Pakistan Veterinary Journal

ISSN: 0253-8318 (PRINT), 2074-7764 (ONLINE)
DOI: 10.29261/pakvetj/2020.101

RESEARCH ARTICLE

Global Long Intergenic Non-coding RNA Expression Profiles in Response to Clinical Ketosis in Transition Dairy Cows

Zhou-Lin Wu[§], Kun Du[§], Shi-Yi Chen, Xianbo Jia, Jie Wang and Song-Jia Lai[*]

Farm Animal Genetic Resources Exploration and Innovation Key Laboratory of Sichuan Province, Sichuan Agricultural University, Chengdu 611130, China
[§]These Authors contributed equally to this work.
*Corresponding author: laisj5794@163.com

ARTICLE HISTORY (20-347)

Received:	July 01, 2020
Revised:	October 27, 2020
Accepted:	November 01, 2020
Published online:	December 14, 2020

Key words:
dairy cow
clinical ketosis
transcriptome
long intergenic non-coding RNA

ABSTRACT

Many studies showed that long intergenic non-coding RNAs (lincRNAs) have multifaceted roles in various biological processes. However, roles and profiles of lincRNAs in clinical ketosis of dairy cows are still poorly understood. In this study, blood samples of both nine cows affected with clinical ketosis (i.e., case group) and ten healthy ones as control group at 2 weeks before as well as 2 weeks after parturition were collected for genome-wide analysis of lincRNA transcriptome. A total of 2154 novel intergenic transcripts were obtained from all the 38 RNA-Seq libraries, with 795 lincRNAs were identified. By comparing to protein coding transcripts, the characteristics reported in long non-coding RNAs were observed in our identified lincRNAs, such as shorter in transcript length, lower expression level, and fewer exons. Furthermore, we found thirteen differentially expressed lincRNAs (DELs) between case and control group. The *cis*- and *trans*-regulatory analysis obtained 124 target genes for these 13 DELs that were functionally involved in metal ion binding, cation binding, biological stress response, and energy metabolism. This study provides a comprehensive expression profile of lincRNAs associated with the progression and onset of clinical ketosis in transition dairy cows, which will facilitate further researches on the functional roles of non-coding RNAs on cow ketosis biology.

To Cite This Article: Wu ZL, Du K, Chen SY, Jia X, Wang J and Lai SJ, 2020. Global long intergenic non-coding RNA expression profiles in response to clinical ketosis in transition dairy cows. Pak Vet J. http://dx.doi.org/10.29261/pakvetj/2020.101

INTRODUCTION

The transition from late pregnancy to early lactation is known to be the most challenging period for dairy cows. During which, reduced feed intake and, at the same time, increased nutrient requirement for milk production induces a physiological state of negative energy balance (Esposito *et al*., 2014). One of the main causes of negative energy balance is lipid mobilization from stored body fat, and which may predispose the animals to fatty liver and ketosis (García *et al*., 2011). It is widely known that ketosis is associated with hepatic lipid accumulation and increased concentration of biofluid hydroxybutyrate while blood glucose concentration is decreased. In addition, inappetence, rapid weight loss, and dry manure also characterize the ketotic state (Gordon *et al*., 2013). Recently, alterations in transcriptome, metabolome and proteome between ketotic and non-ketotic cows have been described, although the mechanism of the incidence and progression of ketosis is not completely understood. Several publications so far have focused on the detection of the associated coding genes and gene regulatory networks involved in ketosis (Kroezen *et al*., 2018; Parker Gaddis *et al*., 2018). However, few have focused specially on non-coding genes.

Previous study has revealed that the vast majority of eukaryotic transcripts are non-coding RNAs (ncRNAs) (Lander *et al*., 2001). Generally, ncRNAs are arbitrarily defined by the length of mature transcripts, which conventionally divides ncRNAs into long non-coding RNAs (lncRNAs) that exceed 200 bp in length and the remaining ones that are called short ncRNAs (Ma *et al*., 2013). Based on the location in the genome compared to their nearby protein coding RNAs, lncRNAs can be categorized into senses, antisense, intronic, and intergenic lncRNAs (lincRNAs) (Ma *et al*., 2013). Of the various

lncRNAs classes, lincRNAs do not cover protein-coding genes or other types of genes. It is easier to identify and study their functions than those of transcripts with regions shared by other gene classes (Ulitsky and Bartel, 2013). A rapidly growing number of lincRNAs are reported to play key roles, including epigenetic modification, mRNA splicing, transcriptional regulation, and other post-transcriptional regulation (Dykes and Emanueli, 2017).

