

# RNA-Seq 技术在瘤胃微生物研究中的应用进展

廖 奇<sup>1</sup> 刘旭川<sup>2</sup> 李 清<sup>1,2</sup> 樊月圆<sup>2</sup> 张春勇<sup>1,2</sup>

杨舒黎<sup>1,2</sup> 毛华明<sup>1,2</sup> 冷 静<sup>1,2\*</sup>

(1.云南省动物营养与饲料科学重点实验室,昆明 650201;2.云南农业大学动物科学技术学院,昆明 650201)

**摘 要:**反刍动物瘤胃内多种多样的微生物组成一个复杂的微生态系统,它们在瘤胃内物质消化过程中发挥重要作用。RNA-Seq 技术的出现为反刍动物瘤胃微生物的研究提供了前所未有的机遇。本文主要从新一代高通量测序技术出发,阐述 RNA-Seq 技术在瘤胃微生物代谢酶特性、瘤胃微生物的多样性、瘤胃微生物新功能基因方面的应用,以期对瘤胃微生物的研究提供参考。

**关键词:** RNA-Seq 技术;瘤胃微生物;多样性;新功能基因

**中图分类号:** Q51

**文献标识码:** A

**文章编号:** 1006-267X(2015)04-1061-07

瘤胃特殊的内环境为瘤胃微生物提供了优越的栖息场所,但也导致了体外纯培养瘤胃微生物的困难<sup>[1]</sup>。RNA-Seq 技术,又称转录组测序(transcriptome sequencing)技术或全转录鸟枪法测序(whole transcriptome shotgun sequencing)技术,是基于新一代测序技术的转录组学研究方法,能够从群体水平上研究瘤胃微生物功能基因的表达水平以及在不同环境条件下的转录调控规律,这是传统分子生物学技术所欠缺的<sup>[2]</sup>。2008 年 6 月, Nagalakshmi 等<sup>[3]</sup>、Wilhelm 等<sup>[4]</sup>分别在《科学》和《自然》杂志发表了利用 RNA-Seq 技术论述有关酿酒酵母(*Saccharomyces cerevisiae*)与裂殖酵母(fission yeast)转录组学的研究成果,为后续环境微生物研究奠定了 RNA-Seq 技术的应用基础。目前 RNA-Seq 技术在土壤<sup>[5]</sup>、植物<sup>[6]</sup>、人体<sup>[7]</sup>、昆虫<sup>[8]</sup>、水产<sup>[9]</sup>、食品<sup>[10]</sup>、海洋<sup>[11]</sup>微生物研究领域得到了广泛的应用,但在反刍动物瘤胃微生物研究领域的应用较少,本文就 RNA-Seq 技术及其在瘤胃微生物研究中的应用进展作一介绍。

## 1 RNA-Seq 技术原理、测序流程及测序平台的特点

### 1.1 原理及测序流程

RNA-Seq 技术是建立在高通量测序的基础上发展而来的,基于 Solexa 测序平台测序原理是采用可逆性末端边合成边测序反应。具体流程是通过将样本中总 RNA 的提取,使用相关分子生物学技术将 mRNA 分离纯化,利用超声波处理 mRNA 使其片段化,然后进行反转录生成 cDNA,同时补平双链 cDNA 黏性末端并且磷酸化,在 3'端加上腺嘌呤“A”,应用 Solexa 测序平台的测序接头连接在 cDNA 片段两端,采用桥式 PCR 扩增形成一个簇,构建 cDNA 文库,此时的文库 cDNA 由待测序列、测序接头、引物互补序列及 index 标签序列(如非 index 文库则无)组成,最后上机测序,根据 4 种不同的荧光信号确认碱基种类<sup>[12]</sup>。

### 1.2 测序平台的特点

目前, RNA-Seq 技术主要是建立在第 2 代高通量测序平台基础上的,已实现商品化的测序平

收稿日期:2014-10-31

基金项目:云南省自然科学基金重点项目(2014FA033);国家自然科学基金项目(31060314);云南省自然科学基金项目(2010CD059);云南省中青年学术和技术带头人后备人才项目(2011HB025)

