

马氏珠母贝 (*Pinctada fucata*) 血细胞全长转录组测序与分析

王菁, 李桂英, 陈志炜, 郭焕娣, 龚晓晴, 王忠良* (广东海洋大学水产学院, 广东湛江 524088)

摘要 [目的] 开发马氏珠母贝 (*Pinctada fucata*) 基因组资源, 挖掘功能基因。[方法] 以马氏珠母贝血细胞为试材, 利用单分子实时技术 (single-molecule real time, SMRT) 进行全长转录组测序, 并对所获得 unigenes 进行功能注释和基因结构分析。[结果] 共获得 277 064 条全长非嵌合序列和 82 381 个基因。经过比对 nr、SwissProt、KEGG 和 KOG 数据库进行注释和功能分类后, 共得到 59 621 个注释基因。同时, 在 8 493 条基因中共发现 11 219 个 SSR 位点, 其中以二核苷酸重复基元类型最高 (50.9%)。生物信息学分析鉴定得到 20 013 个 lncRNA、2 004 个转录因子、10 522 个可变剪切分析位点。[结论] 使用 SMRT 技术能够深入挖掘马氏珠母贝全长转录组数据, 为进一步探讨马氏珠母贝功能基因的挖掘、免疫响应及遗传机制的研究提供可靠的基因组资源。

关键词 马氏珠母贝; 全长转录组; SSR 标记; 可变剪切

中图分类号 S917.4 **文献标识码** A

文章编号 0517-6611(2022)06-0086-05

doi: 10.3969/j.issn.0517-6611.2022.06.019



开放科学 (资源服务) 标识码 (OSID):

Full-length Transcriptome Sequencing from Hemocytes of *Pinctada fucata*

WANG Jing, LI Gui-ying, CHEN Zhi-wei et al (Fisheries College, Guangdong Ocean University, Zhanjiang, Guangdong 524088)

Abstract [Objective] To deep analysis of the genomic resources of the pearl oyster (*Pinctada fucata*). [Method] The single-molecule real time technology (SMRT) was performed to sequence the full-length transcriptome from hemocytes, and functional annotation and the structure of the unigenes were analyzed. [Result] The results showed that a total of 277 064 full-length non-chimeric sequences were generated, and 82 381 unigenes were finally obtained. By comparing the nr, SwissProt, KEGG and KOG databases for annotation and functional classification, a total of 59 621 unigenes were annotated. GO function annotation showed that most genes were mainly enriched in cellular process, metabolic process and single biological process of biological process. Annotated genes in cell component were mainly enriched in cell and cell part, while in molecular function, they were mainly concentrated in binding and catalytic activity. Meanwhile, 11 219 SSR sites were identified in 8 493 unigenes, among which the proportion of dinucleotide repeat motif was the highest (50.9%). Additionally, the results also identified 20 013 lncRNAs, 2 004 transcription factors and 10 522 alternative splicing sites. [Conclusion] These results may improve knowledges on transcriptome data of *P. fucata* using SMRT, and provide reliable genomic resources for further exploring the development of functional genes, growth traits related gene markers, immune response and genetic mechanisms.

Key words *Pinctada fucata*; Full-length transcriptome; SSR; Alternative splicing

马氏珠母贝 (*Pinctada fucata*) 是我国海水珍珠养殖的主要母贝, 主要分布于我国南海沿岸^[1]。目前, 关于马氏珠母贝的研究普遍集中于生长性状^[2-3]、砂化功能^[4]、免疫响应^[5]等方面。分析不同性状形成的机理、性状相关基因分子标记的开发以及遗传机制的研究, 都需要借助于马氏珠母贝的基因组资源^[6]。随着基因组和转录组数据资源的不断开发, 全长转录组测序技术能够提高转录组注释的准确性, 提供一种更加全面的功能基因注释方法^[7]。转录组测序不仅可以反映细胞内基因的类型和数量, 还可以在分子水平上揭示细胞的生理生化过程^[8], 具有识别参与机体生长、发育和免疫调节等生物过程的相关基因以及分析基因表达调控的重要功能^[9]。第三代测序技术 PacBio 利用单分子实时测序技术 (single-molecule real time, SMRT) 直接测序, 无需经过组装过程即可直接生成包含 5'-UTR、3'-UTR 和 poly-A 尾巴的完整转录本^[10]。相较于第二代测序技术, PacBio 平台所得到

