.001 739 717	<i>Strongylocentrotus purpuratus</i>
c63920.graph_c0	ZIP Zinc transporter	0.005 190 047	<i>Crassostrea gigas</i>

珍珠呈橙红色；严俊等^[6]通过比较白、粉红、紫3种颜色的淡水珍珠发现，白色和紫色珍珠的

Ti、Fe、Mg与Cu的含量比粉色珍珠低，白色珍珠中Mn的含量比粉色高；彭艳华^[40]发现在海水

珍珠中, 金黄色珍珠中的Sr明显高于其他浅色海水珍珠。

差异表达基因的KEGG注释 56个差异表达基因在KEGG数据库中得到注释, 这些基因显著富集的通路有31个, 其中基因数量排名前10位的KEGG通路包括泛素介导蛋白降解、黏着、内质网中的蛋白质过程、类NOD受体信号通路、细胞凋亡、RNA转运、精氨酸及脯氨酸代谢、过氧化物酶体、ECM受体交互作用和Wnt信号通路。这些代谢通路与生物体的蛋白质合成、信号传导过程有关, 参与贝壳基质蛋白的合成和转运, 这些过程是否与珍珠和贝壳珍珠层的颜色形成有关尚待进一步的研究。

3.4 珍珠颜色部分候选基因功能分析

珍珠的颜色是评价珍珠质量的一个重要指标, 珍珠的成分与贝壳珍珠层成分基本一致, 因此研究贝壳珍珠层的颜色即为研究珍珠的颜色。张根芳等^[8]研究发现贝壳珍珠层颜色主要由遗传决定, 因此筛选珍珠颜色相关基因尤为重要。本实验通过对贝壳珍珠层颜色分别为紫色和白色的三角帆蚌中央膜转录组测序分析, 筛选与珍珠颜色形成相关的基因。有机色素和金属离子是影响珍珠体色的主要因素, 因此本实验主要从筛选有机色素和金属离子相关的基因入手。通过比较2种颜色蚌的转录组文库, 从差异表达基因中筛选出33个色素及金属离子代谢相关基因。

差异表达基因中有4个酪氨酸酶基因, 分别与长牡蛎的*tyr-1*和*tyr-3*相似, Bai等^[36]的转录组数据中也筛选出酪氨酸酶基因, Chen等^[12]还依据Bai等^[36]的转录组文库克隆出了2个三角帆蚌酪氨酸酶基因*HcTyr*和*HcTyp-1*, 但Bai等^[36]筛选出的酪氨酸酶基因与阿根廷鱿鱼(*Illex argentinus*)相似。酪氨酸酶是合成黑色素的关键酶, 黑色素是自然界分布最广泛的色素, 在软体动物中^[41]也有发现。酪氨酸酶能够催化酪氨酸羟基化形成L-多巴, L-多巴经过一系列的反应生成对动物的眼睛、皮肤和头发不可缺少的黑色素^[42-43]。Speiser等^[44]研究认为酪氨酸酶表达模式改变是影响石鳖(*Acanthopleura granulata*)眼睛进化的关键环节, 即影响石鳖对眼色素的选择。蒋燕玲等^[45]研究发现在橘色双冠丽鱼(*Amphilophus citrinellus*)3个典型体色褪黑过程中, 各组织TYR基因表达量变化

显著, 表明其与橘色双冠丽鱼体色变化有关系。这4个基因也可能参与三角帆蚌珍珠颜色形成。

血红素是高等动物血液和肌肉中的一种含铁的红色色素, 是血红蛋白、肌红蛋白、细胞色素和过氧化物酶等众多活性生物大分子的辅基^[46], 存在于机体内所有组织中。血红素不单独以色素模式存在, 但是很多色素蛋白包含血红素, 血红素通过这种方式在动物体中呈现颜色^[47]。差异表达基因Pfam注释发现, 有10条Unigene包含动物血红素超氧化酶结构域, 1条Unigene含有血红素转运蛋白D结构域, 对差异表达基因进行GO功能分类发现还有3个Unigene具有与血红素结合的分子功能, 这14个基因的产物均包含血红素结构域, 但其是否与珍珠颜色形成的调控有关还有待进一步研究。

