

高通量测序技术及其在贝类转录组研究中的应用

刘敏, 叶晟, 杨凤, 吴立新* (大连海洋大学水产与生命学院, 辽宁大连 116023)

摘要 高通量测序是传统 DNA 测序的一次重大技术创新, 它为贝类养殖物种的研究带来了新的机遇。着重对 454、Solexa 和 SOLiD 高通量测序平台的技术原理、数据分析及其在贝类转录组的研究应用进行综述, 同时对高通量测序未来发展方向做了总结和展望。

关键词 高通量测序; 转录组; 贝类; 应用

中图分类号 S188 **文献标识码** A **文章编号** 0517-6611(2017)07-0130-04

High-throughput Sequencing Technology and Its Application in the Study of Shellfish Transcriptome

LIU Min, YE Sheng, YANG Feng, WU Li-xin* (College of Fisheries and Life Science, Dalian Ocean University, Dalian, Liaoning 116023)

Abstract High-throughput sequencing is a major technology innovation of traditional DNA sequencing, which brings new opportunities to the research of shellfish species. This review focused on their technical principles, data analysis of the 454, Solexa and SOLiD high-throughput sequencing platforms and their application in the study of shellfish transcriptome. At the same time, the future development direction of high-throughput sequencing was summarized and forecasted.

Key words High-throughput sequencing; Transcriptome; Shellfish; Application

核酸测序技术是一种重要的分子生物学分析方法, 被广泛地应用在生物科学研究中。20 世纪 70 年代, Sanger 提出了具有里程碑意义的双脱氧核苷酸末端终止法^[1], 该方法在“人类基因组计划”实施中发挥了巨大作用。随着测序技术的不断发展, 第一代测序 (Sanger 测序) 因其成本高、通量低和速度慢等缺点而不能满足研究者的需求^[2]。自 2000 年以后, 高通量测序技术平台相继出现, 主要包括罗氏的 454 测序平台、Illumina 的 Solexa 测序平台和 ABI 的 SOLiD 测序平台, 它们以其通量高、速度快、准确度高、成本低等特点而被广泛应用于非模式生物测序研究中。高通量测序技术在贝类物种的转录组研究中发挥了重要作用, 主要应用在贝类分子标记筛选、功能基因的挖掘和非编码 miRNA 方面的研究中。笔者综述了高通量测序技术的原理、数据分析及其在贝类转录组的研究应用, 并展望了高通量测序今后的研究发展方向。

1 高通量测序的原理及测序步骤

高通量测序 (High-throughput sequencing) 又叫作深度测序或下一代测序, 该技术是对 Sanger 测序的一次重大技术创新, 在一次反应中能对数百万条的核酸序列进行测定。目前, 高通量测序技术主要包括 Roche/454^[3], Illumina/Solexa^[4] 和 ABI/SOLiD 测序平台^[5], 下面主要针对这 3 种测序技术的原理进行介绍。

1.1 Roche/454 454 测序平台是采用边合成边测序的原理, 其测序步骤如下: ①DNA 文库构建, 将所测核酸序列打断成几百 bp 的小片段后, 在两端分别加上特异性的接头。②乳液 PCR, 上述步骤带接头的 DNA 片段会与一个磁珠结合, 磁珠被扩增试剂乳化之后, 形成油包水的微反应器系统。

测序 PCR 反应就发生在液滴包裹的微反应器里, 每一个油滴体系中只包含 1 个 DNA 模板。扩增以后, 每一个 DNA 分子都能进行大量富集, 最后构成一个克隆集落。454 测序仪的测序通道体积狭窄, 仅仅能容纳一个磁珠。③转移至 PTP 载板 (Pico Titer Plate), 将每个磁珠移至 GS FLX 测序仪载样板的每个小孔内, 为测序做准备。④焦磷酸测序, 在 PTP 孔内添加被激活液处理的磁珠, 测序系统会把底物的聚合与荧光信号释放偶联起来。当发生碱基配对时, 释放的焦磷酸基团在酶的作用下产生级联反应, 获得测序荧光信号。通过捕捉的荧光信号, 就能够达到核酸测序的目的。与其他高通量测序技术相比, 454 测序平台具有测序读长长和速度快的优点, 适合基因组的重测序、从头测序和宏基因组学等。然而, 454 测序平台也存在一些缺点, 如成本高、准确率低, 对相同碱基聚合区域难以处理。454 测序平台在性能上很难有升级和突破, 所以罗氏公司从 2016 年起逐步停止焦磷酸测序仪的生产^[6]。