的转录本更长, 转录本注释率也得到提高^[11]。使用 SMRT 技术对花鲈 (*Lateolabrax maculatus*) 进行测序分析增加了对硬骨鱼类盐度适应机制的理解^[12]。Katneni 等^[7]对印度对虾 (*Penaeus indicus*) 全长转录组数据进行分析, 为进一步研究相关经济性状的基因提供宝贵的基因组资源。使用 SMRT 技术对金乌贼 (*Sepia esculenta*) 进行测序并获得高质量的数据库, 为探索生长发育、组织分化等相关性状提供了基础资源^[13]。Chen 等^[9]对达氏鲟 (*Acipenser dabryanus*) 进行全长转录组测序, 鉴定得到 3 个 *TRIM* 基因, 并证明其在抗细菌感染的免疫反应中起着关键作用。Liu 等^[14]对平胸龟 (*Platysternon megacephalum*) 肝脏进行单分子实时测序, 系统分析线粒体 DNA 转录组的基因重排和控制区重复线性现象, 加深了对线粒体基因组转录调控机制的理解。此外, 全长转录组数据可用于分析选择性剪接事件 (alternative splicing, AS), 有助于进一步理解 RNA 的加工过程^[6,15], 而综合分析转录组基因的剪接异构体, 有助于完善现有基因模型的功能注释^[12]。

该研究使用 SMRT 技术对马氏珠母贝血细胞进行全长转录组测序、功能基因挖掘、可变剪切分析, 旨在提高现有马氏珠母贝基因组资源功能注释的准确性及多样性, 对于进一步开展马氏珠母贝功能基因的挖掘及其免疫机制的研究提供参考。

1 材料与方法

1.1 马氏珠母贝的暂养及总 RNA 提取 马氏珠母贝 (壳长

基金项目 广东省自然科学基金面上项目 (2019A1515011875 号); 2021 年度广东省普通高校特色创新类项目 (2021KTCX044); 广东省省级科技计划项目 (国际科技合作领域) (2019A050510044 号); 广东海洋大学 2019 年“冲一流”省财政专项资金建设项目; 广东海洋大学 2021 年度本科生创新团队项目 (CXTD2021001 号); 国家级大学生创新创业训练计划项目 (CXXL2020002 号); 广东海洋大学“南海学者计划” (2017 年度)。

作者简介 王菁 (1997—), 女, 蒙古族, 河南南阳人, 硕士研究生, 研究方向: 水产经济动物生物学。* 通信作者, 副教授, 从事水产经济动物功能基因与基因组学研究。

收稿日期 2021-10-29

约 65 mm) 购自广东省湛江市雷州养殖场, 暂养于室内海水桶, 1 d 换水 1 次; 水温 25 °C, 暂养期间保持充气, 定时投喂螺旋藻粉。室内暂养 7 d 后随机选择 3 个马氏珠母贝于闭壳肌处采集血淋巴, 混合为 1 个样品, 4 °C、1 200 r/min 条件下离心 10 min 收集血细胞, 进一步提取总 RNA 及进行全长转录组测序。

1.2 全长转录组文库构建及测序 按照制造商说明书进行 SMRTbell 文库构建, 采用 MagBead Loading 上机测序获得原始数据, 使用 SMRT Link v6.0 对原始数据进行分析提取 CCS (Circular Consensus Sequence) 序列, 用 ICE (Interactive Clustering and Error Correction) 算法和 Quiver 算法对 FLNC (full-length non-chimeric) 进行聚类 and 校正, 得到的全长转录组数据已提交至 NCBI 数据库 (SRR9644402)。

1.3 基因比对及功能注释 Unigene 在 Nt、Nr、Swiss-Prot、KEGG、COG 及 Pfam 数据库中比对并获得基因注释信息 (E 值 $< 1 \times 10^{-5}$)。使用 Blast2GO 软件得到序列的 GO 注释信息, 根据 KEGG 注释信息进一步得到通路的注释信息, 利用 iTAK 软件预测转录因子 (TF)。

1.4 SSR、lncRNA 及可变剪切分析 使用 MISA (<http://pgrc.ipk-gatersleben.de/misa/>) 软件对转录组数据中所有 unigenes 进行检索并鉴定 unigenes 中的串联重复单元 (Simple sequence repeat, SSR)。对未注释到数据库的全长序列进行 lncRNA 分析, 使用 cnci 软件^[16] 和 CPC 软件^[17] 进行编码能力预测得到 lncRNA。使用 Cogent^[18] 软件组装出编码序列, 然后以组装好的序列作为参考, 使用 SUPPA^[19] 软件进行可变剪切分析。