作者简介:廖 奇(1989—),男,湖北咸宁人,硕士研究生,研究方向为反刍动物肠道微生态。E-mail: 532318552@qq.com

\* 通信作者:冷 静,教授,硕士生导师,E-mail: lenrui-1997@sohu.com

台包括: Illumina 公司的基因组分析仪(原 Solexa 技术)、罗氏(Roche)公司的 GS-FLX 454 基因组测序仪和 Life Technology 公司的 SOLID 序列分析仪<sup>[13-16]</sup>。其中 Roche 公司的 GS-FLX 454 基因组测序仪平台读长最长,平均读长 400 bp,最长可达 1 000 bp,可用于基因组的从头测序和转录组的拼接,但缺点是通量低,费用高,易由于碱基插入缺失引起测序错误; Illumina 公司的基因组分析仪采用双端测序,平均读长在 100 bp,通量最高,一次测序数据可达 200 Gb/run,可用于 ChIP-Seq 与 RNA-Seq 定量测定基因的表达水平,缺点是测序序列读长短; Life Technology 公司的 SOLID 序列分析仪也是采用双端测序,读长最短,但是测序错误率最低,可用于基因组反复测序以检测变异位点。

## 2 RNA-Seq 技术测序的数据处理与分析方法

### 2.1 数据处理

不同的高通量测序平台产生的数据文件格式不同,FASTQ 格式是高通量数据分析常用的文件格式<sup>[17]</sup>。在进行数据分析之前,需要根据不同的测序文件格式对数据进行标准化处理,以常见的 FASTQ 格式为例,数据预处理包括质量控制、读段(reads)清理、转录组组装、转录组定量和标准化 5 个过程<sup>[18]</sup>。质量控制可以使用基于 R 语言的 bioconductor 程序中的 ShortRead 安装包<sup>[19]</sup>或者 FastQC 与 StatsDB 工具<sup>[20-21]</sup>,reads 清理主要是清理 reads 两端低质量的区域(测序质量分数小于 Q20)、3 端测序接头、测序中出现的未知核苷酸,可以使用 Fastx\_toolkit 与 FastQC 工具<sup>[21-22]</sup>共同完成;转录组的组装包括有参考基因序列的组装、从头组装以及 2 种方法的结合这 3 种方式,有参考基因序列的组装可以使用 Tophat 与 Cufflinks 工具<sup>[23]</sup>,从头组装可以使用 Trinity 工具<sup>[24]</sup>,转录组定量与标准化可以使用 RPKM(reads per kilo bases per million reads)指标,标准化后的数据可以直接用作后续的数据分析工作。

### 2.2 分析方法

转录组数据的分析与计算机、网络技术联系紧密,形式多种多样<sup>[25]</sup>。比如将序列比对到 NCBI 的子数据库进行 BLAST 同源比对的数据库分析形式,应用 Bowtie 与 Trinity 软件组装及拼接 reads

的单机程序分析形式,应用在线程序如 CpGPlot 预测核酸序列 CpG 岛的在线分析形式。根据参考基因组的有无<sup>[26]</sup>,RNA-Seq 数据分析的内容与方法有所差异。对于没有参考基因组的从头组装数据,分析的内容主要包括基因预测与注释、KOG/COG 分析、GO 分析、代谢通路分析、基因差异表达分析、差异基因富集分析与 SSR/SNP 分析;对于有参考基因组的重测序数据,分析内容主要包括基因表达差异分析、差异基因富集分析、新转录本预测、非翻译区(UTR)分析、可变剪接分析、SSR/SNP 分析、UTR 分析、Operon 分析与 Non-coding RNA 分析。

## 3 RNA-Seq 技术在反刍动物瘤胃微生物研究中的应用

### 3.1 瘤胃微生物代谢酶特性

RNA-Seq 技术能够从转录组的水平定性研究瘤胃微生物酶基因的表达情况,结合原核表达、真核表达和酶动力学试验等能够进一步定量分析酶学特性。Dai 等<sup>[27]</sup>对荷斯坦奶牛瘤胃微生物进行转录组研究,发现分别约有 1% 和 0.1% 的基因能够编码糖苷水解酶(GH)和碳水化合物降解酶,这些基因编码的酶有 98% 来自 GH 蛋白家族,并且发现饲料中含有大量木聚糖时瘤胃瘤胃球菌属(*Ruminococcus*)、纤维杆菌属(*Fibrobacter*)和普氏菌属(*Prevotella*)细菌 GH48 家族中外切葡聚糖酶(cellobiohydrolases, CBH)基因表达量最高,占纤维素酶基因的 7.65%。研究表明,底物木聚糖能够诱导 GH48 基因的表达,是提高饲料中碳水化合物利用率的重要因素。然而,在 Dodd 等<sup>[28]</sup>的试验中发现一些能够抑制碳水化合物酶活性的营养成分,试验中用小麦阿拉伯糖基木聚糖(arabinoxylan)为唯一碳源的培养基(WAX)和以木糖(xylose)加果胶糖(arabinose)为碳源的培养基(XA)富集培养瘤胃布氏普雷沃氏菌(*Prevotella bryantii*),提取 mRNA 进行转录组学研究发现,与 XA 培养环境比较,在 WAX 培养环境下 GH 基因高度表达,其表达量是 XA 培养环境下的 4~16 倍,因而推测饲料中的木糖和果胶糖能够抑制木聚糖酶基因的表达。进一步研究发现,影响酶活的因素不仅仅是底物的营养成分,不同酶之间的协同作用也是影响酶活的重要原因。例如,在 WAX 培养环境下发现一些未知功能的新基因,命名为

