

转录组学技术在水产动物研究中的运用

罗 辉^{1,2}, 叶 华², 肖世俊³, 郑曙明², 王晓清^{1*}, 王志勇^{1,3*}

(1. 湖南农业大学动物科学技术学院,湖南省水产高效健康生产协同创新中心,湖南 长沙 410128;

2. 西南大学鱼类繁育与健康养殖研究中心,重庆 402460;

3. 集美大学水产学院,农业部东海海水健康养殖重点实验室,福建 厦门 361021)

摘要:近年来,转录组学技术及其在水产动物中的研究备受研究者的广泛关注。转录组学技术主要有基于杂交技术和测序技术为基础的两大类技术;两类技术在水产动物的转录组学研究中均得到了广泛运用。尤其是二代测序技术的出现使得水产动物的转录组学研究达到了前所未有的热度。本实验综述了近年来水产动物在免疫应答、生长发育、生物进化和毒理学方面的转录组学研究进展。在综述已有成果的基础上,分析了转录组学技术自身存在的局限性和现阶段转录组学在水产动物研究中面临的问题与挑战,并对未来水产动物转录组学研究趋势进行了展望。

关键词:水产动物;转录组学;基因表达;生物信息学

中图分类号: Q 75; S 917.4

文献标志码: A

近年来,转录组学、基因组学和蛋白质组学等各种组学技术在揭示水产动物抗病免疫、生长发育、系统进化和生物毒理过程及相应机理方面的研究中越来越重要。通过组学研究,可以深刻理解水产动物各种生命活动规律的内在联系和分子机制,并根据相应结果进一步运用到抗病育种、药物筛选、种质资源保护和环境监测等多个研究领域。在各种组学技术中,转录组学技术是伴随着后基因组学的到来而率先发展起来并得到广泛应用的科学^[1]。转录组学是研究特定细胞、组织或器官在特定生长发育阶段或某种生理状况下所有转录本的科学^[2]。这所有的转录本就称之为转录组,包括编码蛋白质的 mRNA 和非编码 RNA (rRNA, tRNA 和其他 ncRNA)^[3]。与基因组相对稳定不同的是,转录组是随着生长发育阶段、生理状态和外界环境的改变而变化的。因此,转录组分析成为研究生物生长发育、应激生理、抗病免疫等作用机制的有力工具。通过对转录组图谱的分析,可以得到丰富的信息,包括基因表达信息、可

变剪切、反义转录本、基因结构、新基因、差异表达基因和单核苷酸多态性等。即使所研究的物种没有现成的基因组信息可供参考,同样可以进行 de novo 测序研究,将所得序列信息通过 de novo 组装拼接,得到该物种的单一基因序列集数据。进一步可以对基因进行功能注释、基因差异表达分析、GO 富集分析、Pathway 富集分析、蛋白互作网络分析和单核苷酸多态性分析等^[1,3-5]。转录组的复杂性必然要求有相应高通量的转录组学技术作为研究工具。

依据转录组学技术原理的不同,可以将其划分为两类技术(图 1)^[6]。一种是基于杂交的转录组学技术,如利用 cDNA 微阵列(cDNA microarray)和 DNA 宏阵列(DNA macroarray)进行检测的转录组学技术;一种是基于测序的转录组学技术,如 cDNA 文库或表达序列标签(expressed sequence tags, EST)文库测序技术,基因表达系列分析(serial analysis of gene expression, SAGE)技术和大规模平行测序

收稿日期:2014-11-02 修回日期:2014-12-12

资助项目:国家自然科学基金(U1205122,31402302);中央高校基本业务费专项(XDJK2015C034, XDJK2014C152)

通信作者:王晓清,E-mail:wangxiao8258@126.com;王志勇,E-mail:zywang@jmu.edu.cn

(massively parallel signature sequencing, MPSS) 技术,以及近年来发展起来的下一代高通量测序技

术(next generation sequencing, NGS),即 RNA 测序(RNA sequencing, RNA-seq)技术等。

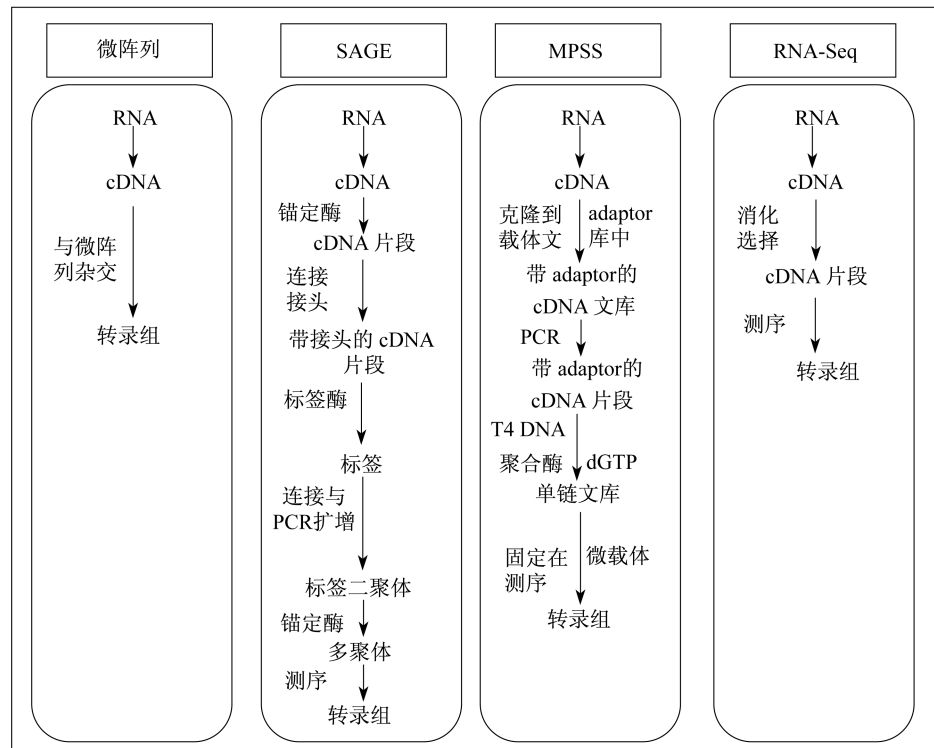


图 1 主要转录组测序技术的工作流程^[7]

