

水牛泌乳期与非泌乳期乳腺组织 miRNAs 表达谱鉴定及差异表达分析

蔡小艳^{1,2}, 鲍正攀¹, 古景开¹, 邓凯¹, 张晓溪¹, 任艳萍¹, 沈朋雷¹, 石德顺^{1*}, 刘庆友^{1*}

(1. 广西大学 亚热带农业生物资源保护与利用国家重点实验室, 南宁 530004;

2. 广西壮族自治区畜牧研究所, 南宁 530001)

摘要: 旨在鉴定分析水牛泌乳期和非泌乳期 miRNA 的表达谱和差异作用机制。采集水牛泌乳期和非泌乳期乳腺组织, 利用 Solexa 测序技术构建 2 个时期 miRNAs 表达谱, 鉴定前体 miRNAs、成熟 miRNAs 和新 miRNAs, 分析 miRNAs 的第一偏好性核苷酸和染色体分布, 发掘水牛泌乳期和非泌乳期高表达及差异表达 miRNAs。结果显示, 本研究构建了水牛泌乳期和非泌乳期乳腺组织 2 个 miRNA 表达谱, 分别获得 12 768 110 和 12 569 467 条 18~31 nt 的高质量序列。测序确认了归属 259 个 miRNAs 家族的 359 个成熟 miRNAs 和 363 个 pre-miRNAs 及 5 个水牛特有的 miRNAs。U 是 19 和 25 nt miRNAs 的 5' 端最普遍的核苷酸。bbu-let-7b、bbu-let-7a、miR-26a 和 miR-21 等 miRNAs 在 2 个时期均呈现高表达; bbu-miR-148a、143、200c、200a 和 bbu-let-7f 等 miRNAs 在非泌乳期特异性高表达。bbu-miR-125b、29a 和 bbu-let-7c 等 miRNAs 在泌乳期特异性高表达。bbu-miR-148a、143、200a、141 和 30a-5p 等 miRNAs 在泌乳期的表达量下降为非泌乳期的 1/2 以下。而 bbu-miR-26a、29a、125b、99a 和 bbu-let-7c 等 miRNAs 在泌乳期表达量大于或等于 2 倍非泌乳期丰度。本研究构建了水牛泌乳期以及非泌乳期乳腺组织 2 个 miRNAs 表达谱, 测序确认了 259 个水牛 miRNAs 家族的 359 个成熟 miRNAs 和 363 个前体 miRNAs, 5 个水牛特有 miRNAs 和 10 个高差异表达 miRNAs, 为进一步阐明奶水牛泌乳关键 miRNAs 作用机制奠定了基础。

关键词: 水牛; 泌乳期和非泌乳期; 乳腺组织; miRNA 表达谱; 信号通路

中图分类号: S823.83

文献标志码: A

文章编号: 0366-6964(2017)09-1635-13

Identify and Analyze the Expression Profile and Mechanism of Buffalo Mammary Gland miRNAs in the Lactation and Non-lactation Periods

CAI Xiao-yan^{1,2}, BAO Zheng-pan¹, GU Jing-kai¹, DENG Kai¹, ZHANG Xiao-xi¹, REN Yan-ping¹, SHEN Peng-lei¹, SHI De-shun^{1*}, LIU Qing-you^{1*}

(1. State Key Laboratory of Subtropical Bioresource Conservation and Utilization, Guangxi University, Nanning 530004, China;

2. Guangxi Institute of Animal Science, Nanning 530001, China)

Abstract: This study aimed to identify and analyze the expression patterns and mechanism of buffalo mammary gland miRNAs in the lactation and non-lactation periods. Buffalo mammary gland tissues in lactation and non-lactation periods were collected, their miRNAs expression profiles were analyzed by Solexa sequencing technology, pre-miRNAs, mature miRNAs and novel miRNAs were identified, respectively, the first preferred nucleotide and chromosome distribution were analyzed. The highly and differentially expressed miRNAs in the lactation and non-lactation periods were detected. The results showed that 2 miRNAs expression profiles from buffalo lacta-

收稿日期: 2017-01-09

基金项目: 国家 863 重点项目课题(2013AA102504); 广西科技厅项目(桂科合 14125008-2-13; 桂科 AD17129043; 桂科 AA16380026)

作者简介: 蔡小艳(1982-), 女, 山西兴县人, 博士生, 主要从事牧草及草食动物遗传育种与推广应用, E-mail: caixiaoyan282@163.com

* 通信作者: 石德顺, 博士, 研究员, E-mail: ardsshi@gxu.edu.cn; 刘庆友, 博士, 研究员, 主要从事动物生殖生理研究, E-mail: qyliu2002@126.com

tion and non-lactation mammary glands were constructed. 12 768 110 and 12 569 467 high-quality reads between 18 and 31 nt were obtained separately. Three hundred fifty-nine mature miRNAs and 363 pre-miRNAs belonged to 259 miRNAs families, and 5 buffalo-specific miRNAs were confirmed by sequencing. U was the most common nucleotide at the 5' end of 19 and 25 nt miRNAs. The *bbu-let-7b*, *bbu-let-7a*, *miR-26a* and *miR-21* showed high expression in both periods. *bbu-miR-148a*, *143*, *200c*, *200a* and *bbu-let-7f* were specifically highly expressed in non-lactation period, the *bbu-miR-125b*, *29a* and *bbu-let-7c* were specifically highly expressed in lactation period. The *bbu-miR-148a*, *143*, *200a*, *141* and *30a-5p* decreased the expression in the lactation period to less than 1/2 of the non-lactation period, *bbu-miR-26a*, *29a*, *125b*, *99a* and *bbu-let-7c* were highly expressed in the lactation period more than or equal to 2 times of the non-lactation period. The result indicate that 2 miRNAs profiles from buffalo lactation and non-lactation mammary gland tissues were constructed. 359 mature miRNAs and 363 pre-miRNAs belonging to 259 buffalo miRNAs families were identified. 5 buffalo genome-specific miRNAs and 10 highly differentially expressed miRNAs were obtained, which lay the foundation for further elucidating the mechanism of key miRNAs in lactating buffalo.