In the past few years, lncRNAs have been discovered in large diversity of species, including *Homo sapiens*, rat, chicken, goat, pig, and bovine. It is likely that lncRNAs contribute significantly to economic production traits and phenotypes in livestock, such as milk performance (Cai *et al*., 2018), metabolic efficiency (Nolte *et al*., 2019) and fat metabolism (Yu *et al*., 2017). However, limited studies that have addressed the associations between global lincRNAs and cow ketosis. Therefore, in this study, we employed RNA-Seq analysis from blood samples collected at prepartum and postpartum to explore not only potential mechanisms that are associated with postpartum onset of clinical ketosis, but also the profile changes from prepartum to postpartum for ketotic cows, which would provide a useful resource for lincRNAs involved in ketosis and lay the foundation to investigate their functions.

MATERIALS AND METHODS

Animals management: The animals and experimental design were previously described by us (Wu *et al*., 2020). Briefly, seventy-four pregnant Holsteins at 3 weeks before the expected parturition were raised in the same environment. Clinical ketosis case was determined as whose β-hydroxybutyrate (BHBA) concentration exceed 2.60 mM and control with under 1.0 mM at 2 weeks after parturition, respectively. Out of the 74 animals, only 9 were ultimately developed into clinical ketosis and 10 remained controls at 2 weeks after parturition. For which, blood samples from both prepartum (2 weeks before parturition) and postpartum (2 weeks after parturition) were collected and then classified into four different groups: case (CK; BHBA=2.79±0.12mM; N=9) and whose blood at prepartum (PCK; BHBA=0.36±0.05mM; N = 9) and control (HC; BHBA=0.65±0.22mM; N = 10) and whose blood at prepartum (PHC; BHBA=0.42±0.08 mM; N = 10), respectively.

Blood collection, RNA-Seq library construction and sequencing: Before the morning feeding, blood was sampled via coccygeal vein using vacutainer tubes containing EDTA-K2. RNA sample was isolated from blood using TRIzol Reagent (TaKaRa, Dalian, China) and then purified using RNeasy Midi Kit (Qiagen, Valencia, CA, USA) according to manufacturer's protocol. The sequencing libraries were created from approximately 1 μg high quality total RNA using Illumina NEBNext Ultra RNA Library Prep Kit (NEB, Ipswich, MA, USA). The library quality was assessed on an Agilent Bioanalyzer 2100 System and the sequencing was performed on an Illumina Hiseq X Ten platform that generate >15G 150bp paired-end raw reads per sample.

Reads mapping and transcript assembly: The adapter sequences removal and low-quality filtering for raw reads were performed to get the clean reads using fastp software (Chen *et al*., 2018). Clean reads were then aligned to the bovine reference genome (ARS-UCD1.2.95) using HISAT2 (Kim *et al*., 2015). Next, StringTie software (Pertea *et al*., 2016) was enrolled to assemble the mapped reads, and Cuffmerge was used to merge the assembled transcripts together to build a consensus set of transcripts across samples. To identify new transcripts, the merged transcripts were compared to the reference annotation GTF file, by which, only the novel transcripts (those annotated by "u") were subjected to further analyses. Fragments per kilobase per million mapped reads (FPKM) was applied to measure transcript expression level.

Screening of lincRNAs and differential expression analysis: The novel transcripts were filtered and assembled in order to get the putative lincRNAs, the multi-step pipeline was followed as Yang *et al*. (Yang *et al*., 2018) described with little modification. (1) Only novel intergenic transcripts were retained; (2) transcripts with only one exon or shorter than 200 nt were discarded; (3) transcripts with protein-coding potency were removed with the CPC, Pfam, CNCI, and CPAT software; (4) transcripts whose maximal FPKM <0.5 among all samples were removed; The remaining transcripts were considered as lincRNAs and then were compared to the NONCODE version v5.0 database (http://www.noncode.org/index.php) using BLASTN with P-value < 1e-06. The differentially expressed lincRNAs (DELs) between two groups were identified using DESeq2 R package (Love *et al*., 2014), under the criteria of |log2 fold change| >1 and Bonferroni-corrected p-value <0.05.

