Circular RNA of cattle casein genes are highly expressed in bovine mammary gland

ChunLei Zhang,¹ Hui Wu, YanHong Wang, ShiQi Zhu, JunQiang Liu, XingTang Fang, and Hong Chen
Institute of Cellular and Molecular Biology, School of Life Science, Jiangsu Normal University, Xuzhou, Jiangsu 221116, China

ABSTRACT

Recent studies have revealed that, in addition to hormones and other protein factors, noncoding RNA molecules play an important regulatory role in milk protein synthesis. Circular RNAs (circRNAs) are universally expressed noncoding RNA species that have been proposed recently to regulate the expression of their parental genes. In the present study, the deep RNA-sequencing technique known as RNA-seq was used to compare expression profiles of circRNAs from 2 pooled RNA samples from cow mammary gland on d 90 and 250 postpartum and to identify the key circRNAs involved in lactation. A total of 4,804 and 4,048 circRNAs were identified in the cow mammary gland on d 90 and 250 postpartum, respectively, of which only 2,231 circRNAs were co-expressed at both lactation stages, suggesting high stage specificity in the circRNAs. The enrichment of some Gene Ontology terms for the circRNA parental genes differed between lactation stages. Among the top 10 enriched Gene Ontology terms, vesicle, endoplasmic reticulum, and mitochondrial lumen were more common on lactation d 90. All 4 casein-coding genes (*CSN1S1*, *CSN1S2*, *CSN2*, and *CSN3*) produced circRNAs in the cattle mammary gland. In total, 80 circRNAs were identified from these 4 genes; circRNAs from *CSN1S1* had very high abundance, and 3 of them accounted for 36% of all the circRNAs expressed in the mammary gland on lactation d 90. Three circRNAs from *CSN1S1*, 1 circRNA from *CSN1S2*, and 1 circRNA from *CSN2* were all more highly expressed on lactation d 90 than on lactation d 250, as confirmed by quantitative PCR. These circRNAs had several target sites for the microRNA miR-2284 family and were predicted to target *CSN1S1* and *CSN2* mRNA, suggesting their potential involvement in regulating expression of the casein genes. **Key words:** lactation, circular RNA, casein biosynthesis

INTRODUCTION

The well-balanced amino acids provided by milk proteins are essential for maintaining mammalian health and well-being, especially for the infant (Pellegriano et al., 2013). In addition, milk proteins are of high biological value to human health in that bioactive peptides derived from fermentation or digestion of milk proteins can be used to treat and prevent metabolic syndromes such as obesity, atherogenesis, arterial hypertension, and type 2 diabetes (Schaafsma, 2000; Kitts and Weiler, 2003; Hartmann and Meisel, 2007; Ricci-Cabello et al., 2012). Cattle milk proteins can be grouped into 2 categories: caseins and whey proteins. Caseins are organized in micelles and account for about 80% of milk proteins. Casein micelles have evolved to concentrate, stabilize, and transport essential nutrients, mainly calcium and protein, to the neonate (DeKruif and Holt, 2003). In effect, casein micelles, which act as a natural nano-delivery system for nutrients, offer a promising nano-delivery vehicle for drugs (Elzoghby et al., 2011). Therefore, a better understanding of the factors regulating casein synthesis should provide clues to strategies and technologies for enhancing milk production in dairy cows and facilitate the availability of high-quality milk protein to the global population for both nutritional and functional uses.

It is well established that milk protein production in the mammary gland depends on the hormonal and developmental cues that modulate the activity of specific transcription factors and change the chromatin structure in the mammary epithelial cells, but the molecular mechanisms underlying the regulation processes are still not clear (Rhoads and Grudzien-Nogalska, 2007). In addition to hormones and other protein factors, increasing numbers of noncoding RNAs (ncRNA) have been shown to have important regulatory functions in milk protein synthesis. Inhibition of microRNA (miRNA) miR-29s increased the methylation levels of the promoter of α_{S1} -casein (*CSN1S1*), and reduced the secretion of milk protein (Bian et al., 2015). Interestingly, a transgenic calf engineered to express 2 tandem miRNA displayed an absence of β -lactoglobulin and a

Received September 11, 2015.

Accepted February 20, 2016.

¹Corresponding author: clzhang@jsnu.edu.cn

concurrent increase in all casein milk proteins (Jabed et al., 2012).