Key words: buffalo; lactation and non-lactation periods; mammary gland tissue; miRNA expression pattern; signaling pathway

水牛作为发展中国家重要的乳、肉、役兼用畜种,被国际粮农组织(FAO)认为是最具开发潜力和开发价值的家畜^[1]。水牛奶因其脂肪、乳糖、蛋白质和灰分的含量比普通牛奶高而备受到人们的青睐^[2],并享有“奶中之王”的美称^[3]。水牛奶的钙、磷、铁和维生素含量均高于荷斯坦奶牛,其他微量元素也比较均衡。水牛奶的脂肪球和酪蛋白胶粒都比较大,加工干酪优势明显^[4]。但因产奶量和繁殖率低,制约了其奶业发展的步伐^[5]。

microRNAs(miRNAs)是在动物和植物中发现的一类通过转录后抑制基因表达的内源性非编码RNA^[6-8]。越来越多的证据表明,miRNA具有控制基因表达^[9],细胞分化、增殖和凋亡^[10]等功能,可能对基因功能研究和疾病防治探索有重要意义。近些年,人们在各种动物乳汁中发现了miRNA,由此扩展了人们对于miRNA调控功能范围的理解,表明miRNA有可能通过哺乳从母体传递给新生儿,进而调控新生儿发育过程中的某些关键生理生化途径,有助于新生个体的正常发育^[11-12]。众所周知,乳腺在发育和泌乳过程中受到激素、生长因子和一些蛋白质严格且复杂的调控。随着生物信息学的发展,miRNA以其独特的调控模式和表达特性迅速成为生物学领域的研究热点^[13]。但纵观这几年miRNA的研究,其主要集中于乳腺癌^[14-15],而关于正常生理状态的乳腺发育和泌乳研究较少^[16]。

miRNA由于其自身序列短、同源性高,利用传统的克隆法、基因芯片技术等检测miRNA非常困难。随着测序技术的发展,高通量测序技术相继诞生并逐渐成熟^[17]。该技术具有高通量、成本低、灵敏度高等特点,可以获得低丰度的表达基因,不局限于已知的基因组序列信息,适用于未知基因组序列的物种,不需要克隆,操作简单,应用领域广^[18]。目前,高通量测序技术已成功运用于人^[19]、奶牛、奶山羊等生物miRNA的研究,发现奶牛中有885条pre-miRNAs编码921条成熟体miRNAs,这些成熟体中有884条是唯一序列,并且成熟体中有544条在2个时期都表达。这些唯一序列中有283条已知miRNAs,96条是与其他物种同源的miRNAs,还有505条新发现的miRNAs^[20]。利用Solex测序技术检测出山羊泌乳中期乳腺上皮细胞中miRNA的读数约占小RNAs的47.7%,共鉴定出1143个miRNAs,山羊和牛特有的miRNAs为116个,并发现17个差异表达miRNAs的靶基因主要参与乳腺发育、细胞增殖和乳脂合成等重要代谢过程^[21]。但是水牛乳腺组织miRNA测序与表达谱的相关研究较少。

本试验通过对本地水牛泌乳期和非泌乳期乳腺组织miRNA进行Solexa测序,探究乳腺特异性表达miRNA,并对其进行生物信息学分析,为研究水牛乳腺发育和泌乳提供理论参考。

1 材料与方法

1.1 试验材料

从广西大学牛场随机选取 2 头健康经产水牛, 其中 1 头 8 岁, 泌乳 3 个月作为泌乳期样品来源 (Lactation), 另 1 头 10 岁, 非泌乳未妊娠时期作为非泌乳期样品来源 (Non-lactation)。分别从 2 头水牛乳腺随机采集乳腺组织样品, 迅速放入无 RNase 的冷存管中, 放入液氮带回实验室转入 $-80\text{ }^{\circ}\text{C}$ 超低温冰箱中保存, 用于 Solexa 高通量测序。

1.2 主要试剂、仪器和软件

1.2.1 主要试剂 按照 J. Sambrook 等^[22]的方法配制常规试剂, 以重蒸水为溶剂, 15bf/in2 ($1.034 \times 10^5\text{ Pa}$) 为高压灭菌条件, 蒸汽灭菌 20 min。

内切酶和 T4 连接酶 (TaKaRa); TRIzol、脂质体 Lipofectamine TM2000 (Invitrogen 公司); AMV 逆转录酶、pMD-18T 和荧光定量用 SYBR[®] Premix Ex Taq TM (TaKaRa 公司); RIPA 裂解液; 无特别说明的其他生化试剂均为国产分析纯。

1.2.2 使用仪器 Millipore 超纯水器、生化培养箱、离心机、定量 PCR 仪 (ABI 7500)、Eppendorf 移液器、EC250-90 水平电泳仪、Gel Doc 2000 凝胶成像系统、PCR 仪 (AB 2720)、Boxun 5W-CJ-2F 超净工作台、BIO-RAD 凝胶成像系统。

1.2.3 主要使用软件 使用到的相关分子生物学软件有 DNAMAN、Vector NTI、OLIGO6.0、BioEdit 等。

1.3 试验方法

1.3.1 组织 RNA 的提取 将所采集的组织在液氮中磨碎, 每 50~100 mg 组织加入 1 mL TRIzol, 用匀浆仪进行匀浆处理。样品体积不超过 TRIzol 体积的 10%, 将匀浆样品在室温 ($15\sim 30\text{ }^{\circ}\text{C}$) 放置 5 min, 使核酸蛋白复合物完全分离, 从上清液中提取总 RNA, RNA 的数量和浓度用 NanoDrop 分光光度计进行确认, 只有 RNA 完整指数 ≥ 7 时方可被用来进一步测序和 miRNA 的分离鉴定, 只有完整指数 ≥ 8 时 RNA 方可被用做 mRNA 测序和分析。

1.3.2 Solexa 测序 Solexa 高通量测序属于新一代测序方法 (二代), 主要原理是合成和测序同时进行 (Sequencing by synthesis)^[23]。

Solexa 测序流程: 经过屏蔽适配器过滤被 rRNA、tRNA、mRNA、snRNA、snoRNA 污染的序列, 评估序列质量和计算小 RNA 序列的长度分布,

干净的序列用计算机分析处理, 最后, 干净序列在 miRBase 的 version 17.0 中进行比对。

Solexa 序列用 SOAP 放在牛基因库中 (Btau 4.0)^[23], 完美匹配的序列被保留作为下一步分析对象。为了进一步分析包含完美匹配 Solexa 序列的 RNA 二级结构, 提取这些序列每边的基因组侧翼 100 个核苷酸, 用 RNAfold 对二级结构进行预测, 用默认设置的 MIREAP 计算机工具进行分析, MIREAP 是特别设计用于鉴别小 RNA 文库中的 miRNA 的, 能够完整分析生物种源、测序深度和 miRNAs 的结构特征, 以此提高 miRNA 鉴定的敏感性和特异性。只有完全满足以下条件的茎环发夹结构才被认为是 miRNA: (1) 成熟 miRNA 必须是在发夹前的一个臂上, 缺乏大量内环或凸起; (2) 发夹的二级结构必须稳定, 混合淡化技术自由能低于 $75.35\text{ kJ}\cdot\text{mol}^{-1}$; (3) 发夹位于内含子的基因间区。最后, 在 miRBase database 17.0 中用 BLAST 查询候选 miRNA 序列, 容许 2 个错配。

1.3.3 miRNAs 靶基因的预测 在 Ensembl genome browser 中使用查询功能将每个确认的 miRNA 定位于基因组上 (http://www.ensembl.org/Bos_taurus)。靶基因的预测用 MIREAP 完成。MIREAP 的基本预测规则: (1) 在 sRNA 和靶基因间 (G-U 计算为 0.5 错配) 不大于 4 个错配; (2) 在 miRNA/target duplex 不大于 2 个邻近错配; (3) 在 miRNA/target duplex 的 2~12 位置上没有邻近错配; (4) 在 miRNA/target duplex 的 10~11 位置上没有错配; (5) 在 miRNA/target duplex 的 1~12 位置上 ≤ 2.5 个错配; (6) miRNA/target duplex 的最小自由能应该 \geq 完全互补 miRNA 的最小自由能的 75%。执行 Blastx 查询, 用靶基因序列和 EMBL 数据库预测可能靶基因的功能。