*PbXyn5A*,通过薄层色谱(TLC)分析发现 *PbXyn5A* 有木聚糖内切酶(endoxylanase)活性,与 *PbAra43A* 共同存在时对 WAX 的降解率远远高于二者单独存在时。在瘤胃真菌方面,Wang 等<sup>[29]</sup> 利用 RNA-Seq 技术对瘤胃 *Neocallimastix patriciarum* 进行转录组学研究,发现潜在的 219 个编码 GH 家族蛋白的叠连群(contigs),这些 contigs 被细分到 25 个 GH 蛋白家族,在对应的 25 个 GH 蛋白家族中发现 4 种酶共同参与水稻、紫狼尾草和甘蔗渣的降解,这些酶具有葡萄糖苷酶、木聚糖酶和葡聚糖酶活性,组装后的 contigs 进一步通过真核表达试验成功克隆获得 19 个纤维素酶基因,当培养基中底物是紫狼尾草时,纤维素内切酶的活性最强,活性为 0.6 U/mg,纯化后的活性为 1.15 U/mg;当底物是水稻时,葡聚糖外切酶(exoglucanase,EXG)与  $\beta$ -葡萄糖苷酶( $\beta$ -glucosidases,BGLUs)的活性最强,其中 EXG 的活性为 16.93 U/mg,BGLUs 的活性为 0.81 U/mg,纯化后的活性分别为 78.97 和 2.17 U/mg,推测这 2 种酶具有潜在的工业生产价值,紫狼尾草、水稻和甘蔗渣中不同营养组分对纤维素酶的诱导作用的差异性可能导致了酶活性的不同。此外,Lawley 等<sup>[30]</sup> 利用 RNA-Seq 技术对 3 头牛瘤胃乳酸杆菌(*L. ruminis*)转录组测序发现了苔聚糖酶和木质纤维素酶基因,这些基因编码的苔聚糖酶和木质纤维素酶可以将饲料中的葡聚糖降解成[3-O- $\beta$ -cellotriosyl-D-glucose (LDP4) 和 4-O- $\beta$ -laminaribiosyl-D-cellobiose(CDP4)],当底物中 LDP4 或 CDP4 含量增加,操纵子正调控作用增强,促进苔聚糖酶和纤维素酶基因的表达。以上研究表明,瘤胃微生物所产生的酶的活性受饲料组分或培养基底物影响,可能是饲料中底物的营养成分或其代谢的中间产物对酶的分泌与活性有一定的诱导作用,从而促使相应操纵子的调控,进而影响酶基因的表达,这种调控作用可能促进也可能抑制酶的代谢,RNA-Seq 技术应用于了解酶与酶、酶与底物之间的调控关系,有利于提高饲料的利用效率。

## 3.2 瘤胃微生物多样性

### 3.2.1 瘤胃未知微生物

Fouts 等<sup>[31]</sup> 利用 RNA-Seq 技术获得 12 头牛瘤胃微生物转录组数据,对每头牛瘤胃细菌部分 reads(5 520 个)中的 16S rRNA 的 V1~V3 保守

区域进行分析,发现 1 903~2 432 个归类到种水平的操作性分类单元(operational taxonomic unit, OTU),其中 80% 来自梭菌目(Clostridiales)、拟杆菌目(Bacteroidales)、丹毒丝菌目(Erysipelotrichales)和未知分类情况的 unclassified TM7,同时发现 13 个 OTU 可以归类到古细菌(archaeal),并且其中有 2 种未知古细菌,分别与甲烷短杆菌属(*Methanobrevibacter*)和热裸单胞菌属(*Thermogymnomonas*)最为相似;在对真菌 18S rRNA 保守区归类分析发现 71 个 OTU,与 GenBank 数据库比对后发现 53 个 OTU 与数据库中已知序列的相似度在 97% 以上,通过 BLAST(同源检索)与 LCA(聚类分析)后发现未知曲霉属(*Aspergillus*)真菌,目前在 GenBank 数据库中的基因编号是 HQ393873.1,而在这之前 GenBank 数据库中没有报道过。