Fig. 1 Outline protocols of the main transcriptome platforms^[7]

以杂交为基础的技术通常是通过原位合成或微阵列方法将众多序列已知的核酸探针固定于支持物上,然后与带荧光标记的样品进行序列杂交,检测杂交信号位置和强弱以分析样品中的基因序列及其表达^[7]。自该技术于 1995 年首次运用开始,微阵列技术就广泛运用于转录组的研究中^[6]。微阵列技术在鱼类生物学的研究中同样得到了广泛的运用,其研究历史已超过 10 年^[8]。然而该技术也存在一些固有的局限性,例如该技术常常依赖于已知的基因组或 cDNA 序列的信息;一些非特异杂交会导致高背景噪音;一些高丰度转录本会导致信号饱和;另一些低丰度的转录本则难以检测到;另外很难比较不同研究下的表达水平等^[9]。除此之外,微阵列技术运用于鱼类转录组学研究中还存在一个商业化微阵列平台很少的问题。由于微阵列的开发多数依赖于基因组或 EST 序列信息,而大多数鱼类的相关信息还很少,因此商家通常只为少数的模式水产动物生产微阵列平台^[10]。

以测序为基础的技术则是通过直接测定

cDNA 的序列而获得转录组信息。这种方法依赖于样本中产生可代表转录本的特定 EST 标签,在经过连接、扩增、克隆等步骤后得到测序片段,最后对标签序列进行测定即可得到转录组,如常见的 SAGE 和 MPSS 等^[6,11]。该类技术不仅可用于转录组测序,也可对不同发育阶段和生理条件下的组织表达进行定量分析,研究基因表达差异和挖掘新基因。此类技术也在水产动物转录组研究中有所运用。Zheng 等^[12]就采用高通量 RNA-SAGE 测序技术研究了雌雄斑马鱼 (*Danio rerio*) 的肝脏转录组,探索雌雄鱼肝脏性别差异的分子基础。然而这类技术是以昂贵的 sanger 测序法为基础,实验过程复杂,克隆步骤也费时费力^[13],分析结果还常常带有序列偏好性,难以区分不同亚型和等位基因的表达等^[2],因此在应用上受到一定的限制。

近年来,高通量测序技术的迅速发展不仅给基因组领域带来革命性的突破,同时也给转录组检测方法带来重大革新。新一代高通量测序仪可以通过测定细胞或组织全部转录产物序列,通过

序列拼接得到最后的转录组,该技术即为 RNA 测序技术(RNA-seq),这一技术几乎克服了以上技术的所有局限(表 1),因此目前已经广泛运用于转录组表达谱分析,差异表达基因鉴定和可变剪切分析等研究之中。其研究流程包括文库构建、测序,生物信息学分析和后期实验验证等步骤。即先从样本中提取 RNA,再将其反转录成 cDNA,断裂成小片段后进行 PCR 扩增,再利用合适的测序平台对扩增出来的转录本进行直接测

序。目前常用的测序平台有 Illumina/Solexa 和 AB/SOLiD 等。测序获得的大量序列信息需要经过比对、拼接、注释等生物信息学分析过程进而获得相关研究结果。RNA 测序技术由于其不依赖于基因组序列信息,对于未知基因组序列信息的非模式生物来说尤其具有应用价值。因此近年来 RNA 测序技术在水产动物的研究中得到了大量应用^[6,14]。

表 1 RNA-SEQ 技术与其他转录组学技术的比较^[2]
Tab. 1 Comparison of RNA-Seq with other methods for surveying transcriptome^[2]

| 技术 technology | 微阵列技术 tiling microarray | SAGE/MPSS/ EST 等 | RNA-Seq |
|---|----------------------------|---------------------|---------|
| 原理 principle | 杂交 | Sanger 测序 | 高通量测序 |
| 分辨率 resolution | 数个 ~ 100 bp | 单碱基 | 单碱基 |
| 信号 signal | 荧光模拟信号 | 数字化信号 | 数字化信号 |
| 通量 throughput | 高 | 低 | 高 |
| 周转时间 turnaround time | 长 | 长 | 短 |
| RNA 样本量 amount of RNA samples | 多 | 多 | 少 |
| 分析成本 cost of analysis | 高 | 高 | 低 |
| 背景噪声 background noise | 高 | 低 | 低 |
| 是否依赖现有基因组数据 reliance on existing genomic sequence | 是 | 否 | 否 |
| 发现未知转录区的能力 ability to discover unknown transcribed regions | 有限 | 能 | 能 |
| 区分不同亚型和等位基因的能力 ability to distinguish different isoforms and allelic expressions | 有限 | 有限 | 能 |
| 发现突变的能力 ability to discover mutations | 有限 | 能 | 能 |
| 确定剪切位点的能力 ability to determine splicing sites | 有限 | 能 | 能 |
| UTRs 的识别 identification of UTRs | 有限 | 有限 | 能 |

水产动物是人类的优质蛋白源,在人类食谱中占据重要位置,对水产动物的研究历来受到研究者的重视。近年来许多研究者在水产动物上开展了转录组学研究。本文将从水产动物免疫学、发育生物学、进化生物学和水产生物毒理学的角度综述其相应的转录组学研究进展,并分析转录组学在水产动物研究中现阶段面临的问题与挑战以及未来的发展趋势。