1.3.4 miRNAs 的差异表达分析 为了比较泌乳期和非泌乳期 2 个文库 miRNA 的表达差异, 首先, miRNAs 在 2 类样品中的表达被标准化到每百万转录本中的表达量。如果一个 miRNA 的标准化表达 (NE) 值标准化为 0, 那么它的校正表达值为 0.01; 如果一个 miRNA 的 NE 值在 2 个文库中都小于 1, 用公式计算标准化表达的倍数和 P-value。

标准化表达 (Normalized expression) = (实际 miRNA 测序序列读数 / 总干净序列读数) $\times 1\ 000\ 000$;

差异倍数 (Fold change) = \log_2 (泌乳期标准化表达值 / 非泌乳期标准化表达值);

P-value:

$$p(x | y) = \left(\frac{L}{N}\right) \frac{(x+y)!}{x!y! \left(1 + \frac{L}{N}\right)^{(x+y+1)}}$$

$$C(y \leq y_{\min} | x) = \sum_{y=0}^{y \leq y_{\min}} p(y | x)$$

$$D(y \geq y_{\max} | x) = \sum_{y \geq y_{\max}}^{\infty} p(y | x)$$

N 和 x 分别代表非泌乳期 sRNA 文库的一个给定 miRNA 的干净序列总读数 and 标准化表达水平, L 和 y 分别代表泌乳期 sRNA 文库的一个给定 miRNA 的干净序列总读数和标准化表达水平。

1.3.5 miRNAs 表达簇的分级聚类 Permut-Matrix 软件的 Pearson distance^[24] 用来执行 miRNA 表

达簇的分析, 可将每个 miRNA 在这个库中的相对克隆频率计算出来, 作为序列读数的总量分流一个特定库的每个 miRNA 的序列数量。

2 结果

2.1 水牛泌乳期和非泌乳期 miRNA 表达谱分析

为了鉴定水牛泌乳期和非泌乳期差异表达的 miRNAs, 成功构建了 2 个用于 Solexa 高通量测序的 sRNA 文库, 泌乳期和非泌乳期乳腺组织文库分别测序获得了 13 667 160 和 13 150 323 条 reads。将低质量的序列移除, 少量标记和测序, 最终获得非泌乳期和泌乳期 18~31 nt 的 12 569 467 和 12 768 110 条高质量 reads(表 1)。

表 1 两个文库序列的表达谱

Table 1 Expression profile of sequenced reads in 2 libraries

时期 Period	总读数 Total reads	不同序列 Distinct reads	百分数/% Percent	高质量序列 High-quality reads	百分数/% Percent
非泌乳期 Non-lactation	总读数 Total reads	840 949	100	12 569 467	100
	完美匹配至基因组读数 Reads perfectly matched genome	275 964	32.82	8 456 705	67.28
泌乳期 Lactation	总读数 Total reads	922 005	100	12 768 110	100
	完美匹配至基因组读数 Reads perfectly matched genome	254 130	27.56	7 665 749	60.04

用 SOAP2.20 aligner: SOAP, huge -M0 -v0 -r2 与牛基因组比对

Aligned with *Bos taurus* genome using SOAP2.20 aligner: SOAP, huge-M 0-v 0-r 2

所有水牛基因组的测序读数长序列用 SOAP 程序与牛基因组 (Btau_4.0) 进行比对^[25], 结果表明, 非泌乳期文库 12 569 467 条高质量序列中有 8 456 705 条序列与牛基因组匹配, 匹配率达 67.28%。泌乳期文库 12 768 110 条序列中有 7 665 749 条与牛基因组匹配 (表 2), 匹配率为 60.04%。统计显示, 水牛泌乳期和非泌乳期的 sRNA 分布的宽度模式介于 18~31 nt 之间, 在 22 nt 达到一个高数值。在泌乳期乳腺组织中, 22 nt 序列占总 sRNA 数量的 33.4% (图 1)。

2.2 水牛乳腺组织新 miRNA 的鉴定

为了鉴定这些 sRNA 是否是真正的 miRNA, 用 MIREAP 软件查阅牛基因组包括候选 miRNA 的发夹结构(水牛的基因组还未完全测序获得), 利用 Solexa 深度测序从 sRNA 文库来区分已知和未知 miRNAs。在 miRbase17.0 库中全部成熟

microRNAs 和前体 microRNAs 分别是 676 和 662 个, 本次水牛乳腺组织测序确认的成熟 microRNAs 和前体 microRNAs 分别有 359 和 363 个; 数据库中共有 500 个 microRNA 基因家族, 本次水牛乳腺组织测序获得的 miRNAs 归属 259 个 miRNAs 家族 (表 2)。

经过移除与蛋白质外显子编码区重叠且自由能小于 -75.35 kJ · mol⁻¹ 的 miRNAs, 剩余的有新发夹结构的 262 位点被认为是 miRNA 的候选基因, 其中 230 个 miRNAs 是水牛乳腺组织中新发现的 miRNAs。在新发现的候选 miRNAs 中, 35 个新候选 miRNAs 中可以鉴定出反义 miRNAs, 为它们作为候选 miRNA 提供了新的证据(miRNA*)。5 个 miRNAs 被认为是水牛特有的, 因为在 miRNA-Abase17.0 中没有其同源序列。

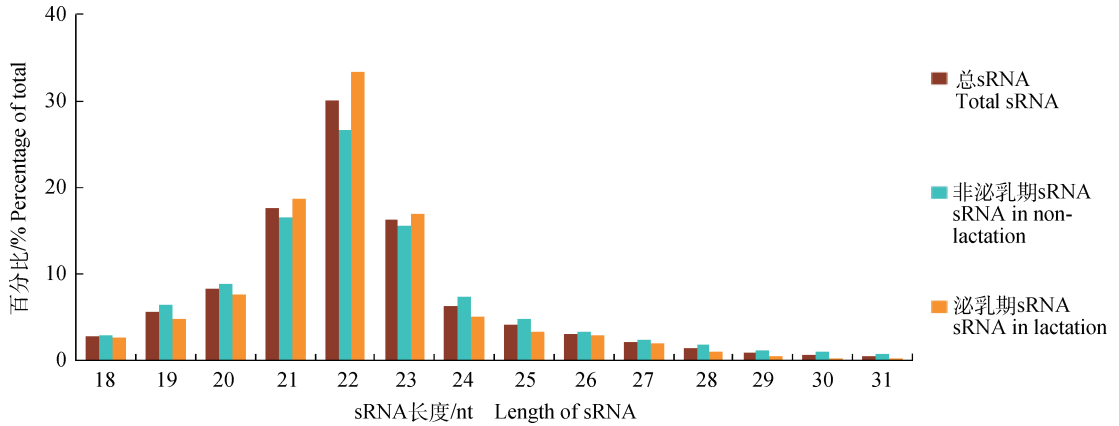


图 1 总 sRNAs 和 2 个文库序列长度的分布和频度百分数

Fig. 1 Length distribution and frequency percent of sequences in total sRNAs and the 2 libraries

