基因组 De novo 组装结合比较 RNA-Seq 策略在 VBNC 状态乳杆菌转录组学研究中的应用

徐振波[1,2]，侯玉超[1]，刘君彦[1]，邓阳[1]，张霞[1]，李冰[1]，李琳[1]
（1.华南理工大学轻工与食品学院，广东广州 510640）（2.美国马里兰大学微生物病理系，巴尔的摩 MD21201）

摘要：本文应用有参考基因组组装策略对一株提取自啤酒中的乳杆菌的两种生理状态进行转录组学分析。利用基因组测序技术得到此株乳杆菌的全部基因组信息之后，把此株乳杆菌用低温诱导至 VBNC 状态，对正常状态和 VBNC 状态乳杆菌提取总 RNA。对 RNA 质量进行验证，将符合要求的 RNA 进行逆转录形成 cDNA，之后进行文库构建和 Solexa 测序。通过相关指标对测序结果进行评估，之后以此前获得的全基因组序列为对照，进行转录组与基因组的序列比对和基因表的达差异分析。结果显示比对率大于 95%，表达上调的基因有 21 个，表达下降的基因有 7 个。本文所应用的基因组 De novo 组装结合比较 RNA-Seq 策略与常规 De novo RNA-Seq 策略相比，省却 De novo 拼接与组装的繁琐步骤，具有更高的准确性，为今后食品微生物在不同生理状态中分子调控机制的研究提供重要基础。

关键词：乳杆菌；转录组；VBNC
文章篇号：1673-9078(2016)07-124-129 DOI: 10.13982/j.mfst.1673-9078.2016.7.020

Application of Genome De novo Assembly and Reference-based Transcriptome Assembly Strategy for Studying Lactobacilli in the VBNC State

XU ZHEN-bo[1,2], HOU YU-chao[1], LIU JUN-yan[1], DENG Yang[1], ZHANG Xia[1], LI Bing[1], LI Lin[1]
(1.College of Light Industry and Food Sciences, SouthChinaUniversity of Technology, Guangzhou 510640, China)
(2.Department of biomedical science of university of Maryland, Baltimore 21201)

Abstract: A new reference-based assembly strategy was applied for the transcriptome analysis of a *Lactobacillus* strain in two physiological states, isolated from beer. First, the genomic information was collected using the Illumina HiSeq 2500 sequencing system and de novo assembly. Next, this *Lactobacillus* strain was induced into the viable but nonculturable (VBNC) state and total RNA were extracted in the normal and VBNC states. The quality of total RNA was evaluated and RNA with acceptable quality was reverse-transcribed into cDNA. A library was constructed and Solexa sequencing was performed. Based on five relevant indices, the sequencing quality was evaluated. Subsequently, the whole genome sequence obtained previously was used as a control for sequence alignment between the transcriptome and genome, as well as to analyze differences in gene expression. The results showed that the total mapping ratio was greater than 95% after sequence comparison, and 21 significantly up-regulated and seven down-regulated genes were detected. Compared to conventional de novo RNA-Seq technology, this application of genome de novo assembly and reference-based transcriptome assembly strategy can help avoid the tedious steps required in de novo assembly, shows high accuracy, and provides an important foundation for studying the molecular regulation mechanisms of food microorganisms in different physiological states.

Key words: *Lactobacillus*; transcriptome; viable but nonculturable (VBNC)

收稿日期：2015-07-15

基金项目：国家 973 计划项目“食品加工过程安全控制理论与技术的基础研究”（2012CB720800）；国家自然科学基金青年基金项目“食品加工过程中金黄色葡萄球菌生物被膜行为的分子机制研究”（31201362）；广东省优秀博士学位论文作者资助项目“食品体系中典型病原微生物关键致毒与有害因子监控平台研究”（K3140030）；中国博士后科学基金面上项目（2014M552204）