### 3.2.2 瘤胃微生物组成

瘤胃中拥有数量庞大、种类繁多的微生物,其中有大量的纤维、淀粉和脂肪降解菌,这些微生物共同参与营养物质的代谢。有研究通过 RNA-Seq 技术发现,与饲喂标准饲料粮相比较,分别摄入富含淀粉与脂肪的饲料粮后,瘤胃内容物中的细菌种类没有显著变化,但是细菌菌种的主体地位变化显著,尤其是在饲料粮中添加脂肪的环境下瘤胃脂肪降解菌脂解厌氧弧菌(*Anaerovibrio lipolytica*)的数量明显增加<sup>[32]</sup>。另据报道,脂肪能够附着在饲料颗粒表面阻碍非脂肪降解细菌对营养物质的摄取,导致其他细菌的生长速率下降,推测瘤胃中脂肪能够影响瘤胃微生物的组成<sup>[33]</sup>。进一步研究发现,增加饲料淀粉含量时,瘤胃中 *Barnesiella*、*Oribacterium* 与 *Olsenella* 的数量显著增加,同时丁酸弧菌属(*Butyrivibrio*)—假丁酸弧菌属(*Pseudobutyrvibrio*)与 *Rikenellaceae*\_RC 的数量显著减少<sup>[32]</sup>,原因可能是 *Barnesiella*、*Oribacterium* 与 *Olsenella* 能够产生高效降解淀粉的复合酶,通过降解淀粉提供的营养物质促进自身的生长,这与 Fava 等<sup>[34]</sup> 研究中认为 *Barnesiella*、*Oribacterium* 与 *Olsenella* 能够分泌高活性淀粉酶的结果一致。此外,Li<sup>[35]</sup> 利用 RNA-Seq 技术研究 5 头被去势的公牛瘤胃微生物的菌落组成,结果表明瘤胃中变形菌门(Proteobacteria)、拟杆菌门(Bacteroidetes)、厚壁菌门(Firmicutes)、螺旋体门(Spirochaetes)和互养菌门(Syn-



ergistetes) 为优势细菌, 其中 Proteobacteria 在瘤胃微生物中占 45.7%, 远大于 Bacteroidetes 的 23.7%, 广古菌门 (Euryarchaeota) 在所有古细菌中占 94.2%, 研究表明, 尽管这些牛在个体上有差异, 但是瘤胃微生物组成差异不大。在 Fouts 等<sup>[31]</sup>的研究中发现, 仅在瘤胃液中存在并且占优势的细菌是 *Prevotella* (7.3%) 和 *Tannerella* (1%), 而仅在瘤胃内容物中存在并且占优势的细菌是 *Butyrivibrio* (2.5%) 和 *Blautia* (0.5%), 研究表明细菌在瘤胃液与瘤胃内容物中的组成差异显著, 但古细菌和真菌 (fungi) 在瘤胃液与瘤胃内容物中的组成没有显著差异, 而瘤胃液与瘤胃内容物中的营养成分相似, 进而推测, 瘤胃中营养物质存在的环境可能是导致细菌差异分布的重要原因。以上研究表明, 瘤胃微生物的组成与饲料的营养成分、形态有关, 且 RNA-Seq 技术是研究瘤胃微生物多样性的一种有效的方法, 在瘤胃微生物的组成以及挖掘未知瘤胃微生物资源方面有很大的前景。

