

研究报告

Research Report

基于高通量测序的凡纳滨对虾的转录组分析

曾地刚 陈秀荔 谢达祥 赵永贞 杨春玲 马宁 李咏梅 陈晓汉*

广西壮族自治区水产研究所, 广西水产遗传育种与健康养殖重点实验室, 南宁, 530021

* 通讯作者, gxnnscs@126.com

摘要 凡纳滨对虾是世界养殖最广泛的甲壳动物, 但是目前凡纳滨对虾的基因组及转录组数据还比较缺乏。为了获得凡纳滨对虾的转录组信息, 本研究应用 454 高通量测序技术对凡纳滨对虾肝胰腺的转录组进行测序。获得了 500 177 条凡纳滨对虾 EST, 平均长度 363 bp。拼接获得了 20 225 条 unigene, 长度范围 50~8 980 bp, 平均长度 507 bp。所有 unigene 与 NCBI 的非冗余蛋白质数据库(Nr)进行相似搜索(E 值 $<10^{-5}$) 结果一共有 13 676 条 unigene (68%)与数据库中的已知基因同源。此外, 还对 unigene 进行了 GO、COG 和 KEGG 的功能注释、分类或通路分析。我们通过高通量测序, 获得了丰富的凡纳滨对虾转录组信息, 为凡纳滨对虾的新基因克隆和基因组学研究提供了有价值的信息。

关键词 高通量测序, 凡纳滨对虾, 转录组

Deep Sequencing-based Transcriptome Analysis of *Litopenaeus vannamei*

ZengDigang ChenXiuli XieDaxiang ZhaoYongzhen YangChunling MaNing LiYongmei ChenXiaohan*

Guangxi Key Laboratory of Aquatic Genetic Breeding and Healthy Aquaculture, Guangxi Institute of Fisheries, Nanning, 530021

*Corresponding authors, gxnnscs@126.com

DOI: 10.3969/gab.032.000308

Abstract *Litopenaeus vannamei* shrimp is the most extensively farmed crustacean species in the world. However, the genome and transcriptome data of *Litopenaeus vannamei* is still relatively lacking. To obtain transcriptome information of *Litopenaeus vannamei*, the transcriptome of *Litopenaeus vannamei* hepatopancreas was sequenced by 454 high-throughput sequencing technology. A total of 500 177 ESTs were produced, with an average length of 363 bp. De novo assembly of ESTs data generated 20 225 unigenes between 50 and 8 980 bp in length, with an average length of 507 bp. Similarity searches against the NCBI-Nr database revealed that 13 676 (68%) of these unigenes have significant matches (E -value $<10^{-5}$). In addition, the unigenes were similarity search against the GO, COG and KEGG databases, and were annotated with gene functional descriptions, gene ontology terms, or pathways. By high-throughput sequencing, we obtained abundant transcriptome information that could contribute to novel gene identification and genome research in *Litopenaeus vannamei*.

Keywords High-throughput sequencing, *Litopenaeus vannamei*, Transcriptome

凡纳滨对虾(*Litopenaeus vannamei*) ,又称南美白对虾, 是全球对虾养殖量最大的品种之一(Zhou et al., 2012) ,也是我国南方主要的对虾养殖品种。凡纳滨对虾具有壳薄体肥、肉质鲜嫩、生长迅速、群体增长均匀、抗病力强等优点, 而深受国内外市场的青睐(马宁等, 2008)。由于凡纳滨对虾巨大的经济价值和重要的进化地位, 吸引了越来越多的学者对它的生长、发育、繁殖、免疫和遗传等进行研究(Liu et al., 2010)。但

是, 目前凡纳滨对虾的基因组及转录组数据还比较缺乏, 影响了这方面的研究。基因表达序列标签(expressed sequence tags, EST)可以用来进行转录组的研究。迄今为止, GenBank 已有 162 933 条凡纳滨对虾 EST, 这些 EST 数据大多数是通过传统的 Sanger 测序方法获得的。这些 EST 数据可以用于克隆功能基因, 挖掘分子标记, 设计 cDNA 芯片。但现有数据仍然难以满足研究需求。此外, 这些凡纳滨对虾 EST 只