作者简介：徐振波（1982-），男，博士，讲师，研究方向：食源性微生物安全控制与致毒机理研究

转录组是指一个特定组织或细胞在某一发育阶段或功能状态下转录出来的所有 RNA 集合，其中以 mRNA 为关键。作为新型组学技术之一，转录组学研究反映了不同生理状态、不同组织类型以及不同环境条件下表达的基因[1]，是深入阐明该生物体相关状态以及分子机制的重要手段。近年来，食品安全问题日益引起社会的重视与关注，其中由微生物引起的食物中毒、腐败与变质以及各种潜在危害正逐渐引起相关专家与学者的重视。微生物通过形成各种特定生理状态能逃逸常规检测与消毒灭菌，而其调控机理与分子机制大多尚未阐明，通过转录组学技术将对相关研究提供重要参考。

本文应用一种新型的转录组学研究策略，即通过对微生物进行基因组 De novo 组装结合比较 RNA-Seq 测序技术，对一株提取自啤酒中乳杆菌的正常状态与活但不可培养（VBNC）状态进行转录组学分析。乳杆菌在啤酒中形成 VBNC 状态，其生长缓慢，难以培养，从而在常规检测中易于显示为假阴性。论文在获得其全基因组信息并组装后，通过转录组测序技术对此株菌的正常状态与 VBNC 状态进行 RNA-Seq 分析，将转录组信息与基因组信息进行比对，筛选差异表达基因。论文为研究食品微生物不同生理状态的分子调控机制提供一种新型的转录组学研究策略。

1 实验材料与方法

1.1 实验材料

乳杆菌 BM-LC14617，并将其诱导至 VBNC 状态，从啤酒中获得；M.R.S 培养基，OXOID LTD., BASINGSTOKE, HAMPSHIRE, ENGLAND；DNA 快速提取试剂盒，广东东盛生物科技有限公司；细菌总 RNA 提取试剂盒，广东东盛生物科技有限公司；LIVE/DEAD Bacterial Viability Kit 活菌染色试剂盒，Molecular Probes 公司；dNTP mix, TIANGEN 公司；逆转录酶，TIANGEN 公司；RT Primer，TIANGEN 公司；RNAase inhibitor (40 U/μl)，美国 Thermo Scientific 公司。

1.2 实验方法

1.2.1 乳杆菌培养与 VBNC 状态诱导

以乳杆菌 BM-LC14617 菌株为实验对象，将保存在-20 ℃的 BM-LC14617 菌株用 MRS 培养基进行隔夜培养，培养完毕后，上下反复颠倒，吸取 50 μL 菌液移入 5 mL 新的 MRS 培养基中培养 8 小时。将培养过后的乳杆菌培养物分成等体积的两份，一份记为 BM-LC 14617 CK，此为 BM-LC14617 菌株的正常状态样本，保存于-80 ℃。另外一份乳杆菌培养物放入 0 ℃冰箱内保存，每隔 7 d 进行一次平板计数。以正常状态下的乳杆菌为对照，直至 MRS 检测平板在 26 ℃厌氧培养条件下 14 d 仍无单菌落出现。此时用活菌染色试剂盒进行染色，通过荧光显微镜进行观察，若有较多的绿色活菌存在，且数量多余红色死菌，则判定此乳酸杆菌培养物已经处于 VBNC 状态，记为 BM-LC14617 VBNC，保存于-80 ℃。

1.2.2 DNA 与 RNA 提取

对两种状态的乳杆菌各吸取 1.5 mL 菌液于 2 mL 离心管中，根据 DNA 快速提取试剂盒与细菌总 RNA 提取试剂盒说明书，按照相关操作分别提取基因组 DNA 与总 RNA。通过超微量分光光度计进行检测，测量其 A_{260}/A_{280} 值；同时通过琼脂糖电泳对所提取的 DNA 和 RNA 质量和纯度进行评价。

1.2.3 Solexa 基因组测序与 De novo 组装

对获得的 DNA 通过 Illumina HiSeq 2500 平台进行 Solexa 测序，测序完成后，对获得的序列数据进行处理和过滤后，运用生物信息学技术进行 De novo 拼接和组装，从而建立全基因组信息。