### 3.3 瘤胃微生物新功能基因

RNA-Seq 技术不仅可以高效挖掘瘤胃微生物新的基因, 同时可以挖掘这些基因相关的功能。Dassa 等<sup>[36]</sup>利用基于新一代高通量测序平台的 RNA-Seq 技术在瘤胃微生物中首次发现黄色瘤胃球菌 (*Ruminococcus flavefaciens*) 和白色瘤胃球菌 (*Ruminococcus albus*) 的 strains FD-1、007c、17 和 strains 7、8、SY3 基因, 这些基因能够编码锚定蛋白。Qi 等<sup>[37]</sup>利用 RNA-Seq 技术对加拿大麝牛瘤胃微生物总 mRNA 测序获得 2.8 G 的 reads, 组装拼接后获得 59 129 个 contigs, 并且从 2 500 个编码蛋白质的功能 contigs 中鉴定出 1 000 个与 GH、糖脂酶以及多聚糖裂解酶活性有关的功能基因, 其中发现 54 条编码 GH48 家族蛋白的基因序列, 它们均来自瘤胃中的厌氧真菌。李劲亭等<sup>[38]</sup>在瘤胃细菌中发现 455 条编码 GH48 家族蛋白的基因序列, 以 89% 相似度作为区分其种类的界定标准, 发现这些基因序列分别编码 66 种不同的 GH48 家族蛋白。此外, 有研究通过真核表达试验成功克隆到瘤胃 *Neocallimastix patriciarum* GH 家族基因, 在这些被克隆到的基因里, 还包含了以往瘤胃宏基因组学中很少提及到的 GH6 和 GH48 家族基因<sup>[29]</sup>。基因敲除试验证明, GH48 家族蛋白中的纤维素酶对梭

热杆菌 (*Clostridium thermocellum*) 和 *Ruminococcus albus* 的纤维素降解能力有重要的影响, 将 GH48 基因敲除后前者降解纤维素的能力严重下降, 后者则完全丧失<sup>[39]</sup>。而在 Sukharnikov 等<sup>[40]</sup>的研究中发现放线菌门 (Actinobacteria)、Firmicutes 和绿弯菌门 (Chloroflexi) 类群的 GH48 基因可能源于一个共同的祖先, 推测 GH48 家族基因可能存在一定的保守性, 并且这种保守性与宿主种类没有关系。Xu 等<sup>[41]</sup>对瘤胃细菌进行转录组学研究, 发现 76 个与纤维素降解有关的碳水化合物活性酶 (CAZymes), 之后 Schellenberg 等<sup>[42]</sup>利用 Roche/454 WGS 测序技术在瘤胃 *Clostridium stercorarium* 的转录组中发现 17 个与碳水化合物降解酶有关的功能基因, 其中包括纤维素酶的 *celyZ*, 纤维二糖磷酸化酶的 *cepAB*, 木聚糖酶的 *xynABC*, 木糖苷酶的 *xylAB*、*bxlAB*、*bglZ*, 阿拉伯糖苷酶的 *arfAB*,  $\alpha$ -半乳糖苷酶的 *agaA*, 果胶酸裂解酶的 *pelA* 和  $\alpha$ -鼠李糖苷酶的 *ramA*, 但仍有大量 CAZymes 基因未被挖掘, 并且这些基因所编码的碳水化合物酶的功能有待更深入的研究。与反刍动物不同, 有研究发现白蚁体内半纤维素的消化由肠道内的共生菌完成, 而纤维素的消化却是由白蚁与其共生菌协同完成<sup>[43]</sup>, 白蚁自身是否能够产生纤维素酶有待进一步试验验证。此外, 对猪和人类肠道研究也发现了一些功能基因。Poroyko 等<sup>[44]</sup>通过对全部转录组进行 RNA-Seq 技术测序, 研究对新生仔猪饲喂母乳 (MF) 和人工乳 (FF) 时肠道微生物的功能基因表达的差异, 发现 2 种处理下基因的表达模式是相似的, 所有样品均表达了大量与碳水化合物及蛋白质代谢相关酶类的转录本。Booijink 等<sup>[45]</sup>运用 RNA-Seq 技术对人类肠道微生物的宏转录组进行研究, 通过对 2 个尚未断奶的婴儿的排泄物总 RNA 测序分析, 阐明肠道微生物的原位基因表达, 并通过与 COG (cluster of orthologous genes) 数据库比对发现有 26% 的功能基因簇与新陈代谢有关。以上研究表明, RNA-Seq 技术在挖掘肠道微生物新基因和基因所编码的酶上有巨大的潜力, 有助于人类探知复杂环境中的微生物之间的关联。