Circular RNA (**circRNA**) is a type of RNA molecule that recently joined the growing list of known ncRNA types. These RNA are circularized through 5–3′ linkages as well as through 2–5′ linkages formed by branch-point nucleophilic attack during splicing (Lasda and Parker, 2014). Circular RNAs have been known about for at least 30 yr, but they were considered molecular flukes or products of aberrant RNA splicing because of their low levels of expression (Sanger et al., 1976; Cocquerelle et al., 1993). However, recent research has revealed large numbers of endogenous circRNAs that are abundant, stable, and ubiquitously expressed, and some circRNAs have shown to play special regulatory roles. Two circRNAs in mammals have been shown recently to function as miRNA sponges: CDR1as is a sponge for miR-7, whereas the circRNA generated from the *Sry* gene (circSry) is a sponge for miR-138 (Hansen et al. 2013; Memczak et al. 2013). Circular RNAs can also enhance the expression of their parental genes in the nucleus (Zhang et al., 2013; Li et al., 2015b). Expression of circRNAs is developmentally regulated and varies between tissues (Salzman et al., 2013; Fan et al., 2015; You et al., 2015). A recent study found that the majority of mouse brain circRNAs were also expressed as circRNAs in the human brain, and some of them were detected in *Drosophila* brains (Rybak-Wolf et al., 2015). Taken together, these findings indicate that the function of circRNAs is conserved among different animal species.

Considering the universal expression features of circRNAs and the key role they play in regulating gene expression, we hypothesized that these molecules would also be expressed in the mammary gland to regulate milk protein synthesis. Therefore, the objective of the present study was to identify circRNAs from the mammary gland of cows at d 90 and 250 of lactation, and to screen any candidate circRNAs for their role in regulating milk protein synthesis.

MATERIALS AND METHODS

Animals and Management

Four healthy mastitis-free Holstein cows in their third lactation at the Xuzhou Lvjian Dairy Farm (Xuzhou, China) were selected for the study. The animals were kept in freestall housing, were fed a TMR diet, and had access to water ad libitum. The cows were milked twice daily at 0500 and 1700 h in the milking parlor. The average milk yield of the cows at d 90 and 250 was 35 ± 3 kg and 15 ± 5 kg, respectively. Percutaneous

biopsies from each cow were obtained from the right or left rear quarter of the mammary gland at 90 and 250 d postpartum. The right rear quarter from 2 of the cattle and the left rear quarter from the other 2 cattle were biopsied at 90 d postparturition; the remaining unbiopsied right or left rear quarters were biopsied at 250 d postparturition. The surgical procedure followed is detailed in Bionaz and Looor (2007). The biopsy samples were frozen in liquid nitrogen for subsequent RNA extraction. All procedures were conducted under protocols approved by the Institutional Animal Care and Use Committee of Jiangsu Normal University, China.

RNA Preparation

Total RNA isolated with Trizol (Invitrogen, Carlsbad, CA), according to the manufacturer's instructions, was quantified spectrophotometrically (NanoDrop, Wilmington, DE). To remove residual genomic DNA, each RNA sample was digested with DNase I (Takara, Dalian, China) for 15 min at 37°C following the manufacturer's protocol. The DNA contamination in RNA was assessed by comparing the amplicon length of a housekeeping gene (ribosomal protein S9, *RPS9*). The 2 primers (Supplemental Table S1; <http://dx.doi.org/10.3168/jds.2015-10381>) used for amplification are located in 2 different exons of *RPS9*. The RNA quality was assessed with an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA); RNA with an RNA integrity number >6 was used for further analysis.

Library Synthesis and Illumina HiSeq 2000 Sequencing

Samples of RNA (each 5 µg) from each lactation stage were pooled. The DNA-free RNA were depleted of ribosomal RNA (rRNA) using a RiboMinus kit (A15026, Invitrogen). The rRNA-depleted RNA was used for library preparation using a NEBNext Ultra Directional RNA Library Prep Kit for Illumina (New England Biolabs, Ipswich, MA), and was then subjected to paired-end sequencing by HiSeq 2000 (Illumina, San Diego, CA) to yield 2×150 nucleotide reads.

Bioinformatics Analysis

Raw sequencing data were processed using TrimGalore Version 0.4.1 (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) and cutAdapt version 1.9 (<https://code.google.com/p/cutadapt/>) to trim poor-quality bases and remove adapter sequences. Filtered data were mapped to the bovine genome (Bos_taurus_UMD_3.1.1) using TopHat2 (Kim et al.,

2013). Transcript assembly and abundance estimation was performed using Cufflinks (Trapnell et al., 2010). RNA-sequencing read coverage was visualized for selected genes in the Integrative Genomics Viewer (Robinson et al., 2011).

A novel circRNA identification tool, CIRC, was used to identify and characterize the mammary gland circRNAs (Gao et al., 2015). Briefly, the CIRC software was run on a Linux platform using the default parameters of CIRC (Gao et al., 2015).

The DAVID tool (available from the National Institute of Allergy and Infectious Diseases, Bethesda, MD; <https://david.ncifcrf.gov>) was used for enrichment of GO terms (Huang et al., 2009) for the circRNAs that produced coding genes using the bovine genome (Bos_taurus_UMD_3.1.1) as background.

FASTA files with bovine miRNAs were obtained from miRBase release 20.0 (<http://www.mirbase.org/>). Only mature miRNAs were considered for circRNA target analysis. The miRNAs were aligned with the circRNAs. The putative target site of a miRNA comprises a 6-nucleotide-long sequence in the genome that is the reverse complement of nucleotides 2 to 7 of the mature miRNA sequence. Targetscan software (<http://www.targetscan.org/>; August 2015 release) was used to identify miRNA target genes (Agarwal et al., 2015).