1.2.4 cDNA 的合成

获得正常生长状态与 VBNC 状态下的乳杆菌 RNA 后，以其为模板，逆转录获得 cDNA。反应体系：2 μL 5×Buffer, 1.6 μL 2.5mM dNTP mix，0.2 μL 40 U/μL RNase inhibitor, 0.2 μL 200 U/μL 逆转录酶，1 μL 5μM RT Primer，1 μL RNA template，用 DEPC 处理水定容至 10 μL。将反应体系混合均匀，42 ℃ 60 min，72 ℃ 10 min 终止反应。反应后获得的 cDNA，进行 Illumina HiSeq 2500 平台进行 RNA 测序与后续转录组分析。

1.2.5 文库构建和 Solexa 测序

应用新一代高通量测序平台 Illumina HiSeq 2500 测序系统对文库进行 Solexa 测序。Solexa 所采用的是边合成边测序的方式（Sequencing by Synthesis），采用了可逆阻断技术（即在 dNTP 上连接可剪切的阻断基团）[2]。运用一定的物理化学方法将样品 cDNA 随机片段化，形成几百个碱基片段或是更短的片段，并在两个末端加上接头（adapter）。将这这些片段按一定浓度进行稀释，准备进行样品扩增。

样品的扩增将在 Cluster Station 中的微流量池（Flow cell）中进行。flow cell 由 8 个单元通道（Lane）构成[3~4]，其表面固定有一层与 adapter 互补的单链引物（Primer）[5]。经处理后的碱基片段可以与 flow cell 表面的 primers 杂交，而被间接的“固定”在 flow cell

上。在聚合酶的作用下，以碱基片段为模板，以固定的 primer 为扩增引物，形成双链结构。双链结构经过变性、洗脱之后，只留下了固定在 flow cell 上的复制链。复制链一端固定在 flow cell 上，另外一端可以随机和附近的另一 primer 互补，被“固定”住，形成“桥”（bridge）结构[6]。这种单链桥以与之杂交的 primer 为扩增引物，在 flow cell 表面进行扩增，形成双链桥结构。双链桥经过变性形成单链结构，又可以与附近的 primer 形成 bridge，并作为下一次扩增的模板。经过反复的循环，每个大分子都得到了大量的扩增，在 flow cell 表面形成了大量的单克隆“DNA 簇群”（DNA Clusters）[7]。

DNA Clusters 在 Genome Analyzer 综合分析仪上进行序列分析，通过 SBS（sequencing by synthesis）法，进行测序[3]。每合成一个碱基读取一个数据，通过拍照的方法进行数据捕捉[8]。每反应一次，拍照四次（A、T、G、C 各一次），对取得的序列片段进行处理和拼接。（Solexa 扩增的整个过程如图 1）

1.2.6 序列分析

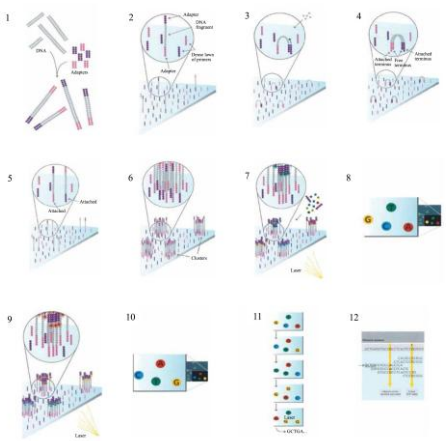


图 1 Solexa 测序基本流程

Fig.1 Solexa sequencing workflow

表 1 组装结果统计表

Table 1 Results of gene assembly

Term	Scaffold	Contig
总数量(Total Num, #)	196	398
总长度(Total Length, bp)	3,045,511	3,037,033
N50 值(N50, bp)[a]	26,583	12,514
N90 值(N90, bp)[b]	7,984	4,014
最大长度(Max Length, bp)	128,069	43,972
最小长度(Min Length, bp)	502	502
GC 含量(Sequence GC, %)	46.12	46.25

注：a, b：把组装出的 contigs 或 scaffolds 从大到小排列，当其累计长度刚刚超过全部组装序列总长度 50%（90%）时，最后一个 cont 或 scaffold 的大小即为 N50（N90）的大小。N50，N90 对评价基因测序的完整性有重要意义。