基金项目: 本研究由国家自然科学基金项目(31160531)资助

有不到 1 万条已经被拼接和注释,很大程度上限制了该 EST 数据的使用。目前已经有一些凡纳滨对虾的转录组相关的研究报告(Robalino et al., 2007; Clavero-Salas et al., 2007),但是主要是使用传统的 Sanger 测序或 cDNA 芯片的方法进行的。传统的 Sanger 测序方法需要先构建一个 cDNA 文库,然后用对挑选的克隆进行测序,不但实验步骤比较繁琐,而且花费的时间和成本比较高。cDNA 芯片方法虽然是一种快速和高通量的方法,但是需要预先知道基因的序列,而目前凡纳滨对虾对虾的基因组资源还比较少,限制了该方法的应用。

高通量测序是在芯片上并行地对数百万计的 DNA 分子进行测序,一次可以产生巨大数据量的测序结果,它显示了测序技术划时代的变革,因此被称为新一代测序技术。同时高通量测序能够对一个物种的基因组或转录组进行深入的分析,因此又被称为深度测序(Margulies et al., 2005)。下一代测序技术又分为第二代和第三代测序技术,第二代测序技术的核心思想是边合成边测序,即通过捕捉新合成的末端的标记来确定 DNA 的序列。第三代测序技术则是基于纳米孔的单分子测序。目前第二代测序技术的技术平台主要包括 Roche 454、Illumina Solexa 和 Applied Biosystems Solid (Liu et al., 2012)。这三个技术平台各有优点,其中 Roche 454 的测序片段比较长,目前高质量的读长能达到 400 bp 以上,与 ABI 的 Solid 和 Illumina 公司的 Solexa 相比,由于读长较长,尤其适合没有基因组参考序列的物种的基因组或转录组测序(You et al., 2011)。本研究我们采用 454 焦磷酸测序技术,对凡纳滨对虾的肝胰腺进行高通量转录组测序,并对所得的 EST 进行数据分析。在本研究中,我们获得了超过 50 万条高质量的 EST 序列,并进行了序列拼接、功能注释和 KEGG 通路分析等研究,这些测序结果和分析为凡纳滨对虾的基因组学研究提供了丰富的资源,促进了我们对凡纳滨对虾转录组的了解。

1 结果与分析

1.1 测序和序列拼接

我们用凡纳滨对虾肝胰腺提取的 mRNA 构建了 cDNA 文库,对文库进行 454 焦磷酸测序。所得高通量测序数据提交至 NCBI sequence read archive (SRA) 登录号 SRX181883。去除原始序列中的低质量、短序列和接头序列后,结果获得了 500 177 条凡

纳滨对虾 EST 序列,长度范围从 41 bp 至 620 bp,平均长度 363 bp。拼接获得了 20 225 条单基因序列(unigene),长度范围从 50 bp 至 8 980 bp,平均长度 507 bp (图 1)。

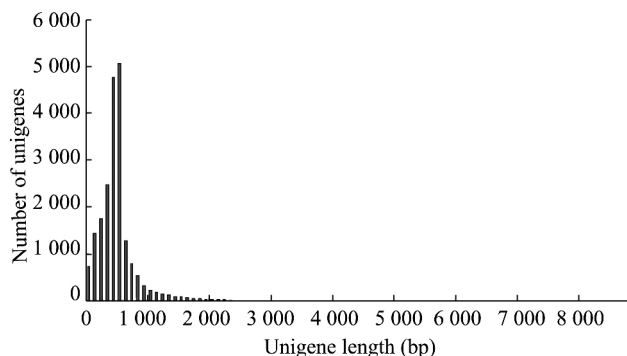


图 1 Unigene 的长度分布

注: X 轴表示 unigene 的长度(bp); Y 轴表示相应的基因数量

Figure 1 Length distribution of unigenes

Note: The X axis shows the unigene length (bp); the Y axis shows the number of the corresponding unigenes

1.2 基因功能注释

所有的 unigene 与 NCBI (National Center of Biotechnology Information) 的非冗余蛋白质数据库(Nr)进行相似搜索(E 值 $<10^{-5}$) 结果一共有 13 676 个 unigene (68%)与数据库中已知基因同源,而其余 6 549 个 unigene (32%)与数据库中的已知基因同源性较低,可能属于新基因。