Validation of circRNAs by Sanger Sequencing

An aliquot of total RNA, incubated for 15 min at 37°C with or without (mock) 3 U/μg of RNase R (Epicenter Biotechnologies, Chicago, IL), was subjected to Trizol (Invitrogen) extraction. Next, the aliquot of RNase R mock-treated RNA and the aliquot of RNase R-digested RNA were reverse transcribed by SuperScript II (Invitrogen) using random hexamers according to the manufacturer's instructions, and then PCR amplified with the outward-facing primer sets listed in Supplemental Table S1; <http://dx.doi.org/10.3168/jds.2015-10381>). These primer pairs can only amplify circular RNA molecules and cannot amplify linear RNA because the primers are divergently priming (Jeck et al., 2013). The PCR products from the individual circRNAs were subjected to Sanger sequencing.

cDNA Synthesis and Reverse Transcription Quantitative PCR

To further confirm the circular structures of the circRNA candidate molecules, quantitative (q)PCR was conducted using a SybrGreen reaction mix (Takara) and the StepOnePlus real-time PCR system (Applied Biosystems, Foster City, CA) to check the sensitivity of

specific RNA targets with potential circRNA properties to RNase R.

Complementary DNA synthesis for each total RNA sample from mammary gland was conducted using SuperScript II (Invitrogen) with random hexamers. Quantitative PCR was used to analyze the expression changes of the circular transcripts from the parental casein genes between the 2 lactation stages. The primer sets used for qPCR, which were designed according to CircleSeq (Jeck et al., 2013), are listed in Supplemental Table S1 (<http://dx.doi.org/10.3168/jds.2015-10381>). The genes *UXT*, *RPS9*, and *RPS15* were used as reference genes as recommended by Bionaz and Looor (2007), and the geometrical average of the cycle threshold (Ct) value from these 3 genes was used for normalization.

Statistical Analyses

To compute the circRNA expression, head-to-tail spliced read counts for each circRNA was normalized to sequencing depth by the total number of reads that map to all protein-coding exons. The calculated value was multiplied by one million, and changed to the expression value representing the counts of a circular RNA that appeared in one million mapped reads (TPM). The edgeR software package, which is based on a negative binomial model, was used to examine the differential expression of the circRNAs (Robinson et al., 2010). The TPM values of the transcripts were used for comparison by computing fold changes (with absolute value ≥ 4) and using Fisher's exact test ($P < 0.01$). Because differentially expressed gene analyses generate large multiplicity problems in which thousands of hypothesis are tested simultaneously, correction for false-positive and false-negative errors was performed using the false discovery rate (FDR) method (Benjamini and Yekutieli, 2001). We set " $FDR \leq 0.01$, $P \leq 0.01$ and $|\log_2 \text{fold-change}| \geq 2$ " as the threshold for significantly differential circRNA expression. The qPCR data for each differentially expressed miRNA were compared between the 2 lactation periods by Student's *t*-tests.

RESULTS

Prediction and Validation of circRNA

We obtained 99 million and 66 million clean reads from the mammary gland tissues that could be mapped to the bovine genome at lactation d 90 and 250, including 138,970 and 34,758 head-to-tail spliced reads, respectively. In total, 62,477 head-to-tail spliced reads were assigned to casein genes at lactation d 90, which accounted for ~45% of the total head-to-tail spliced

Figure 1. Validation of circular RNA (circRNA) species. Quantitative PCR validation of RNase R enrichment in 11 novel back-splice junctions of circRNA candidates. Noncircular RNA (*UXT*, *RPS9*, *RPS15*, *CSN1S1*, *CSN1S2*, and *CSN2*) were depleted by RNase R treatment. The RNA with a lowercase letter “c” before their parental gene names are circular RNA. The numbers in parentheses represent different circular RNA isomers from one parental gene. The data are presented as the mean $\pm$ standard error of the mean; ** $P < 0.01$. Mock mean RNA was only treated with buffer for RNase R and without the enzyme. Significant difference was compared between the treatments.

reads. We detected 4,048 distinct circRNAs from the d 250 lactation sample and 4,804 circRNAs from the d 90 lactation sample; 2,231 circRNAs were common to both samples.

We validated the authenticity of the identified circRNAs by 2 independent methods. First, because the circRNAs have a special head-to-tail junction site that differs from canonical splicing linear transcripts, we were able to validate 22 of 24 (92%) circRNA candidates by targeted amplification of the head-to-tail junction region unique to the circRNAs using reverse transcription-PCR. To do this we used outward-facing primer sets followed by standard Sanger sequencing (Supplemental Table S1; <http://dx.doi.org/10.3168/jds.2015-10381>).

