



# Genome wide expression analysis of circular RNAs in mammary epithelial cells of cattle revealed difference in milk synthesis

Syed Mudasir Ahmad<sup>1,\*</sup>, Basharat Bhat<sup>1,\*</sup>, Zainab Manzoor<sup>2</sup>, Mashooq Ahmad Dar<sup>1,2</sup>, Qamar Taban<sup>1</sup>, Eveline M. Ibeagha-Awemu<sup>3</sup>, Nadeem Shabir<sup>1</sup>, Mohd Isfaqul Hussain<sup>4</sup>, Riaz A. Shah<sup>1</sup> and Nazir A. Ganai<sup>1</sup>

<sup>1</sup>Division of Animal Biotechnology, Faculty of Veterinary Sciences, Sher-e-Kashmir University of Agricultural Sciences and Technology of Kashmir, India, Srinagar, Jammu and Kashmir, India

<sup>2</sup>Department of Clinical Biochemistry, University of Kashmir, Srinagar, Jammu and Kashmir, India

<sup>3</sup>Sherbrooke Research and Development Centre, Agriculture and Agri-Food Canada, Sherbrooke, Canada

<sup>4</sup>Division of Veterinary Microbiology, SKUAST-Kashmir, Srinagar, Jammu and Kashmir, India

\*These authors contributed equally to this work.

## ABSTRACT

Milk is an excellent source of nutrients for humans. Therefore, in order to enhance the quality and production of milk in cattle, it is interesting to examine the underlying mechanisms. A number of new investigations and research have found that, circRNA; a specific class of non-coding RNAs, is linked with the development of mammary gland and lactation. In the present study, genome wide identification and expression of the circRNAs in mammary epithelial cells of two distinct cattle breeds viz Jersey and Kashmiri at peak lactation was conducted. We reported 1554 and 1286 circRNA in Jersey and Kashmiri cattle, respectively, with 21 circRNAs being differentially expressed in the two breeds. The developmental genes of the established differentially expressed circRNAs were found to be largely enriched in antioxidant activity, progesterone, estradiol, lipid, growth hormone, and drug response. Certain pathways like MAPK, IP3K and immune response pathways were found significantly enriched in KEGG analysis. These results add to our understanding of the controlling mechanisms connected with the lactation process, as well as the function of circRNAs in bovine milk synthesis. Additionally, the comparative analysis of differentially expressed circRNAs showed significant conservation across different species.

Submitted 12 November 2021

Accepted 8 February 2022

Published 1 March 2022

Corresponding authors

Syed Mudasir Ahmad,  
mudasirbio@gmail.com,  
mudasirbio@skuastkashmir.ac.in  
Basharat Bhat, basharatb-  
hat09@gmail.com, bb284@snu.edu.in

Academic editor

Kenta Nakai

Additional Information and  
Declarations can be found on  
page 11

DOI 10.7717/peerj.13029

© Copyright  
2022 Syed et al.

Distributed under  
Creative Commons CC-BY 4.0

OPEN ACCESS

**Subjects** Agricultural Science, Bioinformatics, Genomics, Veterinary Medicine

**Keywords** CircRNA, Mammary gland, Bovine, RNA-Seq, Identification, Differential expression

## INTRODUCTION

Milk is a highly nutritious liquid produced in mammary glands of mammals to support and feed their new born babies during their initial months of life. Milk is secreted by epithelial alveolar cells that are organized into isolated pockets called lobules. Nutritional value of milk is very high and contains almost every nutrient that our body requires. Cow's milk is a rich and economical source of proteins (casein and whey proteins), casein alone accounts

for approximately 80% of milk protein. It is also a good source of fat, carbohydrates, hormones, vitamins and minerals (*Elzoghby, El-Fotoh & Elgindy, 2011*).

As milk production is an economic trait in dairy animals, an increase in the milk synthesis is a desirable goal for the dairy industry (*Bionaz, Hurley & Loor, 2012*). A deeper understanding of the factors controlling milk synthesis can therefore provide guidance on the methods and processes to enhance milk production. As reported earlier, hormones, protein factors and non-coding RNAs (ncRNAs) play a significant role in controlling the milk production (*Bionaz, Hurley & Loor, 2012; Sandhu et al., 2016; Hansji et al., 2014*).

Circular RNAs (circRNAs), discovered more than three decades ago (*Hsu & Coca-Prados, 1997*), have gained substantial attention in the last couple of years because of their potential function in eukaryotic cells. CircRNAs are one of the ncRNAs that are typically produced when pre-mRNA (pre-messenger RNA) is alternatively spliced which involves joining of upstream splice acceptor with downstream splice donor in a process called backsplicing forming a unique covalently closed circular structure lacking 3' and 5' ends (*Morris & Mattick, 2014; Barrett, Wang & Salzman, 2015*). This mechanism protects it from RNA exonucleases, stabilising it and making it less prone to degradation (*Schindewolf, Braun & Domdey, 1996; Jeck & Sharpless, 2014; Memczak et al., 2013*). CircRNAs contain exonic or intronic sequences of its parental genes and range in size from 100 nt to 4 kb (*Szabo & Salzman, 2016*). Mainly, three types of circRNAs have been identified: exonic circular RNAs (ecircRNAs), intronic circular RNAs (ciRNAs) and exon-intron circular RNAs (ElicRNAs). With the advances in RNA sequencing (RNA-Seq) technologies and bioinformatics, diverse properties of circRNAs are being explored and have been found to be expressed in all cells and tissues just like linear mRNAs (*Mumtaz et al., 2020*). Some circRNAs contain micro RNA response elements (MREs) that serve as competitive endogenous RNA and play a special regulatory role by functioning as micro RNA (miRNA) sponge, thereby controlling the expression of parent genes in nucleus (*Zhang et al., 2013*). Studies also suggest that circRNAs act as miRNA sponge and thus are capable of enhancing the target gene expression (*Hansen et al., 2013*). Certain circRNAs build protein complexes, regulate cell cycle and protein translation (*Du et al., 2016*). CircRNAs also function as protein scaffolds particularly for RNA-binding proteins (RBPs) including RNA Polymerase II (*Zwierzchowski, 2005*) and Argonaute (AGO) proteins. Moreover, circRNAs are involved in disease pathogenesis by directly regulating certain relevant pathways and maintaining mRNA homeostasis. Because of their vital role in cellular regulation, protein signaling and disease pathogenesis, circRNAs are becoming a new research hotspots in RNA biology (*Liu et al., 2020*).