## 4 小 结

RNA-Seq 技术的瘤胃微生物转录组学研究

已有数年,新一代高通量测序技术的发展促进了该领域的进步,对于有参考基因组数据的瘤胃微生物,转录组测序数据可以很好地验证基因组数据的注释,在没有参考基因组序列的瘤胃微生物中,转录组测序可以丰富和完善该物种的遗传数据库,这些完善的数据资源有助于对该物种进行后续的分子生物学研究。通过与蛋白质组、基因组、代谢组、基因芯片技术等其他分子生物学研究手段的结合,使人们更加深入地了解复杂瘤胃微生物菌群功能基因表达与调控的方式,如何将 RNA-Seq 技术与其他分子生物学技术结合探讨瘤胃微生物相关基因在特殊的瘤胃内环境中的表达与调控、瘤胃微生物之间的关系、瘤胃微生物与宿主之间的关系、环境对瘤胃微生物多样性的影响是今后研究的热点。然而,问题和挑战并存,首先,瘤胃环境复杂,RNA 纯化试剂盒并不能除掉所有杂质,尤其是内容物中携带的腐殖酸,而获得高纯度的 RNA 样品是试验的关键;其次,在对瘤胃微生物群体或个体进行 cDNA 序列分析时,转录组中 mRNA 的富集是主要的难题;再次,瘤胃中大量未知微生物的相关功能基因在基因注释数据库中没有注释信息,这些缺陷导致寻找功能基因存在很大困难;最后,RNA-Seq 技术获得的遗传信息是海量的,生物信息学分析将面临巨大的挑战,传统生物信息学分析工具很难满足大数据要求,适用于常规实验室使用的新算法和分析软件的开发成为当前的迫切需求。随着测序技术本身以及配套技术的不断发展完善,相信这类技术难题最终会得到解决,测序成本将更加低廉。

## 参考文献:

- [ 1 ] BUDDLE B M, DENIS M, ATTWOOD G T, et al. Strategies to reduce methane emissions from farmed ruminants grazing on pasture [ J ]. *The Veterinary Journal*, 2011, 188 ( 1 ) : 11-17.
- [ 2 ] 秦楠, 栗东芳, 杨瑞馥. 高通量测序技术及其在微生物学研究中的应用 [ J ]. *微生物学报*, 2011, 51 ( 4 ) : 445-457.
- [ 3 ] NAGALAKSHMI U, WANG Z, WAERN K, et al. The transcriptional landscape of the yeast genome defined by RNA sequencing [ J ]. *Science*, 2008, 320 ( 5881 ) : 1344-1349.
- [ 4 ] WILHELM B T, MARGUERAT S, WATT S, et al. Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution [ J ]. *Nature*, 2008, 453 ( 7199 ) : 1239-1243.
- [ 5 ] CARVALHAIS L C, DENNIS P G, TYSON G W, et al. Application of metatranscriptomics to soil environments [ J ]. *Journal of Microbiological Methods*, 2012, 91 ( 2 ) : 246-251.
- [ 6 ] LU T T, LU G J, FAN D L, et al. Function annotation of rice transcriptome at single nucleotide resolution by RNA-Seq [ J ]. *Genome Research*, 2010, 20 ( 9 ) : 1238-1249.
- [ 7 ] DAVID L A, MAURICE C F, CARMODY R N, et al. Diet rapidly and reproducibly alters the human gut microbiome [ J ]. *Nature*, 2013, 505 ( 7484 ) : 559-563.
- [ 8 ] ROSENTHAL A Z, MATSON E G, ELDAR A, et al. RNA-Seq reveals cooperative metabolic interactions between two termite-gut spirochete species in co-culture [ J ]. *The ISME Journal*, 2011, 5 ( 7 ) : 1133-1142.
- [ 9 ] BOMAR L, GRAF J. Investigation into the physiology of *Aeromonas veronii* *in vitro* and inside the digestive tract of the medicinal leech using RNA-Seq [ J ]. *Biological Bulletin*, 2012, 223 ( 1 ) : 155-166.
- [ 10 ] BISANZ J E, MACKLAIM J M, GLOOR G B, et al. Bacterial metatranscriptome analysis of a probiotic yogurt using an RNA-Seq approach [ J ]. *International Dairy Journal*, 2014, 39 ( 2 ) : 284-292.
- [ 11 ] BAKER B J, SHEIK C S, TAYLOR C A, et al. Community transcriptomic assembly reveals microbes that contribute to deep-sea carbon and nitrogen cycling [ J ]. *The ISME Journal*, 2013, 7 ( 10 ) : 1962-1973.
- [ 12 ] MARDIS E R. Next-generation DNA sequencing methods [ J ]. *Annual Review of Genomics and Human Genetics*, 2008, 9 ( 1 ) : 387-402.
- [ 13 ] GLENN T C. Field guide to next-generation DNA sequencers [ J ]. *Molecular Ecology Resources*, 2011, 11 ( 5 ) : 759-769.
- [ 14 ] 沈圣, 屈彦纯, 张军. 下一代测序技术在表观遗传学研究中的重要应用及进展 [ J ]. *遗传*, 2014, 36 ( 3 ) : 256-275.
- [ 15 ] WALL P K, LEEBENS-MACK J, CHANDERBALI A S, et al. Comparison of next generation sequencing technologies for transcriptome characterization [ J ]. *BMC Genomics*, 2009, 10 ( 1 ) : 347.
- [ 16 ] CRAWFORD J E, GUELBEIGO W M, SANOU