1.3 COG 功能注释及分类

COG 是一个用来归类基因产物的数据库,其中每一个蛋白质被假定为来自同源的祖先蛋白质,整个数据库是根据细菌、藻类和真核细胞具有完整的基因组编码蛋白以及清晰的系统演化关系进行构建的(Tatusov et al., 2000)。将所有 unigene 与 COG 数据库进行比对,结果有 4 645 个 unigene 得到注释,并被分为 25 个 COG 类(图 2)。其中,一般功能预测(general function prediction only)所占的比例最大,其次是翻译后修饰,蛋白质折叠,分子伴侣(posttranslational modification, protein turnover, chaperones)和信号传导机制(signal transduction mechanisms)。而核结构(nuclear structure)、防御机制(defense mechanisms)和细胞活动(cell motility)等类别所占的比例最小。

1.4 GO 功能注释及分类

GO 是一个采用动态更新的标准词汇表来描述基因和其产物功能的数据库,目前被广泛应用于生

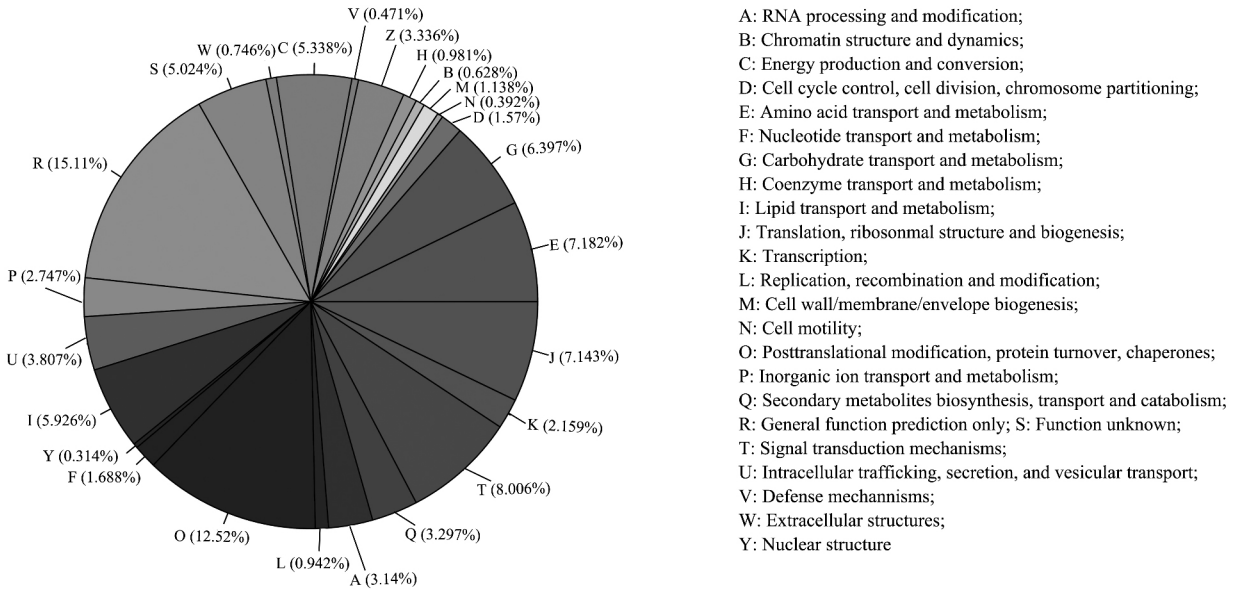


图 2 Unigene 的 COG 分类

注: 图中的百分数代表各个 COG 功能分类所占的比例

Figure 2 COG classification of the unigenes

Note: The percentages in the figure represent the proportion of different COG categories

物的转录组数据分析研究中(Conesa and Gotz, 2008), GO 总共分为 3 大功能类, 分别描述基因的分子功能(molecular function)、所处的细胞位置(cellular component)和参与的生物过程(biological process)。所有的 unigene 与 GO 数据库进行比对的结果表明, 一共有 4 067 条 unigene (20%)被注释和分类到 49 个功能亚类中(图 3)。结果显示, 凡纳滨对虾的细胞杀伤(cell killing)、节律过程(rhythmic process)、突触(synapse)、通道调节活动(channel regulator activity)、金属分子伴侣活动(metallo chaperone activity)、核酸结合的转录因子的活动(nucleic acid binding transcription factor activity)和转录调节活动(translation regulator activity)等 8 个功能亚类的 Unigene 很少, 而其余的功能亚类的 Unigene 分布比较均衡。