1.5 RT-PCR 验证 为了验证通过 SMRT 技术获得的转录本准确性, 提取马氏珠母贝血细胞总 RNA, 反转录为 cDNA 第一链, 基于 SMRT 获得的 unigenes 中随机选择 5 个基因进

行特异性引物设计 (表 1), 使用 RT-PCR 及产物回收送至上海生工生物技术公司进行测序。

表 1 SMRT 测序数据统计

Table 1 Statistics of sequenced data by SMRT

类别 Category	数目 Number	平均长度 Average length//bp	N50 长度 N50 length//bp
subread	7 640 496	1 838	2 563
CCS	346 737	2 639	—
FLNC	277 064	2 296	—
ICE consensus	169 961	2 384	—
unigenes	82 381	2 512	2 861
annotation gene	59 621	—	—

2 结果与分析

2.1 SMRT 测序数据统计 利用 SMRT 测序技术, 共获得 7 640 496 个 subreads, 平均长度为 1 838 bp, N50 为 2 563 bp (表 1)。使用 Permissive CCS2 软件进行序列提取和质量过滤后, 共得到 346 737 条循环一致序列 (CCS)。其中 277 064 条为全长非嵌合序列 (FLNC), 平均长度为 2 296 bp。对上述 FLNC 进行聚类 and 校正后, 共获得 169 961 条一致性序列 (ICE consensus), 平均长度为 2 384 bp。序列去冗余后, 最终获得 82 381 条校正序列 (unigenes), 平均长度为 2 512 bp, N50 为 2 861 bp (表 1)。

2.2 基因注释及功能分类 使用 BLAST 工具书将所有的 unigene 与 NR、NT、Pfam、Siss-Prot 等蛋白序列数据库进行对比, 结果共得到 59 621 个注释基因 (表 1)。GO 功能注释结果表明, 基因在生物过程中主要富集在细胞进程、代谢过程、单生物体进程, 在细胞组分中主要富集在细胞和细胞部分, 在分子功能中主要富集在结合和催化活性 (图 1)。KEGG 通路注释显示代谢途径、次生代谢物的生物合成、吞噬体等通路较为富集 (表 2)。

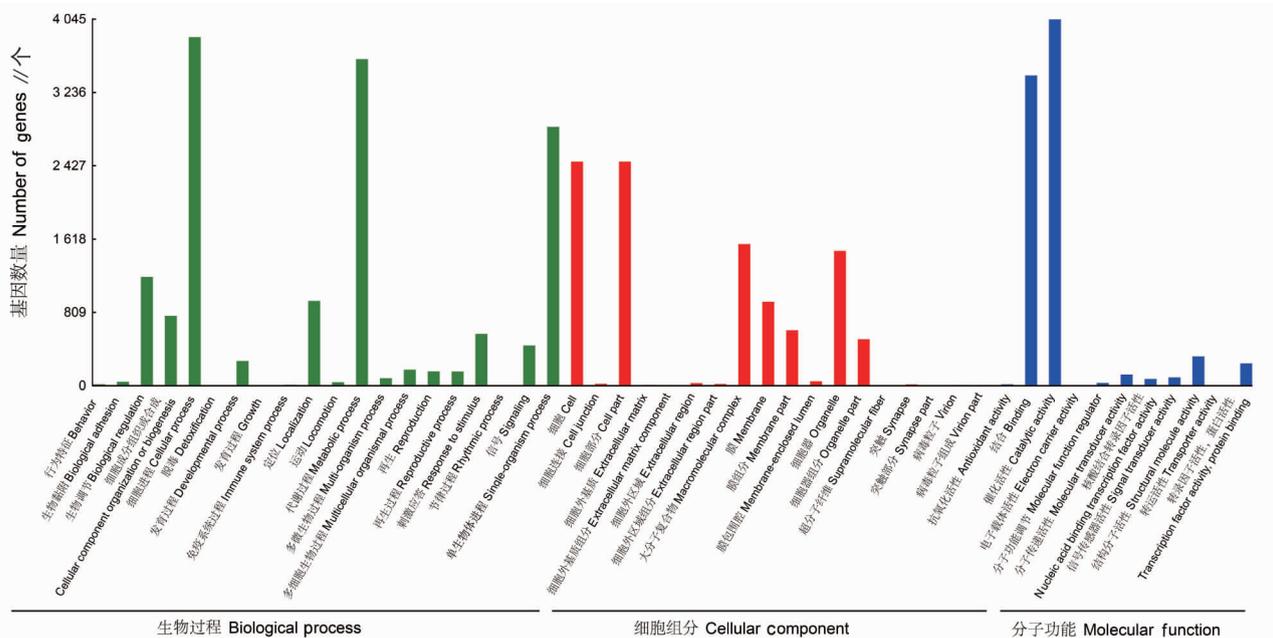


图 1 GO 功能注释

Fig.1 GO function annotation

表 2 马氏珠母贝转录组 unigene 的 KEGG 功能分类统计(前 10 个)