Accounting the universal expression characteristics of circRNAs, these molecules have been found to be involved in the milk synthesis as well as mammary gland development of various species, like sheep (*Wang et al., 2021; Hao et al., 2020a; Hao et al., 2020b*), cattle (*Zhang et al., 2016a; Zhang et al., 2016b*), goat (*Ma et al., 2019*) and rat (*Zhang et al., 2015*). In the present study we performed a systematic examination and expression profile of circRNA by RNA-Seq in mammary gland epithelial cells (MECs) at peak lactation in two cattle breeds with distinct milk production properties *i.e.*, high lactating Jersey and low lactating Kashmiri cattle. We describe the type, length and chromosomal distribution of

the circRNAs and their functional enrichment analysis with the aim to understand their role in lactation of dairy cows.

## MATERIAL AND METHODS

As a global trend, we concluded that the larger replicate numbers increase the power of differential expression analysis, except for low-expressed genes. We utilized RnaSeqSampleSize v2.0.0 (Zhao *et al.*, 2018) for the estimation of samples required for the analysis. RnaSeqSampleSize recommended five samples (biological replicates) per condition for the identification of majority of differentially expressed circRNAs.

In this study, a total of eight samples (four Jersey and four Kashmiri cows) at the peak lactation was collected from our published data in the SRA database (accession ID. [SRR6324372](#)), including one sample from each breed as technical replicate. The details of experiments followed are mentioned below.

### Sampling and isolation of MECs

Clean and fresh milk samples (1.5 l/cow) were collected from healthy and unrelated cows of each cattle breed by hand milking. The animals were kept in free stall sheds, provided with a complete healthy nutritious and balanced diet, and had unrestricted access to clean water and grass. The detailed sample collection including sample preparation and RNA sequencing were explained in the previous study (Bhat *et al.*, 2020).

### Quantification of circRNAs

A total of eight samples (4 Jersey and 4 Kashmiri) at peak stage of lactation were selected for the identification of circRNAs mediating different milk traits in both breeds. The FASTQC program v0.11.2 was used to perform quality control on all raw reads (Chanumolu, Albahrani & Otu, 2019). The Cutadapt software v2.10 (Martin, 2011) and Sickle tool v1.33 (Joshi & Fass, 2011) were used to filter out adapter sequences as well as decreased and low quality reads. Using the HISAT2 program v2.03, the filtered reads were mapped to the reference bovine genome (ARS-UCD1.2) (Kim *et al.*, 2019).

### Identification of differentially expressed circRNAs

Two-way approach was employed for the identification of differentially expressed (DE) circRNAs. At first seekCRIT tool v1.0.0.b (Chaabane *et al.*, 2020) were run on filtered reads to identify DE circRNAs. Another strategy of identification of DE circRNA involves quantification of circRNAs (using CIRI and CIRC-explorer) and running edgeR on the quantified dataset.

Based on the expression of circular - RNAs, the quasi-likelihood F-test method of the edgeR program (McCarthy, Chen & Smyth, 2012) was used for identifying circRNAs that are differentially expressed between two breeds. Differentially expressed circRNAs between two contrast groups were screened using  $P$ -value  $<0.05$  and  $\log_2FC >1$  threshold. The final set of DE circRNAs contains circRNAs common between seekCRIT dataset and edgeR dataset. CircRNAs whose  $\log_2FC >1$  and  $p$ -value  $<0.05$  were considered as DE in two breeds. To identify the conserved circRNAs, all DE circRNAs were compared against the

known circRNAs from *Homo sapiens*, *Mus musculus*, *Sus scrofa*, *Gallus gallus* and *Canis lupus*.

### RNA extraction, cDNA synthesis and quantitative real-time PCR

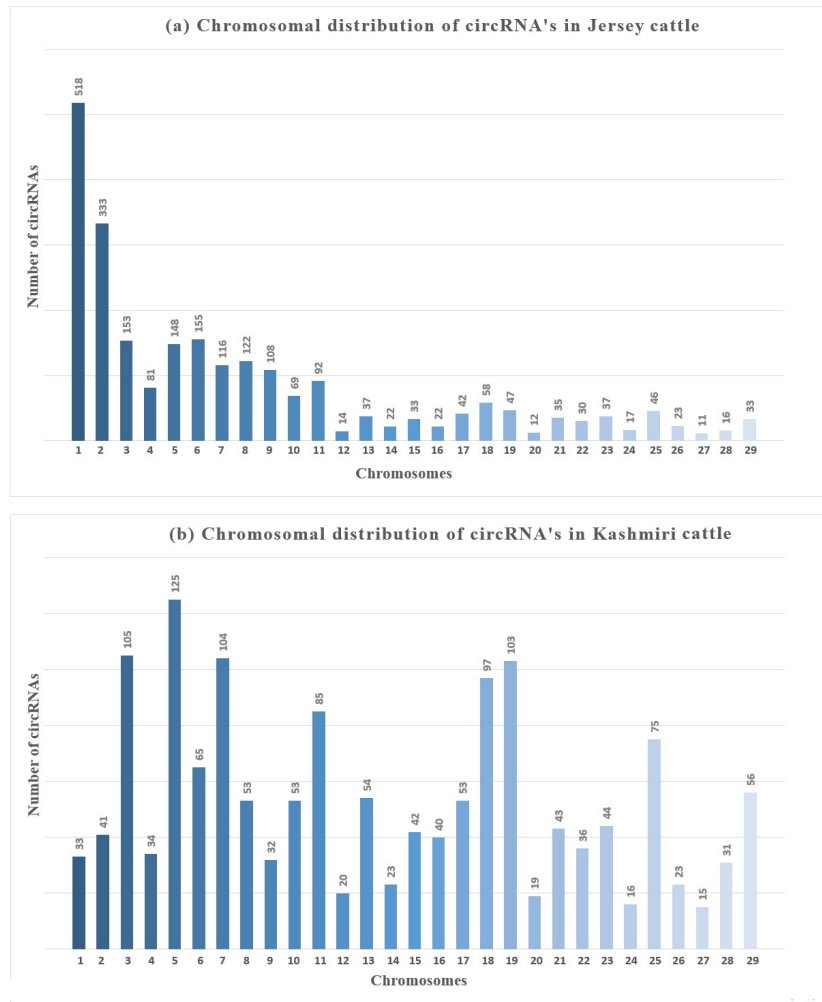
Total RNA previously used for RNA-Seq was used for the construction of cDNA. Total RNA that was previously used for RNA-Seq was extracted by Trizol method following the manufacturer's protocol followed by cDNA synthesis. circRNAs were enriched in the total RNA pool by using RNA exoribonucleases (RNase R). The circRNA enriched samples with depleted ribosomal RNA were reverse transcribed into cDNA with random hexamer primers using iScript RT Supermix Kit (Bio-Rad) following the manufacturer's procedure. cDNA synthesized using specific divergent sets of primers were subjected to RT-qPCR. Experiment was performed in triplicate for RT-qPCR. A total reaction of 20  $\mu$ l was made with 0.5  $\mu$ l (10  $\mu$ M) of each forward and backward divergent primer, 10  $\mu$ l of SYBR Green PCR master mix (Roche), 1  $\mu$ l (70 ng/  $\mu$ l) of cDNA and 8  $\mu$ l of nuclease free water. Glyceraldehyde-3-Phosphate (GAPDH) was used as an internal control. The experimental validation was selected for the group of five circRNAs, including circ\_03409, circ\_87295, circ\_03409a, circ\_10119b and circ\_25279. For the selected circRNAs, five divergent primers were designed using Primer3 Plus software. The details of the primers are given in (Table S1). The relative quantification was determined using the  $2^{-\Delta\Delta Ct}$  method (Livak and Schmittgen, 2001).