Target gene prediction and functional enrichment analysis: To reveal the potential functions of our lincRNAs, the protein coding genes located in 100kb upstream and downstream of a lincRNA were screened and selected as DELs potential *cis*-regulated genes, and the LncTar bioinformatics software (Li *et al*., 2014) was enrolled to predict *trans*-target genes of the DELs. Gene Ontology (GO) and KEGG enrichment analysis were performed using KOBAS 3.0 online tool (http://kobas.cbi.pku.edu.cn/).

RESULTS

Summary of sequencing data: A total of 626.15 G raw sequencing data were produced among all samples, which ultimately produced 581.22 G clean data with an average of 15.30 G per sample after quality filtering. Approximately 92.68% of the total reads were mapped to the bovine reference genome by HISAT2. Of these, 88.77% (range: 83.16-91.91%) were uniquely mapped reads and 3.90% (3.03-7.05%) were multi-mapped reads (Supplementary Table 1). All the mapped reads were considered in further analysis.

Identification and characterization of lincRNAs: Transcripts of each sample were assembled and then merged using StringTie merge functionality, resulting in 84,528 nonredundant transcript isoforms. After removing the annotated transcripts in the reference annotation GTF file, a total of 2154 intergenic transcripts were determined.

Table 1: The differentially expressed lincRNAs between CK and HC group

LincRNA transcript	Chromosome	Start	End	Base Mean	Log2 (Fold Change)		pvalue
MSTRG.13328.3	29	43,671,609	43,678,849	148.75	27.18	7.73E-17	1.64E-13
MSTRG.5513.3	17	5,124,476	5,221,483	7.97	23.22	2E-08	9.92E-06
MSTRG.13328.5	29	43,671,610	43,678,827	22.24	21.92	1.7E-07	6.09E-05
MSTRG.5513.6	17	5,126,634	5,203,388	9.26	15.34	0.000254	0.036089
MSTRG.19978.8	X	94,560,195	94,610,650	303.97	12.84	5.24E-05	0.010176
MSTRG.12485.4	26	43,825,465	43,834,484	398.87	2.42	9.8E-05	0.017454
MSTRG.16598.1	5	61,167,398	61,174,123	101.86	2.29	1.64E-06	0.000465
MSTRG.3049.3	12	84,535,837	84,547,749	847.74	2.14	0.000292	0.039645
MSTRG.5427.2	16	75,638,996	75,640,384	21.08	-22.50	7.21E-19	2.38E-15
MSTRG.16046.11	5	101,472,480	101,516,247	39.20	-22.38	2.22E-11	2.53E-08
MSTRG.19167.1	NKLS02000125.1	1,983	153,026	14.46	-16.81	5.38E-05	0.010316
MSTRG.3623.3	13	65,178,885	65,184,020	346.82	-9.87	1.23E-05	0.002993
MSTRG.17342.1	7	19,551,335	19,552,741	25.52	-4.57	0.000346	0.045324