1.2 Illumina /Solexa Solexa 测序平台的原理是边合成边测序, 它是基于桥式 PCR 和可逆终止子的反应, 测序步骤如下: ①构建文库, 首先把 DNA 序列随机打断成小片段, 并在两端加上特定的接头。②锚定桥接, 将第 1 步获得的 DNA 片段变性为单链, 并和接头引物结合, 形成后续扩增的桥状结构。放大反应在带有 8 个通道的 flow cell 玻璃管内进行, 每一个通道的内表面有许多个固定的单链接头。③预扩增, 添加 Taq 酶和不带标记的底物 dNTP 进行桥式 PCR 扩增。在变性时, 单链桥型片段锚定到附近的固相表面。经过连续的扩增反应, 上百万条成簇的双链 DNA 待测片段将会在 flow cell 的固相表面上。④单碱基延伸测序, 在 flow cell 玻璃管中添加 DNA 聚合酶、4 种不同的荧光标记 dNTP 和引物进行扩增, 当互补链合成时, 就会释放出带有荧光标记 dNTP 的荧光, 系统根据捕捉的荧光信号就可以推断出待测的序列信息。

1.3 ABI/SOLiD SOLiD 测序平台是利用连接酶在连接过程中进行测序的原理, 即边连接边测序。SOLiD 测序文库的

基金项目 国家海洋局近岸海域生态环境重点实验室开放基金项目 (201511)。

作者简介 刘敏 (1990—), 女, 吉林四平人, 硕士研究生, 研究方向: 分子生态学。* 通讯作者, 教授, 博士, 从事水产动物营养学和水产养殖学方面的研究。

收稿日期 2017-01-22

构建和微乳滴 PCR 扩增与 454 技术类似,只是 SOLiD 的微珠更小,只有 1 μm 。它的特点是采用双碱基编码技术(two-base encoding)进行误差校正,其测序步骤如下:①制备单链 DNA 文库,将 DNA 待测序列打断成很小的片段,并在两头添加不同的接头,连接载体,获得单链 DNA 文库。②微乳 PCR(emulsion PCR),将带有接头的单链 DNA 固定在磁珠表面,进行 PCR 扩增,并对扩增产物进行 3'端修饰。③转移至上样玻片,将磁珠放置在 SOLiD 上样玻片(slide),准备测序。④进行双碱基编码测序,在体系中加入通用引物、DNA 连接酶和带有 8 个碱基单链荧光探针混合物(3'-XXnnnzzz-5'),此探针 3'端的第 1 位和第 2 位的碱基是确定的,5'端是带有荧光标记的(根据种类不同在 6~8 位 zzz 上加不同的荧光)。SOLiD 测序包括 5 轮测序反应,每一轮反应又由多个连接反应构成。第 1 轮的第 1 次测序反应中,当探针 3'端的第 1 和 2 位在连接酶的作用下进行配对连接时,就会发出荧光信号。记录荧光颜色信息后,除去 6~8 位碱基和淬灭 5'末端荧光信号,也就是说,实际上每一次测序的位置都相差 5 个碱基,第 1 轮得到第 1 和第 2 位的颜色信息。以此类推,第 2 次测得第 6 和 7 位的颜色信息,第 3 次测得第 11 和 12 位的颜色信息,第 4 次测得第 16 和 17 位的颜色信息……第 1 轮测序之后,将新合成的链变性、洗脱,引物重置,进行下一轮测序。由于之后的每一轮引物都要比前一轮引物向前移动 1 位,所以第 2 轮将会测得第 0~1 位、5~6 位、10~11 位和 15~16 位等的荧光颜色信息。5 轮反应结束后,可以得到全部位置的荧光颜色信息,同时每个位点都会被检测 2 次,准确率达 99.94%。SOLiD 测序仪每次循环可以测 75~110 bp,与 Solexa 技术类似,后续的序列拼接工作需要足够运算能力的计算机(主要指 CPU 和内存)和具有一定生物信息学操作经验的处理人员^[7]。

454、SOLiD 和 Solexa 技术的测序原理、运行成本和数据输出量均不相同,但其共同之处是不需要克隆,可以直接进行高通量测序。此外,这 3 种测序技术也都要进行文库构建、片段扩增、测序、信号捕获等环节。研究者可以根据研究目的、试验经费、测序深度来选择测序平台。近年来,利用高通量测序对贝类进行测序研究刚刚兴起,加之传统的 Sanger 测序成本昂贵,使得贝类的基因组资源缺乏。对于分析贝类这样的非模式生物,需要找出邻近物种的基因组数据或较完整的转录组信息,若是没有搜寻到相关信息,在经费允许的情况下 454 测序可作为首要选择,从而为研究获得较长的序列片段。如果试验经费有限,Solexa 测序也可以在考虑范围之内,主要体现在 2 个方面:一方面是可以产生很大的数据量,从而增加测跨区域片段的概率;另一方面是功能强大的软件不断地被开发出来,可能使拼接序列的读长变得稍长些。从这两方面来看,Solexa 测序在某种程度上也削减了短序列带来的缺陷。