表 2 已知 miRNA 检测分析

Table 2 Known microRNA detection

分类 Classification	Bta (Bbu)
全部成熟 microRNA Total mature microRNA	676
测序确定的成熟 microRNA Confirmed mature microRNA by sequencing	359
miRNAs 家族 miRNAs families	259
全部前体 microRNA Total pre-microRNA	662
测序确定的前体 microRNA Confirmed pre-microRNA by sequencing	363
miRNAs 家族 miRNAs family	259

在 mirBase17.0 中有 500 个牛 microRNA 家族

There are 500 bta microRNA families in mirBase 17.0

2.3 候选 miRNAs 的第一个核苷酸偏好性分析

对所有鉴定 miRNAs 的第一个核苷酸偏好性分析显示(图 2), U 是 19 和 25 nt miRNAs 5'端最普遍的第一核苷酸(94.15%和 97.90%)。而 A 是泌乳期乳腺组织 20、23 和 24 nt 新 miRNAs 5'端最普遍的第一核苷酸, 分别为 96.71%、50.72%和 62.51%。在非泌乳期乳腺组织中, U 是 19 (94.80%)、25(93.35%)和 28 nt (100%)新 miRNAs 5'端的最普遍第一核苷酸, 而 A 是 23 (94.59%) 和 24 nt (97.62%)新 miRNAs 5'端的最普遍第一核苷酸。

2.4 miRNA 在染色体上的分布

本研究分析了已知 miRNA 的染色体位点 (Btau 4.0 为参考基因组序列), 通过 BLAST 分析寻找水牛已知和新的前体 miRNAs, 结果见表 3。发现, 68.77%的已知 miRNA 和 84.69%的新 miRNA 位于常染色体上, 成功定位于乳腺组织染色体上的基因间隔区。已知 miRNA 主要分布在 21 号和

X 染色体, 分别为 74 和 37 个, 占总数量(635)的 11.65%和 5.83%, X 染色体共有 37 个已知 miRNAs (表 3)。

表 4 显示, miRNA 主要分布在常染色体和 X 染色体上, 总 miRNA 密度为 0.09~1.05 个·Mb⁻¹不等。21 号染色体是 miRNAs 主要表达的染色体, 分布最多, 总 miRNAs、泌乳期表达 miRNAs 和非泌乳期 miRNAs 分别为 1.05、1.34 和 1.35 个·Mb⁻¹。泌乳期和非泌乳期 miRNAs 在相同染色体上的分布密度基本相同。

2.5 水牛泌乳期与非泌乳期 miRNAs 的差异表达

Solexa 测序方法通过测定序列频度提供了一个检测 miRNAs 表达谱的重要方法。2 个文库中 miRNAs 的序列频度见图 3, 基于 miRNAs 相对克隆频度的分层表明, 水牛泌乳时 miRNA 表达是不同的。

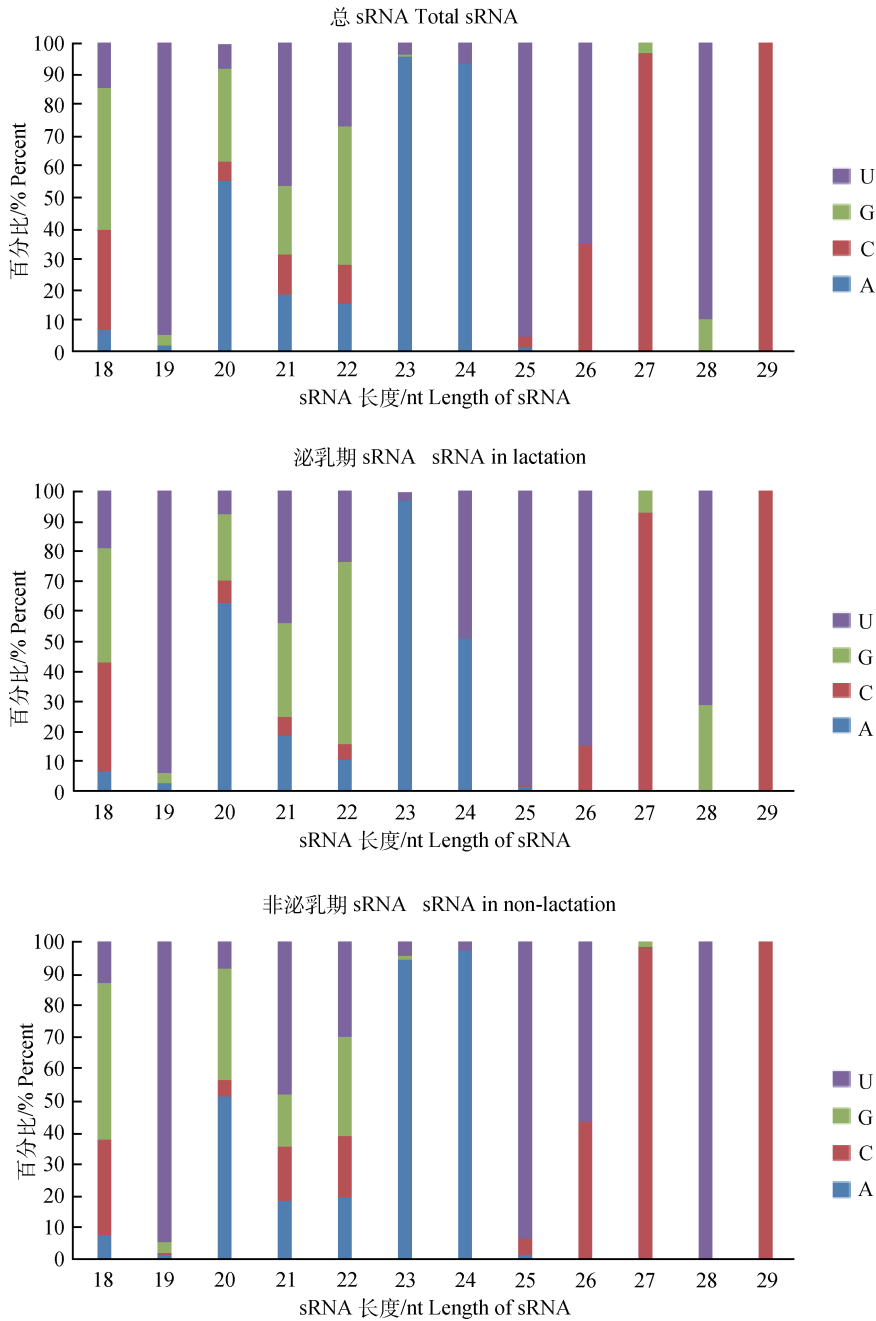


图 2 候选 18~29 nt miRNAs 的第一核苷酸偏好性
 Fig. 2 The first nucleotide bias of the 18-29 nt candidate miRNAs

在非泌乳期, *bbu-miR-148a*, *bbu-let-7b*, *bbu-let-7a*, *bbu-miR-21*, *bbu-miR-143*, *bbu-miR-200c*, *bbu-miR-26a*, *bbu-miR-200a* 和 *bbu-let-7f* 共 9 个 miRNAs 是主要表达的 miRNAs, 大于 20 000 个序列, 组成了总已知 miRNA 序列的 53.8%, 表明它们在非泌乳期组织中高丰度表达。57 个 miRNAs 序列频度(比如 *bbu-miR-92b*, *bbu-miR-421* 和 *bbu-miR-129* 等)较低($1 \leq \text{序列数} \leq 10$)。