Second, compared with linear RNAs, circRNAs are strongly resistant to the exonuclease RNase R. The 11 circRNA candidates were also resistant to RNase R treatment, confirming their circularized characteristics (Figure 1, Supplemental Figure S1; <http://dx.doi.org/10.3168/jds.2015-10381>).

Properties of Cattle Mammary Gland circRNAs

The majority (87%) of the circRNAs consisted of protein-coding exons only, whereas the smaller fractions also contained introns, noncoding RNA gene regions, and unannotated intergenic regions (Figure 2A). We found that the majority of the circRNAs were

composed of internal exons, but not the first and last exons within the same parental gene, and their lengths were usually <1 kb (Figure 2B). Most circRNAs were formed from multiple exons; only a few were formed from a single exon. One parental gene may produce several different circRNAs. About 40% (1,265/3,239) of the parental genes have the ability to produce more than one circRNA. We found that the *CSN1S1* gene had 34 predicted circRNAs, 16 of which had more than 5 head-to-tail RNA-seq reads, as shown in Figure 2C.

circRNA Parent Gene Function Analysis

In cow mammary gland, a total of 3,239 parental genes were predicted to produce circular transcripts. Gene ontology (GO) analysis of these genes showed strong enrichment for terms related to Golgi apparatus, membrane-enclosed lumen, cellular protein catabolic process, protein localization, nucleotide binding, macromolecule biosynthetic process, and identical protein binding, thereby suggesting the potential roles played by these circRNAs in these functional areas ($P < 1.00 \times 10^{-6}$, FDR < 0.01).

The enrichment of some GO terms differed between the lactation stages. Among the top 10 enriched GO terms, vesicle, endoplasmic reticulum, and mitochondrial lumen were more enriched on lactation d 90 than on d 250 (Figure 3).

Expression Patterns of circRNAs from Casein Genes

Interestingly, all 4 casein-coding genes (*CSN1S1*, *CSN1S2*, *CSN2*, and *CSN3*) produced circRNAs in the cattle mammary gland. In total, 80 circRNAs were predicted to be produced from these 4 genes. All circRNAs were derived from exons, and most of them had multiple exons (Table 1). Circular RNAs derived from *CSN1S* were highly abundant, and 3 of these circRNAs accounted for 36% of all the circRNAs expressed in the mammary gland at lactation d 90. The 3 *CSN1S1* circRNAs were more abundant on lactation d 90 than on lactation d 250 (Table 1). The expression levels of the top 5 expressed casein circRNAs were further validated by RNase R enrichment (Figure 1). Three circRNAs from *CSN1S1*, 1 circRNA from *CSN1S2*, and 1 circRNA from *CSN2* were all expressed at significantly higher levels on lactation d 90 than on lactation d 250 (Figure 4).

miRNA Binding Sites of Casein circRNAs

The casein circRNAs had a high density of miRNA target sites. For 14 of these circRNAs, the number of

target sites in the mir-2284 family varied from 4 to 37 with an average density of 6.42/100 bp. One of the circRNAs had 14 miRNA target sites from the mir-2284 family in its 63-nucleotide length. Interestingly, *CSN1S1*, *CSN2*, and the *Ltf* whey protein genes were the putative targets in the mir-2284 family. To check the protective function of casein circRNAs on *CSN1S1* and *CSN2* mRNAs, associations in the expression levels among the top 3 expressed circRNAs and *CSN1S1* and *CSN2* mRNA were analyzed. The results showed a slight positive correlation (Figure 5).

DISCUSSION

We used a recently published novel algorithm, CIRI, to identify circRNAs from RNA-seq data. Systematic filtering by this algorithm ensures a low false-positive rate without sacrificing the sensitivity of small circRNA and nonexonic circRNA detection (Gao et al., 2015). In the present study, 685 circRNAs <200 bp were identified, 58 of which were from the 4 casein genes, accounting for over 70% of the circular casein transcripts.

It has been proposed that circRNA biogenesis is facilitated by inverted repeats flanking the circularized exon(s), and additional genome-wide computational sequence analysis suggests this complementary sequence-mediated mechanism may be widespread in nature (Dubin et al., 1995; Jeck et al., 2013; Liang and Wilusz, 2014; Zhang et al., 2014; Ivanov et al., 2015). However, we did not detect any typical inverted repeats flanking the circularized exons of the bovine casein genes. Early in vitro studies have shown that complementary sequences flanking an exon do not appear to be necessary for short exon (95 bp) circularization (Pasman et al., 1996). The circular exon length for *CSN1S1* and *CSN1S2* varied from 21 to 54 bp, so their circularization is possibly not attributable to the inverted repeats in neighboring introns. Interestingly, a recent study published on *Drosophila* discovered that there are no noticeable reverse complementary motifs flanking circularized exons besides the canonical splice site sequences (Westholm et al., 2014), suggesting that the mechanism of alternative circularization is very complicated and might be regulated by other factors, such as RNA binding proteins (Conn et al., 2015).