2 高通量测序数据分析步骤

2.1 数据处理 测序获得的原始读段(raw reads),结果以 fastq 格式储存,并被提交到 NCBI SAR 数据库中。数据获得

后要对测序读段进行统计,通过统计测序质量值分布和碱基分布情况,评估数据量是否满足后续分析要求。之后对原始数据进行质量评价,过滤掉低质量序列、接头序列和核糖体,获得高质量的数据进行接下来的组装分析。

2.2 数据组装 转录组测序序列组装包括有参基因组装和无参基因组装(从头组装)。有参基因组装:利用 Scripture 或 Cufflinks 软件,将测序读段比对到基因组上,再通过重叠信息得到 graph,将 graph 中的信息进行转换,最后获得转录本。有参基因组装具有内存需求小、灵敏度高、污染小和拼接的转录本序列较完整等优点,但也存在一些缺点,如依赖参考序列和软件错误比对等。无参基因组装:又称从头组装,常用软件为 Trinity。以 Trinity 软件对样本组装情况为例,根据序列之间的重叠信息组装获得 contig,再通过序列的 contig 和双末端信息之间的相似性比对,然后进行 contig 聚类,而后在局部组装得到 transcript^[8]。通常把最主要的 transcript 当成 Unigene。获取的 Unigene 进行后续的注释、定量、基因结构分析、基因的表达式差异和差异表达分析。与有参基因组装相比,从头组装不依赖参考序列、比对软件,能较好地重建。但是,同样存在缺点,如要有较大内存和更高的测序深度、敏感性强、高相似度的转录本可能被合并等。

2.3 比对及注释 利用比对软件将组装得到的 Unigene 序列和公共数据库 COG^[9]、Swiss-Prot^[10]、NR^[11]、KEGG^[12] 进行相似性搜查,若是 Unigene 序列被比对到高优先级的数据库,就不用进行下一轮比对,否则仍会与之后的 database 比对完。在 BLAST 结果中,取比对到相似性最高的蛋白,最后得到这个 Unigene 的功能注释信息。将 NR 比对的结果输入到 Blast2GO 软件^[13]中,就可以获得 Unigene 序列的 GO 注释信息。获得 GO 注释后,通过 WEGO 软件^[14]可以统计 GO 功能分类,从而了解物种的基因功能。目前,采用的蛋白质数据库包括 PDB、TrEMBL 和 Swiss-Prot 等。

3 高通量测序在贝类转录组研究中的应用

转录组是指细胞或特定组织在某一状态下转录出来的所有 RNA 集合。以真核生物基因组研究为例,转录组测序比全基因组测序针对性更强,仅仅研究被转录的部分,既减小了研究范围又降低了成本。此外,通过基因表达谱量的变化确定是否有某个特定基因,转录组是作为衡量功能基因组研究的一个重要指标。近些年,高通量测序技术被广泛应用在贝类转录组研究中,同时大量的贝类转录组研究文章被发表,取得了很大成果。

3.1 高通量测序在贝类分子标记开发中的应用 高通量测序技术作为一种快速开发分子标记的有效方法,越来越受到研究人员的青睐。通过高通量测序开发的分子标记主要包括简单重复序列(SSR)和单核苷酸多态性(SNP)。简单重复序列是指由 2~6 个核苷酸为重复基序的串联的 DNA 序列组成,具有高多态性、稳定性强和共显性遗传等特点,主要用于遗传分析、遗传图谱构建、亲缘关系鉴定和分子标记辅助育种。单核苷酸多态性是指由单个核苷酸变异所引起的 DNA 序列多态性,通常用于鉴别生物体的遗传变异。对 SSR

标记而言,一般采用 MISA 软件进行 SSR 标记筛查。MISA 软件是一个可以大量识别和定位 SSR 序列的筛查工具,同时,它还提供了一个和引物设计 Primer3 的接口工具,利用这个工具可以将识别的 SSR 转成 Primer3 的格式,为设计引物提供了便利条件。Hou 等^[15]利用 454 GS FLX 测序系统对虾夷扇贝 (*Patinopecten yessoensis*) 进行转录组测序分析,检测到了超过 49 000 SNPs 和 2 700 SSRs。Wang 等^[16]对栉孔扇贝 (*Chlamys farreri*) 进行高通量测序,并与之前研究的虾夷扇贝测序进行比较分析,38 个假定的快速演变基因和大量的 SNP 被辨别。Niu 等^[17]在对缢蛏 (*Sinonovacula constricta*) 的转录组研究中也检测到大量的 SSR 和 SNP。Franchini 等^[18]对中间鲍 (*Haliotis midae*) 进行了转录组测序,研究获得了数以千计的分子标记 SSR (4 707 个 contigs 鉴定出 7 381) 和 SNP (4 380 个 contigs 鉴定出 11 934),同时为大量的 SSR 设计出高质量的 PCR 引物。Chen 等^[19]利用 Illumina 测序平台对河蚬 (*Corbicula fluminea*) 进行了转录组测序,鉴定出 2 151 条 SSR 序列。此外,还有其他贝类物种利用高通量测序技术去开发分子标记,如虾夷扇贝 (*Patinopecten yessoensis*)^[20]、大珠母贝 (*Pinctada maxima*)^[21] 和褶纹冠蚌 (*Cristaria plicata*)^[22] 等。从这些例子看出,利用 454 和 Solexa 测序技术可以大量挖掘分子标记 SSR 和 SNP,它成为快速开发分子标记的一条有效途径。同时,这些分子标记非常有助于进一步分析群体遗传结构和构建遗传图谱。