1 转录组学技术在水产动物免疫学中的运用

水产动物的免疫系统习惯上被分成先天性免疫(非特异性免疫)和获得性免疫(特异性免疫)。先天性免疫通常先于获得性免疫而激活并决定获得性免疫的性质和特点,二者相互配合,共同维持机体的内稳态。因此在机体抵抗疾病的过程中先

天性免疫发挥着更加重要的作用^[15]。目前转录组学在水产动物免疫方面的研究主要集中在细菌、病毒和寄生虫感染前后一些免疫相关基因的差异表达和分子标记的发现等方面。

1.1 细菌性疾病转录组学研究

在人工高密度养殖环境下,细菌性疾病是对养殖鱼类危害较大的主要疾病。由于细菌性鱼病具有发病频率高、症状复杂且病程长等特点,给水产养殖业带来了巨大的经济损失。因此在转录组学研究中也是备受关注和研究最多的一类。

Mu 等^[16]首先构建了嗜水气单胞菌攻毒后 12 小时的大黄鱼(*Larimichthys crocea*)脾脏转录组,并以此为参考序列研究了嗜水气单胞菌攻毒后 48 h 的大黄鱼脾脏数字基因表达谱。分析发现,与正常脾脏相比,攻毒后脾脏中有 1 996 个基因表达存在差异,其中 727 个基因表达明显上调,

489 个基因表达明显下调。这些基因主要集中在 Toll-like 通路、JAK-STAT、MAPK 通路和 T 细胞受体信号通路。同时发现 IL-8、TNF、TLR、Vasp 等多个免疫相关基因在大黄鱼感染嗜水气单胞菌后的免疫调节中起到重要作用。Wang 等^[17]运用混合分群结合转录组测序技术研究了斑点叉尾鲷 (*Ictalurus punctatus*) 和长鳍叉尾鲷 (*Ictalurus furcatus*) 抗肠道败血症相关候选基因和疾病相关的表达差异显著的单核苷酸多态性 (SNP)。结果发现在抗病鱼群和易感鱼群有 1 255 个差异表达的基因,在抗病鱼群中有 130 个基因上调表达,94 个基因下调表达;而易感鱼群中有 771 个基因上调表达,469 个基因下调。这些差异表达的基因主要是一些急性期反应基因,如 CC 趋化因子, Toll 样受体,补体成分蛋白, MHC 和 TNF 等。另外在易感鱼群和抗病鱼群的 4 304 个非冗余基因序列上发现了 56 419 个显著的 SNP。Cui 等^[18]用含有微球菌、溶藻弧菌和毕赤酵母等 3 种病原菌的水体浸泡中华绒螯蟹 (*Eriocheir sinensis*) 幼体 1 h 后,将所有幼体混合制备样品,研究其转录组图谱。研究者将测序得到的 reads 组装成 100 252 个基因序列,非冗余基因为 65 535 个,其中 17 097 个非冗余基因在 NCBI 的非冗余蛋白数据库中得到注释。此外有 23 188 个基因序列得到了 GO 功能富集分析,15 071 条和 8 574 条基因序列分别在 COG 和 KEGG 中得到注释。进一步分析发现,有许多基因与多种免疫通路有关,包括 Toll、免疫缺陷、JAK 信号转导、转录活化因子和 MAPK 通路。其中一些基因,如 TRAF6、FGF、PTP、JIP1 等均是首次在中华绒螯蟹中被确认。Li 等^[19]则用微球菌、溶藻弧菌和毕赤酵母等 3 种病原菌的混合液注射到成熟雌性中华绒螯蟹体内,8 h 后取其肝胰脏用于转录组分析。结果获得了 70 300 条 Unigenes,其中有 52 074 条非冗余 Unigenes。非冗余基因中有 17 617 条在 Nr 中得到注释。随后的功能分类和通路分析中,18 734 条 Unigenes 归入 GO 分类中,12 243 条 Unigenes 归类到 COG 中的 25 类基因功能中,8 983 条 Unigenes 归类到 KEGG 通路中。进一步分析发现,分别有 24、14、47 和 132 条 Unigenes 被归入免疫相关的 Toll、免疫缺陷、JAK-STAT 和 MAPK 通路中。Moreira 等^[20]通过构建菲律宾蛤仔 (*Ruditapes philippinarum*) 含有血细胞免疫相关序列的微阵

列来研究菲律宾蛤仔感染溶藻弧菌后血细胞的转录基因谱。结果发现,579 条差异表达的转录本。研究者将这些差异表达的转录本进行 GO 功能注释和富集分析,发现蛤仔一些随着感染时间增加而不同的反应机制。即在感染 3 h 后,与信号传递、转录和细胞凋亡相关的基因如 IL-17D、NF- κ B 或钙调蛋白等都有典型的表达。而作为与免疫密切相关的基因如 PGRPs、FREPs 和防卫蛋白等则在感染后 8 h 出现。这一免疫触发反应在感染 24 h 后被激活,并影响大量的免疫相关过程。功能研究也表明这一时段细胞凋亡、细胞坏死和细胞迁移增多。而 72 h 后免疫水平恢复正常。超过 50% 的基因下调并以一种负反馈的方式降低先前的活性反应。

1.2 病毒性疾病转录组学研究

制约水产养殖健康发展的另一个主要病害是水产动物的病毒病。病毒个体微小,可在宿主体内复制并长期潜伏,从而导致病毒病症状复杂多变,传染性强,死亡率高且难以防控,严重危害水产动物的健康。但目前对于水产动物机体抗病毒机制方面研究不多。转录组学技术的飞速发展给这一领域带来了发展机遇,转录组技术可以从基因水平研究机体抗病毒机制。