Although still under debate, there is growing evidence that circRNAs can regulate expression of their parental genes. Some circRNAs can interact with the polymerase II machinery and modulate parental transcription activity (Zhang et al., 2013). Subsequently, another research group revealed that the circRNAs containing intron sequences interacted with U1 small nuclear ribonucleoproteins and enhanced the transcription of

Figure 2. Detection of circular RNA (circRNA) species in cattle mammary gland. (A) Genome distribution of circRNA; (B) length distribution of the circRNA in cattle mammary gland at lactation d 90 and 250; (C) example (*CSN1S1*) showing that one parental linear RNA gene can generate multiple circRNA species with different joints. All the exons are indicated and numbered along the gene. The arcs between the different exons within the circle show the end-joining events of the circRNA.

their parental genes in a cis-acting manner (Li et al., 2015b). Circular RNAs can also function as sponges for the miRNAs that target their parental genes, thereby maintaining the expression of these genes (Hansen et al., 2013; Memczak et al., 2013; Taulli et al., 2013; Li et al., 2015a). Vesicle, endoplasmic reticulum, and mitochondrial lumen were the top enriched GO terms of the parental genes at lactation d 90. During lactation, caseins synthesized in the endoplasmic reticulum require ATP from mitochondria, and are further processed in

the Golgi apparatus, and then transported by vesicles before secretion. This suggests that circRNA might regulate the function of the endoplasmic reticulum, mitochondrion, and vesicle to promote milk protein synthesis and secretion.

MicroRNAs function primarily by imperfect base pairing with the 3'-untranslated region of target mRNAs to negatively affect their protein output by translational repression followed by mRNA deadenylation and decay (Djuranovic et al., 2012). Recently

Figure 3. Top Gene Ontology (GO) terms for the parental genes from which the circular RNA (circRNA) species were generated. (A) Parental genes from lactation d 90; (B) parental genes from lactation d 250. The number in parentheses is the number of enriched genes.

Table 1. Identification of differentially expressed circular RNA (circRNA) from casein genes between lactating mammary gland in cows at d 90 and 250 postpartum, as assessed by the edgeR program¹

circRNA ID ²	Gene symbol	Day 90	Day 250 ³	Log ₂ ratio ⁴	P-value	FDR	Exon number
chr6:87145973 87146475 ⁵	<i>CSN1S1</i>	383.00	0.01	18.50	<0.001	<0.001	2
chr6:87152493 87156336 ⁵	<i>CSN1S1</i>	314.00	0.01	18.30	<0.001	<0.001	4
chr6:87149981 87153200 ⁵	<i>CSN1S1</i>	259.00	0.01	18.00	<0.001	<0.001	4
chr6:87147119 87150809	<i>CSN1S1</i>	78.90	0.01	16.30	<0.001	<0.001	3
chr6:87148458 87150013	<i>CSN1S1</i>	29.30	0.65	5.50	<0.001	<0.001	2
chr6:87266978 87270975 ⁵	<i>CSN1S2</i>	18.20	0.01	14.20	<0.001	<0.001	8
chr6:87152501 87154445	<i>CSN1S1</i>	14.50	0.01	13.80	<0.001	<0.001	3
chr6:87150786 87155366	<i>CSN1S1</i>	13.80	0.01	13.70	<0.001	<0.001	6
chr6:87185064 87185090 ⁵	<i>CSN2</i>	12.40	0.01	13.60	<0.001	<0.001	1
chr6:87266978 87268830	<i>CSN1S2</i>	10.70	1.84	2.54	<0.001	<0.001	5
chr6:87147742 87150013	<i>CSN1S1</i>	8.72	0.01	13.10	<0.001	<0.001	3
chr6:87154404 87156336	<i>CSN1S1</i>	8.66	0.01	13.10	<0.001	<0.001	3
chr6:87274095 87276280	<i>CSN1S2</i>	6.45	0.19	5.09	<0.001	<0.001	3
chr6:87267518 87268830	<i>CSN1S2</i>	4.22	0.01	12.00	<0.001	<0.001	3
chr6:87269874 87272398	<i>CSN1S2</i>	3.40	0.01	11.70	<0.001	<0.001	5
chr6:87152493 87155366	<i>CSN1S1</i>	2.87	0.01	11.50	<0.001	<0.001	4
chr6:87380887 87386904	<i>CSN3</i>	1.87	0.01	10.90	<0.001	<0.001	2
chr6:87268337 87269897	<i>CSN1S2</i>	1.76	0.01	10.80	<0.001	<0.001	3
chr6:87378408 87386900	<i>CSN3</i>	0.98	0.01	9.94	<0.001	<0.001	3

¹Differential expression standard: $|\text{Log}_2 \text{Ratio}| \geq 2$; $P\text{-value} \leq 0.01$, false discovery rate (FDR) ≤ 0.01 ; the RNA sequencing data were from 2 pooled RNA samples (each from 4 different mammary gland samples obtained during lactation).