1.5 KEGG Pathway 注释

KEGG 数据库记录细胞中基因产物的功能以及基因产物的相互作用网络, 基于 KEGG 通路分析有助于我们进一步了解基因的生物学功能(Ogata et al., 1999)。分析结果, 一共有 4 104 条 unigene (20%)被注释, 并被归类到 176 个 KEGG 通路。图 4 显示了基因数量排名前 20 位的 KEGG 通路, 包括代谢途径(metabolic pathways)、吞噬体(phagosome)、粘着(focal adhesion)、紧密连接(tight junction)、粘着连接(adherens junction)、次生代谢物合成(biosynthesis of secondary metabolites)、溶酶体(lysosome)、核糖体(ri-

bosome)、氧化磷酸化(oxidative phosphorylation)、蛋白质消化吸收(protein digestion and absorption)、酪氨酸代谢(tyrosine metabolism)、过氧化物酶体(peroxisome)、细胞色素 P450 介导的外源物代谢(metabolism of xenobiotics by cytochrome P450)、氨基糖和核苷酸代谢(amino sugar and nucleotide sugar metabolism)、内质网中的蛋白质过程(protein processing in endoplasmic reticulum)、脂肪酸代谢(fatty acid metabolism)、核黄素代谢(riboflavin metabolism)、氨基苯甲酸降解(aminobenzoate degradation)、异喹啉类生物碱生物合成(isoquinoline alkaloid biosynthesis)和 RNA 运输(RNA transport)。

2 讨论

转录组是在一个或多个细胞中表达的 RNA 转录物的集合(Etebari et al., 2011)。通过转录组分析可以帮助我们在整体水平上研究细胞中基因转录的情况及转录调控规律。过去的转录组研究通常采用 Sanger 测序和 cDNA 芯片方法, 而现在的转录组研究越来越多采用高通量测序技术。和 cDNA 芯片方法相比, 高通量转录组测序不需要预先针对已知序列设计探针, 可以对任意物种的转录组进行测序, 提供更高的检测通量和花费更低成本和时间的, 是研究转录组更理想的方法。目前已经有一些甲壳动物, 例如罗氏沼虾(*Macrobrachium rosenbergii*)、中华绒螯蟹