### Functional enrichment analysis

To gain insight about the function of mRNAs encoded by similar DE circRNAs, we performed Gene Ontology (GO) and pathways enrichment analysis. The enrichment analysis was performed utilizing KOBAS 3 server (Wu et al., 2006). *Bos taurus* gene-set were used as background (Bhat et al., 2019a; Bhat et al., 2019b).

## RESULTS

An average of 2986 circRNAs were reported from the samples analyzed, majority of which were ecircRNAs ( $\approx 72.16\%/2154$ ) whereas ciRNAs were less common ( $\approx 27.83\%/831$ ). Most of the circRNAs were distributed on chromosome (chr) 5 ( $\approx 20.26\%/605$  and chr 6 ( $\approx 7.23\%/216$ ) while fewest circRNA were distributed on chr 12, 24, 26 and 27. The number of circRNAs obtained from mammary glands of Jersey and Kashmiri were 1554 (966 exonic, 588 intronic) and 1286 (1047 exonic, 239 intronic) respectively. In Jersey, a higher number of circRNAs were derived from chr 4 (125 circRNAs) followed by chr 3 (105 circRNAs), chr 7 (104) and chr 19 (103). While as in Kashmiri cattle 518 circRNAs were derived from chr 1 followed by chr 2 (333), chr 6 (155), chr 3 (153) and chr 7 (Fig. 1). The most abundant and highly expressed circRNAs in both Kashmiri and Jersey cattle were circ“-\_003409, circ\_010119, circ\_081923, circ\_006590, circ\_021268, circ”-\_023981 and circ\_013402 which were derived from CSN2, CSN1S1, PAEP, CSN1S2, RPS5, NUPR1 and TPT1 respectively. A total of 150 circRNAs were obtained from casein protein that includes 87 CSN1S1, 31 CSN1S2, 29 CSN2, and 03 CSN3 circRNAs. Two circRNAs, LALBA and BLG were derived from whey protein coding genes.



**Figure 1** Chromosomal distribution of circRNAs in mammary gland epithelial cells of (A) Jersey and (B) Kashmiri cattle at peak lactation.

Full-size DOI: [10.7717/peerj.13029/fig-1](https://doi.org/10.7717/peerj.13029/fig-1)

### Analysis of differentially expressed circRNAs

There were 21 differentially expressed circRNAs between Kashmiri and Jersey breeds with  $\log_2FC > 1$  and  $P$ -value  $< 0.05$ , with 14 up-regulated circRNAs in Kashmiri cattle and 7 up-regulated circRNAs in Jersey. All differentially expressed circRNAs were found to be exonic, only one circ\_60634 derived from RHBDD2 gene (Start: 34128469, Stop: 34128556, Strand: +, Exon count: 01) present on chr 25 and upregulated in Jersey breed was intronic. The DE circRNA were compared against the known circRNAs from *Homo sapiens*, *Mus musculus*, *Sus scrofa*, *Gallus gallus* and *Canis lupus*. A total of 8 out of 21 DE circRNAs were mapped with the threshold of e-value  $< 0.00001$  and similarity  $> 85\%$  (Table S2). Seven DE circRNAs were derived from casein genes (03 circRNAs from CSN2 gene, 03 circRNAs from CSN1S1, 01 circRNA from CSN1S2) (Table S3). The circRNAs that were upregulated in Kashmiri cattle included circ\_03409 (gene CSN2 on chr 6), circ\_03409A

(gene CSN2 on chr 6), circ\_03409B (gene CSN2 on chr 6), circ<sup>+</sup>\_81923A (gene PAEP on chr 11), circ<sup>-</sup>\_81923B (gene PAEP on chr 11) circ\_10119A (gene CSN1S1 on chr 6), circ\_10119B (gene CSN1S1 on chr 6), circ\_10119C (gene CSN1S1 on chr 6) circ\_63604 (gene CSN1S1 on chr 25), circ\_06590 (gene CSN1S2 on chr 6), circ\_21268 (gene RPS5 on chr 8), circ\_61308 (gene ATRX on chr X), circ\_81923 (gene PAEP on chr 11), and circ\_70874 (gene PICALM on chr 29). CircRNAs upregulated in Jersey breed with respect to Kashmiri cattle include circ\_87295 (gene USP3 on chr 10), circ\_23981 (gene NUPR1 on chr 25), circ\_82741 (gene NDEL1 on chr 19), circ\_13402 (gene TPT1 on chr 12), circ\_25279 (gene ITGB7 on chr 5), circ\_20009 (gene CD14 on chr 7) and circ\_34753 (gene KANSL1 on chr 19). Among these DE circRNAs the three highly abundant circRNAs upregulated in Kashmiri cattle were: circ\_81923A, circ\_03409A and circ\_10119B. Most abundant circRNAs upregulated in Jersey breed with respect to Kashmiri cattle include circ\_87295 and circ\_25279. To confirm the authenticity of the differentially expressed circRNAs, we performed RT-qPCR of only five circRNAs *i.e.*, 03 DE circRNAs from Kashmiri cattle (circ\_03409, circ\_03409A and circ\_10119B) and 02 DE circRNAs from Jersey cattle (circ\_87295 and circ\_25279). The different suffixes indicate the circRNAs derived from same genes. The findings demonstrated that the head-to-tail backsplicing of circRNAs was consistent with results of RNA-Seq, verifying the reliability and integrity of the circRNAs discovered in the analysis.