Table 2 KEGG functional classification statistics for *P.fucata* (Shown top ten contents)

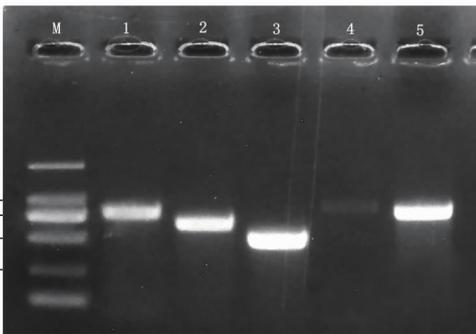
序号 No.	通路 Pathway	注释基因数目 Annotated genes number	百分比 Percentage//%
1	代谢途径	4 289	24.90
2	次生代谢物的生物合成	1 642	9.53
3	吞噬体	1 519	8.82
4	内吞作用	1 380	8.01
5	抗生素的生物合成	1 185	6.88
6	不同环境中的微生物代谢	1 143	6.64
7	RNA 转运	1 028	5.97
8	核糖体	991	5.75
9	溶酶体	987	5.73
10	内质网中的蛋白质加工	884	5.13

2.3 转录本验证 为了验证 SMRT 挖掘基因的准确性,随

表 3 PCR 反应中使用基因序列及其引物序列

Table 3 Special primers of random genes in PCR

序列 id Squence id	上游引物 Forward primer (5'—3')	下游引物(5'—3')Reverse primer
Isoform0058059	F1:AGCAACACCTGACCCGC	R1:ACTGTTTTGAATCGTTGAATCG
Isoform0068805	F2:CAACCAGAAATGTCTCAAATCCT	R2:CACACTGTTTTGAATCGTTGAATC
Isoform0000002	F3:ATAAAGCATCTGTCCACATCAA	R3:GTTTCAAAGCAGGTCTGATGGA
Isoform0000048	F4:TCGGGTAGACCAGAGACAAG	R4:TTTGTCTTTTCTCTCTCTCTGTA
Isoform0010136	F5:GGTAATGAGATAGACAAGACGGT	R5:CAATGGGTAGGGGAAGATC



注:1~5.随机挑选基因; M.DNA Marker

Note:1-5.Random selected genes;M.DNA Marker

图 2 PCR 反应验证基因长度

Fig.2 Validation of gene length by PCR reaction

机选择 5 个基因(特异性引物见表 3)进行 PCR 反应并进行纯化回收测序。结果(图 2)显示,基于全长转录组测序获得的基因序列,通过设计特异性引物,均能扩增出目标大小的 DNA 片段,且 PCR 产物经上海生工生物技术公司测序也验证了序列的正确性。

2.4 转录因子预测及 lncRNA 鉴定 转录因子是一类能够调控转录起始、激活或抑制转录过程的蛋白质,能够对细胞环境变化作出应答^[20]。利用 iTAK 软件预测得到 2 004 个转录因子(TF),其中预测得到的转录因子家族中,zf-C2H2 有 677 个、TF_bZIP 有 167 个、Homeobox 有 118 个、bHLH 有 116 个,可见这些转录因子较为丰富(图 3)。通过 cnci 软件和 CPC 软件进行编码能力预测,共鉴定到 20 013 个 lncRNA(图 4)。

2.5 串联重复单元(SSR) 全长转录组结果显示,在 8 493 条序列中共鉴定得到 11 219 个 SSRs,其中 1 973 条序列含有超过 1 个 SSR 位点。串联重复序列中二核苷酸和三核苷酸是主要的重复单元,其密度分别占总密度的 50.9%和 22.5%(表 4)。二核苷酸重复类别中,密度最大为 AT/AT,可见在马氏珠母贝体内重复序列主要为短的 2~6 bp 的微卫星重复(图 5)。

2.6 可变剪切分析 可变剪切分析结果显示,SMRT 测序共鉴定得到 10 522 个 AS,对 AS 类型进行分析显示,SMRT 技术鉴定的主要 AS 类型为 RI(图 6)。进一步分析结果显示,只有 1 条转录本的基因数目为 1 165 个(占 7.79%),含有 2 条转录本的基因数目为 3 268 个(21.86%),而含有>10 条转录本的基因数目为 2 187 个(14.63%)(图 7)。