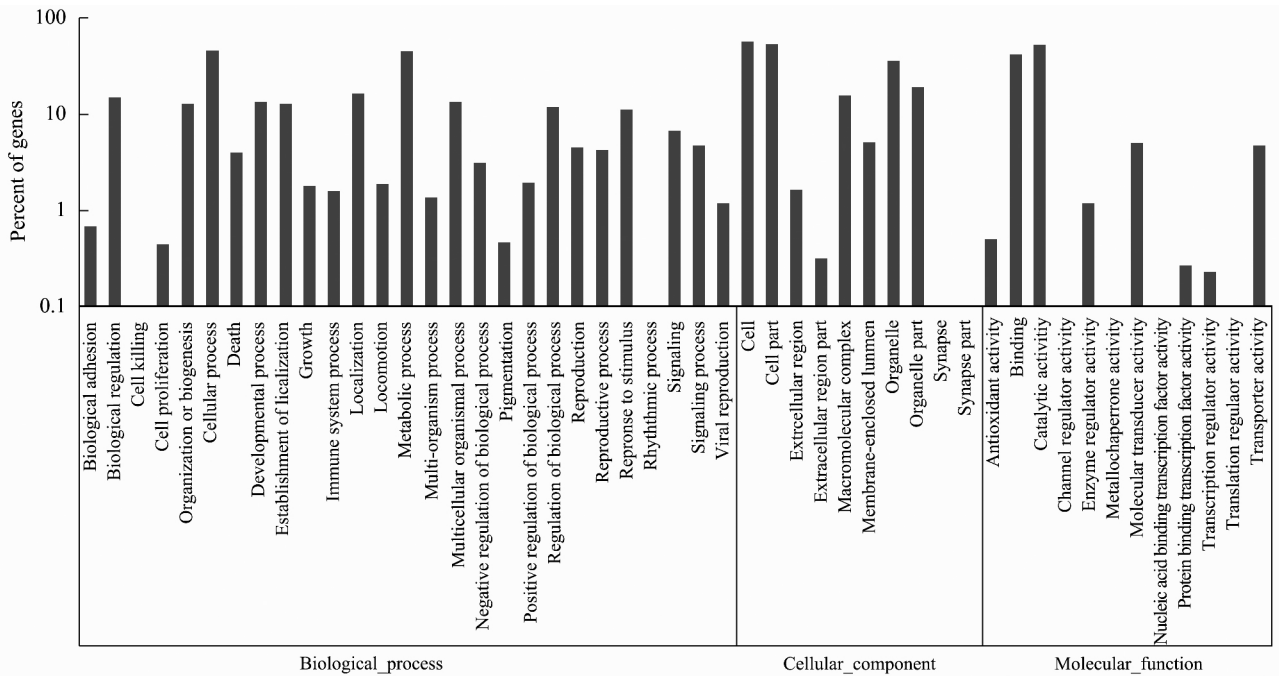


图3 凡纳滨对虾基因的 GO 功能分类

注: X 轴表示 GO 功能的亚类; Y 轴表示注释基因的百分比

Figure 3 Gene Ontology (GO) classification of putative functions of unigenes from *Litopenaeus vannamei*

Note: X axis: Subgroups of molecular functions from GO classification; Y axis: The percentage of the matched unigenes

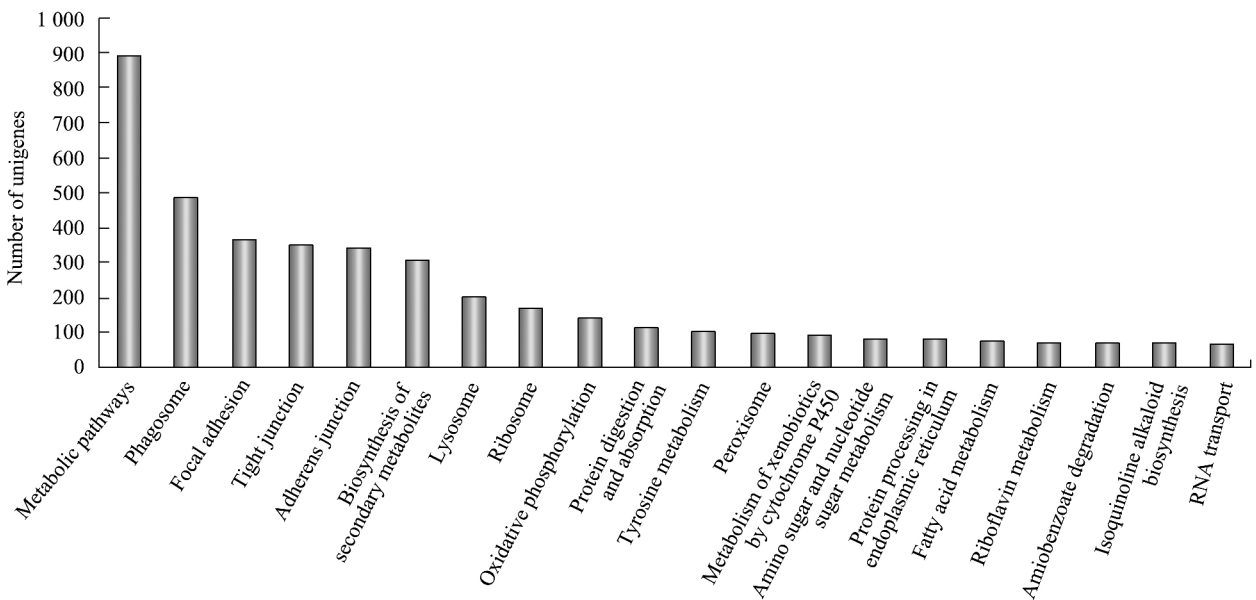


图4 Unigene 的 KEGG 分析, 显示基因数量排名前 20 位的 KEGG 通路

注: X 轴表示 KEGG 分类, Y 轴表示匹配的基因数量

Figure 4 KEGG Classification of the unigenes, the top 20 most abundant KEGG pathways are shown

Note: The x axis shows pathways from KEGG classification and the y axis shows the number of the matched unigenes