3.2 高通量测序在挖掘贝类功能基因或新基因中的应用 根据试验方案:①取不同发育时期或经过胁迫(重金属、菌、温度或盐度胁迫等)处理的生物个体组织,然后提取 RNA。②转录组测序后,利用组装的序列和数据库的序列进行相似性比对,研究者可以寻找到功能基因、新基因或差异表达基因等。

Moreira 等^[23]利用 454 测序技术对实施体外刺激的菲律宾蛤仔 (*Ruditapes philippinarum*) 进行转录组测序,产生了平均读长为 250 bp 的 974 976 个高质量读段,利用这些读段组装得到 51 265 个 contigs,其中有 22 915 个 contigs 被注释。在注释的结果中,发现 35 个 contigs 包括大量的免疫相关基因,并且详尽地分析展示出几个免疫通路和进程的假定成员。同时,在补体途径中发现了双壳类从未被描述的分子序列。和 454 测序技术相比,尽管 Solexa 序列要比 454 所测读长短,但是它的应用性和测序深度要远大于 454 技术。Meng 等^[24]利用 Illumina 测序平台对镉暴露的虾夷扇贝 (*Mizuhopecten yessoensis*) 进行测序,14 240 条序列被注释。通过 GO 注释和 KEGG 通路分析,辨别出 720 个基因涉及刺激反应和 302 个基因涉及免疫反应通路。此外,利用数字基因表达图谱 DGE 系统对鳃和消化腺检测转录组变化,在这 2 个组织分别检测到了 7 556 和 3 002 个差异表达基因。此外,Meng 等^[25]为了阐明铜暴露在免疫系统上的毒理学机制,以之前研究的虾夷扇贝 (*Mizuhopecten yessoensis*) 为材料,利用 Illumina HiSeq 2000 研究其鳃组织中的差异表达基因和转录本丰度。总共鉴定了 1 312 个上调基因和 2 237 个下调基因。同

时,显著富集分析确定了 9 个 GO 术语和 38 个涉及铜暴露反应的途径。Zhang 等^[26]利用 Illumina 测序技术对美洲牡蛎 (*Crassostrea virginica*) 进行转录组分析揭示其先天免疫相关基因的丰富性。测序组装得到 66 229 个 contigs,它涵盖了 89.4% 的已发表的 EST 和 97.9% 的线粒体基因,并且 39 978 个 contigs 和 unigenes 被辨别和注释。仍有 26 251 个 contigs 没有出现在比对的数据库中,说明 Illumina 测序可以胜任转录组深度测序的工作。

Qin 等^[27]利用 454 测序技术在 8 个不同的早期发育阶段对葡萄牙牡蛎 (*Crassostrea angulata*) 进行 *de novo* 转录组测序,提供了深度覆盖的 EST 数据库,并注释了大量的序列数据,同时鉴定和量化了 6 个独特序列涉及生长和早期发育的一些功能基因,例如肾上腺素能受体、多巴胺受体、表皮生长因子受体 (EGFR)、胰岛素样生长因子 1 受体 (IGFR)、胰岛素诱导的蛋白质、卵泡抑素前体。Huan 等^[28]对文蛤 (*Meretrix meretrix*) 的 4 个不同幼虫期进行转录组测序,总共获得 704 671 个读数,并组装成 124 737 个独特序列 (35 205 个重叠群和 89 532 个单独序列)。进一步分析显示,这些序列中的 94.66% (118 075 个) 是低表达水平的转录本。通过对 UniProt 蛋白质数据库进行搜索,15 000 215 个 (12.20%) 序列被注释。通过分析每个独特序列的深度,进行初步定量分析,大量序列显示了 4 个幼虫阶段之间的转录差异,其与发育、生长、壳形成和免疫应答等相关。Song 等^[29]利用 Hiseq 2500 测序平台对脉红螺 (*Rapana venosa*) 的 6 个不同发育时期进行转录组测序和分析,对 6 个涉及神经分泌功能序列进行了量化和鉴定。对 unigenes 涉及的各种生物过程中的功能注释可以刺激对该物种早期发育机制的研究,同时了解早期发育和变态的机制将有利于该物种的水产养殖。