细胞色素P450(cytochrome P450, CYP450)是一类含血红素的单链蛋白质, 广泛存在于生物体内^[48]。细胞色素P450代谢通路与细胞色素代谢密切相关, 牛晓艳等^[49]利用转录组测序分析不同毛色獭兔皮肤组织中基因的表达差异, 发现细胞色素P450代谢通路中有8个差异表达基因, 推测这些基因可能在毛色形成中发挥作用。Sun等^[50]对具有非对称色素模式的虾夷扇贝(贝壳左边外壳颜色为红褐色, 右边为白色)的左右外套膜进行转录组测序并进行差异表达基因分析, 筛选出4个未知基因和一个细胞色素P450基因, 通过实时荧光定量PCR实验发现细胞色素P450基因在左边外套膜的表达量极显著高于右边外套膜, 推测细胞色素P450基因可能参与红褐色贝壳颜色形成。差异表达基因中有6个基因包含细胞色素P450结构域, 还有几个基因含有细胞色素b561和b560结构域, 这些差异表达基因也可能参与珍珠颜色形成过程。

本实验从三角帆蚌中央膜转录组差异表达基因中筛选出2个色氨酸加氧酶基因, 1个ABC转运蛋白, 1个GTP环化水解酶I基因及4个金属代谢基因, 包括2个多铜氧化酶、1个镁离子转运蛋白和1个锌离子转运蛋白基因。眼黄素和喋啶色素是昆虫眼睛呈色的重要色素, 色氨酸加氧酶是眼黄素合成途径中的第一个作用酶, 3-羟基犬尿氨酸是中间产物, 3-羟基犬尿氨酸进入色素颗粒内部需通过色素颗粒膜上ABC转运蛋白的转运才能进行, 3-羟基犬尿氨酸进入色素颗粒内部后在吩噁嗪酮合成酶催化作用下生成眼黄素^[51-52]。

GTP环化水解酶 I 参与合成昆虫的眼色素喋啶色素，反应为GTP环化水解酶 I 催化GTP形成三磷酸二氢新蝶呤^[53]。色氨酸加氧酶参与合成眼黄素，GTP环化水解酶 I 及ABC转运蛋白参与喋啶色素的合成与转运。江琰等^[54]研究发现三角帆蚌黄、白、黑珍珠的部分微量元素含量有明显差异，表现为白色珍珠Mn的含量相对丰富，黄色珍珠Cu、Zn、Mg的含量相对丰富，黑色珍珠Fe的含量相对丰富。这些基因参与色素的合成转运及金属离子的转运，可能在珍珠颜色形成过程中发挥作用。

感谢浙江省金华市威旺养殖新技术有限公司提供实验样本三角帆蚌，感谢浙江大学转录组测序平台 Illumina HiSeq™ 2500 对测序的技术支持及北京百迈客生物科技有限公司对转录组测序数据的注释分析。