²Where chr6 represents chromosome 6, the number represents the genome location of the circRNA, and “|” denotes a circularization site; the location referred to is the latest cattle genome assembly and annotation release (Bos_taurus_UMD_3.1.1).

³Expression levels were normalized to per million mapped reads.

⁴The mammary gland at 250 d postpartum was used as a reference.

⁵Validated by quantitative real-time PCR and Sanger sequencing.

Figure 4. Expression patterns of representative circular RNA (circRNA) species from casein genes. The RNA with a lowercase letter “c” before their parental gene names are circRNA. The number in parentheses represents the different circRNA isomers from one parental gene. Quantitative PCR data were derived from 4 biological replicates of the mammary gland samples at 2 different time points. *UXT*, *RPS9*, and *RPS15* were used as reference genes, and the geometrical average of the cycle threshold (Ct) value from these 3 genes was used for normalization. The data are presented as the mean $\pm$ standard error of the mean; $**P < 0.01$. Significant differences were compared between lactation d 90 and lactation d 250.

Figure 5. Correlations between circular RNA (circRNA) expression levels against their parental gene mRNA as measured by quantitative (q)PCR. The qPCR data represent 4 biological replicates of the mammary gland samples at lactation d 90. *UXT*, *RPS9*, and *RPS15* were used as reference genes, and the geometrical average of the cycle threshold (Ct) value from these 3 genes was used for normalization. $R^2 = 0.35$, $P < 0.001$. The solid line is the linear regression line, and the two dashed lines are confidence bands.

accumulated evidence indicates that circRNAs can function as miRNA sponges or competing endogenous RNAs (Hansen et al., 2013; Memczak et al., 2013; Li et al., 2015b); therefore, some circRNAs may inhibit or relieve repression of miRNA on translation. We found that the cattle casein circRNAs had high-density binding sites for the miR-2284 family and were predicted to target *CSN1S1* and *CSN2* mRNAs. The abundance of the top expressed circRNAs of *CSN1S1* and *CSN2* on lactation d 90 was significantly higher than on lactation d 250, which is in agreement with the decrease in milk casein production observed during lactation. Previous studies found that the concentrations of the 4 casein variants, α_{S1} -, α_{S2} -, β -, and κ -casein, in ruminant milk did not match the mammary gland expression of their respective mRNA transcripts, *CSN1S1*, *CSN1S2*, *CSN2*, and *CSN3*, whereas the translation efficiencies of *CSN1S1* and *CSN2* were several-fold higher than those of *CSN1S2* and *CSN3* (Bevilacqua et al., 2006; Cosenza et al., 2011; Bai et al., 2013). The miR-2284 family was highly expressed in bovine mammary epithelial cells (Jin et al., 2014); hence, casein circRNAs may act as miR-2284 sponges for regulating casein translation.

CONCLUSIONS

We found that circular RNAs were abundant in cow mammary gland, and their expression levels varied at different lactation stages. Circular RNAs derived from *CSN1S1* were highly abundant, and 3 of these molecule types accounted for 36% of all the circRNAs expressed in the mammary gland at lactation d 90, a percentage significantly higher than that at lactation d 250. The circRNAs contained many miR-2284 family binding sites and possibly act as miR-2284 sponges that can regulate casein translation. Circular casein transcripts might be a novel type of factor involved in milk protein synthesis regulation; hence, they warrant further study.

ACKNOWLEDGMENTS

This work was partly funded by the National Natural Science Foundation of China (31472077, 31101703), Natural Science Foundation of Jiangsu Province (BK2011206), Jiangsu Overseas Research and Training Program for University Prominent Young and Middle-aged Teachers and Presidents Project, and a Project

Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions.