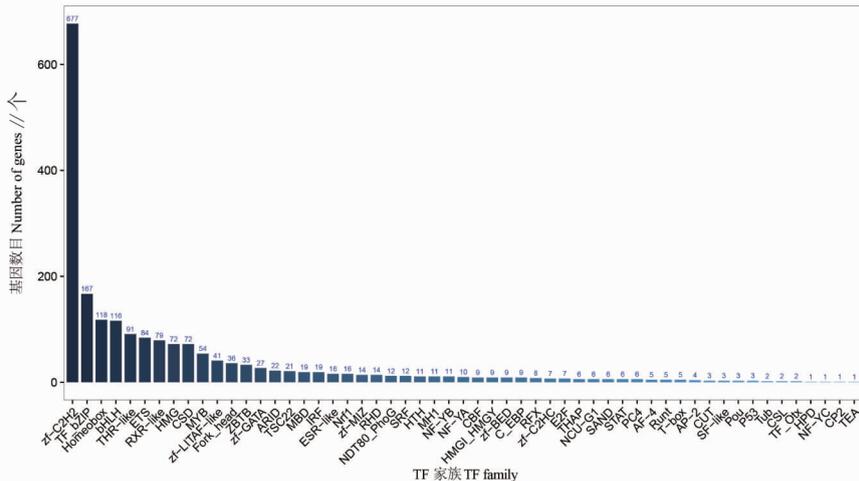


图 3 马氏珠母贝 PacBio 转录本中的 TF 家族

Fig.3 TF families in the PacBio of *P.fucata*

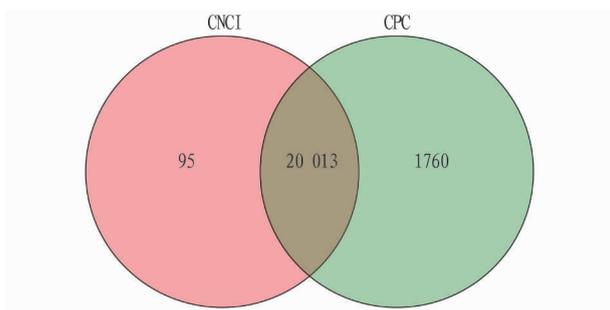


图 4 CNCI 和 CPC 预测结果的维恩图

Fig.4 Venn diagram of lncRNAs predicted by CPC and CNCI methods

表 4 马氏珠母贝全长转录组鉴定到的 SSRs

Table 4 Statistics of identified SSRs of *P.fucata*

序号 No.	类别 Category	数量 Number//个
1	鉴定的序列总数	82 381
2	已识别的 SSRs	11 219
3	包含 SSR 的序列	8 493
4	包含超过 1 个 SSR 的序列	1 973
5	二核苷酸	5 711
6	三核苷酸	2 520
7	四核苷酸	2 485
8	五核苷酸	334
9	六核苷酸	169