研究者利用高通量测序技术在对其他双壳贝类转录组研究中也发现了功能基因,如栉孔扇贝^[16]、合浦珠母贝 (*Pinctada martensii*)^[30-31]、河蚬^[19]、南极帽贝 (*Laternula elliptica*)^[32]、中国蛤蜊 (*Macra chinensis*)^[33] 和紫贻贝 (*Mytilus galloprovincialis*)^[34] 等。

3.3 高通量测序在贝类 MicroRNA 发现和功能研究中的应用 微小 RNA (miRNA) 是一类长度为 20~25 个核苷酸 (nt) 的小非编码 RNA,其在动物和植物中进行转录后的基因表达调节^[35]。随着更多 miRNA 的调节靶基因和其功能被发现,miRNA 指导的基因调节的广度和重要性开始备受关注。近年来,利用高通量测序有很多关于贝类的 miRNA 研究,其研究表明 miRNA 控制很多细胞功能,包括代谢、增殖、凋亡和免疫等,控制了不同生物代谢通路多个基因的表达,在贝类生长和发育中扮演重要角色。

对于小 RNA (miRNA) 测序而言,Illumina 和 SOLiD 测序平台要优于 454 测序平台,因前两者有更大的单次数据产出量,可以提供一定的测序深度,从而提高其测序覆盖度。Bao 等^[36]对泥蚶 (*Tegillarca granosa*) 血细胞 miRNA 进行测序,2 个小 RNA 文库被构建 (对照组和 Cd 处理组),经过与 miR-Base 数据库进行比对,总共鉴定出 215 个保守的 miRNA,同

时在没有泥酣参考基因组序列的情况下,利用生物分析发现 39 个泥酣特有的 miRNA。在 Cd 刺激的文库中有 5 个 miRNA 表达显著上调,11 个 miRNA 表达显著下调,上调 miRNA 主要涉及发育和细胞进程,下调 miRNA 在胚胎发育、炎症反应和免疫应答中发挥重要作用。Zhao 等^[37]利用 2 种牡蛎物种香港牡蛎(*Crassostrea hongkongensis*)和长牡蛎(*C. gigas*)进行低盐度胁迫,并对鳃组织构建的 4 个小 RNA 文库进行高通量测序。在长牡蛎中鉴定出 137 个保守的和 65 个新的 miRNA,同时在香港牡蛎中也鉴定 85 个保守的和 2 个新的 miRNA。此外,在香港牡蛎物种中,显示相反表达类型的 5 个 miRNA-mRNA 互作对被辨别。对这些辨别的 miRNA 进一步研究,可能会揭示其在应答胁迫中的重要作用,从而为阐明牡蛎耐盐的分子机理提供线索。Jiao 等^[38]利用 Illumina/Solexa 深度测序对合浦珠母贝 MicroRNA 进行描述和辨别,研究鉴定出 258 个 MicroRNA,其中,205 个是跨物种保守的,53 个是其特有的。一些 pm-miRNA 被预测参与生物矿化,例如 miR-2305 和 miR-0046。Zhou 等^[39]利用 Illumina 平台对来自牡蛎血细胞的 3 个小 RNA 文库进行测序,以研究细菌攻击和热胁迫后 miRNA 的潜在免疫调节作用。对数据进行分析,鉴定出 199 个牡蛎 miRNA(71 种已知的和 128 种新型的)。研究显示免疫攻击可以诱导表达 miRNA,其可以调节免疫应答,例如吞噬作用和凋亡。此外,一些免疫相关的表达 miRNA 也可以通过热应激来调节改善牡蛎对环境的适应。Chen 等^[40]也对刺激后的牡蛎血细胞 miRNA 进行了免疫调节研究。关于 miRNA 的研究还有中间鲍^[41]和长牡蛎^[42]。这些在某阶段特异表达的 miRNA 可能会在一些生物进程中发挥重要作用,因此,miRNA 的测序分析将会促进靶基因功能研究。

4 展望

近些年,高通量测序技术被广泛应用在贝类转录组的功能基因研究中。对于贝类这样的非模式生物,大多是无参考基因组序列的物种,通过高通量测序可以在短时间内产生海量数据,这些数据信息将会丰富贝类物种的遗传数据库,并且有利于这些物种进行深入的分子生物学研究。

高通量测序虽然在转录组学和分子生物学的研究中发挥了重要作用,但是在大量序列的拼接上仍然有不足之处。后续的拼接(序列的准确性和完整性)工作对于分子标记的开发和获取有意义的信息非常重要。在分子标记开发方面,错误拼接会导致标记扩增失败,同时太短序列的拼接也会增添设计引物的困难。研究者可以着手对拼接方法和测序工艺进行改进,从而增加序列的拼接长度。面对这些庞大的数据分析,如何得到有用的生物信息学信息将成为之后研究的热点。今后高通量测序技术会在各领域有更广泛的应用,为分子标记和功能基因的挖掘提供一条便利的途径,从而极大地促进遗传分析和分子育种方面的研究。

参考文献

[1] SANGER F, NICKLEN S, COULSON A R. DNA sequencing with chain-terminating inhibitors[J]. Proceedings of the national academy of sciences, 1977, 74(12): 5463-5467.