(*Eriocheir sinensis*)和藤壶(*Balanus amphitrite*)等用高通量测序技术进行了转录组研究(de Gregoris et al., 2011; He et al., 2012; Jung et al., 2011)。本研究采用 454 焦磷酸高通量测序技术,对凡纳滨对虾肝胰腺的

转录组进行测序和功能分析,拼接共得到 20 225 条 unigene,与 NCBI 非冗余蛋白质数据库(Nr)比对的结果显示,有 6 549 条 unigene 与数据库中的已知基因同源性较低,占总数的 32%,可能是未知基因。其余

的 13 676 条 unigene (68%)与数据库中的已知基因同源性比较高,其中和数据库已有的凡纳滨对虾序列匹配的仅占 5.11%,而大部分是和其它物种匹配。这说明数据库中已有的凡纳滨对虾基因序列还很少,我们获得的 unigene 大大丰富了现有数据库中的凡纳滨对虾基因资源。本研究测序获得的序列平均长度为 363 bp,这已经和传统 Sanger 测序的读长相差不大。比较长的测序读长可以有效减少拼接的错误,提高拼接重叠群的长度,对于没有参照基因组的物种的测序尤为重要(Franssen et al., 2011)。截至目前,凡纳滨对虾还没有全基因组测序的报道。我们在没有参照基因组的情况下,对测序所得的序列进行从头拼接(de novo assembly)获得了 20 225 条 unigene,平均长度为 507 bp,通过抽样与 GenBank 中的序列比对发现,拼接的质量很好,没有发现拼接错误的情况。我们测序的结果表明 454 测序是一种高效率、低成本和高通量的转录组分析方法。

GO、COG 和 KEGG 注释对于深入了解基因的功能很重要。我们获得的 unigene 只有大约 20%被注释到 GO、COG 和 KEGG 数据库,这主要是因为目前国际公共基因数据库中收录甲壳动物基因还比较少,本研究获得的很多凡纳滨对虾 unigene 都搜索不到同源基因序列。尽管这样,这 3 个数据库注释结果可以帮助我们了解更多凡纳滨对虾生物学特性。通过这些注释,我们可以了解基因的分子功能、所处细胞位置、参与的生物过程、所处的代谢途径或信号通路等等,这为今后发掘凡纳滨对虾功能基因、研究相关生理功能提供了数据。例如,我们发现几百个基因和细胞免疫功能相关,这些基因序列将来可以用来制作基因表达谱芯片,用来检测凡纳滨对虾的免疫水平,作为凡纳滨对虾抗病品系选育的生化指标。

本研究通过高通量测序,获得了大量的凡纳滨对虾的转录组信息,为凡纳滨对虾的基因克隆、分子标记发掘和基因组学研究等提供了有价值的信息。今后,还将进行比较转录组的研究,包括不同发育阶段、不同性状、病毒感染和环境因子胁迫的凡纳滨对虾的转录组研究。部分工作目前正在进行中。

3 材料与方法

3.1 材料

实验对虾采自广西南美白对虾遗传育种中心的无特定病原(SPF)凡纳滨对虾家系,共 20 尾,平均体重约 11 g。实验对虾暂养在 25~26℃ 水温度的海水中。采样时取虾的肝胰腺组织,保存在液氮中备用。

3.2 方法

3.2.1 RNA提取

使用 Trizol 试剂盒(Qiagen)按照说明书的步骤提取凡纳滨对虾肝胰腺的总 RNA。用分光光度计测定 RNA 浓度,并通过 1.5% (重量/体积)琼脂糖凝胶电泳检测 RNA 的完整性。RNA 提取后,使用 PolyA Tract mRNA 分离系统(Promega)纯化 mRNA,并使用 Reasy RNA 试剂盒(Qiagen)收集。然后将 20 个样品的 mRNA 进行等量混合,用于 cDNA 合成。