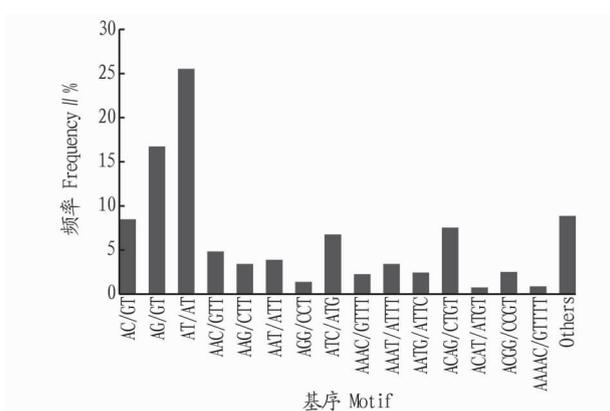


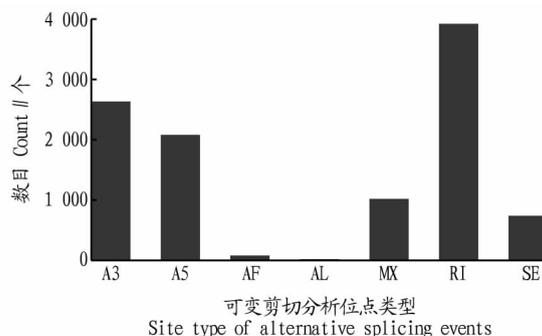
图 5 不同串联重复单元类型的 SSR 在总 SSR 中所占比例

Fig.5 SSR of different tandem repeating types in total SSR

3 讨论

目前, SMRT 测序手段已在多种水产动物中展开了广泛研究, 包括印度对虾^[7]、花鲈^[12]、斑节对虾 (*Penaeus monodon*)^[11]、金乌贼^[21] 等物种。该研究通过 SMRT 技术获得 346 737 个 CCS, 其中 277 064 个被鉴定为 FLNC。通过分析测序数据共获得 82 381 个 unigenes, 其中通过比对四大数据库注释得到 59 621 个注释基因。相较于第二代转录组测序获得 74 007 个 unigenes, 29 615 个注释基因的结果^[22], PacBio 三代全长转录组测序获得了更加全面的功能基因注释及分析研究。使用全长转录组测序进行分析可获得较长的序列, 同时能够避免二代测序中短序列拼接带来的序列错误^[23]。通过使用特异性引物进行 PCR 扩增及产物纯化测序, 可进一步证实 SMRT 测序在基因测序上的准确性。以上结果表

明, 通过 SMRT 技术进行全长转录组测序挖掘基因组序列是一种可靠的技术手段, 全长转录组数据的获得对于马氏珠母贝基因组注释和功能研究具有极其重要的作用。



注: SE. 外显子跳跃; MX. 外显子互斥; A5. 可变性 5' 剪切位点; A3. 可变性 3' 剪切位点; RI. 内含子保留; AF. 第一外显子可变剪接; AL. 最后外显子可变剪接

Note: SE. Skipped exon; MX. Mutually exclusive exon; A5. Alternative 5' splice site; A3. Alternative 3' splice site; RI. Retained intron; AF. First exon is alternative Splice; AL. Last exon is alternative splice

图 6 AS 数量及类型

Fig.6 Numbers and types of Alternative Splicing events

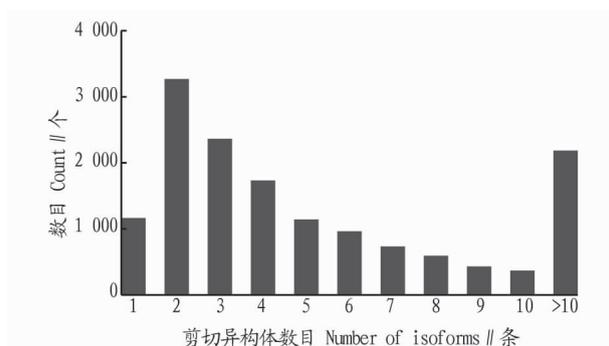


图 7 SMRT 测序转录本数目

Fig.7 Number of transcripts in SMRT sequence

对全长转录组测序获得的基因进行 GO 聚类后, 发现其在生物过程中主要富集于细胞进程、代谢过程、单生物体进程; 在细胞组分中主要富集在细胞和细胞部分; 在分子功能中主要富集在结合和催化活性。前期马氏珠母贝转录组数据中注释到 KEGG 数据库的序列有 7 156 个^[22], 而全长转录组注释序列增加至 37 185 个, 这表明 SMRT 全长转录组与二代短片段转录组测序相比, 在基因功能注释方面显著改善。KEGG 通路分析结果显示, 注释比例前三的通路为代谢途径、次生代谢物的生物合成、吞噬体, 暗示这些基因可能共同作用以维持马氏珠母贝的正常生命活动。

SSR 作为一种常用的遗传标记, 具有重复性好、共显性高、稳定性强、遗传信息多态性丰富等特点^[24]。该研究在马氏珠母贝全长转录组中共鉴定得到 11 219 个 SSRs, 分布在 8 493 个 unigenes 中, SSR 位点发生频率为 10.31%。通过转录组测序平台开发 SSR 标记已在多种水产经济物种中展开广泛研究, 其 SSR 发生频率与栉江珧 (*Atrina pectinata*) (8.2%)^[25]、缢蛏 (*Sinonovacula constricta*) (8.89%)^[26]、泥蚶

(*Tegillarca granosa*) (14.83)^[27] 较为接近。SSR 的发生频率与检索标准、数据库丰富程度以及物种差异等因素有关^[28]。前期研究中使用二代转录组从马氏珠母贝中分析发现 SSR 的发生频率为 13.34%^[29],而在该研究中发生频率下降。张秀英等^[30] 研究发现,栉孔扇贝 (*Chlamys farreri*) SSR 出现频率随检索标准的提高而降低,与该研究结论一致。基于全长转录组鉴定得到的 SSR 重复类型中以二核苷酸最为丰富,其次为三核苷酸重复基元,且二核苷酸重复基元类型约为 50.9%,这和虾夷扇贝 (*Mizuhopecten yessoensis*)^[31] 和栉江珧^[25] 的 SSR 中二核苷酸重复基元类型占主导地位的研究结果类似。