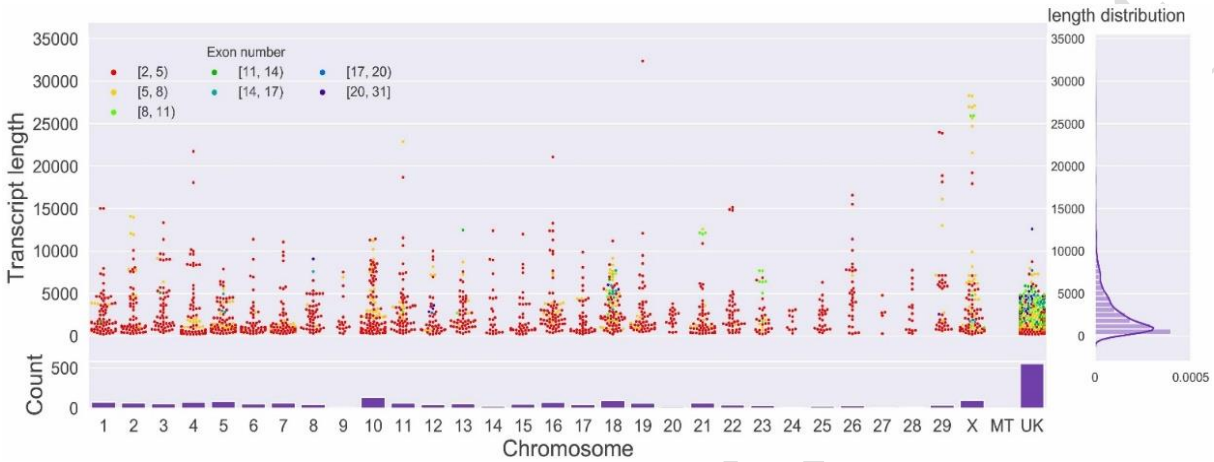

Fig. 1: Genome characterization of novel intergenic transcripts.

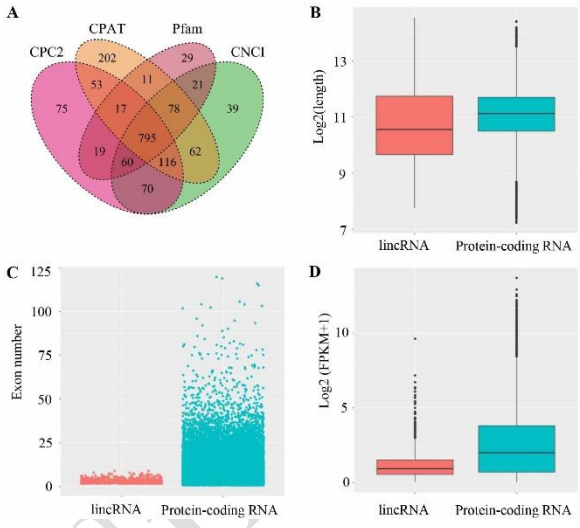

Fig. 2: Identification and characterization of long intergenic non-coding RNA (lincRNA). (A) Venn diagram showing the lincRNAs identified by four softwares of CPC2, CPAT, Pfam and CNCI; (B) Boxplot indicating transcript length of lincRNAs and protein-coding RNAs by log2(length); (C) The exon number of lincRNAs compared to protein-coding RNAs; (D) Expression level of lincRNAs and protein-coding RNAs based on log2(FPKM+1).

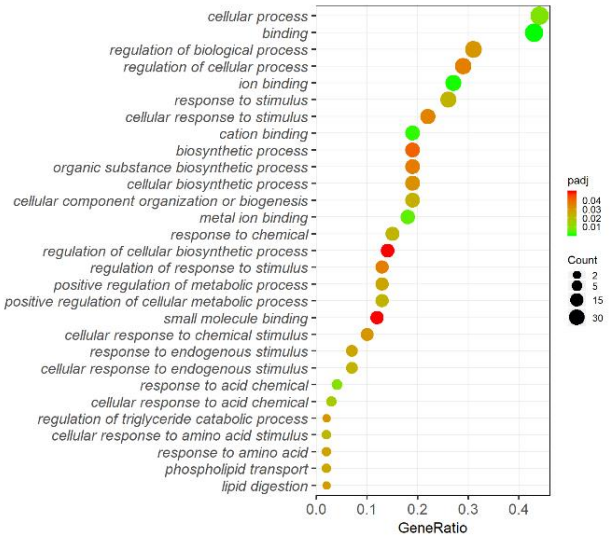

Fig. 3: GO annotation of target genes for differentially expressed lincRNAs (DELs) between case and control group.

The genome characterization of those transcripts was shown in Fig. 1. The results revealed that the majority of these transcripts (71.82 %) have 2-5 exons. More than half of these transcripts are in lengths shorter than 1750 bp, and about 30% of such transcripts are in lengths shorter than 1000 bp. Moreover, chromosome 10 contains the highest number of novel intergenic transcripts (137), while chromosome 27 contains lowest number of novel intergenic transcripts (5) (Fig. 1).