[2] SHENDURE J, JI H. Next-generation DNA sequencing[J]. Nature biotechnology, 2008, 26(10): 1135-1145.

[3] EMRICH S J, BARBAZUK W B, LI L, et al. Gene discovery and annotation using LCM-454 transcriptome sequencing[J]. Genome research, 2007, 17(1): 69-73.

[4] MORIN R D, BAINBRIDGE M, FEJES A, et al. Profiling the HeLa S3 transcriptome using randomly primed cDNA and massively parallel short-read sequencing[J]. Biotechniques, 2008, 45(1): 81-94.

[5] CLOONAN N, FORREST A R R, KOLLE G, et al. Stem cell transcriptome profiling via massive-scale mRNA sequencing[J]. Nature methods, 2008, 5(7): 613-619.

[6] VAN DIJK E L, AUGER H, JASZCZYSZYN Y, et al. Ten years of next-generation sequencing technology[J]. Trends in genetics, 2014, 30(9): 418-426.

[7] STRICKLER S R, BOMBARELY A, MUELLER L A. Designing a transcriptome next-generation sequencing project for a nonmodel plant species[J]. American journal of botany, 2012, 99(2): 257-266.

[8] 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.

[9] TATUSOV R L, GALPERIN M Y, NATALE D A, et al. The COG database: A tool for genome-scale analysis of protein functions and evolution[J]. Nucleic acids research, 2000, 28(1): 33-36.

[10] APWEILER R, BAIROCH A, WU C H, et al. UniProt: The universal protein knowledgebase[J]. Nucleic acids research, 2004, 32(S1): 115-119.

[11] 邓泱泱, 蔡建琦, 吴松峰. nr 数据库分析及其本地化[J]. 计算机工程, 2006, 32(5): 71-74.

[12] KANEHISA M, GOTO S, KAWASHIMA S, et al. The KEGG resource for deciphering the genome[J]. Nucleic acids research, 2004, 32(S1): 277-280.

[13] CONESA A, GÖTZ S, GARCÍA-GÓMEZ J M, et al. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research[J]. Bioinformatics, 2005, 21(18): 3674-3676.

[14] YE J, FANG L, ZHENG H K, et al. WEGO: A web tool for plotting GO annotations[J]. Nucleic acids research, 2006, 34(S2): 293-297.

[15] HOU R, BAO Z M, WANG S, et al. Transcriptome sequencing and *de novo* analysis for Yesso scallop (*Patinopecten yessoensis*) using 454 GS FLX[J]. PLoS One, 2011, 6(6): 21560.

[16] WANG S, HOU R, BAO Z M, et al. Transcriptome sequencing of Zhikong scallop (*Chlamys farreri*) and comparative transcriptomic analysis with Yesso scallop (*Patinopecten yessoensis*) [J]. PLoS One, 2013, 8(5): 63927.

[17] NIU D H, WANG L, SUN F Y, et al. Development of molecular resources for an intertidal clam, *Sinonovacula constricta*, using 454 transcriptome sequencing[J]. PLoS One, 2013, 8(7): 67456.

[18] FRANCHINI P, VAN DER MERWE M, ROODT-WILDING R. Transcriptome characterization of the South African abalone *Haliotis midae* using sequencing-by-synthesis[J]. BMC Research Notes, 2011, 4: 59.

[19] CHEN H H, ZHA J L, LIANG X F, et al. Sequencing and *de novo* assembly of the Asian clam (*Corbicula fluminea*) transcriptome using the Illumina GAIIx method[J]. PLoS One, 2013, 8(11): 79516.

[20] DING J, ZHAO L, CHANG Y Q, et al. Transcriptome sequencing and characterization of Japanese scallop *Patinopecten yessoensis* from different shell color lines[J]. PLoS One, 2015, 10(2): 116406.

[21] DENG Y W, LEI Q N, TIAN Q L, et al. *De novo* assembly, gene annotation, and simple sequence repeat marker development using Illumina paired-end transcriptome sequences in the pearl oyster *Pinctada maxima* [J]. Bioscience, biotechnology, and biochemistry, 2014, 78(10): 1685-1692.