在泌乳期, 7 个 miRNAs (*bbu-let-7b*, *bbu-let-7a*, *bbu-miR-26a*, *bbu-miR-125b*, *bbu-miR-21*, *bbu-miR-29a* 和 *bbu-let-7c*) 每个有 20 000 多条序列表达, 丰度较高。可见 *bbu-let-7b*, *bbu-let-7a*, *miR-26a* 和 *miR-21* 这 4 个 miRNAs 在 2 个时期均呈现高表达。有 124 个 miRNAs(比如 *bbu-miR-34c*, *bbu-miR-2284c* 和 *bbu-miR-383*) 显示出较低的序列频度(< 10 reads)。比较 2 个文库, *bbu-miR-148a*,

表 3 已知和新 miRNA 在染色体上的分布

Table 3 Distribution of known and new miRNAs in chromosomes

染色体 Chromosome	染色体 大小/Mb Size of the chromosome	已知 miRNAs 数量/个 Number of known miRNAs	泌乳期已知 miRNAs 数量/个 Number of known miRNAs in lactation	非泌乳期已知 miRNAs 数量/个 Number of known miRNAs in non-lactation	泌乳期新 miRNA 数量/个 Number of novel miRNAs in lactation	非泌乳期新 miRNAs 数量/个 Number of novel miRNAs in non-lactation
chr1	164	18	12	11	49	49
chr2	143	16	7	7	34	34
chr3	130	21	11	10	33	33
chr4	127	28	17	15	38	38
chr5	128	44	27	26	49	50
chr6	125	20	5	4	20	22
chr7	114	33	21	21	34	34
chr8	119	28	20	20	23	23
chr9	110	11	5	5	31	31
chr10	108	10	5	5	39	39
chr11	112	16	12	11	31	31
chr12	87	18	11	11	26	27
chr13	86	13	9	8	31	32
chr14	83	12	7	6	26	26
chr15	86	24	17	15	23	23
chr16	79	20	15	15	23	24
chr17	78	8	3	3	26	26
chr18	67	24	9	10	17	17
chr19	66	47	31	31	24	24
chr20	77	12	5	6	14	14
chr21	70	74	56	57	38	38
chr22	63	21	13	13	16	16
chr23	54	10	5	4	21	21
chr24	66	7	4	4	16	16
chr25	45	24	10	10	5	5
chr26	53	14	5	5	19	19
chr27	50	7	2	2	18	18
chr28	47	6	3	2	13	14
chr29	53	12	5	5	21	21
Un					105	103
chrX	90	37	32	32	32	32
Total	2 678	635	384	374	895	900

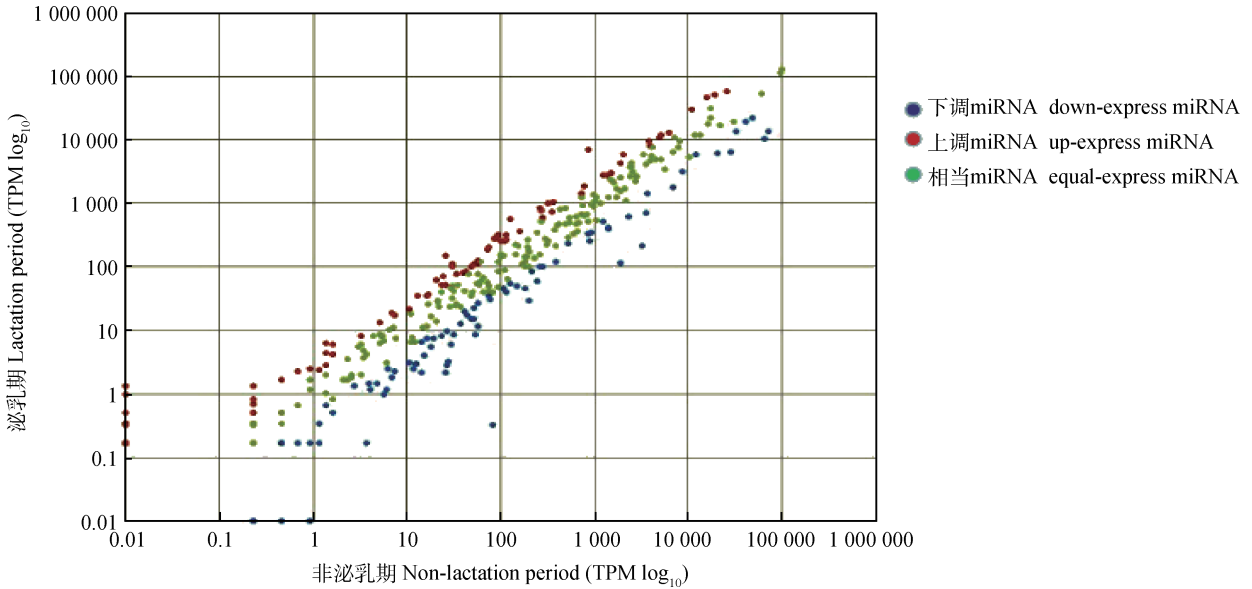
Un. 未知染色体, 表 4 同

Un. Unknown chromosome, the same as Table 4

表 4 染色体上 miRNA 的表达和密度分析

Table 4 Expression number and density of miRNAs in chromosome

染色体 Chromosome	泌乳期表达 miRNAs 数量/个 Number of expressed miRNAs in lactation period	非泌乳期表达 miRNAs 数量/个 Number of expressed miRNAs in non-lactation period	总 miRNAs 密度(单条 染色体每 Mb 的 前体 miRNAs 数量)/ (个·Mb ⁻¹) Density for the total miRNAs (number of pre-miRNAs per Mb on each chromosome)	泌乳期表达 miRNAs 密度/(个·Mb ⁻¹) Density for the expressed miRNAs in lactation period	非泌乳期表达 miRNAs 密度/ (个·Mb ⁻¹) Density for the expressed miRNAs in non-lactation period
chr1	61	60	0.11	0.37	0.37
chr2	41	41	0.11	0.29	0.29
chr3	44	43	0.16	0.34	0.33
chr4	55	53	0.22	0.43	0.42
chr5	76	76	0.34	0.59	0.59
chr6	25	26	0.16	0.20	0.21
chr7	55	55	0.29	0.48	0.48
chr8	43	43	0.24	0.36	0.36
chr9	36	36	0.10	0.33	0.33
chr10	44	44	0.09	0.41	0.41
chr11	43	42	0.14	0.38	0.37
chr12	37	38	0.21	0.43	0.44
chr13	40	40	0.15	0.47	0.47
chr14	33	32	0.15	0.40	0.39
chr15	40	38	0.28	0.46	0.44
chr16	38	39	0.25	0.48	0.49
chr17	29	29	0.10	0.37	0.37
chr18	26	27	0.36	0.39	0.40
chr19	55	55	0.71	0.83	0.83
chr20	19	20	0.16	0.25	0.26
chr21	94	95	1.05	1.34	1.35
chr22	29	29	0.33	0.46	0.46
chr23	26	25	0.18	0.48	0.46
chr24	20	20	0.11	0.30	0.30
chr25	15	15	0.54	0.33	0.33
chr26	24	24	0.27	0.46	0.46
chr27	20	20	0.14	0.40	0.40
chr28	16	16	0.13	0.34	0.34
chr29	26	26	0.23	0.49	0.49
Un	105	103			
chrX	64	64	0.41	0.71	0.71
Total	1 279	1 274	0.24	0.48	0.48