Shi 等^[21]利用转录组学技术研究了草鱼 (*Ctenopharyngodon idella*) 感染出血病病毒后机体在 2、24、48、72、96 和 120 h 时肠道、鳃、肝脏和脾脏的转录组图谱。结果表明,与对照组相比,草鱼感染病毒后的 4 种组织和 6 个时间点上共有 9 060 条 unigenes 差异表达。进一步 GO 功能分析,发现 4 种组织中均发生了免疫反应,并且发现在草鱼感染病毒后每一个器官中的脂代谢和糖代谢均出现代谢失调。KEGG 代谢通路分析发现补体系统和细胞免疫在病毒感染过程中发挥着重要作用。肝脏和肠道组织易感出血病病毒可能就是由于病毒降低了这两种组织中补体途径的活性。进一步分析还发现,在 4 种组织中 MHC I 和 MHC II 均存在差异表达,其中 MHC I 上调表达而 MHC II 则下调表达。许宝红^[22]也对草鱼出血病进行了转录组学研究,结果发现草鱼感染出血病后与对照草鱼存在 32 621 条差异表达的 unigenes,其中 20 826 条上调表达,11 795 条下调表达。进一步 KEGG 代谢通路分析发现,差异表达的 unigenes 中有 9 916 条富集在 217 条通路中。

FDR \leq 0.05的通路有32条,差异最显著的4条与凝血和免疫紧密相关的通路是凝血-补体通路,造血细胞谱系通路,细胞吞噬体通路和细胞-细胞互作通路,其中凝血-补体通路和细胞因子受体互作通路是上调表达因子显著富集的两条通路。Mu等^[23]利用Ploy I:C刺激大黄鱼,构建了刺激后12 h的脾脏转录组,共获得15 192条unigenes。进一步KEGG通路分析中有5 389个基因得到注释并得到许多免疫相关通路。这些免疫相关基因和通路主要是抗病毒免疫因子,如模式识别受体、信号转导、干扰素和干扰素刺激基因、炎症细胞因子和受体、补体成分、B细胞和T细胞抗原活化分子等。用RT-PCR进一步验证发现,Toll样受体信号通路、RIG-I样受体信号通路、JAK-STAT信号通路和T细胞受体信号通路的部分基因在Ploy I:C刺激后发生了相应的变化,表明这些信号通路可能被病毒类似物Ploy I:C所调节。Pereiro等^[24]分别对5组大菱鲆(*Scophthalmus maximus*)肌肉注射诺达病毒(MrNV),腹腔注射出血性败血症病毒(VHSV),肌肉注射pMCV(一种用于构建DNA疫苗的表达载体,空的pMCV可作为双链DNA病毒刺激物),免疫Pmcv-G860(一种DNA疫苗)后分别在不同时间点取脑、鳃、肝脏、脾脏和头肾等组织用于转录组测序,获得915 216条高质量的reads,组装获得55 404个contigs。进一步分析发现,大菱鲆在上述处理后免疫相关基因表达水平增加,在补体系统、Toll样受体信号通路、B细胞信号通路、T细胞信号通路、细胞凋亡和细胞程序性死亡、细胞因子和其他一些免疫分子都有显著表达差异。

1.3 寄生虫性疾病转录组学研究

水产动物集约化高密度养殖过程中的疾病暴发给水产养殖业带来了巨大的经济损失,其中寄生虫病也占据相当大的比例。寄生于水产动物机体的寄生虫种类繁多,大多寄生于体表及体内各种器官上,导致水产动物生长缓慢,繁殖力降低,抵抗继发感染能力下降,严重时可造成水产动物大量死亡。以前对水产动物寄生虫病的研究主要集中在危害、流行性、区系发生及药物防治等方面,对寄生虫致病机制及宿主机体抗寄生虫机制尚未深入研究。随着转录组技术的出现,使得研究者可以从基因水平深入开展寄生虫致病机制及

宿主抵抗寄生虫相关免疫应答的研究。

Sutherland等^[25]利用微阵列技术研究了大西洋鲑(*Salmo salar*)、大麻哈鱼(*Oncorhynchus keta*)和细鳞大麻哈鱼(*O. gorbuscha*)在感染大麻哈鱼虱(*Lepeophtheirus salmonis*)后头肾和皮肤的转录组变化。研究结果表明,大麻哈鱼虱感染后3种鱼类头肾和皮肤转录组均有大量差异表达的基因存在。在感染鱼虱后6 d,大西洋鲑、大麻哈鱼和细鳞大麻哈鱼头肾中分别有112、145和128个基因上调表达,34、145和121个基因下调表达。其中差异表达的免疫相关基因有补体、模式识别受体、急性期反应和炎症相关基因,如C3、C7、C型凝集素家族4-M、多聚免疫球蛋白受体、肥大细胞免疫受体信号传导分子等。Tadiso等^[26]也利用微阵列技术,研究了大西洋鲑感染大麻哈鱼虱后皮肤和脾脏的转录组变化。结果发现,鱼虱感染1 d后,大西洋鲑即发生大规模高度复杂的转录组反应,感染15 d,大西洋鲑的皮肤和脾脏分别有2 438和922个差异表达基因。这些差异表达基因涉及细胞免疫、体液免疫和炎症反应的相关基因,如鼠李糖结合凝集素WCL1,甘露糖受体C-2,C1q样蛋白4,急性期血清淀粉样蛋白A,转铁蛋白等。Ribas等^[27]采用插管技术将粘孢子虫接种到大菱鲆肠道,然后分别于4、7、14、25和34 d取其脾脏、头肾、胸腺、肝脏和幽门盲囊,分别构建对照组和接种组两个cDNA文库用Sanger法测序,共获得3 043个序列。结合已发表的大菱鲆Sanger法测序所得序列进行注释和进一步分析得到许多免疫相关基因,并发现几种新的免疫相关基因,如急性期反应和炎症相关的成分或家族成员,应激或疾病抵抗反应的成分等。