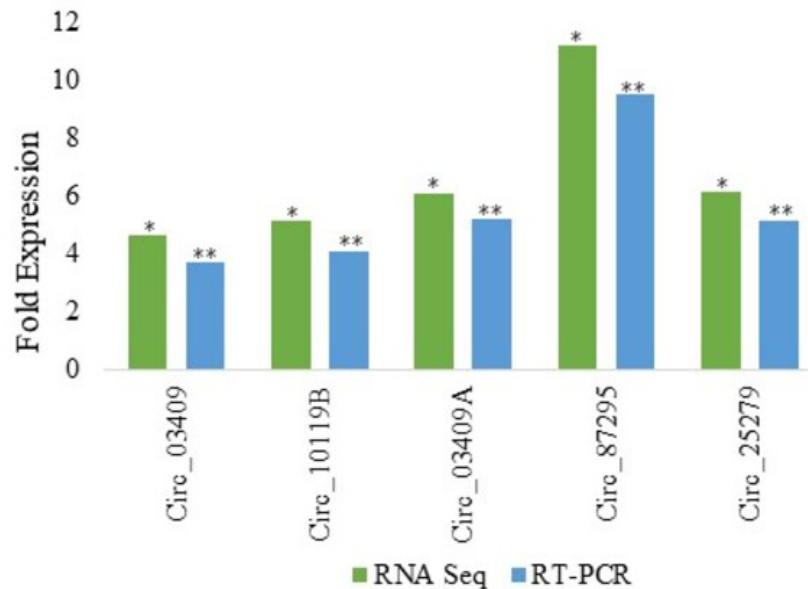
### Enrichment analysis of differentially expressed genes

We analyzed the cluster analysis of gene ontology (GO) and KEGG pathway enrichment of differentially expressed circRNAs to further investigate the relationship between circRNAs and regulation of gene expression. The parental genes for 21 differentially expressed circRNAs were primarily enriched in 09 biological processes including response to hormones like progesterone, estradiol, dehydroepiandrosterone, 11- deoxycorticosterone and growth hormone, response to lipids, drugs, cellular oxidant detoxification, 01 molecular function showing antioxidant activity and 05 cellular components showing strong enrichment for terms related to golgi lumen, extracellular region, golgi apparatus, extracellular space and endomembrane system. The KEGG pathway revealed that the developmental genes of circRNAs were concentrated in a variety of pathways including intestinal immune network for IgA production, pertussis pathway, prolactin pathway, salmonella infection, Toll like receptor pathway, ECM receptor interaction, MAPK signalling pathway and PI3K-Akt signaling pathway, NF kappa B signalling pathway, and cell adhesion molecules ([Tables S4](#) and [S5](#)).

### Validation of identified circRNAs by Real Time-qPCR

Among various DE circRNAs, RT-qPCR of five differentially expressed circRNA from Kashmiri and Jersey breed was performed to confirm the reliability as well as accuracy and reliability of RNA-Seq data. From Kashmiri cattle we selected 03 DE circRNAs (circ\_03409, circ\_03409A and circ\_10119B) and in case of Jersey cattle circ\_25279 and circ\_87295 were selected for validation of RNA-Seq results. As already mentioned above, the different suffixes indicate the circRNAs derived from same genes. The results of RT-qPCR revealed





**Figure 2** Real time PCR validation of differentially expressed circRNAs in Jersey and Kashmiri cattle. The y-axis represents the log<sub>2</sub> fold change of circRNAs expression; the x-axis shows the circRNA IDs employed for validation. RNA seq: \* $P < 0.05$ ; RT-qPCR: \*\* $P < 0.05$ . Suffixes A and B represent circRNAs derived same gene having same ensemble ID.

Full-size DOI: 10.7717/peerj.13029/fig-2

that the levels of expression of DE circRNAs and RNA-Seq are highly accurate and precise (Fig. 2).

## DISCUSSION

Hormones, growth factors, and the availability of nutrients closely control the production and functional initiation of lactation in the mammary gland. Previous studies have concluded that mammary glands milk production and lactation success in mammals is influenced by a substantial number of ncRNAs and various functional genes (Miglior *et al.*, 2017; Wang *et al.*, 2012). Analysis of the molecular mechanism involved in mammary gland development and lactation is thus essential for promoting milk yield and quality in livestock. CircRNAs are stable molecules as they lack free 3' and 5' ends. Therefore, in the context of circRNAs function in mammary glands, we studied the circRNAs expression levels in two dairy cattle breeds viz Kashmiri and Jersey, to investigate how circRNAs play a role in controlling the expression of milk-related genes.

Previous studies suggest that in bovine MECs, circular RNAs were abundant, and their levels of expression differed at various lactation stages (Zhang *et al.*, 2016a; Zhang *et al.*, 2016b). CircRNAs function as sponges of miR-2284 that control the translation of casein genes and thus regulate the synthesis of milk protein (Zhang *et al.*, 2016a; Zhang *et al.*, 2016b). The disparity in the production of circRNA between species may be due to their different lactating periods. In our study, 1554 and 1286 circRNAs were obtained at peak lactation in Jersey and Kashmiri cattle respectively. The majority of circRNAs

*i.e.*, 1780 originated from multiple exons, which accounted for 59.63%, and 1205 were found to be single exonic. Similar results were obtained from mammary gland of sheep, where the majority of circRNAs originated from multiple exons (Li *et al.*, 2015). A total of 150 circRNAs were produced from 7 casein genes and 2 from whey protein genes. The most abundant circRNAs were derived from alpha s1-casein gene (CSN1S1). These results were similar to the previous studies in sheep (Hao *et al.*, 2020a; Hao *et al.*, 2020b) and bovine (Memczak *et al.*, 2013). Studies have found that high milk yield and protein content were correlated with the CSN1S1 gene (Van Eenennaam & Medrano, 1991). Beta-casein gene (CSN2) is essential for the structure of casein micelle and high milk yield (Zwierzchowski, 2005; Jann *et al.*, 2002). A milk protein that is encoded by kappa casein gene (CSN3) is essential for the formation, stability and structural characteristics of casein micelles and is a key protein in the development of cheese (Bingham & Groves, 1979). In addition to this, LALBA gene codes for alpha-lactalbumin protein which functions as lactose synthases essential regulatory subunit, in mammary gland alters the substrate specificity of galactosyltransferase, rendering glucose a suitable acceptor substrate for this galactosyltransferase enzyme. In the current study, we found two circRNAs, LALBA and BLG were derived from whey protein coding genes. Findings from all these previous studies suggest that the circRNAs obtained from these casein and whey genes may be involved in enhancing the milk production and quality in bovines. Further, these circRNAs may also be associated with growth and development of mammary gland but this presumption requires further verification.