参考文献：

- [1] Li W J, Shi Z Y, He X J. Study on immune regulation in *Hyriopsis cumingii* Lea: Effect of pearl-nucleus insertion in the visceral mass on immune factors present in the hemolymph[J]. Fish & Shellfish Immunology, 2010, 28(5-6): 789-794.
- [2] 罗文, 杨受保, 丁盈, 等. 三角帆蚌(*Hyriopsis cumingii*)珍珠形成初期钙代谢的特征分析[J]. 海洋与湖沼, 2010, 41(6): 895-900.
Luo W, Yang S B, Ding Y, et al. Calcium metabolism in *Hyriopsis cumingii* during early pearl-forming stages[J]. *Oceanologia et Limnologia Sinica*, 2010, 41(6): 895-900(in Chinese).
- [3] 中华人民共和国国家质量监督检验检疫总局, 中国国家标准化管理委员会. GB/T 18781-2008珍珠分级[S]. 北京: 中国标准出版社, 2009.
General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, Standardization Administration of the People's Republic of China. GB/T 18781-2008 Cultured pearl grading[S]. Beijing: Standards Press of China, 2009 (in Chinese).
- [4] 杨明月, 郭守国, 史凌云, 等. 淡水养殖珍珠的化学成分与呈色机理研究[J]. 宝石和宝石学杂志, 2004, 6(2): 10-13.
Yang M Y, Guo S G, Shi L Y, et al. Study on compositions and colouring mechanism of freshwater cultured pearls[J]. Journal of Gems & Gemmology, 2004, 6(2): 10-13(in Chinese).
- [5] Shi L, Liu X, Mao J, et al. Study of coloration mechanism of cultured freshwater pearls from mollusk *Hyriopsis cumingii*[J]. *Journal of Applied Spectroscopy*, 2014, 81(1): 97-101.
- [6] 严俊, 胡仙超, 王巨安, 等. 不同颜色的淡水养殖珍珠呈色机理研究[J]. 岩矿测试, 2013, 32(2): 263-268.
Yan J, Hu X C, Wang J A, et al. Investigation on the coloring mechanism of freshwater cultured pearls with different color[J]. Rock and Mineral Analysis, 2013, 32(2): 263-268(in Chinese).
- [7] 王照旗, 韩学凯, 白志毅, 等. 三角帆蚌紫色选育系1龄阶段内壳色及生长性状的遗传参数估计[J]. 水产学报, 2014, 38(5): 644-650.
Wang Z Q, Han X K, Bai Z Y, et al. Estimates of genetic parameters for inner shell color and growth traits during one year old stage in the purple strain of *Hyriopsis cumingii* using microsatellite based parentage assignment[J]. *Journal of Fisheries of China*, 2014, 38(5): 644-650(in Chinese).
- [8] 李清清, 白志毅, 刘晓军, 等. 三角帆蚌生长性状和内壳色与所产无核珍珠质量的相关性分析[J]. 水产学报, 2015, 39(11): 1631-1639.
Li Q Q, Bai Z Y, Liu X J, et al. Correlation analysis of non-nucleated pearl quality parameters with growth traits and inner shell color of *Hyriopsis cumingii*[J]. *Journal of Fisheries of China*, 2015, 39(11): 1631-1639 (in Chinese).
- [9] Miyashita T, Takagi R. Tyrosinase causes the blue shade of an abnormal pearl[J]. *Journal of Molluscan Studies*, 2011, 77(3): 312-314.
- [10] Li X L, Bai Z Y, Luo H R, et al. Cloning, differential tissue expression of a novel *hcApo* gene, and its correlation with total carotenoid content in purple and white inner-shell color pearl mussel *Hyriopsis cumingii*[J]. *Gene*, 2014, 538(2): 258-265.
- [11] 罗红瑞. 三角帆蚌贝壳珍珠层颜色形成相关基因的克隆与表达分析[D]. 上海: 上海海洋大学, 2015.
Luo H R. Cloning and expression analysis of genes involved in the nacre color of shell in *Hyriopsis cumingii*[D]. Shanghai: Shanghai Ocean University, 2015 (in Chinese).
- [12] Chen X J, Liu X J, Bai Z Y, et al. *HcTyr*, and *HcTyp-1*, of *Hyriopsis cumingii*, novel tyrosinase and tyrosinase-