对原始数据进行质量评估和可信度分析，对数据进行处理与过滤。之后进行重叠分析，得到完整的基因组序列信息。测序得到的原始序列长度有数十个碱基，需要通过生物信息学分析将这些原始序列组装成较长的重叠群（contig），脚手架结构（scaffolds）。用“N”填补部分 read pairs 间存在的 gaps，直至最终的拼接序列不能再延伸为止。根据转录组整体拼接过程中不同测序材料 read 类型的不同，获得相材料的转录组序列[9]。

2 结果与分析

2.1 RNA 质量检测

用超微量分光光度计对本实验两个样本的总 RNA 进行质量检测，测量其 A_{260}/A_{280} 值。结果显示两个样本总 RNA 的 A_{260}/A_{280} 值均在在 2.0~2.2 之间，表明供试材料总 RNA 完整性良好，纯度较高，满足下一步测序建库的要求。

2.2 全基因组测序与 De novo 组装

经过 Solexa 测序与 De novo 组装之后得到预测基因 2526 个，基因总长度为 2,492，112 bp。其中 GC 含量为 47.15%，平均基因长度为 986.58 bp。其组装结果如表 1 所示。

2.3 测序文库的质量检测

对测序数据（Raw Data）去接头与低质量的 reads 处理后，所得数据记为 Total data，随后除去已知核糖体 RNA（28 S rRNA、18 S rRNA 等），获得统计数据为 BM-LC 14617 菌株的正常状态与 VBNC 状态的 clean reads、rRNA reads 与 effective ration 分别为 24,381,122 bp、12,984 bp 与 99.95%（正常状态）以及 18,957,492 bp、18,010 bp 与 99.90%（VBNC 状态）。

本次试验质量检测所用的软件为 FastQC（版本：0.10.1），对原始数据（reads）进行质量评估。质量评价决定测序结果的可靠性，由于测序采用双向测序（Paired-End Sequencing），即把每份样本的两个结果分为 read 1，read 2，将 BM-LC 14617 CK read 1，BM-LC 14617 CK read 2，BM-LC 14617 VBNC read 1 与 BM-LC 14617 VBNC read 2 分别记为 A、B、C 和 D。其主要指标分述如下，

Per base sequence quality 即 Fred 值，Fred 值 $=-10*\log_{10}(p)$，p 为测序错误的概率。例如，一条 reads 某位置出错的概率为 0.1，则其 Per base sequence

quality 为 10。图 2，Per base sequence quality 的横轴代表位置，纵轴代表数值，红色横线代表其的中位数，蓝线代表平均值，黄色区域是 25%~75%区间，触须之间为 10%~90%的区间。在测序质量检测过程中，若某一位置的下四分位数低于 10 或中位数低于 25，报"WARN"；若某一位置的下四分位数低于 5 或中位数低于 20，报"FAIL"。从图 2 可得，A，B，C，D 四次测序中，每个位置碱基的测序质量没有"WARN"和"FAIL"的情况出现，而且其下分位数都在 30 以上，表明测序质量可靠。

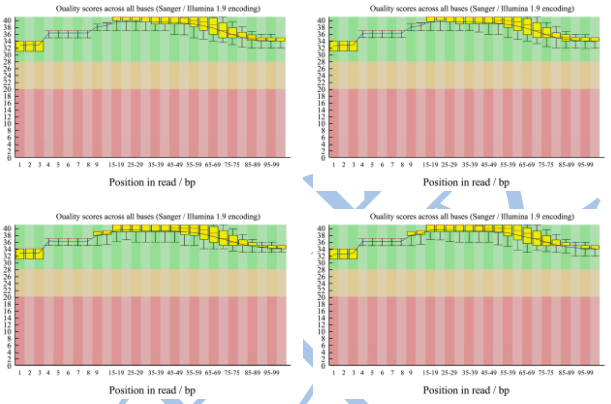
Per sequence quality scores 的横轴代表 reads 序列质量值，纵轴代表 reads 数目。当峰值对应的横坐标小于 27 时，报"WARN"；当其小于 20 时报"FAIL"。一般认为 reads 序列质量值峰值对应的横坐标大于 35 时，为高质量测序。而 A、B、C、D 四次测序恰好符合要求，其峰值横坐标均为 38，为高质量测序。