A rigorous filtering listed in Materials and Methods was used to detect lincRNAs. Among them, 795 lincRNAs were finally identified from the 2154 novel intergenic transcripts (Figure 2A). Next, a total of eleven low expressed lincRNAs were filtered out due to the fact that extremely low expression was generally considered to be transcriptional noise, and therefore, the remaining 784 lincRNAs were robustly expressed (Supplementary Table 2). Using the BLASTN algorithm, a total of 489 lincRNAs were found in the bovine NONCODE database and 295 novel lincRNAs were obtained. To further examine structural characteristics of our identified

lincRNAs, the transcript length, exon number and expression level were compared between the lincRNAs and protein-coding RNAs. As shown in Figure 2, the bovine lincRNAs were shorter in length than protein-coding RNAs, and the lincRNAs had, on average, fewer exons per transcript (~3.04) than protein-coding RNAs (~11.73). Further, the average expression of protein-coding RNAs was more than 7.10-fold that of lincRNAs.

Differentially expressed lincRNAs between cows diagnosed with clinical ketosis and controls: To study the key lincRNAs involved in the bovine clinical ketosis, differentially expressed lincRNAs (DELs) between cows diagnosed with clinical ketosis and controls, and ketotic cows at prepartum and postpartum were examined. A total of 13 DELs were identified between the CK and HC group, with 8 up-regulated and 5 down-regulated in the CK group (Table 1). On the other hand, we found 14 DELs between the CK and PCK group, which would help to reveal the profile changes from prepartum to postpartum for ketotic cows. Complete list of those DELs can be found in Supplementary Table 3.

Target gene prediction and function analysis: To investigate the potential function of these DELs, their potential target genes were predicted in *cis*- and *trans*-roles. The potential target genes for those 13 DELs from case and control group that corresponded to 124 protein-coding genes (Supplementary Table 4). Our GO annotation indicated that target genes were significantly enriched in cellular process, metal ion binding, cation binding, biological stress response (such as "cellular response to stimulus", "cellular response to chemical stimulus", "regulation of response to stimulus", "cellular response to amino acid stimulus and endogenous stimulus", etc.), and energy metabolism (such as "primary metabolic process", "lipid digestion", "phospholipid transport", and "triglyceride catabolic process") (Figure 3 and Supplementary Table 5). As for the pathways, target genes for those lincRNAs were involved in several KEGG pathways such as cell adhesion molecules (CAMs), MAPK, PI3K-Akt, NF-kappa B, metabolic pathways, and disease-related pathways (Supplementary Table 6).

Meanwhile, we predicted the potential target genes for the DELs from CK vs. PCK, and found that those 14 DELs corresponded to 122 protein-coding genes (Supplementary Table 7). GO functional annotation showed that there target genes were enriched in different categories, such as biological processes of positive regulation of protein kinase B signaling, biological regulation, cellular process, biosynthetic process, molecular function of transporter activity, ion binding, metal ion binding, calcium ion binding, cation binding, cellular component of hemoglobin complex, vesicle membrane, intracellular (Supplementary Table 8). In addition, these target genes were involved in AMPK, FoxO, NF-kappa B, TCA cycle, RIG-I-like receptor signaling pathway (Supplementary Table 9).

DISCUSSION

In the current study, a total of 795 lincRNAs were identified, which are distinct from protein-coding RNAs as to length profiles, exon numbers and expression levels. Our data showed that bovine lincRNAs were shorter in length than protein-coding RNAs, consistent with a previous study conducted in bovine mammary gland tissue (Zheng *et al*., 2018). We also found that the lincRNAs had fewer exons per transcript and lower expression level than protein-coding RNAs. These values are similar to those estimated for bovine and the other mammalian lncRNA (Gao *et al*., 2017; Du *et al*., 2020). Our results provided supplemental evidence for future research into bovine lincRNA.