- A, et al. *De novo* transcriptome sequencing in *Anopheles funestus* using Illumina RNA-Seq technology [J]. PLoS One, 2010, 5(12): e14202.
- [17] COCK P J A, FIELDS C J, GOTO N, et al. The sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants [J]. Nucleic Acids Research, 2010, 38(6): 1767–1771.
- [18] 高山, 欧剑虹, 肖凯. R 语言与 Bioconductor 生物信息学应用 [M]. 天津: 天津科技翻译出版有限公司, 2014: 173–184.
- [19] MORGAN M, ANDERS S, LAWRENCE M, et al. ShortRead: a bioconductor package for input, quality assessment and exploration of high-throughput sequence data [J]. Bioinformatics, 2009, 25(19): 2607–2608.
- [20] RAMIREZ-GONZALEZ R H, LEGGETT R M, WAITE D, et al. StatsDB: platform-agnostic storage and understanding of next generation sequencing run metrics [J]. F1000Research, 2013, 2: 248.
- [21] LOHSE M, BOLGER A M, NAGEL A, et al. RobiNA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics [J]. Nucleic Acids Research, 2012, 40: W622–W627.
- [22] WANG L, WANG S, LI W. RSeQC: quality control of RNA-Seq experiments [J]. Bioinformatics, 2012, 28(16): 2184–2185.
- [23] TRAPNELL C, WILLIAMS B A, PERTEA G, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation [J]. Nature Biotechnology, 2010, 28(5): 511–515.
- [24] VIJAY N, POELSTRA J W, KÜNSTNER A, et al. Challenges and strategies in transcriptome assembly and differential gene expression quantification. A comprehensive *in silico* assessment of RNA-Seq experiments [J]. Molecular Ecology, 2013, 22(3): 620–634.
- [25] CARDENAS E, TIEDJE J M. New tools for discovering and characterizing microbial diversity [J]. Current Opinion in Biotechnology, 2008, 19(6): 544–549.
- [26] 岳桂东, 高强, 罗龙海, 等. 高通量测序技术在动植物研究领域中的应用 [J]. 中国科学: 生命科学, 2012, 42(2): 107–124.
- [27] DAI X, TIAN Y, LI J T, et al. Metatranscriptomic analyses of plant cell wall polysaccharide degradation by microorganisms in the cow rumen [J]. Applied and Environmental Microbiology, 2015, 81(4): 1375–1386.
- [28] DODD D, MOON Y H, SWAMINATHAN K, et al. Transcriptomic analyses of xylan degradation by *Prevotella bryantii* and insights into energy acquisition by xylanolytic bacteroidetes [J]. The Journal of Biological Chemistry, 2010, 285(39): 30261–30273.
- [29] WANG T Y, CHEN H L, LU M Y, et al. Functional characterization of cellulases identified from the cow rumen fungus *Neocallimastix patriciarum* W5 by transcriptomic and secretomic analyses [J]. Biotechnology for Biofuels, 2011, 4: 24.
- [30] LAWLEY B, SIMS I M, TANNOCK G W. Whole-transcriptome shotgun sequencing (RNA-Seq) screen reveals upregulation of cellobiose and motility operons of *Lactobacillus ruminis* L5 during growth on tetrasaccharides derived from barley  $\beta$ -glucan [J]. Applied and Environmental Microbiology, 2013, 79(18): 5661–5669.
- [31] FOUTS D E, SZPAKOWSKI S, PURUSHE J, et al. Next generation sequencing to define prokaryotic and fungal diversity in the bovine rumen [J]. PLoS One, 2012, 7(11): e48289.
- [32] ZENED A, COMBES S, CAUQUIL L, et al. Microbial ecology of the rumen evaluated by 454 GS FLX pyrosequencing is affected by starch and oil supplementation of diets [J]. FEMS Microbiology Ecology, 2013, 83(2): 504–514.
- [33] HENDERSON C. The effects of fatty acids on pure cultures of rumen bacteria [J]. The Journal of Agricultural Science, 1973, 81(1): 107–112.
- [34] FAVA F, GITAU R, GRIFFIN B A, et al. The type and quantity of dietary fat and carbohydrate alter faecal microbiome and short-chain fatty acid excretion in a metabolic syndrome ‘at-risk’ population [J]. International Journal of Obesity, 2013, 37(2): 216–223.
- [35] LI F Y, SUN X, HENDERSON G, et al. Comparative analyses of the bovine rumen microbiota using RNA and targeted DNA-based sequencing approaches [C]//2014 ADSA-ASAS-CSAS Joint Annual Meeting. [s.n.]: ASAS, 2014.
- [36] DASSA B, BOROVOK I, RUMY-ISRRAELI V, et al. Rumen cellulomics: divergent fiber-degrading strategies revealed by comparative genome-wide analysis of six ruminococcal strains [J]. PLoS One,