Per base sequence content 反应了四种碱基在各个位置上的分布。如图 2，其横轴代表 reads 位置，纵轴代表百分比。一般来说，由于碱基互补配对，A 与 T 含量相同，G 与 C 含量相同。正常情况下四种碱基的出现频率应该是接近的，且没有位置差异，因此质量高的样本中四条线应该平行且接近。当某一位置的 A/T 比例与 G/C 比例相差超过 10%，报"WARN"；当某一位置的 A/T 比例与 G/C 比例相差超过 20%，报"FAIL"。如图 2，A，B，C，D 四次测序中均报"FAIL"，这主要是因为测序过程中，通常前几位的 Per base sequence content 较低，导致碱基含量也会不正常，不过这种情况并不影响数据质量。

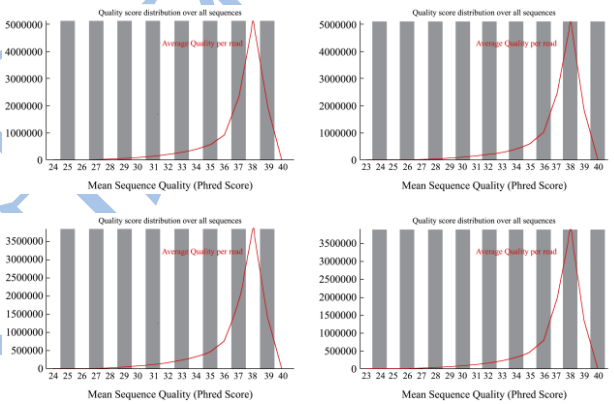
Per sequence GC content 表示 reads 的平均 GC 含量的分布。如图 2，Per sequence GC content 的横轴代表平均 GC 含量，纵轴是数量。红线是实际分布，蓝线是理论分布。若曲线有形状上的偏差，往往是文库污染或是部分 reads 构成的子集有偏差（overrepresented reads）导致；若形状接近正态分布但偏离理论分布的情况，提示我们可能有系统偏差。偏离理论分布的 reads 超过 15%时，报"WARN"；偏离理论分布的 reads 超过 30%时，报"FAIL"。从图 2 可得，A、B、C、D 四次测序中，实际情况和理论分布形状大体一致，位置一样，表明测序良好。

Per base N content 就是用来统计所有位置 N 碱基的比率。如果测序仪器不能辨别 reads 的某个位置到底是什么碱基时，就会产生 N 碱。如图 2 所示，Per base N content 的横轴代表 reads 的位置，纵轴代表百分比。正常情况下未知碱基含量是很小的，所以图上常常看到一条接近横轴的直线。当纵轴在 0-100%的范围内能看到凸起时，说明测序系统出了问题。当任意位置未知碱基的含量超过 5%时，报"WARN"；当任意位置的未知碱基的含量超过 20%时，报"FAIL"。从图 2 可得，四次测序中置 N 碱基含量一直保持低水平，甚至接近 0，表明极少有 N 碱基出现的状况