3.2.2 cDNA 文库构建和测序

用 RNA Fragment reagent 试剂盒(illumina)将混合的 mRNA 碎裂成 300~800 bp 的小片段。片段化完成后,用 Reasy RNA cleaning 试剂盒(Qiagen)对片段化产物进行回收,回收后用随机引物和 MMLV 进行反转录合成第一链,合成完第一链后用 DNA Polymerase 和 RNase H 合成第二链。过程如下: (1)将回收后的 14 μ L RNA 加 2 μ L 400 μ mol/L N6 Radom Primer, 68℃ 温浴 10 min (2)向上述 RNA 管中加 2 μ L dNTP, 2 μ L DTT, 6 μ L 5 \times First Strand Buffer, 2 μ L MMLV RTase 和 2 μ L H₂O, 42℃ 反应 30 min (3)向上述反应液中加 5 μ L 10 \times DNA Polymerase Buffer, 1 μ L DNA Polymerase (10 U), 1 μ L RNase H (6 U), 加水至 50 μ L, 16℃ 温浴 2 h (4)反应完成后用 MinElute DNA Cleaning 试剂盒(Qiagen)回收产物 (5)接头连接:取上述回收产物 15 μ L, 向其中加 1 μ L Adaptor 和 2 μ L 10 \times DNA Ligase Buffer 和 2 μ L T4 DNA Ligase, 25℃ 温浴 10 min, 温浴完毕用 MinElute DNA Cleaning 试剂盒回收产物,在 454 GS FLX 系统(Roche)进行测序。

3.2.3 序列拼接、功能注释、分类和通路分析

用 SeqClean 程序(<http://compbio.dfc.harvard.edu/tgi/software>)去除接头的序列和低质量的序列。用 iAssembler 程序(<http://bioinfo.bti.cornell.edu/tool/i-Assembler>)进行序列拼接(Zheng et al., 2011)。拼接所得 unigene 分别与 NCBI Nr、SwissProt、GO、COG、KEGG 等数据库进行比对(E 值 $<10^{-5}$),并选取最佳注释(Zhang et al., 2010)。Nr 的注释工具采用 blastx, COG 和 SwissProt 的注释采用 blastx, GO 注释采用 blast-2GO, KEGG pathway 注释采用 blastx 和 GenMAPP 2.1 (<http://www.genmapp.org/>)。

作者贡献

曾地刚负责数据分析和论文撰写,陈秀荔、谢达

祥、赵永贞、杨春玲、马宁和李咏梅负责样品处理和实验工作, 陈晓汉负责总体设计及指导研究工作。

致谢

感谢广西南美白对虾遗传育种中心提供实验动物。本论文得到国家自然科学基金项目(31160531)资助, 在此表示感谢!