2 转录组学技术在水产动物发育生物学中的运用

发育生物学的目的是要阐明生物体从精子和卵子发生、受精、发育、生长到衰老、死亡的过程及其机理。其中有关水产动物生殖细胞发生、受精等过程的研究是水产动物人工繁殖、遗传育种、动物胚胎与生殖工程等生产应用技术发展的理论基础。有关细胞分化机理、基因表达调控与形态模式形成及生物功能的关系研究,是解决水产动物重大疾病的基础,也是水产动物基因工程发展的基础。因此近年来许多研究者运用转录组学技术

开展了水产动物发育生物学各个领域的研究工作。

Reading 等^[28]采用 454 焦磷酸测序法研究了条纹鲈(*Morone saxatilis*)的卵巢转录组学。研究者分别采集成熟条纹鲈不同阶段的卵巢组织(卵母细胞生长阶段,包括早期生长、后期生长和卵黄生长阶段,卵母细胞成熟和闭孔阶段以及排卵阶段),并将这些组织样品混合建库测序,共获得 230 151 个序列表达标签,过滤后组装获得 11 208 个 contigs,其中有 2 984 个 contigs 大于 500 bp。Blastx 比对获得 5 482 个直系同源基因,4 120 个得到 GO 注释,其中超过 1 300 个基因与繁殖和发育过程相关的。Reading 等^[29]在此研究基础上,结合公共数据库的数据制备了包含 233 个基因的微阵列,并将其运用于研究条纹鲈母体卵巢基因表达与受精卵质量和胚胎发育能力的关系。结果表明,利用母体卵巢转录组预测受精卵质量时,并不能从单个基因的表达情况来预测受精卵质量,而是要根据转录组中一系列基因的集体协调表达来预测。相应的候选基因相关性分析发现,泛素-26s 蛋白酶体, COP9 信号小体和随后对细胞周期控制的功能障碍导致了胚胎发育失能。Sun 等^[30]采用 Illumina 测序技术研究了斑点叉尾鲷精巢的转录组,测序得到 269.6 M 高质量的 reads, 组装得到 193 462 条 N50 为 806 bp 的 contigs, 进一步比对得到 25 307 条 unigenes, 其中包括 167 条未在斑点叉尾鲷上鉴定的 unigenes。在精巢和雌核(双单倍体雌性)表达基因的比较分析中得到 5 450 个优先在精巢中表达的基因。研究者将这部分基因视为雄性相关基因池,进一步的分析发现,这些雄性相关基因中的许多基因涉及性腺发生、精子发生、配子发育、性腺分化和性别分化。

3 转录组学技术在水产动物进化生物学中的运用

动物的遗传进化和对环境的适应性一直以来都是进化生物学中的研究热点,而要清楚地阐述遗传进化和生物对环境适应性的分子基础却是进化生物学的一个巨大挑战。基因型的多样性通常被认为是生物表型差异及对环境适应性差异的潜在机制之一。转录组学技术的发展为这一领域的研究带来了巨大的机遇,运用该技术可以在没有

参考基因组数据的背景下研究生物体在不同状态下基因的差异表达情况,进而可以深入地研究生物的遗传进化和环境适应性的分子基础。

Schunter 等^[31]采用 RNA-seq 法研究了德氏三鳍鲷(*Tripterygion delaisi*)交配策略中雄性二态性的分子基础。在繁殖季节,一部分雄性德氏三鳍鲷可以改变体色并筑巢来吸引雌性(筑巢型),另一部分雄性则并不改变体色也不筑巢,依靠偷袭来繁殖(偷袭型)。研究者通过对筑巢型、偷袭型雄鱼和雌性大脑的转录组进行研究,共获得 38 056 381 条 reads,从头组装获得 328 565 条 contigs,进一步分析发现两种雄性间的显著差异表达基因达到 600 条,其中筑巢型雄鱼比偷袭型雄鱼高表达的基因有 360 条,而偷袭型雄鱼比筑巢型雄鱼高表达的基因有 240 条。另外也发现两种雄性间显著差异表达的基因数量远多于雄性与雌性间的差异表达基因数量,这表明在德氏三鳍鲷繁殖季节中,表型可塑性比雌雄相异性在表型间产生了更大的基因表达模式差异。Wang 等^[32]利用 RNA-seq 技术研究了不同体色表型性状瓯江彩鲤(*Cyprinus carpio* var. *color*)的皮肤转录组,分别从体色为红色带大黑点的鲤(*Cyprinus carpio*)(RB)皮肤转录组中获得 408 884 条 reads,从白色鲤鱼(WW)皮肤转录组中获得 328 641 条 reads。结合 NCBI 中已登录的普通鲤转录组数据(2 119 118 条 reads)组装获得 546 676 条 Singletons 和 43 923 条 Isotigs。进一步注释获得超过 10 000 条基因。在与斑马鱼的 147 条着色基因比对后发现有 593 条 Isotigs 比对上了斑马鱼的 99 条着色基因,有 501 条 Singletons 比对上了 95 条着色基因。这 501 条 Singletons 中有 291 条来自 WW,210 条来自 RW 型鲤。同时发现有 77 条着色基因同时出现在 Isotigs 和 Singletons 之中,有 22 条仅存在于 Isotigs 中,18 条仅存在于 Singletons 中。将 RB 和 WW 中比对上着色基因的 reads 进行比较发现有 187 条 isotigs 至少有 2 倍以上的表达差异,其中有 106 条在 WW 鲤中高表达,81 条在 RB 鲤中高表达。Jeukens 等^[33]利用 RNA-seq 技术研究了短小型和正常鲱形白鲑(*Coregonus clupeaformis* spp.)的适应性转录组变化,结果发现,正常的鲱形白鲑过表达基因主要与蛋白质合成有关,而短小型过表达基因主要与免疫, DNA 复制和修复以及能量代谢有关。