- related protein genes involved in nacre color formation[J]. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, 2017, 204: 1-8.
- [13] Maher C A, Kumar-Sinha C, Cao X H, et al. Transcriptome sequencing to detect gene fusions in cancer[J]. *Nature*, 2009, 458(7234): 97-101.
- [14] 王晓锋, 何卫龙, 蔡卫佳, 等. 马尾松转录组测序和分析[J]. 分子植物育种, 2013, 11(3): 385-392.
- Wang X F, He W L, Cai W J, et al. Analysis on transcriptome Sequenced for *Pinus massoniana*[J]. *Molecular Plant Breeding*, 2013, 11(3): 385-392(in Chinese).
- [15] Zhao Y C, Bai Z Y, Fu L L, et al. Comparison of growth and pearl production in males and females of the freshwater mussel, *Hyriopsis cumingii*, in China[J]. *Aquaculture international*, 2013, 21(6): 1301-1310.
- [16] Grabherr M G, Haas B J, Yassour M, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome[J]. *Nature Biotechnology*, 2011, 29(7): 644-652.
- [17] Altschul S F, Madden T L, Schäffer A A, et al. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs[J]. *Nucleic Acids Research*, 1997, 25(17): 3389-3402.
- [18] 邓泱泱, 荔建琦, 吴松峰, 等. nr数据库分析及其本地化[J]. 计算机工程, 2006, 32(5): 71-73, 76.
- Deng Y Y, Li J Q, Wu S F, et al. Integrated nr database in protein annotation system and its localization[J]. *Computer Engineering*, 2006, 32(5): 71-73, 76(in Chinese).
- [19] Apweiler R, Bairoch A, Wu C H, et al. UniProt: The Universal Protein knowledgebase[J]. *Nucleic Acids Research*, 2004, 32(D1): D115-D119.
- [20] Ashburner M, Ball C A, Blake J A, et al. Gene ontology: Tool for the unification of biology[J]. *Nature Genetics*, 2000, 25(1): 25-29.
- [21] Tatusov R L, Galperin M Y, Natale D A. The COG database: A tool for genome-scale analysis of protein functions and evolution[J]. *Nucleic Acids Research*, 2000, 28(1): 33-36.
- [22] Koonin E V, Fedorova N D, Jackson J D, et al. A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes[J]. *Genome Biology*, 2004, 5(2): R7.
- [23] Kanehisa M, Goto S, Kawashima S, et al. The KEGG resource for deciphering the genome[J]. *Nucleic Acids Research*, 2004, 32(D1): D277-D280.
- [24] Eddy S R. Profile hidden Markov models[J]. *Bioinformatics*, 1998, 14(9): 755-763.
- [25] Finn R D, Bateman A, Clements J, et al. Pfam: The protein families database[J]. *Nucleic Acids Research*, 2014, 42(D1): D222-D230.
- [26] Langmead B, Trapnell C, Pop M, et al. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome[J]. *Genome Biology*, 2009, 10(3): R25.
- [27] Li B, Dewey C N. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome[J]. *BMC Bioinformatics*, 2011, 12(1): 323.
- [28] Anders S, Huber W. Differential expression analysis for sequence count data[J]. *Genome Biology*, 2010, 11(10): R106.
- [29] Dermauw W, Van Leeuwen T. The ABC gene family in arthropods: Comparative genomics and role in insecticide transport and resistance[J]. *Insect Biochemistry and Molecular Biology*, 2014, 45: 89-110.
- [30] Yue X, Nie Q, Xiao G Q, et al. Transcriptome analysis of shell color-related genes in the clam *Meretrix Meretrix*[J]. *Marine Biotechnology*, 2015, 17(3): 364-374.
- [31] Joubert C, Piquemal D, Marie B, et al. Transcriptome and proteome analysis of *Pinctada margaritifera* calcifying mantle and shell: Focus on biomineralization[J]. *BMC Genomics*, 2010, 11: 613.
- [32] Feng D D, Li Q, Yu H, et al. Comparative transcriptome analysis of the pacific oyster *Crassostrea gigas* characterized by shell colors: Identification of genetic bases potentially involved in pigmentation[J]. *PLoS One*, 2015, 10(12): e0145257.
- [33] Huang R L, Zheng Z, Wang Q H, et al. Mantle branch-specific RNA sequences of moon scallop *Amusium pleuronectes* to identify shell color-associated genes[J]. *PLoS One*, 2015, 10(10): e0141390.
- [34] Harney E, Dubief B, Boudry P, et al. *De novo* assembly and annotation of the European abalone *Haliotis tuberculata* transcriptome[J]. *Marine Genomics*, 2016, 28: 11-16.
- [35] Chen X M, Li J K, Xiao S J, et al. *De novo* assembly and characterization of foot transcriptome and microsatellite marker development for *Paphia textile*[J]. *Gene*, 2015, 576(1): 537-543.
- [36] Bai Z Y, Zheng H F, Lin J Y, et al. Comparative analys-