X 和 Y 轴分别表示泌乳期和非泌乳期乳腺组织中 miRNA 的表达量。红点代表 $\text{ratio} > 2$ 的 miRNA, 绿点代表 $1/2 < \text{ratio} < 2$ 的 miRNA, 蓝点代表 $\text{ratio} < 1/2$ 的 miRNA。比值 = 处理的标准化表达/对照的标准化表达

The expression level of miRNAs in non-lactation (X-axis) and lactation (Y-axis) periods in mammary gland tissues. Red symbols represent miRNAs with $\text{ratio} > 2$; Green symbols represent miRNAs with $1/2 < \text{ratio} < 2$; Blue symbols represent miRNAs with $\text{ratio} < 1/2$. Ratio = normalized expression of the treatment/normalized expression of the control

图 3 乳腺组织泌乳期和非泌乳期 miRNA 表达

Fig. 3 The expression levels of miRNAs in mammary gland in non-lactation and lactation periods

bbu-miR-143、bbu-miR-200a、bbu-miR-141 和 bbu-miR-30a-5p 等 miRNAs 在泌乳期的表达量降至非泌乳期的一半以下。另外一些 miRNAs(比如 bbu-miR-26a, bbu-miR-29a, bbu-miR-125b, bbu-let-7c 和 bbu-miR-99a)在泌乳期的表达量 ≥ 2 倍的非泌乳期表达量。在 2 个时期的乳腺组织中, 316 个已知 miRNAs 在整个乳腺组织持续表达, 在已知这些 miRNAs 中, 分别发现 19 和 24 个非泌乳期和泌乳期特异表达的 miRNAs。

2.6 靶基因和信号通路预测

使用 MIREAP 软件对已知和新发现 miRNA 的潜在靶基因进行预测。发现了已知 miRNA 有 246 786 个假定的靶基因位点, 乳腺组织中新发现的 miRNA 有 209 655 个假定靶基因位点。利用 KEGG 对 20 个差异高表达 miRNA 的预测靶基因进行功能分类(表 5)。已知 miRNA 的 109 个预测靶基因都被标记在 MAPK 信号通路, 其他重要途径有 Jak-STAT 信号通路、促性腺激素释放激素信号通路(GnRH)和胰岛素信号通路等。

3 讨论

在人类、小鼠、山羊和牛等物种中开展了很多关

于 miRNA 的研究, 但是水牛乳腺组织泌乳期和非泌乳期 miRNA 的系统 and 全面研究工作还未见报道。本研究应用 Solexa 测序技术对水牛泌乳期和非泌乳期乳腺组织进行了高通量测序, 得到 13 667 160 和 13 150 323 个序列, 而荷斯坦奶牛泌乳期和非泌乳期乳腺组织高通量测序共得到 13 711 046 和 170 680 43 个序列。长度分析结果显示, 荷斯坦奶牛小 RNA 中 52.6% 为 22 nt RNA^[20], 而水牛 22 nt 小 RNA 为 33.4%, 相差 1.5 倍左右, 可能因为水牛的基因组尚未公布, 有些序列难以比对。大部分新鉴定的水牛 miRNA 能够同源物种上的已知 miRNA, 甚至有很近的关系, 证明 miRNA 的是具有进化上的保守性^[26]。研究共发现 230 个水牛乳腺组织新 miRNA, 有 5 个在 miRNAbase 17.0 中没有同源序列, 为水牛特有。

不同阶段差异 miRNA 表达模式的分析能够为 miRNA 在乳腺的生长、发育和泌乳方面的作用提供有价值的信息。基于 Solexa 测序结果, 不同 miRNAs 家族展示出显著的差异表达水平。例如, bta-let-7 家族在水牛泌乳期乳腺中的丰度在深度测序中从 12 reads (bta-let-7a*) 到 687 63 reads (bta-let-7b) 不等, 鉴定的 bbu-miR-29 家族有 bbu-miR-29-a、

表 5 水牛乳腺中 20 个差异高表达 miRNAs 的靶基因在 KEGG 信号通路中的富集

Table 5 KEGG pathways enriched for target genes of the 20 differential highly expressed miRNAs in buffalo mammary gland

通路 Pathway	计数 Count	P 值 P-value	本杰明法 Benjamini	miRNA 计数 miRNA count
bta04142:Lysosome	70	1.78E-07	3.43E-05	16
bta04670:Leukocyte transendothelial migration	57	0.003 600 312	0.159 727 074	19
bta04620:Toll-like receptor signaling pathway	49	0.004 013 488	0.143 783 791	16
bta00230:Purine metabolism	72	0.005 367 392	0.158 960 060	18
bta00010:Glycolysis / Gluconeogenesis	31	0.006 432 692	0.144 175 523	15
bta04520:Adherens junction	36	0.008 358 428	0.164 725 718	17
bta04914:Progesterone-mediated oocyte maturation	42	0.009 103 453	0.148 244 435	17
bta04910:Insulin signaling pathway	59	0.021 130 654	0.255 038 622	18
bta04722:Neurotrophin signaling pathway	56	0.022 026 581	0.235 601 105	17
bta04912:GnRH signaling pathway	42	0.053 234 806	0.288 640 630	19
bta04010:MAPK signaling pathway	109	0.060 090 705	0.311 863 487	18
bta04144:Endocytosis	78	0.061 754 452	0.311 199 362	19
bta04530:Tight junction	55	0.064 315 186	0.314 325 632	16
bta04012:ErbB signaling pathway	37	0.065 292 678	0.303 712 062	17
bta04350:TGF-beta signaling pathway	37	0.078 227 293	0.324 990 241	16
bta04062:Chemokine signaling pathway	74	0.078 927 130	0.320 920 700	19
bta04114:Oocyte meiosis	49	0.084 366 496	0.326 724 770	17
bta04660:T cell receptor signaling pathway	47	0.096 963 400	0.360 695 783	17
bta04720:Long-term potentiation	28	0.111 005 643	0.383 178 162	18
bta04360:Axon guidance	50	0.118 831 923	0.386 341 444	16
bta04370:VEGF signaling pathway	33	0.146 771 898	0.421 343 326	16
bta04810:Regulation of actin cytoskeleton	78	0.189 884 471	0.449 908 739	18
bta04270:Vascular smooth muscle contraction	45	0.206 483 147	0.462 034 169	18
bta04510:Focal adhesion	73	0.279 741 039	0.521 173 628	19
bta04120:Ubiquitin mediated proteolysis	54	0.317 569 318	0.559 295 213	18
bta04666:Fc gamma R-mediated phagocytosis	34	0.454 503 651	0.661 438 723	17
bta04630:Jak-STAT signaling pathway	52	0.508 285 967	0.708 949 225	16
bta04115:p53 signaling pathway	24	0.518 987 827	0.710 335 150	16
bta04662:B cell receptor signaling pathway	26	0.614 536 539	0.778 670 643	15
bta04310:Wnt signaling pathway	52	0.660 938 221	0.796 779 247	18
bta04150:mTOR signaling pathway	18	0.764 323 927	0.855 878 623	14

bbu-miR-29-b、bbu-miR-29-c、bbu-miR-29-d 和 bbu-miR-29-e 等 5 个成员,但 miRNA 家族中是由其主要成员发挥调控功能的。一些在泌乳期表达量降至非泌乳期表达量一半以下的 miRNA,是在水牛开始泌乳后表达改变的,可能对泌乳发挥重要调控