In an attempt to further characterize the biological processes of circRNAs that are differentially expressed in two breeds, GO and KEGG enrichment pathway analysis of host genes of DE circRNAs was performed. The differentially expressed circRNAs showed response to progesterone, estradiol, growth hormone, dihydroxyepiandrosterone (DHEA), deoxycorticosterone, oxygen containing compounds, drug response and cellular oxygen detoxification. Previous studies have already shown that progesterone alone with prolactin has a role in generating milk alveoli cells which secrete milk during lactation (Macias & Hinck, 2012). Further studies using microarray assays revealed 124 genes in mammary glands of mammals whose transcription is controlled by oestrogen, the majority of which are associated with cell multiplication and proliferation activation like epidermal growth factor 1 (EGF1), cyclin D1 and insulin growth factor-1 (IGF-1) (Connor *et al.*, 2007). Some studies have also shown that estradiol is capable of regulating transcription of progesterone receptor (Petz *et al.*, 2004). Previous research has shown that progesterone binds to its receptor in the mammary glands to cause Wnt-4 gene transcription, thus mediating role of progesterone in cell proliferation (Briskin & O'Malley, 2010). Two steroid hormones, DHEA and deoxycorticosterone, on the other hand, serve in synthesis of milk. The parental genes of DE circRNAs also showed strong enrichment of antioxidant activity. Golgi lumen, extracellular region, golgi apparatus, extracellular space and endomembrane system were top enriched GO terms of parental genes. This might be because milk proteins are synthesized in endoplasmic reticulum and processed in golgi apparatus where these are released in vesicles before secretion.



The host genes of circRNAs were abundantly expressed in the prolactin pathway, Ras, Mitogen Activated Protein Kinase signalling pathways (MAPK), Phosphatidylinositol 3-Kinase (PI3K) and Akt Kinase signalling pathway in KEGG pathway enrichment. In the epithelial cells of mammary gland, the PI3K-Akt signalling pathway as well as MAPK through Stat5 and mTOR and also prolactin-mediated activation of the Jak2-Stat5 pathway, regulate the majority of metabolic mechanisms and biological pathways including transcriptional activation of genes for proliferation as well as cell cycle and synthesis of milk protein (DaSilva et al., 1996; Liu et al., 1997; Chen et al., 2012; Bhat et al., 2017). Other substantially enriched pathways including the NF- $\kappa$ B signalling pathway, the intestinal immune network for IgA development and production and Toll-like receptor pathway were related to immunity. On the basis of the study done, circRNAs identified in our study might have the capability to act as regulators of milk synthesis and stimulate growth and development of mammary glands.

AMPK plays an essential role in cellular energy sensing (Hardie, Ross & Hawley, 2012) and mTOR activation (Inoki, Zhu & KL, 2003). This signaling pathway is found to be involved in regulating the effect of glucose supply and utilization in the lactating mammary glands (Bionaz, Hurley & Loor, 2012). It is well known that mTOR is an essential regulator in milk protein synthesis, and most studies are concerned with the role of AAs in the regulation of P-mTOR on Ser<sup>2448</sup> in milk protein synthesis (Appuhamy et al., 2012; Apelo, Knapp & Hanigan, 2014). In the present study 150 circRNAs were obtained from casein protein only.

In our previous study (Ahmad et al., 2021) we found in Jersey cattle, the phosphatidylinositol 3-kinase (PI3K)/protein kinase B (Akt) signaling pathway was most enriched. PI3K-Akt pathway is important for normal insulin-mediated glucose metabolism (Taniguchi et al., 2006). Jersey being higher milk producer breed requires more energy source than Kashmiri cattle. Toll like receptor 4 (TLR4) signaling results in the activation of NF- $\kappa$ B pathway and releases pro-inflammatory cytokines (Medzhitov & Kagan, 2006). NF- $\kappa$ B being an essential transcriptional factor, regulates immune response and gene expression of many inflammatory cytokines (Putra et al., 2014). As supported by our earlier findings most of the immune related pathways were found exclusively expressed in Kashmiri cattle (Bhat et al., 2019a; Bhat et al., 2019b; Bhat et al., 2020; Ahmad et al., 2021). This suggests the possible defense mechanism of Kashmiri cattle against mammary gland infections. Based on our results and data from recent studies, circRNAs have been found to have a significant role in milk synthesis and in the development of the mammary gland in cattle.

Although numerous studies have reported that the circRNAs are tissue-specific and conserved across different species (Rybak-Wolf et al., 2015; Chen, 2016; Mumtaz et al., 2020). In the present study we demonstrated that the DE circRNAs showed significant similarity (>85%) across various species like *Homo sapiens* and *Sus scrofa*. The target gene analysis of DE circRNAs in cattle were 100% similar for UP3 and PICALM genes in *Sus scrofa* and RPS5 gene in *Homo sapiens*.

## CONCLUSION

The function of circRNAs remains largely unclear in the lactation physiology of dairy animals. In the present study an association between circRNAs and milk synthesis has been reported using RNA-seq data. The results suggest that circRNAs are abundant in bovine MECs and the expression level of circRNAs varied between Jersey and Kashmiri cattle. The number of circRNAs obtained from MECs of Jersey were high (1,554) compared to Kashmiri cattle (1,047). The results also reveal that among the total circRNAs identified, 21 circRNAs were expressed differentially between Jersey and Kashmiri cattle. Most of the DE circRNAs are derived from casein genes. This signifies the role of circRNAs in lactation of animals. The chromosomal distribution revealed most of circRNAs were assembled on chr 4 in Jersey, while as in Kashmiri cattle most circRNAs were gathered on chr 1. These results suggest breed specific distribution and variation of circRNAs in cattle. The pathway analysis shows that most enriched pathways like MAPK, IP3K and prolactin pathway are involved in milk synthesis. These findings add to our understanding of the functional role of circRNAs in the production of bovine milk. Further analysis revealed that identified circRNAs were highly conserved across various species like *Homo sapiens*, *Sus scrofa* and *Canis lupus*.

### Abbreviations

<b>CircRNA</b>	circular RNA
<b>ncRNA</b>	non-coding RNA
<b>MEC's</b>	mammary epithelial cells
<b>MAPK</b>	Mitogen-activated protein kinases
<b>IP3K</b>	Phosphatidylinositol-3-Kinase
<b>CSN 2</b>	beta- casein
<b>CSN1S1</b>	casein alpha s1
<b>protein (PAEP)</b>	progesterone-associated endometrial
<b>CSN1S2</b>	alpha-S2 casein <i>gene</i>
<b>RPS5</b>	Ribosomal Protein S5
<b>NUPR1</b>	Nuclear protein 1
<b>TPT1</b>	Translationally Controlled Tumor
<b>LALBA</b>	Lactalbumin-alpha
<b>BLG</b>	beta-lactoglobulin
<b>RHBDD2</b>	Rhomboid domain-containing protein 2
<b>USP3</b>	Ubiquitin Specific Peptidase 3
<b>NDEL1</b>	Nuclear distribution protein nudE-like 1
<b>ITGB7</b>	integrin subunit beta 7
<b>KANSL1</b>	KAT8 Regulatory NSL Complex Subunit 1
<b>ECM</b>	extracellular matrix
<b>DHEA</b>	dehydroepiandrosterone

## ACKNOWLEDGEMENTS

The Dairy Farm of Mountain Livestock Research Institute-Manasbal, SKUAST-K is acknowledged for care and management of the animals.