- is of the transcriptome in tissues secreting purple and white nacre in the pearl mussel *Hyriopsis cumingii*[J]. *PLoS One*, 2013, 8(1): e53617.
- [37] Zhang A J, Liu S L, Zhu J Y, et al. Transcriptome analysis of the freshwater pearl mussel, *Hyriopsis cumingii* (Lea) using illumina paired-end sequencing to identify genes and markers[J]. Iranian Fisheries Sciences Research Institute, 2016, 15(4): 1425-1440.
- [38] Sun X J, Yang A G, Wu B, et al. Characterization of the mantle transcriptome of yesso scallop (*Patinopecten yesoensis*): Identification of genes potentially involved in biomineralization and pigmentation[J]. *PLoS One*, 2015, 10(4): e0122967.
- [39] Clark M S, Thorne M A, Vieira F A, et al. Insights into shell deposition in the Antarctic bivalve *Laternula elliptica*: Gene discovery in the mantle transcriptome using 454 pyrosequencing[J]. *BMC Genomics*, 2010, 11: 362.
- [40] 彭艳华. 金黄色海水珍珠颜色研究进展[J]. 中国市场, 2016(17): 162-164.
- [41] Verhecken A. The indole pigments of Mollusca[J]. Annalen van de Koninklijke Belgische Vereniging voor Dierkunde, 1989, 119(2): 181-197.
- [42] Kim Y J, Uyama H. Tyrosinase inhibitors from natural and synthetic sources: Structure, inhibition mechanism and perspective for the future[J]. *Cellular and Molecular Life Sciences*, 2005, 62(15): 1707-1723.
- [43] Hofreiter M, Schöneberg T. The genetic and evolutionary basis of colour variation in vertebrates[J]. *Cellular and Molecular Life Sciences*, 2010, 67(15): 2591-2603.
- [44] Speiser D I, Demartini D G, Oakley T H. The shell-eyes of the chiton *Acanthopleura granulata* (Mollusca, Polyplacophora) use pheomelanin as a screening pigment[J]. *Journal of Natural History*, 2014, 48(45-48): 2899-2911.
- [45] 蒋燕玲, 宋红梅, 刘奕, 等. 橘色双冠丽鱼TYR基因的克隆及其发育时序和组织表达分析[J]. 农业生物技术学报, 2016, 24(5): 697-707.
- Jiang Y L, Song H M, Liu Y, et al. Cloning and expression analysis of the developing sequence and tissue expression of TYR gene in *Amphilophus citrinellus*[J]. *Journal of Agricultural Biotechnology*, 2016, 24(5): 697-707(in Chinese).
- [46] Tsiftsoglou A S, Tsamadou A I, Papadopoulou L C. Heme as key regulator of major mammalian cellular functions: Molecular, cellular, and pharmacological aspects[J]. *Pharmacology & Therapeutics*, 2006, 111(2): 327-345.
- [47] Takeuchi K, Satou Y, Yamamoto H, et al. A genome-wide survey of genes for enzymes involved in pigment synthesis in an Ascidian, *Ciona intestinalis*[J]. *Zoological Science*, 2005, 22(7): 723-734.
- [48] 张立群, 巴拉, 玉斯日古楞, 等. 细胞色素P450 2J亚家族研究进展[J]. 动物医学进展, 2014, 35(8): 84-89.
- Zhang L Q, Ba L, Yusif R G L, et al. Advance in cytochrome P450 2J subfamily[J]. *Progress in Veterinary Medicine*, 2014, 35(8): 84-89(in Chinese).
- [49] 牛晓艳, 任克良, 曹亮, 等. 利用转录组测序方法研究獭兔毛色相关基因[J]. 中国草食动物科学, 2016, 36(2): 1-6.
- Niu X L, Ren K L, Cao L, et al. Transcriptome analysis on coat color related genes in Rex Rabbits[J]. *China Herbivore Science*, 2016, 36(2): 1-6(in Chinese).
- [50] Sun X J, Zhou L Q, Liu Z H, et al. Transcriptome-wide analysis reveals candidate genes responsible for the asymmetric pigment pattern in scallop *Patinopecten yesoensis*[J]. *Invertebrate Survival Journal*, 2016, 13: 298-308.
- [51] Mackenzie S M, Howells A J, Cox G B, et al. Sub-cellular localisation of the white/scarlet ABC transporter to pigment granule membranes within the compound eye of drosophila melanogaster[J]. *Genetica*, 2000, 108(3): 239-252.
- [52] 蒋鹏凌. 褐飞虱红眼突变型的复眼形态结构及其形成机制初探[D]. 杭州: 浙江大学, 2013.
- Jiang P L. Preliminary studies on the morphology and structure of compound eyes and formation mechanisms of the red-eye mutant in the brown planthopper, *Nilaparvata lugens* Stål[D]. Hangzhou: Zhejiang University, 2013 (in Chinese).
- [53] Mackay W J, O'Donnell J M. A genetic analysis of the pteridine biosynthetic enzyme, guanosine triphosphate cyclohydrolase, in *Drosophila melanogaster*[J]. *Genetics*, 1983, 105(1): 35-53.
- [54] 江琰, 刘克武, 李静, 等. 三角帆蚌黄白黑珍珠元素分析及药用价值评价[J]. 化学研究与应用, 2003, 15(3): 433-434.
- Jiang Y, Liu K W, Li J, et al. Elemental analysis and pharmaceutical evaluation for yellow white and black pearls of *Hytiopsis cumingii* (Lea)[J]. *Chemical Research and Application*, 2003, 15(3): 433-434(in Chinese).