作用。bbu-miR-148a、bbu-miR-143、bbu-miR-200c、bbu-miR-200a 和 bbu-let-7f 等为非泌乳期特异性高表达 miRNAs,bbu-miR-125b、bbu-miR-29a 和 bbu-let-7c 等为泌乳期特异性高表达 miRNAs,对不同阶段也许起着重大调控作用。通过比较,非

泌乳期 *bbu-miR-148a*、*bbu-miR-143*、*bbu-miR-200a* 等 miRNAs 差异高表达,泌乳期 *bbu-miR-26a*、*bbu-miR-29a*、*bbu-miR-125b*、*bbu-let-7c* 等 miRNAs 差异高表达,这些特异性 miRNAs 的实际表达量及其对水牛泌乳机制的调控作用值得进一步深入研究。

miRNA 通过调控靶基因翻译来行使功能。目前,研究 miRNA 在正常生理条件下牛乳腺发育和泌乳相关的报道较少。比如 *miR-107* 被证明能够抑制乳腺组织乳腺细胞分裂和阻滞细胞周期,与细胞脂肪代谢有关^[27]。D. P. Bartel^[26] 研究发现,当导管系统开始活跃生长时,*miR-23* 在泌乳期和非泌乳期表现出相反的表达模式。预测 *miR-29*、*miR-15b* 和 *miR-107* 的靶基因 *XDH* (Xanthine dehydrogenase) 活性的增加是在乳腺的早期发育中^[28]。*miR-142*、*miR-23*、*miR-374b*、*miR-30a* 和 *miR-27b* 等通过调节泌乳关键基因 *PRLR*, 从而维持乳腺腺泡的存活、紧密连接和控制奶蛋白和乳糖的合成^[29]。研究表明,*miR-141* 的靶基因是 *STAT5*, *STAT5* 能作为乳腺中泌乳和催乳激素信号转导通路上的节点而对泌乳产生作用^[30]。*miR-141*、*miR-484* 和 *miR-500* 通过靶向 *STAT5* 和 *HK2* 基因对调控泌乳至关重要。本研究发现,泌乳期 *bbu-miR-141* 表达量下降至非泌乳期时的一半以下,对水牛泌乳至关重要,是否也是通过 *STAT5* 和 *HK2* 基因调控泌乳来发挥作用的有待进一步研究。另外,通过对山羊泌乳期和非泌乳期乳腺组织进行高通量测序,得知 *miR-103* 泌乳期表达量为非泌乳期的 4.3 倍,并通过调控靶基因 *PANK3* 来调控奶山羊的乳汁分泌^[31]。T. Tanaka 等^[32] 采用微阵列分析乳腺不同发育阶段细胞中 miRNA 的表达差异,得出 *miR-101* 通过靶向 *cox-2* 影响乳腺发育。研究表明,敲除人正常乳腺上皮细胞的 *miR-29a* 可提高乳腺细胞的生长速度^[33],*miR-29a*、*miR-29b*、*miR-29c* 均能提高奶牛乳腺上皮细胞的活力,促进细胞增殖,并且对奶牛乳腺上皮细胞 β -酪蛋白、甘油三酯及乳糖的分泌均有促进作用^[34],与本研究测序发现水牛 *bbu-miR-29a* 在泌乳期特异性高表达结果一致。本研究中,109 个 miRNAs 都被标记在 MAPK 信号通路,这个通路几乎在乳腺发育的每一步都发挥重要作用,其他重要途径有 Jak-STAT 信号通路,其与乳腺泌乳密切相关,尤其是在催乳素 (PRL) 信号转导中^[35]。还有一条显著的高丰度 miRNA 靶向的通路是胰岛素信号通路,已知其涉及乳糖代谢和奶生产

性能^[36]。

4 结 论

本试验通过对水牛泌乳期和非泌乳期乳腺组织进行高通量测序,研究其 miRNA 的表达谱。测序获得 359 个物种间保守的 miRNAs、5 个水牛特有的 miRNAs、4 个 2 个时期均高表达的 miRNAs、9 个非泌乳时期高丰度表达的 miRNAs、7 个泌乳时期高丰度表达的 miRNAs 和 10 个差异表达的 miRNAs。miRNA 靶基因预测和功能分析得知,miRNA 参与乳腺发育与泌乳相关的细胞信号通路,为进一步研究 miRNA 在乳腺发育与泌乳中的调控作用奠定了理论基础。

参考文献 (References):

- [1] 林浪, 龚云, 王梦, 等. 水牛缝隙连接蛋白 43 基因克隆及表达[J]. 中国畜牧兽医, 2013, 40(9): 1-7.
LIN L, GONG Y, WANG M, et al. Cloning and expression of buffalo Cx43 gene[J]. *China Animal Husbandry & Veterinary Medicine*, 2013, 40(9): 1-7. (in Chinese)
- [2] AHMAD S, GAUCHER I, ROUSSEAU F, et al. Effects of acidification on physico-chemical characteristics of buffalo milk; A comparison with cow's milk[J]. *Food Chem*, 2008, 106(1): 11-17.
- [3] 唐善生. 广西奶水牛业发展现状与思考[J]. 广西农学报, 2014, 29(1): 50-52, 55.
TANG S S. Dairy buffalo industry developmental current situation and considerations in Guangxi[J]. *Journal of Guangxi Agriculture*, 2014, 29(1): 50-52, 55. (in Chinese)
- [4] 简保权, 秦学敏, 龚芳. 世界水牛奶业发展现状和典型模式分析[J]. 世界农业, 2015(3): 115-118.
JIAN B Q, QIN X M, GONG F. Buffalo dairy industry developmental current situation and the analysis of typical model in the world[J]. *World Agriculture*, 2015(3): 115-118. (in Chinese)
- [5] 王锦. 水牛乳腺基因表达谱与生长激素转基因水牛的初步研究[D]. 南宁: 广西大学, 2013.
WANG J. Preliminary studies on buffalo mammary gene expression profile and growth hormone transgenic buffalo [D]. Nanning: Guangxi University, 2013. (in Chinese)
- [6] BARTEL D P. MicroRNAs: genomics, biogenesis,