REFERENCES

- Agarwal, V., G. W. Bell, J. Nam, and D. P. Bartel. 2015. Predicting effective microRNA target sites in mammalian mRNAs. *eLife* 4:e05005.
- Bai, W. L., R. H. Yin, W. Q. Jiang, O. O. Ajayi, S. J. Zhao, G. B. Luo, Z. H. Zhao, and I. G. Imumorin. 2013. Molecular analysis of α s1-, β -, α s2- and κ -casein transcripts reveals differential translational efficiency in yak lactating mammary gland. *Livest. Sci.* 152:74–78.
- Benjamini, Y., and D. Yekutieli. 2001. The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.* 29:1165–1188.
- Bevilacqua, C., J. C. Helbling, G. Miranda, and P. Martin. 2006. Translational efficiency of casein transcripts in the mammary tissue of lactating ruminants. *Reprod. Nutr. Dev.* 46:567–578.
- Bian, Y., Y. Lei, C. Wang, J. Wang, L. Wang, L. Liu, L. Liu, X. Gao, and Q. Li. 2015. Epigenetic regulation of miR-29s affects the lactation activity of dairy cow mammary epithelial cells. *J. Cell. Physiol.* 230:2152–2163.
- Bionaz, M., and J. J. Loo. 2007. Identification of reference genes for quantitative real-time PCR in the bovine mammary gland during the lactation cycle. *Physiol. Genomics* 29:312–319.
- Cocquerelle, C., B. Mascré, D. Hétiuin, and B. Bailleul. 1993. Missplicing yields circular RNA molecules. *FASEB J.* 7:155–160.
- Conn, S. J., K. A. Pillman, J. Toubia, V. M. Conn, M. Salamanidis, C. A. Phillips, S. Roslan, A. W. Schreiber, P. A. Gregory, and G. J. Goodall. 2015. The RNA binding protein quaking regulates formation of circRNAs. *Cell* 160:1125–1134.
- Cosenza, G., A. Paucillo, A. Coletta, A. DiFrancia, M. Feligini, D. Gallo, D. DiBerardino, and L. Ramunno. 2011. Short communication: Translational efficiency of casein transcripts in Mediterranean river buffalo. *J. Dairy Sci.* 94:5691–5694.
- DeKruif, C. G., and C. Holt. 2003. Casein micelle structure, functions and interactions. Pages 233–276 in *Advanced Dairy Chemistry: 1. Proteins*, part A. P. F. Fox and P. L. H. McSweeney, ed. Kluwer Academic/Plenum Publishers, New York, NY.
- Djuranovic, S., A. Nahvi, and R. Green. 2012. miRNA-mediated gene silencing by translational repression followed by mRNA deadenylation and decay. *Science* 336:237–240.
- Dubin, R. A., M. A. Kazmi, and H. Ostrer. 1995. Inverted repeats are necessary for circularization of the mouse testis Sry transcript. *Gene* 167:245–248.
- Elzoghby, A. O., W. S. AboEl-Fotoh, and N. A. Elgindy. 2011. Casein based formulations as promising controlled release drug delivery systems. *J. Control. Release* 153:206–216.
- Fan, X., X. Zhang, X. Wu, H. Guo, Y. Hu, F. Tang, and Y. Huang. 2015. Single-cell RNA-seq transcriptome analysis of linear and circular RNAs in mouse preimplantation embryos. *Genome Biol.* 16:148.
- Gao, Y., J. Wang, and F. Zhao. 2015. CIRI: An efficient and unbiased algorithm for de novo circular RNA identification. *Genome Biol.* 16:4.
- Hansen, T. B., T. I. Jensen, B. H. Clausen, J. B. Bramsen, B. Finsen, C. K. Damgaard, and J. Kjems. 2013. Natural RNA circles function as efficient microRNA sponges. *Nature* 495:384–388.
- Hartmann, R., and H. Meisel. 2007. Food-derived peptides with biological activity: From research to food applications. *Curr. Opin. Biotechnol.* 18:163–169.
- Huang, D. W., B. T. Sherman, and R. A. Lempicki. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4:44–57.
- Ivanov, A., S. Memczak, E. Wyler, F. Torti, H. T. Porath, M. R. Orejuela, M. Piechotta, E. Y. Levanon, M. Landthaler, C. Dietrich, and N. Rajewsky. 2015. Analysis of intron sequences reveals hallmarks of circular RNA biogenesis in animals. *Cell Reports* 10:170–177.
- Jabed, A., S. Wagner, J. McCracken, D. N. Wells, and G. Laible. 2012. Targeted microRNA expression in dairy cattle directs production of β -lactoglobulin-free, high-casein milk. *Proc. Natl. Acad. Sci. USA* 109:16811–16816.
- Jeck, W. R., J. A. Sorrentino, K. Wang, M. K. Slevin, C. E. Burd, J. Liu, W. F. Marzluff, and N. E. Sharpless. 2013. Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA* 19:141–157.
- Jin, W., E. M. Ibeagha-Awemu, G. Liang, F. Beaudoin, X. Zhao, and L. L. Guan. 2014. Transcriptome microRNA profiling of bovine mammary epithelial cells challenged with *Escherichia coli* or *Staphylococcus aureus* bacteria reveals pathogen directed microRNA expression profiles. *BMC Genomics* 15:181.
- Kim, D., G. Pertea, C. Trapnell, H. Pimentel, R. Kelley, and S. L. Salzberg. 2013. TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* 14:R36.
- Kitts, D. D., and K. Weiler. 2003. Bioactive proteins and peptides from food sources. Applications of bioprocesses used in isolation and recovery. *Curr. Pharm. Des.* 9:1309–1323.
- Lasda, E., and R. Parker. 2014. Circular RNAs: Diversity of form and function. *RNA* 20:1829–1842.
- Li, F., L. Zhang, W. Li, J. Deng, J. Zheng, M. An, J. Lu, and Y. Zhou. 2015a. Circular RNA ITCH has inhibitory effect on ESCC by suppressing the Wnt/ β -catenin pathway. *Oncotarget* 6:6001–6013.
- Li, Z., C. Huang, C. Bao, L. Chen, M. Lin, X. Wang, G. Zhong, B. Yu, W. Hu, L. Dai, P. Zhu, Z. Chang, Q. Wu, Y. Zhao, Y. Jia, P. Xu, H. Liu, and G. Shan. 2015b. Exon-intron circular RNAs regulate transcription in the nucleus. *Nat. Struct. Mol. Biol.* 22:256–264.
- Liang, D., and J. E. Wilusz. 2014. Short intronic repeat sequences facilitate circular RNA production. *Genes Dev.* 28:2233–2247.
- Memczak, S., M. Jens, A. Elefantioti, F. Torti, J. Krueger, A. Rybak, L. Maier, S. D. Mackowiak, L. H. Gregersen, M. Munschauer, A. Loewer, U. Ziebold, M. Landthaler, C. Kocks, F. leNoble, and N. Rajewsky. 2013. Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* 495:333–338.
- Pasman, Z., M. D. Been, and M. A. Garcia-Blanco. 1996. Exon circularization in mammalian nuclear extracts. *RNA* 2:603–610.
- Pellegrino, L., F. Masotti, S. Cattaneo, J. A. Hogenboom, and I. Noni. 2013. Nutritional quality of milk proteins. Pages 515–538 in *Advanced Dairy Chemistry. Vol. 1A: Proteins: Basic Aspects*. 4th Edition. P. L. H. McSweeney, and P. F. Fox, ed. Springer, New York, NY.
- Rhoads, R. E., and E. Grudzien-Nogalska. 2007. Translational regulation of milk protein synthesis at secretory activation. *J. Mammary Gland Biol. Neoplasia* 12:283–292.
- Ricci-Cabello, I., M. O. Herrera, and R. Artacho. 2012. Possible role of milk-derived bioactive peptides in the treatment and prevention of metabolic syndrome. *Nutr. Rev.* 70:241–255.
- Robinson, J. T., H. Thorvaldsdottir, W. Winckler, M. Guttman, E. S. Lander, and G. Getz. 2011. Integrative genomics viewer. *Nat. Biotechnol.* 29:24–26.
- Robinson, M. D., D. J. McCarthy, and G. K. Smyth. 2010. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26:139–140.
- Rybak-Wolf, A., C. Stottmeister, P. Glazar, M. Jens, N. Pino, S. Giusti, M. Hanan, M. Behm, O. Bartok, R. Ashwal-Fluss, M. Herzog, L. Schreyer, P. Papavasiliou, A. Ivanov, M. Ohman, D. Refojo, S. Kadener, and N. Rajewsky. 2015. Circular RNAs in the mammalian brain are highly abundant, conserved, and dynamically expressed. *Mol. Cell* 58:870–885.
- Salzman, J., R. E. Chen, M. N. Olsen, P. L. Wang, and P. O. Brown. 2013. Cell-type specific features of circular RNA expression. *PLoS Genet.* 9:e1003777.
- Sanger, H. L., G. Klotz, D. Riesner, H. J. Gross, and A. K. Kleinschmidt. 1976. Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures. *Proc. Natl. Acad. Sci. USA* 73:3852–3856.
- Schaafsma, G. 2000. The protein digestibility-corrected amino acid score. *J. Nutr.* 130:1865S–1867S.