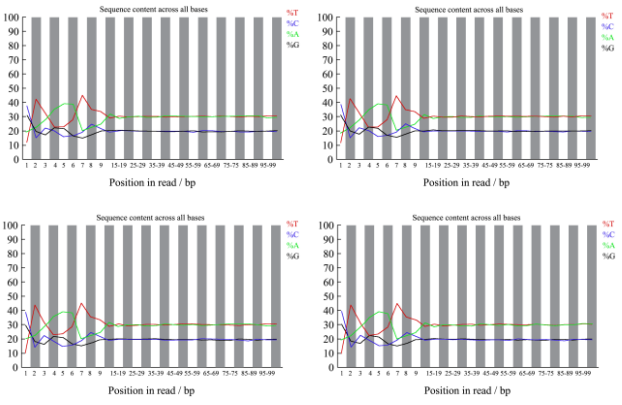
1.Per vase sequence quality



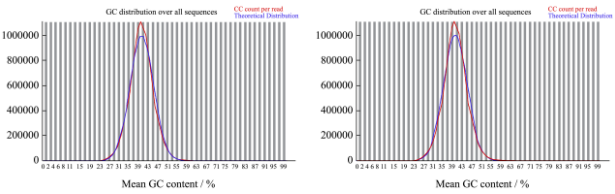
2. Per sequence quality scores

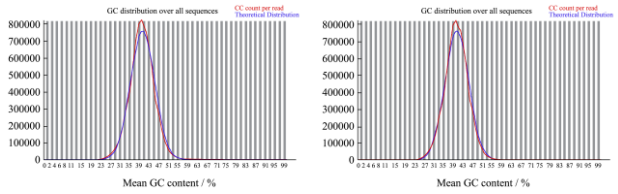


3. Per base sequence content



4. Persequence GCcontent





5. Per base N content

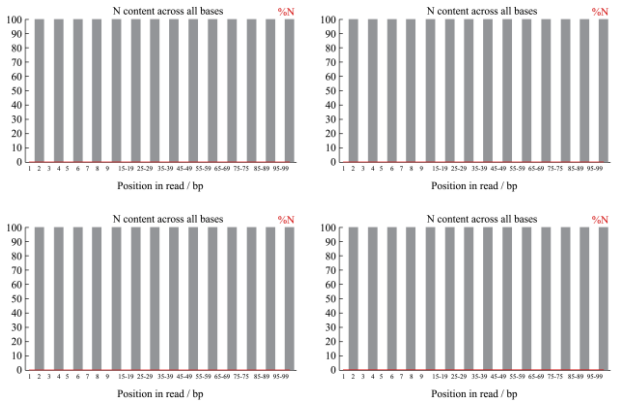


图 2 测序文库质量检测图

Fig.2 Quality analysis of sequencing library

综合上述结果，可以认为测序结果可信度较高，质量优良，可以进行之后的序列比对和表达量差异分析。

2.4 序列比对结果

通过对乳杆菌进行全基因组测序与 De novo 组装（表 1），运用 Tophat 软件将本次实验的转录组信息与基因组信息进行比对，其结果如表 2 所示。能对比到基因组上的序列数量，最低为 95%。Read 1 与 Read 2 成对比对到染色体占 47%以上，Duplication 率为 50%。

2.5 表达差异分析

由表 2 可得，序列组装基因组得到的基因数目是 1824。从表达差异图（图 3）可以看出差异基因的整体分布情况，其横轴 logCPM 代表样品对应的 $\log_2$（counts perl million）的均值，纵坐标 logFC 表示 log2（BM-LC 14617 VBNC/ BM-LC 14617 CK），代表基因在不同样本中的表达倍数。根据 FDR 对比较组 BM-LC 14617 CK vs BM-LC 14617 VBNC 的结果进行筛选。通过差异倍数(｜log2(Fold change)｜>1)和显著水平（FDR<0.05）来挑选样品间的表达差异基因，得到 28 个有显著差异的基因，其中 21 个表达上调，7 个表达下降。

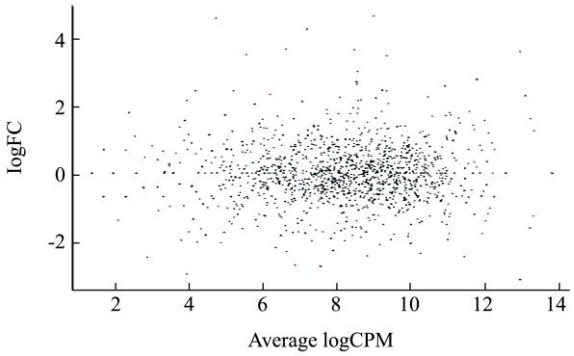


图 3 表达差异图

Fig.3 Plot for differences in gene expression

表 2 比对结果统计表

Table 2 Statistical results of comparison

sample name	BM-LC 14617 CK(percent)	BM-LC 14617 VBNC(percent)
Total reads[a]	24368138(100%)	18939482(100%)
Total mapped[b]	23216282(95.27%)	18223708(96.22%)
Multiple mapped[c]	29973(0.12%)	13647(0.07%)
Uniquely mapped[d]	23186309(95.15%)	18210061(96.15%)
Read 1 mapped[e]	11607869(47.64%)	9111555(48.11%)
Read 2 mapped[f]	11608413(47.64%)	9112153(48.11%)
Reads mapped in proper pairs[g]	11341306(46.54%)	8958329(47.30%)
Duplication[h]	12336153(50.62%)	9506626(50.19%)