[22] PATNAIK B B, WANG T H, KANG S W, et al. Sequencing, *De Novo* assembly, and annotation of the transcriptome of the endangered freshwater pearl bivalve, *Cristaria plicata*, provides novel insights into functional genes and marker discovery[J]. PLoS One, 2016, 11(2): 148622.

[23] MOREIRA R, BALSEIRO P, PLANAS J V, et al. Transcriptomics of *in vitro* immune-stimulated hemocytes from the Manila clam *Ruditapes philippinarum* using high-throughput sequencing[J]. PLoS One, 2012, 7(4): 35009.

[24] MENG X L, LIU M, JIANG K Y, et al. *De novo* characterization of Japanese scallop *Mizuhopecten yessoensis* transcriptome and analysis of its gene expression following cadmium exposure[J]. PLoS One, 2013, 8(5): 64485.

[25] MENG X L, TIAN X, LIU M, et al. The transcriptomic response to copper exposure by the gill tissue of Japanese scallops (*Mizuhopecten yessoensis*) using deep-sequencing technology [J]. Fish & shellfish immunology, 2014, 38(2): 287-293.

3 结论

(1) 仅用发文量指标来确定核心作者不够全面、客观,需要结合被引量,综合运用这 2 项测评指标可以有效测评《现代农业科技》作者群的学术影响力和贡献力。

(2) 2005—2015 年《现代农业科技》的核心作者共 207 位,分布在 25 个省、自治区、直辖市,毗邻或邻近该刊出版地安徽的地区作者较多,包括江苏、河南、浙江、湖北、湖南、山东等地,这些地区的发文量也较多,其中河南省的总发文量和人均发文量均位于前列。这说明这些地区是《现代农业科技》核心作者的主要来源地,有着重要的地区优势。距离较远的地区核心作者人数相对较少,影响较弱,在这些影响薄弱的地区应加强宣传,做好推广工作,使该刊的辐射力和社会影响力进一步扩大。

(3) 在 207 位核心作者中,102 人来自科研院所,75 人来自技术单位,14 人来自高等院校,来自企业和其他机构的人员均为 8 人。其中,科研院所和技术单位是《现代农业科技》最重要的核心作者来源机构。

参考文献

- [1] 王洪江,徐桂珍,鲍勇,等.《现代农业科技》杂志成功转型的实践探索[J]. 农业图书情报学刊,2011,23(6):176-179.
- [2] 王洪江,徐桂珍,鲍勇,等.优质服务:《现代农业科技》杂志发展的核心竞争力[J]. 农业图书情报学刊,2011,23(7):178-180.
- [3] 孙小莉,任福君.《科技导报》科普载文作者群分析[J]. 高等建筑教育,2014,23(6):150-154.
- [4] 朱学武.《图书馆工作与研究》网络计量分析研究[J]. 图书馆工作与研究,2008(12):79-84.
- [5] 贺晓利,张薇.2007-2012 年《情报科学》期刊学术影响力指标分析

- [J]. 情报探索,2014(8):26-32.
- [6] 张瑞麟,吴益伟,于晓庆,等.《浙江农业科学》2009-2011 年载文作者群分析[J]. 浙江农业科学,2012(6):911-913.
- [7] 邱均平,温芳芳. Google Scholar 和 CCD 的引文统计分析功能比较:从期刊被引频次角度分析[J]. 重庆大学学报(社会科学版),2011,17(6):84-89.
- [8] 刘英丽.《农业图书情报学刊》2007-2009 年载文作者群分析[J]. 中国农学通报,2010,26(22):422-424.
- [9] 钟文娟. 基于普赖斯定律与综合指数法的核心作者测评:以《图书馆建设》为例[J]. 科技管理研究,2012,32(2):57-60.
- [10] 廉清.《图书情报工作》核心作者群分析研究[J]. 现代情报,2004,24(11):55-59.
- [11] 李文以.《档案管理》1995-2005 年核心作者群分析[J]. 档案管理,2006(4):48-50.
- [12] 李宗仁. 用综合指数法测评《编辑学报》的核心著者[J]. 编辑学报,2008,20(1):91-92.
- [13] 刘永胜.《情报资料工作》核心著者测评[J]. 情报资料工作,2003(2):49-50.
- [14] 朱亚丽.《现代图书情报技术》核心著者测评[J]. 现代图书情报技术,2004,20(12):83-84.
- [15] 刘江霞. 用综合指数法分析《科技与出版》2002-2012 年核心作者[J]. 科技与出版,2014(6):140-143.
- [16] 徐红星.《中国科技期刊研究》2008-2012 年核心作者群的分析研究[J]. 中国科技期刊研究,2013,24(6):1074-1078.
- [17] 魏强. 2005-2009 年《河北北方学院学报》(自然科学版)载文作者群分析[J]. 河北北方学院学报(自然科学版),2010,26(3):77-82.
- [18] 陈慧萍,聂迎利,冯艳秋. 农业技术类期刊评价指标体系建设初探[J]. 中国科技期刊研究,2013,24(3):512-514.
- [19] 李渝,杜兆恒,刘喻,等.《四川文理学院学报》作者群分析[J]. 科技展望,2016,26(15):322-323.
- [20] 穆丽红. 2000-2003 年《中国环境科学》作者群分析[J]. 科技情报开发与经济,2004,14(12):76-77.
- [21] 王萍,孙明明.《大豆科学》近十年载文与核心作者群分析[J]. 大豆科学,2014,33(3):458-462.