- 2014,9(7):e99221.
- [37] QI M, WANG P, O' TOOLE N, et al. Snapshot of the eukaryotic gene expression in muskoxen rumen-a metatranscriptomic approach[J]. PLoS One, 2011, 6(5):e20521.
- [38] 李劲亭, 苏小运, 田彦, 等. 瘤胃细菌 GH48 家族糖苷水解酶基因多样性[J]. 微生物学报, 2014, 54(1):53-61.
- [39] OLSON D G, TRIPATHI S A, GIANNONE R J, et al. Deletion of the Cel48S cellulase from *Clostridium thermocellum*[J]. Proceedings of the National Academy of Sciences of the United States of America, 2010, 107(41):17727-17732.
- [40] SUKHARNIKOV L O, ALAHUHTA M, BRUNECKY R, et al. Sequence, structure, and evolution of cellulases in glycoside hydrolase family 48[J]. The Journal of Biological Chemistry, 2012, 287(49):41068-41077.
- [41] XU C G, HUANG R R, TENG L, et al. Structure and regulation of the cellulose degradome in *Clostridium cellulolyticum*[J]. Biotechnology for Biofuels, 2013, 6(1):73.
- [42] SCHELLENBERG J J, VERBEKE T J, MCQUEEN P, et al. Enhanced whole genome sequence and annotation of *Clostridium stercoarium* DSM8532T using RNA-Seq transcriptomics and high-throughput proteomics[J]. BMC Genomics, 2014, 15:567.
- [43] TARTAR A, WHEELER M M, ZHOU X G, et al. Parallel metatranscriptome analyses of host and symbiont gene expression in the gut of the termite *Reticulitermes flavipes*[J]. Biotechnology for Biofuels, 2009, 2(1):25.
- [44] POROYKO V, WHITE J R, WANG M, et al. Gut microbial gene expression in mother-fed and formula-fed piglets[J]. PLoS One, 2010, 5(8):e12459.
- [45] BOOIJINK C C G M, BOEKHORST J, ZOETENDAL E G, et al. Metatranscriptome analysis of the human fecal microbiota reveals subject-specific expression profiles, with genes encoding proteins involved in carbohydrate metabolism being dominantly expressed[J]. Applied and Environmental Microbiology, 2010, 76(16):5533-5540.

## RNA-Seq Technology and Its Application in Rumen Microbes

LIAO Qi<sup>1</sup> LIU Xuchuang<sup>2</sup> LI Qing<sup>1,2</sup> FAN Yueyuan<sup>2</sup> ZHANG Chunyong<sup>1,2</sup>  
YANG Shuli<sup>1,2</sup> MAO Huaming<sup>1,2</sup> LENG Jing<sup>1,2\*</sup>

(1. Yunnan Provincial Key Laboratory of Animal Nutrition and Feed Science, Kunming 650201, China;

2. College of Animal Science and Technology of Yunnan Agricultural University, Kunming 650201, China)

**Abstract:** Varied microbes species in the rumen compose a complex microecosystem, which play an important role in the process of rumen digestion. The appearance of RNA-Seq technology has provided huge opportunities for studying on rumen microbes. This review mainly described how to study on enzyme metabolic characteristics of rumen microbes, the diversity of rumen microorganism, new functional genes of rumen microbes using RNA-Seq technology with next-generation sequencing technology. [Chinese Journal of Animal Nutrition, 2015, 27(4):1061-1067]

**Key words:** RNA-Seq technology; rumen microbes; diversity; new functional genes