注：*a*: Total reads，测序序列经过测序数据过滤后的数量统计；*b*: Total mapped，能定位到基因组上的测序序列的数量统计；*c*: Multiple mapped，在参考序列上有多个比对位置的测序序列的数量统计；*d*: Uniquely mapped，在参考序列上只有一个比对位置的测序序列的数量统计；*e*: Read 1 mapped，第一次测序序列比对的数量统计；*f*: Read 2 mapped，第二次测序序列比对的数量统计；*g*: Reads mapped in proper pairs，比对到正确位置的测序序列的数量统计；*h*: Duplication，重复测序序列的数量统计

3 结论

3.1 本次采用的为 velvet 序列组装软件（http://www.ebi.ac.uk/~zerbino/velvet/；版本：1.2.08）进行拼接。Velvet 软件由欧洲生物信息中心（EMBL-EBI）的 Daniel Zerbino 和 Ewan Birney 在 2008 年开发出来的，是一款在 Unix 下运行的 De novo 拼接软件，主要用于 25~500bp 的序列[2, 10]。Velvet 的

执行时基于 de Bruijin 算图（de Bruijin graphs）的算法，在构建算图中会有各种纠错步骤[2]。Velvet 是目前应用最广泛的拼接软件，尤其适用于细菌基因组的拼接。

3.2 通过对转录组的研究，可以得到生物体基因表达与基因变异的信息，还可以揭示生物体特定调控机制的分子学过程。随着科技发展的日新月异，转录组学研究技术也在发生着改变。

3.3 转录组学研究方法可以分为两类，一类是基于杂交的方法，主要为微阵列技术（Microarray）。另外一类是基于测序的方法，包括基因表达系列分析技术（Serial analysis of gene expression, SAGE），大规模平行测序（Massively parallel signature sequencing, MPSS）和 RNA 测序技术（RNA sequencing）等。Microarray 作为研究转录组的早期技术在分子生物学领域具有重大的意义。它可以同时测量样本中成千上万的基因在不同环境和不同状态的表达水平。其原理是在固定支持物表面有规律的固定成千上万的基因探针，这些探针可以和经过荧光标记物或是放射性标记物 DNA 或 cDNA 将结合。通过放射性显影技术读取结果，利用计算机软件进行分析，获得新号强弱及分布模式图，以此反映样本中相关基因的表达谱。Microarray 虽然可以分析大量基因的表达状况，但是其缺点在于必须在实验前得到样本基因序列信息和非特异性杂交。RNA-Seq 是运用高通量测序技术获得 mRNA、small RNA、非编码 RNA 等的序列测信息。全面快速地获取某一物种特定器官或组织在某一状态下的几乎所有转录本的一种技术。该技术先将样本中的 RNA 转录成为 cDNA，将其打成小片段，并加上 adapters。利用高通量测序仪进行测序，得到序列信息[11]。将所得的序列信息通过比对（有基因组信息）或是 De novo 组装形成完整的转录谱。Solexa 测序正是其中的一种，Solexa 测序技术由隶属于剑桥大学的 Solexa 公司发明[12]，2006 年被 Illumina 公司收购，是目前应用最为广泛的测序技术[3, 7]。此项技术通量大且可控制，数据精确，操作简单、所需样品少，能够在极短时间内获得数十亿高精确度的碱基序列信息。近年来，转录组测序技术更是广泛应用于微生物的基因差异表达研究中。

3.4 乳杆菌是一种典型的难培养啤酒易感微生物，其代谢产物能够影响啤酒的风味和口感，并使得其外观品质发生显著变化。作为啤酒易感微生物，耐酸乳杆菌的最大特点为首次检测时间长。在常规乳酸菌检测培养基 MRS 上生长非常缓慢，需 14 天以上才可以培养出肉眼可见的菌落。经研究发现[1]，该菌在外界环境压力下为保证自身存活，容易进入 VBNC（Viable But Non-Culturable）状态。1982 年 Xu 等[13]首次提出了细菌 VBNC 状态，指可培养的不产孢子的细菌在环境压力下，在培养基上形成的菌落数不断下降直至完全消失，但此期间细胞总数基本恒定，这些细胞仍然具有代谢活性，若条件适宜后能够再次回到可培养状态。常规检测方法很难检测到处于 VBNC 状态的细菌，这大大增加了人们的感染几率。研究其转录组对揭示乳杆菌 VBNC 状态调控机制，提升啤酒质量尤为重要