- Tauli, R., C. Loretelli, and P. P. Pandolfi. 2013. From pseudo-ceRNAs to circ-ceRNAs: A tale of cross-talk and competition. *Nat. Struct. Mol. Biol.* 20:541–543.
- Trapnell, C., B. A. Williams, G. Pertea, A. Mortazavi, G. Kwan, and M. J. van Baren. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28:511–515.
- Westholm, J. O., P. Miura, S. Olson, S. Shenker, B. Joseph, P. Sanfilippo, S. E. Celniker, B. R. Graveley, and E. C. Lai. 2014. Genome-wide analysis of *Drosophila* circular RNAs reveals their structural and sequence properties and age-dependent neural accumulation. *Cell Reports* 9:1966–1980.
- You, X., I. Vlatkovic, A. Babic, T. Will, I. Epstein, G. Tushev, G. Akbalik, M. Wang, C. Glock, C. Quedenau, X. Wang, J. Hou, H. Liu, W. Sun, S. Sambandan, T. Chen, E. M. Schuman, and W. Chen. 2015. Neural circular RNAs are derived from synaptic genes and regulated by development and plasticity. *Nat. Neurosci.* 18:603–610.
- Zhang, X. O., H. B. Wang, Y. Zhang, X. Lu, L. L. Chen, and L. Yang. 2014. Complementary sequence-mediated exon circularization. *Cell* 159:134–147.
- Zhang, Y., X. O. Zhang, T. Chen, J. F. Xiang, Q. F. Yin, Y. H. Xing, S. Zhu, L. Yang, and L. L. Chen. 2013. Circular intronic long non-coding RNAs. *Mol. Cell* 51:792–806.