(上接第 133 页)

- [26] ZHANG L L, LI L, ZHU Y B, et al. Transcriptome analysis reveals a rich gene set related to innate immunity in the Eastern oyster (*Crassostrea virginica*) [J]. Marine biotechnology, 2014, 16(1):17-33.
- [27] QIN J, HUANG Z X, CHEN J, et al. Sequencing and *de novo* analysis of *Crassostrea angulata* (Fujian oyster) from 8 different developing phases using 454 GSFLX [J]. PLoS One, 2012, 7(8):43653.
- [28] HUAN P, WANG H X, LIU B Z. Transcriptomic analysis of the clam *Meretrix meretrix* on different larval stages [J]. Marine biotechnology, 2011, 14(1):69-78.
- [29] SONG H, YU Z L, SUN L N, et al. De novo transcriptome sequencing and analysis of *Rapana venosa* from six different developmental stages using Hi-seq 2500 [J]. Comparative biochemistry and physiology, Part D: Genomics and proteomics, 2016, 17:48-57.
- [30] ZHAO X X, WANG Q S, JIAO Y, et al. Identification of genes potentially related to biomineralization and immunity by transcriptome analysis of pearl sac in pearl oyster *Pinctada martensii* [J]. Marine biotechnology, 2012, 14(6):730-739.
- [31] SHI Y H, YU C C, GU Z F, et al. Characterization of the pearl oyster (*Pinctada martensii*) mantle transcriptome unravels biomineralization genes [J]. Marine biotechnology, 2013, 15(2):175-187.
- [32] TRUEBANO M, BURNS G, THORNE M A S, et al. Transcriptional response to heat stress in the Antarctic bivalve *Laternula elliptica* [J]. Journal of experimental marine biology and ecology, 2010, 391(1/2):65-72.
- [33] ZHANG J J, LI H J, QIN Y J, et al. Identification of functional genes involved in Cd²⁺ response of Chinese surf clam (*Macrura chinensis*) through transcriptome sequencing [J]. Environmental toxicology and pharmacology, 2016, 41:113-120.
- [34] GERDOL M, DE MORO G, MANFRIN C, et al. RNA sequencing and *de*

- novo* assembly of the digestive gland transcriptome in *Mytilus galloprovincialis* fed with toxigenic and non-toxic strains of *Alexandrium minutum* [J]. BMC Research Notes, 2014, 7(1):722.
- [35] OBERNOSTERER G, LEUSCHNER P J F, ALENIUS M, et al. Post-transcriptional regulation of microRNA expression [J]. RNA, 2006, 12(7):1161-1167.
- [36] BAO Y B, ZHANG L L, DONG Y H, et al. Identification and comparative analysis of the *Tegillarca granosa* haemocytes microRNA transcriptome in response to Cd using a deep sequencing approach [J]. PLoS One, 2014, 9(4):93619.
- [37] ZHAO X L, YU H, KONG L F, et al. High throughput sequencing of small RNAs transcriptomes in two *Crassostrea* oysters identifies microRNAs involved in osmotic stress response [J]. Scientific reports, 2016, 6:22687.
- [38] JIAO Y, ZHENG Z, DU X D, et al. Identification and characterization of microRNAs in pearl oyster *Pinctada martensii* by Solexa deep sequencing [J]. Marine biotechnology, 2014, 16(1):54-62.
- [39] ZHOU Z, WANG L L, SONG L S, et al. The identification and characteristics of immune-related microRNAs in haemocytes of oyster *Crassostrea gigas* [J]. PLoS One, 2014, 9(2):88397.
- [40] CHEN H, WANG L L, ZHOU Z, et al. The comprehensive immunomodulation of NeurimmiRs in haemocytes of oyster *Crassostrea gigas* after acetylcholine and norepinephrine stimulation [J]. BMC Genomics, 2015, 16(1):942.
- [41] PICONE B, RHODE C, ROODT-WILDING R. Identification and characterization of miRNAs transcriptome in the South African abalone, *Haliotis midae* [J]. Marine genomics, 2017, 31:9-12.
- [42] XU F, WANG X T, FENG Y, et al. Identification of conserved and novel microRNAs in the pacific oyster *Crassostrea gigas* by deep sequencing [J]. PLoS One, 2014, 9(8):104371.