剪切异构体是 mRNA 前体转录后通过不同剪切方式获得的多个成熟 mRNA,在大多数真核生物中普遍存在,具有调节基因表达和蛋白质多样性的重要作用^[12]。可变性剪切是导致基因组功能复杂性的重要调控机制之一,具有调控组织特异性表达的功能^[32]。从凡纳滨对虾 (*Litopenaeus vannamei*) 转录组数据中筛选到 E75 基因具有 6 种可变剪切异构体,通过双链 RNA 干扰方法证明 E75 基因对凡纳滨对虾的蜕皮和生存具有重要作用^[33]。凡纳滨对虾受到细菌感染和低盐度刺激后,具有 AS 的基因表达量显著增加,表明 AS 应激反应基因可能具有响应病原体感染、环境改变等外界压力的功能^[34]。该研究通过 SMRT 测序发现,马氏珠母贝中存在 10 522 个 AS,其主要 AS 类型为 RI,推测其在马氏珠母贝生命活动中扮演着重要角色,但其具体功能仍需进一步研究。

4 结论

该研究对马氏珠母贝血细胞的全长转录组数据进行分析,共获得 82 381 个 unigenes,其中 59 621 个 unigenes 得到功能注释。同时,还筛选到 11 219 个 SSR 位点、10 522 个 AS 和 2 004 个转录因子。以上研究结果将大大丰富马氏珠母贝的基因数据库资源,为进一步研究马氏珠母贝功能基因挖掘、生长性状相关基因标记的开发及免疫响应机制等提供分子基础。

参考文献

- [1] 魏伟,朱国萍,章超桦,等.马氏珠母贝酶解蛋白粉营养成分及其对斑马鱼学习记忆能力的影响[J].广东海洋大学学报,2021,41(3):74-81.
- [2] 郝瑞娟.基于全基因组关联分析解析马氏珠母贝生长和矿化性状关键基因[D].湛江:广东海洋大学,2020.
- [3] 杜晓东,焦钰,邓岳文,等.马氏珠母贝 *Pinctada martensii* 珍珠囊发育的超微结构观察[J].海洋学报(中文版),2010,32(5):160-164.
- [4] 闫怡.基质蛋白 P1Y2 矿化功能研究及珍珠层蛋白复合物的鉴定[D].北京:清华大学,2019.
- [5] 魏雯璐,刘琪,张丞澍,等.不同温度下马氏珠母贝干露耐受能力及其免疫相关基因表达的变化[J].广东海洋大学学报,2021,41(2):69-75.
- [6] 庞丹丹,刘玉飞,孙云南,等.苦茶全长转录组测序及基因结构分析[J].西北农业学报,2021,30(4):563-571.
- [7] KATNENI V K, SHEKHAR M S, JANGAM A K, et al. Novel isoform sequencing based full-length transcriptome resource for Indian white shrimp, *Penaeus indicus* [J]. Front Mar Sci, 2020, 7: 1-4.
- [8] JIA D, WANG Y X, LIU Y H, et al. SMRT sequencing of full-length transcriptome of flea beetle *Agasicta hygrophila* (Selman and Vogt) [J]. Sci Rep, 2018, 8: 1-8.
- [9] CHEN Y Y, WU X Y, LAI J S, et al. Full-length transcriptome sequencing and identification and immune response of TRIM genes in Dabry's sturgeon