## ADDITIONAL INFORMATION AND DECLARATIONS

### Funding

The authors received no funding for this work.

### Competing Interests

Syed Mudasir Ahmad is an Academic Editor for PeerJ.

Eveline M. Ibeagha-Awemu is employed by Agriculture and Agri-Food Canada.

### Author Contributions

- Syed Mudasir Ahmad conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Basharat Bhat and Zainab Manzoor conceived and designed the experiments, prepared figures and/or tables, and approved the final draft.
- Mashooq Ahmad Dar and Qamar Taban performed the experiments, prepared figures and/or tables, and approved the final draft.
- Eveline M. Ibeagha-Awemu performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Nadeem Shabir analyzed the data, prepared figures and/or tables, and approved the final draft.
- Mohd Isfaqul Hussain, Riaz A. Shah and Nazir A. Ganai analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.

### Data Availability

The following information was supplied regarding data availability:

The sequences are available at NCBI GEO: [SRR6324372](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=SRR6324372).

### Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.13029#supplemental-information>.

## REFERENCES

- Ahmad SM, Bhat B, Bhat SA, Yaseen M, Mir S, Raza M, Iquebal MA, Shah RA, Ganai NA. 2021.** SNPs in Mammary gland epithelial cells unraveling potential difference in milk production between jersey and kashmiri cattle using RNA sequencing. *Frontiers in Genetics* **12**:666015 DOI [10.3389/fgene.2021.666015](https://doi.org/10.3389/fgene.2021.666015).
- Apelo SA, Knapp J, Hanigan M. 2014.** Invited review: current representation and future trends of predicting amino acid utilization in the lactating dairy cow. *Journal of Dairy Science* **97**:4000–4017 DOI [10.3168/jds.2013-7392](https://doi.org/10.3168/jds.2013-7392).

- Appuhamy JRN, Knoebel NA, Nayananjalie WD, Escobar J, Hanigan MD. 2012.** Isoleucine and leucine independently regulate mtor signaling and protein synthesis in mac-t cells and bovine mammary tissue slices. *Journal of Nutrition* **142**:484–491 DOI [10.3945/jn.111.152595](https://doi.org/10.3945/jn.111.152595).
- Barrett SP, Wang PL, Salzman J. 2015.** Circular RNA biogenesis can proceed through an exon-containing lariat precursor. *Elife* **4**:e07540 DOI [10.7554/eLife.07540](https://doi.org/10.7554/eLife.07540).
- Bhat B, Ganai NA, Andrabi SM, Shah RA, Singh A. 2017.** TM-Aligner: multiple sequence alignment tool for transmembrane proteins with reduced time and improved accuracy. *Scientific Reports* **7**(1):12543 DOI [10.1038/s41598-017-13083-y](https://doi.org/10.1038/s41598-017-13083-y).
- Bhat BA, Singh G, Sharma R, Yaseen M, Ganai NA. 2019a.** Biological networks: tools, methods, and analysis. In *Essentials of Bioinformatics I*:255–286.
- Bhat SA, Ahmad SM, Ibeagha-Awemu EM, Bhat BA, Dar MA, Mumtaz PT, et al. 2019b.** Comparative transcriptome analysis of mammary epithelial cells at different stages of lactation reveals wide differences in gene expression and pathways regulating milk synthesis between Jersey and Kashmiri cattle. *PLOS ONE* **14**(2):e0211773 DOI [10.1371/journal.pone.0211773](https://doi.org/10.1371/journal.pone.0211773).
- Bhat SA, Ahmad SM, Ibeagha-Awemu EM, et al. 2020.** Comparative milk proteome analysis of Kashmiri and Jersey cattle identifies differential expression of key proteins involved in immune system regulation and milk quality. *BMC Genomics* **21**:161 DOI [10.1186/s12864-02-6574-4](https://doi.org/10.1186/s12864-02-6574-4).
- Bingham EW, Groves ML. 1979.** Properties of casein kinase from lactating bovine mammary gland. *Journal of Biological Chemistry* **254**(11):4510–4515 DOI [10.1016/S0021-9258\(17\)30040-6](https://doi.org/10.1016/S0021-9258(17)30040-6).
- Bionaz M, Hurley W, Loor J. 2012.** Milk protein synthesis in the lactating mammary gland: insights from transcriptomics analyses. In: *Milk protein*. London (U.K.): In Tech Publisher, 285–324 DOI [10.5772/46054](https://doi.org/10.5772/46054).
- Brisken C, O'Malley B. 2010.** Hormone action in the mammary gland. *Cold Spring Harbor Perspectives in Biology* **2**:a003178 DOI [10.1101/cshperspect.a003178](https://doi.org/10.1101/cshperspect.a003178).
- Chaabane M, Andreeva K, Hwang JY, Kook TL, Park JW, Cooper NG. 2020.** seekCRIT: detecting and characterizing differentially expressed circular RNAs using high-throughput sequencing data. *PLOS Computational Biology* **16**(10):e1008338 DOI [10.1371/journal.pcbi.1008338](https://doi.org/10.1371/journal.pcbi.1008338).
- Chanumolu SK, Albahrani M, Otu HH. 2019.** FQStat: a parallel architecture for very high-speed assessment of sequencing quality metrics. *BMC Bioinformatics* **20**:424 DOI [10.1186/s12859-019-3015-y](https://doi.org/10.1186/s12859-019-3015-y).
- Chen CC, Stairs DB, Boxer RB, Belka GK, Horseman ND, Alvarez JV, Chodosh LA. 2012.** Autocrine prolactin induced by the Pten-Akt pathway is required for lactation initiation and provides a direct link between the Akt and Stat5 pathways. *Genes and Development* **26**(19):2154–2168 DOI [10.1101/gad.197343.112](https://doi.org/10.1101/gad.197343.112).
- Chen LL. 2016.** The biogenesis and emerging roles of circular RNAs. *Nature Reviews Molecular Cell Biology* **17**:205–211 DOI [10.1038/nrm.2015.32](https://doi.org/10.1038/nrm.2015.32).