Emerging evidence supports the view that lincRNAs participate in the regulation of numerous foundational biological processes. For example, knockdown of dozens of lincRNAs on mouse embryonic stem cells suggested that lincRNAs are required to maintain pluripotency (Guttman *et al*., 2011), and mouse lincRNA knockout models revealed that lincRNAs play critical roles in peri- and postnatal lethal phenotypes (Sauvageau *et al*., 2013). So far, functionally characterized lincRNAs have been found to work in *cis* or *trans* (Guttman *et al*., 2011; Tan *et al*., 2017). In this study, we found 13 DELs between case and control group, for which, we got their *cis*-regulated target genes based on the threshold of 100 kb and *trans*-regulated target genes using LncTar bioinformatics software. These target genes may provide supplementary evidence for underlying potential mechanisms associated with onset of clinical ketosis. To this end, we performed GO enrichment analysis of the target genes for these 13 DELs from CK vs. HC. We found that these genes were enriched in different GO categories including cellular response to stimulus, cellular response to endogenous stimulus and amino acid stimulus, and those pathways have been shown to be linked with ketosis in our previous work (Wu *et al*., 2020). It was also revealed by GO analysis that lipid digestion, phospholipid transport, and triglyceride catabolic process were enriched. Most recently, Wang *et al* (Wang *et al*., 2016) and Shahzad *et al* (Shahzad *et al*., 2019) have suggested that pathways associated with carbohydrate metabolism, lipid metabolism, amino acid and vitamin metabolism were induced in ketotic cows. Besides, some genes involved in hepatic glucose metabolism, carbohydrate, lipid, carbohydrate, glycan biosynthesis and metabolism are associated with ketosis (Shahzad *et al*., 2019). In the present study, some of these target genes were involved in several KEGG pathways, such as cell adhesion molecules, MAPK, PI3K-Akt, NF-kappa B, metabolic pathways, and disease-related pathways. These results suggest that lincRNAs may play crucial role in the onset of clinical ketosis by mediating various biological responses and pathways, but further works are required to understand the function of both these lincRNAs and whose target genes.

It is clear that the expression of hepatic lipid metabolic genes involved in fatty acid uptake, transport, activation, oxidation, synthesis, and esterification were mainly regulated by PPARα and SREBP-1c (Li *et al*., 2015). More important, the expression disorder of lipogenic genes and triglycerides synthesis genes may participate in the development of ketosis (Zhu *et al*., 2019). However, very few works have been done to determine their role in the development of ketosis. Previous studies have shown that ketosis is a multifactorial

disease, and which is controlled by multiple genetic loci. Genetic evaluations showed that ketosis has a low heritability of 0.02-0.06 (Kroezen *et al*., 2018), which limits the genetic selection for ketosis resistance. Our current study hopefully provides interesting insight into the lincRNAs and their potential target genes related to cow ketosis. Of course, other deeply techniques as well as more researches on ncRNAs will help to further understanding the mechanism of ketosis and finding more candidate genes for ketosis resistance.

Conclusions: In summary, our results comprehensively revealed the lincRNA profiles associated with clinical ketosis in the transition dairy cows. The functional enrichment analysis revealed that the lincRNAs may influence multiple pathways in relation to metal ion binding, cation binding, biological stress response, and energy metabolism. This study would provide a source for bovine genetic study involving in ketosis biology.

Funding: This work was founded by the Sichuan Provincial Science and Technology Development Plan Project [2018NZ0003].

Authors contribution: ZLW: Methodology, Investigation, Manuscript preparation. KD: Investigation, Software. SYC, XJ and JW: Manuscript review. SJL: Supervision and Project administration.

REFERENCES

Cai W, Li C, Liu S, *et al*., 2018. Genome Wide Identification of Novel Long Non-coding RNAs and Their Potential Associations With Milk Proteins in Chinese Holstein Cows. Front Genet 9:281.

Chen S, Zhou Y, Chen, *et al*., 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Y Bioinformatics 34:i884-90.

Du K, Ren AY, Cai MC, *et al*., 2020. Identification of long non-coding RNAs in the early growth stage of Holstein mammary gland. Eur Zool J 87:214-22.

Dykes IM and Emanueli C, 2017. Transcriptional and Post-transcriptional Gene Regulation by Long Non-coding RNA. Genom Prot Bioinform 15:177-86.

Esposito G, Irons PC, Webb EC, *et al*., 2014. Interactions between negative energy balance, metabolic diseases, uterine health and immune response in transition dairy cows. Anim Reprod Sci 144:60-71.