4 转录组学技术在水产动物毒理学中的运用

人类活动导致越来越多的污染物进入水环境,这些污染物在水环境中尽管可以发生一些转化和降解,但这一过程仅能除去一部分污染物,另一部分污染物则存留下来并可能形成新的污染物,从而不可避免地给水产动物带来危害。水中微量或痕量有毒污染物通常不会导致水产动物死亡,但却能改变其生理生化反应,特别对水产动物的生长、发育、繁殖以及健康产生不利影响。水产动物毒理学是研究水产动物和环境中有毒因素之间相互作用的科学,传统方法是从剂量—反应关系中得出污染物对水产动物体的毒性,但只能从表观上研究污染物对生物体的毒害,难以揭示污染物对生物体的内在毒性作用机制,因而存在一定的缺陷。转录组学技术的高速发展为这一领域的研究提供了前所未有的机遇。它可以在分子水平上研究水产动物响应污染物毒性相关基因的差异表达,进一步从基因水平阐述污染物对水产动物的内在毒性机制,进而达到预测、预防和消除污染物对环境的破坏和对水产动物的毒害的目的。

Hussainzada 等^[34]用微阵列技术研究了雄性斑马鱼急性暴露于3种浓度的氯化镍、氯化钴和重铬酸钠3种金属污染物中的全鱼转录组图谱。比较分析发现,斑马鱼暴露于3种污染物中既有相同的差异表达基因,也有各自特有的差异表达基因。与对照组相比,暴露于氯化镍、氯化钴和重铬酸钠的斑马鱼差异表达基因分别有287、461和696个,其中对3种金属污染物响应共有的差异表达基因数量为106条,氯化镍和氯化钴共有的差异表达基因有49条,氯化镍与重铬酸钠共有的有68条,氯化钴与重铬酸钠共有的有99条。进一步运用基因本体论和转录因子富集算法分析表明,金属污染物毒性扰乱了多种生物学过程,相应的转录调节介导了基因表达的变化。金属污染物毒性差异激活的生物学过程主要与核糖体生物合成、蛋白酶体降解和p53信号级联放大相关,以及抑制与氨基酸和脂类代谢相关的氧气产生途径。Bougas 等^[35]采用微阵列技术研究了黄金鲈(*Perca flavescens*)暴露于不同浓度的镍和镉之中的转录组谱。结果发现在不同浓度的镍和镉中分别有287和176条基因显著差异表达。其中二者

共有的差异表达基因有106条。进一步分析发现,两种金属主要影响有关铁代谢、转录和转化过程、维生素代谢、血液凝结和钙转运相关基因的转录水平。Guo 等^[36]采用RNA-seq技术研究了凡纳滨对虾(*Litopenaeus vannamei*)暴露于亚硝酸盐中的转录组图谱。RNA-seq测序共获得52 889 892条高质量reads,组装后获得92 821条contigs,进一步组装成42 336条unigenes。通过与多个数据库比对后获得注释的unigenes有23 532条。进一步分析确定了许多与免疫反应,解毒作用,细胞凋亡途径有关的候选基因。研究者选择了10条与免疫反应和细胞凋亡有关的候选基因采用实时定量PCR进行验证,结果发现暴露于亚硝酸盐后这10条差异表达的基因均上调表达,表明这10条基因与亚硝酸盐应激有关。

5 小结与展望

转录组研究作为一种从整体上研究基因表达、基因结构和功能,揭示特定生物学过程以及疾病发生过程分子机理的高效技术已经在水产动物中有了一定的研究。尤其是二代测序技术为代表的高通量测序技术在转录组学研究中运用后,该领域的研究得到突飞猛进的发展。近年来研究者利用高通量测序技术获得了大量的转录本信息以及基因结构功能和差异表达信息,这些数据对我们深入理解水产动物疾病发生、机体生长发育以及进化演变等具有重要作用。但同时转录组测序技术本身还存在一定的局限性,如随机测序可能导致一些低丰度表达的基因序列被忽略;测序深度和准确率还有待进一步提高,测序读长较短,基因覆盖度不高等。随着测序平台进步和测序成本的迅速降低,测序技术的短板有望在近些年获得突破,届时更高的测序质量和更充分的测序量将是转录组测序新的趋势。同时水产动物普遍缺少参考基因组信息,无法对测序数据进行全面的基因注释与基因组准确定位;另一方面,测序获得了海量信息,这对于生物信息学也是一个巨大的挑战,如何从这些海量的信息中快速简便地挖掘出数据中蕴含的有用信息将是今后相当长一段时间努力的方向。转录组学研究的另一个不可避免的局限在于基因的转录丰度与基因表达产物的活性并不具有必然的联系,蛋白质组学与转录组学研究结果的比较表明二者的相关性是很低的^[37]。

因此,在利用转录组学的的数据解释生物生命活动规律,抗病免疫,生物进化以及应激适应等机制时还需要慎重。针对这个问题,转录组测序技术与蛋白质组学、代谢组学、甲基化组学等方法结合的策略受到越来越多的重视。在可以预见的将来,多组学综合分析将是未来转录组学研究发展的一大特色。

水产动物作为重要的经济养殖动物,在转录组学研究领域占有极为重要的地位,但比较而言,水产动物转录组学研究还相对落后,尤其是在转录组数据的生物信息学分析时存在较大困难。由于水产动物种类繁多,种间差异巨大。在大多数种类的基因组信息还相当缺乏的情况下只能将一些无参考基因组物种的数据与少数模式种比对,进而进行基因注释。这就必然导致一些转录本难以注释,因此需要加快对重要经济养殖品种的基因组测序。就目前水产动物转录组学的研究而言,还主要集中在转录组 *de novo* 测序,新转录本发现,差异基因表达与定量,SNP 和 SSR(简单序列重复, simple sequence repeat) 的寻找等方面。而对于重要表型性状相关联的候选基因,进一步深入挖掘鉴定基因的功能等方面的研究还很少。目前,水产经济动物转录组学研究基本上是用整个组织、器官或某个器官与组织中的混合细胞,而没有区分其中的不同类型细胞;很少应用单细胞转录组测序等技术开展研究。但在有些器官或组织中不同类型细胞行使的功能及其基因表达情况是不一样的。随着单细胞转录组测序技术进步和成本下降,今后可以考虑应用单细胞转录组测序技术,以详细、深入地研究不同细胞的基因表达情况。另外,今后还需要拓展转录组学研究范围,例如将转录组学技术应用到养殖生态学的研究,观察养殖生态环境改变对于基因表达情况的影响,了解生态环境因子对于水产动物生长、发育、免疫等等生理过程影响的机理,为养殖生态环境调控和健康养殖技术的建立提供理论基础。