- Connor EE, Meyer MJ, Li RW, Van Amburgh ME, Boisclair YR, Capuco AV. 2007. Regulation of gene expression in the bovine mammary gland by ovarian steroids. *Journal of Dairy Science* **90**(Suppl 1):E55–E65 DOI [10.3168/jds.2006-466](https://doi.org/10.3168/jds.2006-466).
- DaSilva L, Rui H, Erwin RA, Howard OM, Kirken RA, Malabarba MG, Hackett RH, Lerner AC, Farrar WL. 1996. Prolactin recruits STAT1, STAT3 and STAT5 independent of conserved receptor tyrosines TYR402, TYR479, TYR515 and TYR580. *Molecular and Cellular Endocrinology* **117**(2):131–140 DOI [10.1016/0303-7207\(95\)03738-1](https://doi.org/10.1016/0303-7207(95)03738-1).
- Du WW, Yang W, Liu E, Yang Z, Dhaliwal P, Yang BB. 2016. Foxo3 circular RNA retards cell cycle progression via forming ternary complexes with p21 and CDK2. *Nucleic Acids Research* **44**(6):2846–2858 DOI [10.1093/nar/gkw027](https://doi.org/10.1093/nar/gkw027).
- Elzoghby AO, El-Fotoh WS, Elgindy NA. 2011. Casein-based formulations as promising controlled release drug delivery systems. *Journal of Controlled Release* **153**(3):206–216 DOI [10.1016/j.jconrel.2011.02.010](https://doi.org/10.1016/j.jconrel.2011.02.010).
- Hansen TB, Jensen TI, Clausen BH, Bramsen JB, Finsen B, Damgaard CK, Kjems J. 2013. Natural RNA circles function as efficient microRNA sponges. *Nature* **495**(7441):384–388 DOI [10.1038/nature11993](https://doi.org/10.1038/nature11993).
- Hansji H, Leung E, Baguley BC, Finlay GJ, Askarian-Amiri ME. 2014. Keeping abreast with long non-coding RNAs in mammary gland development and breast cancer. *Frontiers in Genetics* **5**:379 DOI [10.3389/fgene.2014.00379](https://doi.org/10.3389/fgene.2014.00379).
- Hao S, Qu R, Hu C, Wang M, Li Y. 2020b. A Circular RNA derived from golgi glycoprotein 1 mRNA regulates KRAS expression and promotes colorectal cancer progression by targeting microRNA-622. *OncoTargets and Therapy* **13**:12637–12648 DOI [10.2147/OTT.S284032](https://doi.org/10.2147/OTT.S284032).
- Hao Z, Zhou H, Hickford JG, Gong H, Wang J, Hu J, Liu X, Li S, Zhao M, Luo Y. 2020a. Identification and characterization of circular RNA in lactating mammary glands from two breeds of sheep with different milk production profiles using RNA-Seq. *Genomics* **112**(3):2186–2193 DOI [10.1016/j.ygeno.2019.12.014](https://doi.org/10.1016/j.ygeno.2019.12.014).
- Hardie DG, Ross FA, Hawley SA. 2012. AMPK: a nutrient and energy sensor that maintains energy homeostasis. *Nature Reviews Molecular Cell Biology* **13**:251–262.
- Hsu MT, Coca-Prados M. 1997. Electron microscopic evidence for the circular form of RNA in the cytoplasm of eukaryotic cells. *Nature* **280**(5720):339–340 DOI [10.1038/280339a0](https://doi.org/10.1038/280339a0).
- Inoki K, Zhu T, Guan KL. 2003. TSC2 mediates cellular energy response to control cell growth and survival. *Cell* **115**:577–590 DOI [10.1016/S0092-8674\(03\)00929-2](https://doi.org/10.1016/S0092-8674(03)00929-2).
- Jann O, Ceriotti G, Caroli A, Erhardt G. 2002. A new variant in exon VII of bovine beta-casein gene (CSN2) and its distribution among European cattle breeds. *Journal of Animal Breeding and Genetics* **119**:65–68 DOI [10.1046/j.1439-0388.2002.00318.x](https://doi.org/10.1046/j.1439-0388.2002.00318.x).
- Jeck WR, Sharpless NE. 2014. Detecting and characterizing circular RNAs. *Nature Biotechnology* **32**(5):453–461 DOI [10.1038/nbt.2890](https://doi.org/10.1038/nbt.2890).
- Joshi NA, Fass JN. 2011. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files. Version 1.33. Available at <https://github.com/najoshi/sickle>.

- Kim D, Paggi JM, Park C, Bennett C, Salzburg SL. 2019.** Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology* 37(8):907–915 DOI 10.1038/s41587-019-0201-4.
- Liu H, Liu D, Liu Y, Xia M, Li Y, Li M, Liu H. 2020.** Comprehensive analysis of circRNA expression profiles and circRNA-associated competing endogenous RNA networks in IgA nephropathy. *PeerJ* 8:e10395 DOI 10.7717/peerj.10395.
- Li F, Zhang L, Li W, Deng J, Zheng J, An M, Lu J, Zhou Y. 2015.** Circular RNA ITCH has inhibitory effect on ESCC by suppressing the Wnt/ $\beta$ -catenin pathway. *Oncotarget* 6(8):6001–6013 DOI 10.18632/oncotarget.3469.
- Liu X, Robinson GW, Wagner KU, Garrett L, Wynshaw-Boris A, Hennighausen L. 1997.** Stat5a is mandatory for adult mammary gland development and lactogenesis. *Genes and Development* 11(2):179–186 DOI 10.1101/gad.11.2.179.
- Livak KJ, Schmittgen TD. 2001.** Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta CT}$  method. *Methods* 4:402–408 DOI 10.1006/meth.2001.1262.
- Ma D, Zhao X, Yu Y, Zhang S, Cheng M, Cao HF, et al. 2019.** CircRNA as CeRNA mediated by microRNA may be involved in goat lactation. *Small Ruminant Research* 171:63–72 DOI 10.1016/j.smallrumres.2018.12.009.
- Macias H, Hinck L. 2012.** Mammary gland development. *Wiley Interdisciplinary Reviews: Developmental Biology* 1(4):533–557 DOI 10.1002/wdev.35.
- Martin M. 2011.** Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet Journal* 17:10–12.
- McCarthy DJ, Chen Y, Smyth GK. 2012.** Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Research* 40(10):4288–4297 DOI 10.1093/nar/gks042.
- Medzhitov R, Kagan JC. 2006.** Phosphoinositide-mediated adaptor recruitment controls Toll-like receptor signaling. *Cell* 125:943–955 DOI 10.1016/j.cell.2006.03.047.
- Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, Maier L, Mackowiak SD, Gregersen LH, Munschauer M, Loewer A, Ziebold U, Landthaler M, Kocks C, Noble Fle, Rajewsky N. 2013.** Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* 495(7441):333–338 DOI 10.1038/nature11928.
- Miglior F, Fleming A, Malchiodi F, Brito LF, Martin P, Baes CF. 2017.** A 100-Year Review: identification and genetic selection of economically important traits in dairy cattle. *Journal of Dairy Science* 100(12):10251–10271 DOI 10.3168/jds.2017-12968.
- Morris KV, Mattick JS. 2014.** The rise of regulatory RNA. *Nature Reviews Genetics* 15(6):423–437 DOI 10.1038/nrg3722.
- Mumtaz PT, Taban Q, Dar MA, et al. 2020.** Deep insights in circular RNAs: from biogenesis to therapeutics. *Biological Procedures* 22(10): DOI 10.1186/s12575-020-00122-8.
- Petz LN, Ziegler YS, Schultz JR, Kim H, Kemper JK, Nardulli AM. 2004.** Differential regulation of the human progesterone receptor gene through an estrogen response element half site and Sp1 sites. *Journal of Steroid Biochemistry and Molecular Biology* 88(2):113–122 DOI 10.1016/j.jsbmb.2003.11.008.