Transcriptome sequencing analysis on pearl color related genes in *Hyriopsis cumingii*

MAO Yuanyuan¹, ZHENG Rongquan^{1,2,3*}, ZHANG Qipeng¹,
MEI Yiyun¹, KONG Shenshen¹, LIU Zhifang¹

(1. College of Chemistry and Life Sciences, Zhejiang Normal University, Jinhua 321004, China;

2. Key Lab of Wildlife Biotechnology and Conservation and Utilization of Zhejiang Province, Jinhua 321004, China;

3. Xingzhi College, Zhejiang Normal University, Jinhua 321004, China)

Abstract: The transcriptome of *Hyriopsis cumingii* ventral mantle tissue whose nacre colors were purple and white was sequenced by a new generation of high throughput sequencing technology in order to find candidate genes involved in pearl color produced by *H. cumingii*. After quality control and assembly, the sequences acquired were blasted against NR, Swiss-Prot, COG, KOG, KEGG, GO and Pfam databases, and then cluster analysis were performed. In total, 72 800 581 bp of clean data, and 89 529 Unigenes were obtained, and 2644 genes were differentially expressed, among which 1323 were up-regulated and 1321 down-regulated. The Unigenes GO function in the transcriptome library was divided into 3 categories: molecular function, cellular component and biological process with 50 branches; it could be divided into 187 classes taking the KEGG database as a reference. Differentially expressed gene analysis revealed that some genes were related to pigment biosynthesis (such as xanthommatin, melanin and pteridine) and metal ion transportation. And there were 6 cytochrome P450, 3 cytochrome b561 and 1 cytochrome b560 genes. They might play important roles in the formation of pearl color of *H. cumingii*.

Key words: *Hyriopsis cumingii*; pearl color; differential expression; transcriptome sequencing

Corresponding author: ZHENG Rongquan. E-mail: zhengrq@zjnu.cn

Funding projects: Zhejiang Province Twelfth Five-year Special Aquatic Products Breeding (2012C12907-9; 2012C12907-5)