- mechanism, and function[J]. *Cell*, 2004, 116(2): 281-297.
- [7] AMBROS V. The functions of animal MicroRNAs [J]. *Nature*, 2004, 431(7006): 350-355.
- [8] KIM V N. MicroRNA biogenesis: coordinated cropping and dicing[J]. *Nat Rev Mol Cell Biol*, 2005, 6(5): 376-385.
- [9] WANG K, LONG B, JIAO J Q, et al. miR-484 regulates mitochondrial network through targeting Fis1 [J]. *Nat Commun*, 2012, 3: 781.
- [10] WIGHTMAN B, HA I, RUVKUN G. Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*[J]. *Cell*, 1993, 75(5): 855-862.
- [11] WEBER J A, BAXTER D H, ZHANG S L, et al. The microRNA spectrum in 12 body fluids[J]. *Clin Chem*, 2010, 56(11): 1733-1741.
- [12] KOSAKA N, IZUMI H, SEKINE K, et al. microRNA as a new immune-regulatory agent in breast milk [J]. *Silence*, 2010, 1(1): 7.
- [13] 张 抒, 张 强, 张 萌, 等. miRNA 对乳腺发育及泌乳调节作用的研究进展[J]. *中国畜牧兽医*, 2015, 42(3): 663-667.
- ZHANG S, ZHANG Q, ZHANG M, et al. Research progress on regulation effect of microRNA on mammary gland development and lactation[J]. *China Animal Husbandry & Veterinary Medicine*, 2015, 42(3): 663-667. (in Chinese)
- [14] DENIS H, VAN GREMBERGEN O, DELATTE B, et al. MicroRNAs regulate KDM5 histone demethylases in breast cancer cells[J]. *Mol BioSyst*, 2015, 12(2): 404-413.
- [15] PHAN B, MAJID S, URSU S, et al. Tumor suppressor role of microRNA-1296 in triple-negative breast cancer[J]. *Oncotarget*, 2016, 7(15): 19519-19530.
- [16] COCHRANE D R, SPOELSTRA N S, RICHER J K. The role of miRNAs in progesterone action[J]. *Mol Cell Endocrinol*, 2012, 357(1-2): 50-59.
- [17] 熊显荣, 兰道亮, 李 键, 等. 牦牛卵巢小 RNA 高通量测序及生物信息学分析[J]. *畜牧兽医学报*, 2016, 47(1): 55-63.
- XIONG X R, LAN D L, LI J, et al. Solexa sequencing of small RNAs in yak (*Bos grunniens*) ovaries and bioinformatics analysis[J]. *Acta Veterinaria et Zootechnica Sinica*, 2016, 47(1): 55-63. (in Chinese)
- [18] 周 华, 张 新, 刘腾云, 等. 高通量转录组测序的数据分析与基因发掘[J]. *江西科学*, 2012, 30(5): 607-611.
- ZHOU H, ZHANG X, LIU T Y, et al. Data processing and gene discovery of high-throughput transcriptome sequencing[J]. *Jiangxi Science*, 2012, 30(5): 607-611. (in Chinese)
- [19] MORIN R D, O'CONNOR M D, GRIFFITH M, et al. Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells[J]. *Genome Res*, 2008, 18(4): 610-621.
- [20] LI Z, LIU H Y, JIN X L, et al. Expression profiles of microRNAs from lactating and non-lactating bovine mammary glands and identification of miRNA related to lactation[J]. *BMC Genomics*, 2012, 13(1): 731.
- [21] 林先滋. 奶牛羊乳腺组织乳脂代谢相关 miRNAs 的筛选及功能验证[D]. 杨凌: 西北农林科技大学, 2013.
- LIN X Z. Screening and functional investigation of miRNAs regulation milk fat metabolism in dairy goat mammary gland[D]. Yangling: Northwest A & F University, 2013. (in Chinese)
- [22] SAMBROOK J, RUSSELL D W. 分子克隆实验指南[M]. 3 版. 黄培堂, 译. 北京: 科学出版社, 2008.
- SAMBROOK J, RUSSELL D W. *Molecular cloning: a laboratory manual*[M]. 3rd ed. HUANG P T, trans. Beijing: Science Press, 2008. (in Chinese)
- [23] The Bovine Genome Sequencing and Analysis Consortium, ELSIK C G, TELLAM R L, et al. The genome sequence of taurine cattle: a window to ruminant biology and evolution[J]. *Science*, 2009, 324(5926): 522-528.
- [24] CARAUX G, PINLOCHE S. PermutMatrix: a graphical environment to arrange gene expression profiles in optimal linear order[J]. *Bioinformatics*, 2005, 21(7): 1280-1281.
- [25] LI R Q, LI Y R, KRISTIANSEN K, et al. SOAP: short oligonucleotide alignment program[J]. *Bioinformatics*, 2008, 24(5): 713-714.
- [26] BARTEL D P. MicroRNAs: genomics, biogenesis, mechanism, and function[J]. *Cell*, 2004, 116(2): 281-297.
- [27] FINNERTY J R, WANG W X, HÉBERT S S. The miR-15/107 group of microRNA genes: evolutionary biology, cellular functions, and roles in human diseases[J]. *J Cytol Mol Biol*, 2010, 402(3): 491-509.
- [28] HAYDEN T J, BRENNAN D, QUIRKE K, et al. Xanthine oxidase/dehydrogenase in mammary gland of mouse: relationship to mammogenesis and lacto-

- genesis *in vivo* and *in vitro*[J]. *J Dairy Res*, 1991, 58(4): 401-409.
- [29] MCMANAMAN J L, NEVILLE M C. Mammary physiology and milk secretion[J]. *Adv Drug Deliv Rev*, 2003, 55(5): 629-641.
- [30] LIU X, ROBINSON G W, WAGNER K U, et al. Stat5a is mandatory for adult mammary gland development and lactogenesis [J]. *Genes Dev*, 1997, 11(2): 179-186.
- [31] LIN X Z, LUO J, ZHANG L P, et al. MiR-103 Controls milk fat accumulation in goat (*Capra hircus*) mammary gland during lactation[J]. *PLoS One*, 2013, 8(11): e79258.
- [32] TANAKA T, HANEDA S, IMAKAWA K, et al. A microRNA, miR-101a, controls mammary gland development by regulating cyclooxygenase-2 expression [J]. *Differentiation*, 2009, 77(2): 181-187.
- [33] WU Z H, HUANG X N, HUANG X, et al. The inhibitory role of Mir-29 in growth of breast cancer cells[J]. *J Exp Clin Cancer Res*, 2013, 32(1): 98.
- [34] 边艳杰. miR-29 家族对奶牛乳腺上皮细胞泌乳调控机制的研究[D]. 哈尔滨: 东北农业大学, 2015.
- BIAN Y J. Regulation mechanism of miR-29 family to lactation of dairy cow mammary epithelial Cells [D]. Harbin: Northeast Agricultural University, 2015. (in Chinese)
- [35] WATSON C J, BURDON T G. Prolactin signal transduction mechanisms in the mammary gland: the role of the Jak/Stat pathway[J]. *Rev Reprod*, 1996, 1(1): 1-5.
- [36] FREEMAN M E, KANYICKSKA B, LERANT A, et al. Prolactin: structure, function, and regulation of secretion[J]. *Physiol Rev*, 2000, 80(4): 1523-1631.

(编辑 郭云雁)