- Putra ABN, Nishi K, Shiraishi R, Doi M, Sugahara T. 2014.** Jellyfish collagen stimulates production of TNF- $\alpha$  and IL-6 by J774.1 cells through activation of NF- $\kappa$ B and JNK via TLR4 signaling pathway. *Molecular Immunology* **58**:32–37 DOI [10.1016/j.molimm.2013.11.003](https://doi.org/10.1016/j.molimm.2013.11.003).
- Rybak-Wolf A, Stottmeister C, Glazar P, Jens M, Pino N, Giusti S, Hanan M, Behm M, Bartok O, Ashwal-Fluss R, et al. 2015.** Circular RNAs in the mammalian brain are highly abundant, conserved, and dynamically expressed. *Molecular Cell* **58**(5):870–885 DOI [10.1016/j.molcel.2015.03.027](https://doi.org/10.1016/j.molcel.2015.03.027).
- Sandhu GK, Milevskiy MJG, Wilson W, Shewan AM, Brown MA. 2016.** Non-coding RNAs in Mammary Gland Development and Disease. *Advances in Experimental Medicine and Biology* **886**:121–153 DOI [10.1007/978-94-017-7417-8\\_7](https://doi.org/10.1007/978-94-017-7417-8_7).
- Schindewolf C, Braun S, Domdey H. 1996.** In vitro generation of a circular exon from a linear pre-mRNA transcript. *Nucleic Acids Research* **24**:1260–1266 DOI [10.1093/nar/24.7.1260](https://doi.org/10.1093/nar/24.7.1260).
- Szabo L, Salzman J. 2016.** Detecting circular RNAs: bioinformatic and experimental challenges. *Nature Reviews Genetics* **17**:679–692 DOI [10.1038/nrg.2016.114](https://doi.org/10.1038/nrg.2016.114).
- Taniguchi CM, Kondo T, Sajan M, Luo J, Bronson R, Asano T, Farese R, Cantley LC, Kahn CR. 2006.** Divergent regulation of hepatic glucose and lipid metabolism by phosphoinositide 3 kinase via Akt and PKC $\lambda$ /zeta. *Cell Metabolism* **3**(5):343–353 DOI [10.1016/j.cmet.2006.04.005](https://doi.org/10.1016/j.cmet.2006.04.005).
- Van Eenennaam A, Medrano JF. 1991.** Milk protein polymorphisms in California dairy cattle. *Journal of Dairy Science* **74**(5):1730–1742 DOI [10.3168/jds.S0022-0302\(91\)78336-7](https://doi.org/10.3168/jds.S0022-0302(91)78336-7).
- Wang J, Zhou H, Hickford J, Hao Z, Gong H, Hu H, et al. 2021.** Identification and characterization of circular RNAs in mammary gland tissue from sheep at peak lactation and during the nonlactating period. *Journal of Dairy Science* **104**(2):2396–2409 DOI [10.3168/jds.2020-18911](https://doi.org/10.3168/jds.2020-18911).
- Wang M, Moisés S, Khan MJ, Wang J, Bu D, Loor JJ. 2012.** MicroRNA expression patterns in the bovine mammary gland are affected by stage of lactation. *Journal of Dairy Science* **95**(11):6529–6535 DOI [10.3168/jds.2012-5748](https://doi.org/10.3168/jds.2012-5748).
- Wu J, et al. 2006.** KOBAS server: a web-based platform for automated annotation and pathway identification. *Nucleic Acids Research* **34**(suppl\_2):W720–W724 DOI [10.1093/nar/gkl167](https://doi.org/10.1093/nar/gkl167).
- Zhang C, Wu H, Wang Y, Zhao Y, Fang X, Chen C, Chen H. 2015.** Expression patterns of circular RNAs from primary kinase transcripts in the mammary glands of lactating rats. *Journal of Breast Cancer* **18**(3):235–241 DOI [10.4048/jbc.2015.18.3.235](https://doi.org/10.4048/jbc.2015.18.3.235).
- Zhang C, Wu H, Wang Y, Zhu S, Liu J, Fang X, Chen H. 2016a.** Circular RNA of cattle casein genes are highly expressed in bovine mammary gland. *Journal of Dairy Science* **99**(6):4750–4760 DOI [10.3168/jds.2015-10381](https://doi.org/10.3168/jds.2015-10381).
- Zhang Y, Xue W, Li X, Zhang J, Chen S, Zhang J, Yang L, Chen L. 2016b.** The biogenesis of nascent circular RNAs. *Cell Reports* **15**(3):611–624 DOI [10.1016/j.celrep.2016.03.058](https://doi.org/10.1016/j.celrep.2016.03.058).

- Zhang Y, Zhang XO, Chen T, Xiang JF, Yin QF, Xing YH, Zhu S, Yang L, Chen LL. 2013.** Circular intronic long noncoding RNAs. *Molecular Cell* **51(6)**:792–806 DOI [10.1016/j.molcel.2013.08.017](https://doi.org/10.1016/j.molcel.2013.08.017).
- Zhao S, et al. 2018.** RnaSeqSampleSize: real data based sample size estimation for RNA sequencing. *BMC Bioinformatics* **19(1)**:1–8 DOI [10.1186/s12859-017-2006-0](https://doi.org/10.1186/s12859-017-2006-0).
- Zwierzchowski L. 2005.** Cattle genomics–functional polymorphism in milk protein genes and other genes related to milk and meat production. In: *Workshop on genomics and bioinformatics in animal biotechnology* 31.1.–4.2. Jastrzebiec, Poland.