参考文献

- Clavero-Salas A., Sotelo-Mundo R.R., Gollas-Galvan T., Hernandez-Lopez J., Peregrino-Uriarte A.B., Muhlia-Almazan A., and Yepiz-Plascencia G., 2007, Transcriptome analysis of gills from the white shrimp *Litopenaeus vannamei* infected with white spot syndrome virus, *Fish & Shellfish Immunology*, 23(2): 459-472
- Conesa A., and Gotz S., 2008, Blast2GO: A comprehensive suite for functional analysis in plant genomics, *International Journal of Plant Genomics*, 2008: 619832
- De Gregoris T.B., Rupp O., Klages S., Knaust F., Bekel T., Kube M., Burgess J.G., Arnone M.I., Goesmann A., Reinhardt R., and Clare A.S., 2011, Deep sequencing of naupliar-, cyprid- and adult-specific normalised expressed sequence tag (EST) libraries of the acorn barnacle *Balanus amphitrite*, *Biofouling*, 27(4): 367-374
- Etebari K., Palfreyman R.W., Schlipalius D., Nielsen L.K., Glatz R.V., and Asgari S., 2011, Deep sequencing-based transcriptome analysis of *Plutella xylostella* larvae parasitized by *Diadegma semiclausum*, *BMC Genomics*, 12: 446
- Franssen S.U., Shrestha R.P., Brautigam A., Bornberg-Bauer E., and Weber A.P., 2011, Comprehensive transcriptome analysis of the highly complex *Pisum sativum* genome using next generation sequencing, *BMC Genomics*, 12: 227
- He L., Wang Q., Jin X., Wang Y., Chen L., and Liu L., 2012, Transcriptome profiling of testis during sexual maturation stages in *Eriocheir sinensis* using illumina sequencing, *PloS one*, 7(3): e33735
- Jung H., Lyons R.E., Dinh H., Hurwood D.A., McWilliam S., and Mather P.B., 2011, Transcriptomics of a giant freshwater prawn (*Macrobrachium rosenbergii*): De novo assembly, annotation and marker discovery, *PloS One*, 6(12): e27938
- Liu K.F., Chiu C.H., Shiu Y.L., Cheng W., and Liu C.H., 2010, Effects of the probiotic, *Bacillus subtilis* E20, on the survival, development, stress tolerance, and immune status of white shrimp, *Litopenaeus vannamei* larvae, *Fish & Shellfish Immunology*, 28(5-6): 837-844
- Liu L., Li Y., Li S., Hu N., He Y., Pong R., Lin D., Lu L., and Law M., 2012, Comparison of next-generation sequencing systems, *J. Biomed. Biotechnol.*, 2012: 251364
- Ma N., Chen X.H., Zeng D.G., Peng M., and Li Y.M., 2008, SNP analysis of the *Cathepsin L* gene in *Litopenaeus vannamei*, *Wuhan Daxue Xuebao (Lixueban)* (*Journal of Wuhan University: Natural Science Edition*), 54(4): 503-506 (马宁, 陈晓汉, 曾地刚, 彭敏, 李咏梅, 2008, 凡纳滨对虾组织蛋白酶L基因的单核苷酸多态性分析, *武汉大学学报(理学版)*, 54(04): 503-506)
- Margulies M., Egholm M., Altman W.E., Attiya S., Bader J.S., Bemben L.A., Berka J., Braverman M.S., Chen Y.J., Chen Z., Dewell S.B., Du L., Fierro J.M., Gomes X.V., Godwin B.C., He W., Helgesen S., Ho C.H., Irzyk G.P., Jando S.C., Alenquer M.L., Jarvie T.P., Jirage K.B., Kim J.B., Knight J.R., Lanza J.R., Leamon J.H., Lefkowitz S.M., Lei M., Li J., Lohman K.L., Lu H., Makhijani V.B., McDade K.E., McKenna M.P., Myers E.W., Nickerson E., Nobile J.R., Plant R., Puc B.P., Ronan M.T., Roth G.T., Sarkis G.J., Simons J.F., Simpson J.W., Srinivasan M., Tartaro K.R., Tomasz A., Vogt K.A., Volkmer G.A., Wang S.H., Wang Y., Weiner M.P., Yu P., Begley R.F., and Rothberg J.M., 2005, Genome sequencing in microfabricated high-density picolitre reactors, *Nature*, 437(7057): 376-380
- Ogata H., Goto S., Sato K., Fujibuchi W., Bono H., and Kanehisa M., 1999, KEGG: Kyoto encyclopedia of genes and genomes, *Nucleic Acids Research*, 27(1): 29-34
- Robalino J., Almeida J.S., McKillen D., Colglazier J., Trent H.F., Chen Y.A., Peck M.E., Browdy C.L., Chapman R.W., Warr G.W., and Gross P.S., 2007, Insights into the immune transcriptome of the shrimp *Litopenaeus vannamei*: Tissue-specific expression profiles and transcriptomic responses to immune challenge, *Physiological Genomics*, 29(1): 44-56
- Tatusov R.L., Galperin M.Y., Natale D.A., and Koonin E.V., 2000, The COG database: A tool for genome-scale analysis of protein functions and evolution, *Nucleic Acids Research*, 28(1): 33-36
- You F.M., Huo N., Deal K.R., Gu Y.Q., Luo M.C., McGuire P.E., Dvorak J., and Anderson O.D., 2011, Annotation-based genome-wide SNP discovery in the large and complex *Aegilops tauschii* genome using next-generation sequencing without a reference genome sequence, *BMC Genomics*, 12: 59
- Zhang F., Guo H., Zheng H., Zhou T., Zhou Y., Wang S., Fang R., Qian W., and Chen X., 2010, Massively parallel pyrosequencing-based transcriptome analyses of small brown planthopper (*Laodelphax striatellus*), a vector insect transmitting rice stripe virus (RSV), *BMC Genomics*, 11: 303
- Zheng Y., Zhao L., Gao J. and Fei Z., 2011, iAssembler: A package for de novo assembly of Roche-454/Sanger transcriptome sequences, *BMC Bioinformatics*, 12: 453
- Zhou J., Fang W., Yang X., Zhou S., Hu L., Li X., Qi X., Su H., and Xie L., 2012, A nonluminescent and highly virulent *Vibrio harveyi* strain is associated with "bacterial white tail disease" of *Litopenaeus vannamei* shrimp, *PloS One*, 7(2): e29961