Gao X, Ye J, Yang C, *et al*., 2017. Screening and evaluating of long noncoding RNAs in the puberty of goats. BMC Genom 18:164.

García AMB, Cardoso FC, Campos R, *et al*., 2011. Metabolic evaluation of dairy cows submitted to three different strategies to decrease the effects of negative energy balance in early postpartum. Pesq Vet Bras 31:11-7.

Gordon JL, LeBlanc SJ and Duffield TF, 2013. Ketosis Treatment in Lactating Dairy Cattle. Vet Clin North Am Food Anim Pract 29:433-45.

Guttman M, Donaghey J, Carey BW, *et al*., 2011. lincRNAs act in the circuitry controlling pluripotency and differentiation. Nature 477:295-300.

Kim D, Langmead B and Salzberg SL, 2015. HISAT: a fast spliced aligner with low memory requirements. Nat Methods 12:357-60.

Kroezen V, F Schenkel, F Miglior, *et al*., 2018. Candidate gene association analyses for ketosis resistance in Holsteins. J Dairy Sci 101:5240-9.

Lander ES, Linton LM, Birren B, *et al*., 2001. Initial sequencing and analysis of the human genome. Nature 409:860-921.

Li J, Ma W, Zeng P, *et al*., 2014. LncTar: a tool for predicting the RNA targets of long noncoding RNAs. Brief Bioinform 16:806-12.

Li X, Huang W, Gu J, *et al*., 2015. SREBP-1c overactivates ROS-mediated hepatic NF-κB inflammatory pathway in dairy cows with fatty liver. Cell Signal 27:2099-109.

Love MI, Huber W and Anders S, 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550.

Ma L, Bajic VB and Zhang Z, 2013. On the classification of long non-coding RNAs. RNA Biol 10:924-33.

Nolte W, Weikard R, Brunner RM, *et al*., 2019. Biological network approach for the identification of regulatory long non-coding RNAs associated with metabolic efficiency in cattle. Front Genet 10.

Parker K, Gaddis L, Megonigal Jr. JH, Clay JS, *et al*., 2018. Genome-wide association study for ketosis in US Jerseys using producer-recorded data. J Dairy Sci 101:413-24.

Pertea M, Pertea GM, Antonescu CM, *et al*., 2016. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol 33:290-5.

Sauvageau M, Goff LA, Lodato S, *et al*., 2013. Multiple knockout mouse models reveal lincRNAs are required for life and brain development. Elife 2:e01749.

Shahzad K, Lopreiato V, Liang Y, *et al*., 2019. Hepatic metabolomics and transcriptomics to study susceptibility to ketosis in response to prepartal nutritional management. J Anim Sci Biotechnol 10:96.

Tan J Y, Smith AAT, Ferreira da Silva M, *et al*., 2017. cis-Acting complex-trait-associated linc RNA Expression correlates with modulation of chromosomal architecture. Cell Rep 18:2280-8.

Ulitsky I and Bartel DP, 2013. lincRNAs: genomics, evolution and mechanisms. Cell 154:26-46.

Wang Y, Gao Y, Xia C, *et al*., 2016. Pathway analysis of plasma different metabolites for dairy cow ketosis. Ital J Anim Sci 15:545-51.

Wu Z-L, Chen S-Y, Qin C, *et al*., 2020. Clinical Ketosis-Associated Alteration of Gene Expression in Holstein Cows. Genes (Basel). 11:219.

Yang B, Jiao B, Ge W, *et al*., 2018. Transcriptome sequencing to detect the potential role of long non-coding RNAs in bovine mammary gland during the dry and lactation period. BMC Genom 19:605.

Yu L, Tai L, Zhang L, *et al*., 2017. Comparative analyses of long non-coding RNA in lean and obese pig. Oncotarget 8:41440-50.

Zheng X, Ning C, Zhao P, *et al*., 2018. Integrated analysis of long noncoding RNA and mRNA expression profiles reveals the potential role of long noncoding RNA in different bovine lactation stages. J Dairy Sci 101:11061-73.

Zhu Y, Liu G, Du X, *et al*., 2019. Expression patterns of hepatic genes involved in lipid metabolism in cows with subclinical or clinical ketosis. J Dairy Sci 102:1725-35.