综上所述,尽管转录组学研究技术本身目前还面临着许多挑战,但随着技术的进步,尤其是 RNA 测序技术是以定量的方式对转录组进行超高通量分析的直接测序方法,它有望给生物生命活动规律的分子生物学研究带来新的机会。无论是测序技术的进步还是分析手段的更新,都为水产动物的研究提供了有力的工具。

而转录组学与蛋白质组学、基因组学以及代谢组学等多种组学技术的结合研究必然会帮助我们不断拓展和深入理解水产动物生长发育、生物进化、抗病免疫等的分子机制,从而为水产动物分子育种、养殖生产、病害防治等的研究提供理论基础。

参考文献:

- [1] Lockhart D J, Winzeler E A. Genomics, gene expression and DNA arrays [J]. Nature, 2000, 405 (6788): 827 - 836.
- [2] Wang Z, Gerstein M, Snyder M. RNA-Seq: A revolutionary tool for transcriptomics [J]. Nature Reviews Genetics, 2009, 10(1): 57 - 63.
- [3] Lindberg J, Lundeberg J. The plasticity of the mammalian transcriptome [J]. Genomics, 2010, 95 (1): 1 - 6.
- [4] Costa V, Angelini C, Feis I D, *et al.* Uncovering the complexity of transcriptomes with RNA-Seq [J]. Journal of Biomedicine and Biotechnology, 2010; 853916.
- [5] Ruan Y, Le B P, Ng H, *et al.* Interrogating the transcriptome [J]. Trends in Biotechnology, 2004, 22 (1): 23 - 30.
- [6] Qian X, Ba Y, Zhang Q F, *et al.* RNA-Seq technology and its application in fish transcriptomics [J]. OMICS A Journal of Integrative Biology, 2014, 18(2): 98 - 110.
- [7] Shi S B, Chen T, Zhao X M. Transcriptome platforms and applications to metabolic engineering [J]. Chinese Journal of Biotechnology, 2010, 26 (9): 1187 - 1198. [史硕博, 陈涛, 赵学明. 转录组平台技术及其在代谢工程中的应用. 生物工程学报, 2010, 26(9): 1187 - 1198.]
- [8] Hook S. Promise and progress in environmental genomics: A status report on the applications of gene expression based microarray studies in ecologically relevant fish species [J]. Journal Fish Biology, 2010, 77(9): 1999 - 2022.
- [9] Fu X, Fu N, Guo S, *et al.* Estimating accuracy of RNA-Seq and microarrays with proteomics [J]. BMC Genomics, 2009, 10: 161 - 169.
- [10] Douglas S E. Microarray studies of gene expression in fish [J]. Omics, 2006, 10: 474 - 489.
- [11] Chang F, Huang Y. Transcriptome platforms and their application in molecular biology of plant stress resistance [J]. Biotechnology Bulletin, 2011, 6: 40 -