- (*Acipenser dabryanus*) [J/OL]. Aquaculture, 2021, 538 [2021-06-17]. https://doi.org/10.1016/j.aquaculture.2021.736599.
- [10] KUO R I, TSENG E, EORY L, et al. Normalized long read RNA sequencing in chicken reveals transcriptome complexity similar to human [J]. BMC Genom, 2017, 18(1): 1-19.
- [11] POOTAKHAM W, UENGWETWANT T, SONTTHIROD C, et al. A novel full-length transcriptome resource for black tiger shrimp (*Penaeus monodon*) developed using isoform sequencing (Iso-Seq) [J]. Front in Mar Sci, 2020, 7: 1-5.
- [12] TIAN Y, WEN H S, QI X, et al. Characterization of full-length transcriptome sequences and splice variants of *Lateolabrax maculatus* by single-molecule long-read sequencing and their involvement in salinity regulation [J]. Front Genet, 2019, 10: 1-19.
- [13] ZHANG J Y, LIU C L, HE M C, et al. A full-length transcriptome of *Sepia esculenta* using a combination of single-molecule long-read (SMRT) and Illumina sequencing [J]. Mar Genom, 2019, 43: 54-57.
- [14] LIU J J, JIANG H, ZAN J W, et al. Single-molecule long-read transcriptome profiling of *Platysternon megalcephalum* mitochondrial genome with gene rearrangement and control region duplication [J]. RNA Biol, 2018, 15(9): 1244-1249.
- [15] MARDIS E R. The impact of next-generation sequencing technology on genetics [J]. Trends Genet, 2008, 24(3): 133-141.
- [16] SUN L, LUO H T, BU D C, et al. Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts [J]. Nucleic Acids Res, 2013, 41(17): 1-8.
- [17] KONG L, ZHANG Y, YE Z Q, et al. CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine [J]. Nucleic Acids Res, 2007, 35: W345-W349.
- [18] LI J, HARATA-LEE Y, DENTON M D, et al. Long read reference genome-free reconstruction of a full-length transcriptome from *Astragalus membranaceus* reveals transcript variants involved in bioactive compound biosynthesis [J]. Cell Discov, 2017, 3: 1-13.
- [19] ALAMANCOS G P, PAGÈS A, TRINCADO J L, et al. Leveraging transcript quantification for fast computation of alternative splicing profiles [J]. RNA, 2015, 21(9): 1521-1531.
- [20] LATCHMAN D S. Transcription factors: An overview [J]. Int J Biochem Cell Biol, 1997, 29(12): 1305-1312.
- [21] 张金勇, 何暮春, 项子龙, 等. 基于全长转录组测序的金乌贼微卫星位点筛选与特征分析 [J]. 渔业科学进展, 2020, 41(6): 149-155.
- [22] WANG Z L, WANG B, CHEN G, et al. Transcriptome analysis of the pearl oyster (*Pinctada fucata*) hemocytes in response to *Vibrio alginolyticus* infection [J]. Gene, 2016, 575: 421-428.
- [23] 尚晓尧, 周玲芳, 尹芊芊, 等. 蒺藜苜蓿 (*Medicago truncatula*) 全长转录组测序及分析 [J]. 生物技术通报, 2021, 37(8): 131-140.
- [24] TULER A C, CARRIJO T T, NÓIA L R, et al. SSR markers: A tool for species identification in *Psidium* (Myrtaceae) [J]. Mol Biol Rep, 2015, 42(11): 1501-1513.
- [25] 李东明, 杨爱国, 吴彪, 等. 栉江珧 (*Atrina pectinata*) EST-SSR 标记的开发与应用 [J]. 渔业科学进展, 2017, 38(2): 137-142.
- [26] 刘博, 邵艳卿, 滕爽爽, 等. 缢蛏 (*Sinonovacula constricta*) EST-SSR 分布特征及引物开发利用 [J]. 海洋与湖沼, 2012, 43(1): 132-137.
- [27] 周小龙, 朱靖华, 董迎辉, 等. 泥蚶 (*Tegillarca granosa*) 基因组 SSR 和 EST-SSR 的开发及比较研究 [J]. 海洋与湖沼, 2013, 44(2): 467-475.
- [28] VARSHNEY R K, GRANER A, SORRELLS M E. Genic microsatellite markers in plants: Features and applications [J]. Trends Biotechnol, 2005, 23(1): 48-55.
- [29] 王忠良, 丁熹, 许尤厚, 等. 基于转录组数据的马氏珠母贝 EST-SSR 位点的信息分析及其多态性检测 [J]. 海洋与湖沼, 2015, 46(3): 687-693.
- [30] 张秀英, 张晓军, 赵翠, 等. 栉孔扇贝 BES-SSR 的开发及遗传多样性分析 [J]. 水产学报, 2012, 36(6): 815-824.
- [31] 李云峰, 刘卫东, 高祥刚, 等. 虾夷扇贝外套膜和肾脏组织 cDNA 文库构建以及 EST 的初步分析 [J]. 中国水产科学, 2010, 17(3): 578-585.
- [32] MODREK B, LEE C A. A genomic view of alternative splicing [J]. Nat Genet, 2002, 30(1): 13-19.
- [33] 杜江丽, 张晓军, 张小溪, 等. 凡纳滨对虾 E75 基因可变剪切形式的鉴定与分析 [J]. 水产学报, 2019, 43(4): 771-781.
- [34] ZHANG X X, YUAN J B, ZHANG X J, et al. Genome-wide analysis of alternative splicing provides insights into stress response of the Pacific white shrimp *Litopenaeus vannamei* [J]. Front Genet, 2019, 10: 1-11.