3.5 在转录组研究中，通常采用 De novo RNA-Seq 的策略，本文中利用 De novo 基因组测序技术结合比较 RNA-Seq 的策略对一株提取自啤酒中乳杆菌的两种不同状态（正常状态与 VBNC 状态）进行了转录组研究。利用 Illumina/Solexa 高通量测序技术，初步确定了与 VBNC 状态调控相关的基因，为下一步的转录物功能注释及分类奠定了基础。本研究所应用的基因组 De novo 组装结合比较 RNA-Seq 策略，所获得结果可信度高（转录组与基因组高度符合，符合率高于 95%），与常规 De novo RNA-Seq 技术相比具有较大优势，为今后食品微生物在不同生理状态中分子调控机制的研究提供重要基础。该策略一方面能获得乳杆菌的全基因组信息，另一方面在比较 RNA-Seq 中省却 De novo 拼接与组装的繁琐，尤其是在所需要对比的生理状态较多的情况下，能省却每次 RNA-Seq 后的 De novo 拼接与组装，优势更为明显。

参考文献

[1] 刘红亮,郑丽明,刘青青,等.非模式生物转录组研究[J].遗传,2013,35(8):955-970

[2] 黄勇.基于高通量测序的微生物基因组学研究[D].北京:中国人民解放军军事医学科学院,2013
HUANG Yong. The research of microbial genomics based on high-throughput sequencing [D]. Beijing: Academy of military medical sciences, 2013.

[3] Elaine R. Mardis. The impact of next-generation sequencing technology on genetics [J]. Trends in genetics, 2008, 24(3): 133-141

[4] 刘振波. DNA 测序技术比较[J].生物学通报,2012,47(7): 14-17
LIU Zhen-bo. The comparison of DNA sequencing technologies [J]. Bulletin of biology, 2012, 47(7): 14-17

[5] Asha Boodhun, Joe S. Brennan, John A. Bridgham, et al. Accurate whole human genome sequencing using reversible terminator chemistry [J]. Nature, 2008, 456(6): 53-59

[6] H P J Buermans, J T den Dunnen. Next generation

sequencing technology: Advances and applications [J]. Biochimica et Biophysica acta, 2014, 1842: 1932-1941

[7] Jay Shendure, Hanlee Ji. Next-generation DNA sequencing [J]. Nature biotechnology, 2008, 26(10): 1135-1145

[8] Alberto Valdes, Clara Ibanez, Carolina Simo, et al. Recent transcriptomics advances and emerging applications in food science [J]. Trends in Analytical Chemistry, 2013, 52: 142-154

[9] 魏利斌,苗红梅,张海洋.芝麻发育转录组分析[J].中国农业科学,2012,45(7):1246-1256
WEI Li-bin, MIAO Hong-mei, ZHANG Hai-yang. Transcriptomic analysis of sesame development [J]. Scientia Agricultura Sinica, 2012, 45(7): 1246-1256

[10] 朱大强,李存,陈斌等.四种常用高通量测序拼接软件的应用比较[J].生物信息学,2011,9(2):106-112
ZHU Da-qiang, LI Cun, CHEN Bin, et al. Comparison of the widely used high-throughput sequencing assembly software [J]. China Journal of Bioinformatics, 2011, 9(2): 106-112

[11] Eleonora de Klerk, Peter A C t Hoen. Alternative mRNA transcription, processing and translation: insights from RNA sequencing [J]. Trends in Genetics. 2015, 31(3): 128-139

[12] 刘朋虎,林冬梅,林占熇,等.DNA 测序技术及其应用研究进展[J].福建农业学报,2012,27(10):1130-1133

[13] Deng Y, Liu J, Li H, et al. An improved plate culture procedure for the rapid detection of beer-spoilage lactic acid bacteria [J]. Journal of the Institute of Brewing, 2014, 120(2):127-132

现代食品科技