46. [付畅,黄宇.转录组学平台技术及其在植物抗逆分子生物学中的应用.生物技术通报,2011,6:40-46.]
- [12] Zheng W, Xu H, Lam S H, *et al.* Transcriptomic analyses of sexual dimorphism of the zebrafish liver and the effect of sex hormones [J]. PLOS ONE, 2013,8(1):e53562.
- [13] Morozova O, Hirst M, Marra M A. Applications of new sequencing technologies for transcriptome analysis [J]. Annual Review of Genomics and Human Genetics, 2009,10:135-151.
- [14] Eissa N, Wang H P. Transcriptional stress responses to environmental and husbandry stressors in aquaculture species [J]. Reviews in Aquaculture, 2014,6:1-28.
- [15] Magnado'ttir B. Innate immunity of fish (overview) [J]. Fish & Shellfish Immunology, 2006, 20 (2): 137-151.
- [16] Mu Y, Ding F, Cui P, *et al.* Transcriptome and expression profiling analysis revealed changes of multiple signaling pathways involved in immunity in the large yellow croaker during *Aeromonas hydrophila* infection [J]. BMC Genomics, 2010, 11: 506-519.
- [17] Wang R J, Sun L Y, Bao L S, *et al.* Bulk segregant RNA-seq reveals expression and positional candidate genes and allele-specific expression for disease resistance against enteric septicemia of catfish [J]. BMC Genomics, 2013, 14:929-946.
- [18] Cui Z, Li X, Liu Y, *et al.* Transcriptome profiling analysis on whole bodies of microbial challenged *Eriocheir sinensis* larvae for immune gene identification and SNP development [J]. PLOS ONE, 2013, 8(12):e82156.
- [19] Li X, Cui Z, Liu Y, *et al.* Transcriptome analysis and discovery of genes involved in immune pathways from hepatopancreas of microbial challenged mitten crab *Eriocheir sinensis* [J]. PLOS ONE, 2013, 8(7): e68233.
- [20] Moreira R, Milan M, Balseiro P, *et al.* Gene expression profile analysis of Manila clam (*Ruditapes philippinarum*) hemocytes after a *Vibriolginolyticus* challenge using an immune-enriched oligo-microarray [J]. BMC Genomics, 2014, 15:267-282.
- [21] Shi M J, Huang R, Du F K, *et al.* RNA-seq profiles from grass carp tissues after reovirus (GCRV) infection based on singular and modular enrichment analyses [J]. Molecular Immunology, 2014, 61 (1): 44-53.
- [22] Xu B H. Transcriptome analysis of grass carp infected GCHV [D]. Changsha: Hunan Agricultural University, 2012. [许宝红.感病草鱼的脾脏比较转录组分析.长沙:湖南农业大学, 2012.]
- [23] Mu Y, Li M, Ding F, *et al.* De Novo Characterization of the spleen transcriptome of the large yellow croaker (*Pseudosciaena crocea*) and analysis of the immune relevant genes and pathways involved in the antiviral response [J]. PLOS ONE, 2014, 9 (5):e97471.
- [24] Pereira P, Balseiro P, Romero A, *et al.* High-throughput sequence analysis of turbot (*Scophthalmus maximus*) transcriptome using 454-pyrosequencing for the discovery of antiviral immune genes [J]. PLOS ONE, 2012, 7(5):e35369.
- [25] Sutherland B J, Koczka K W, Yasuike M, *et al.* Comparative transcriptomics of Atlantic *Salmo salar*, chum *Oncorhynchus keta* and pink salmon *O. gorbuscha* during infections with salmon lice *Lepeophtheirus salmonis* [J]. BMC Genomics, 2014, 15:200-216.
- [26] Tadiso T M, Krasnov A, Skugor S, *et al.* Gene expression analyses of immune responses in Atlantic salmon during early stages of infection by salmon louse (*Lepeophtheirus salmonis*) revealed bi-phasic responses coinciding with the copepod-chalimus transition [J]. BMC Genomics, 2011, 12:141-157.
- [27] Ribas L, Pardo B G, Fernandez c, *et al.* A combined strategy involving Sanger and 454 pyrosequencing increases genomic resources to aid in the management of reproduction, disease control and genetic selection in the turbot (*Scophthalmus maximus*) [J]. BMC Genomics, 2013, 14:180-200.
- [28] Reading B J, Chapman R W, Schaff J E, *et al.* An ovary transcriptome for all maturational stages of the striped bass (*Morone saxatilis*), a highly advanced perciform fish [J]. BMC Research Notes, 2012, 5: 111-122.
- [29] Chapman R W, Reading B J, Sullivan C V. Ovary transcriptome profiling via artificial intelligence reveals a transcriptomic fingerprint predicting egg quality in striped bass, *Morone saxatilis* [J]. PLOS ONE, 2014, 9(5):e96818.
- [30] Sun F, Liu S, Gao X, *et al.* Male-Biased genes in catfish as revealed by RNA-Seq analysis of the testis transcriptome [J]. PLOS ONE, 2013, 8(7):e68452.
- [31] Schunter C, Vollmer S V, Macpherson E, *et al.*

- Transcriptome analyses and differential gene expression in a non-model fish species with alternative mating tactics[J]. *BMC Genomics*,2014, 15:167 – 180.
- [32] Wang C H, Wachholtz M, Wang J, *et al.* Analysis of the skin transcriptome in two oujiang color varieties of common carp [J]. *PLOS ONE*, 2014, 9 (3) :e90074.
- [33] Jeukens J, Renaut S, St-Cyr J, *et al.* The transcriptomics of sympatric dwarf and normal lake whitefish (*Coregonus clupeaformis* spp. , *Salmonidae*) divergence as revealed by next-generation sequencing[J]. *Molecular Ecology*,2010, 19(24) :5389 – 5403.
- [34] Hussainzada N, Lewis J A, Baer C E, *et al.* Whole adult organism transcriptional profiling of acute metal exposures in male Zebrafish[J]. *BMC Pharmacology and Toxicology*,2014, 15:15 – 29.
- [35] Bougas B, Normandeau E, Pierron F, *et al.* How does exposure to nickel and cadmium affect the transcriptome of yellow perch (*Perca flavescens*)- Results from a 1 000 candidate-gene microarray[J]. *Aquatic Toxicology*,2013, 142 – 143:355 – 364.
- [36] Guo H, Ye C X, Wang A L, *et al.* Transcriptome analysis of the Pacific white shrimp *Litopenaeus vannamei* exposed to nitrite by RNA-seq[J]. *Fish & Shellfish Immunology*,2013, 35(6) :2008 – 2016.
- [37] Deyholos M K. Making the most of drought and salinity transcriptomics [J]. *Plant, Cell and Environment*,2010, 33(4) :648 – 654.

Application of transcriptomics technology to aquatic animal research

LUO Hui^{1,2}, YE Hua², XIAO Shijun³, ZHENG Shuming²,

WANG Xiaoqing^{1*}, WANG Zhiyong^{1,3*}

(1. College of Animal Science & Technology, Hunan Agricultural University, Hunan Collaborative Innovation Center for Aquatic Efficient Health Production, Changsha 410128, China;

2. Fisheries Breeding and Healthy Cultivation Research Centre, Southwest University, Chongqing 402460, China;

3. Key Laboratory of Healthy Mariculture for East China Sea of Agriculture Ministry, PRC, Jimei University, Xiamen 361021, China)

Abstract: There is a growing interest for the progress and application of transcriptomics technology in recent years. Two types of techniques, including hybridization-or sequence-based method, have been developed for transcriptome investigations. With the emergence of next-generation sequencing techniques, those technologies have been widely used in the genetic studies of economic traits and molecular-aided selections for aquatic animals. This review has briefly summarized the progress of transcriptomic researches in aquaculture, involving immune response, growth, development, evolution and toxicology. Meanwhile, we have discussed challenges in both transcriptomic technologies and elaborate applications in aquaculture and analyzed prospects of further development of transcriptome researches on aquatic animals.

Key words: aquatic animal; transcriptomics; gene expression; bioinformatics

Corresponding author: WANG Xiaoqing. E-mail: wangxiao8258@126.com;

WANG Zhiyong. E-mail: zywang@jmu.edu.cn