

Expression of lncRNAs in response to bacterial infections of goat mammary epithelial cells reveals insights into mammary gland diseases

Peerzada Tajamul Mumtaz ^{a,b}, Qamar Taban ^a, Basharat Bhat ^a, Syed Mudasir Ahmad ^{a,*}, Mashooq Ahmad Dar ^a, Zahid Amin Kashoo ^c, Nazir A. Ganie ^a, Riaz Ahmad Shah ^a

^a Division of Animal Biotechnology, Faculty of Veterinary Sciences and Animal Husbandry, Shuhama, SKUAST-K, India

^b Department of Biochemistry, School of Life Sciences Jaipur National University, India

^c Division of Veterinary Microbiology, Faculty of Veterinary Sciences and Animal Husbandry, Shuhama, SKUAST-K, India

ARTICLE INFO

Keywords:

LncRNA
RNA Seq
pMECs
Goat
Bacterial infection
Immune response

ABSTRACT

Mastitis or inflammation of the mammary gland is a highly economic and deadly alarming disease for the dairy sector as well as policymakers caused by microbial infection. Transcriptomic and proteomic approaches have been widely employed to identify the underlying molecular mechanisms of bacterial infections in the mammary gland. Numerous differentially expressed mRNAs, miRNAs, and proteins together with their associated signaling pathways have been identified during bacterial infection, paving the way for analysis of their biological functions. Long noncoding RNAs (lncRNAs) are important regulators of multiple biological processes. However, little is known regarding their role in bacterial infection in mammary epithelial cells. Hence, RNA-sequencing was performed by infecting primary mammary epithelial cells (pMECs) with both gram-negative (*E. coli*) and gram-positive bacteria (*S. aureus*). Using stringent pipeline, a set of 1957 known and 1175 novel lncRNAs were identified, among which, 112 lncRNAs were found differentially expressed in bacteria challenged PMECs compared with the control. Additionally, potential targets of the lncRNAs were predicted in *cis*- and *trans*-configuration. KEGG analysis revealed that DE lncRNAs were associated with at least 15 immune-related pathways. Therefore, our study revealed that bacterial challenge triggers the expression of lncRNAs associated with immune response and defense mechanisms in goat mammary epithelial cells.

1. Introduction

Mastitis is an infectious disease of the mammary gland that is one of the most economically important diseases in the dairy industry for several decades [1]. Mastitis is caused by a wide spectrum of pathogens including Gram-negative *Escherichia coli* (*E. coli*) and Gram-positive *Staphylococcus aureus* (*S. aureus*) [2]. The mastitis caused by *E. coli* and other Gram-negative bacteria is often clinical with acute and severe inflammation, and the immune system may eventually clear the pathogens within days or with antibiotic treatment [3,4]. In contrast, infection with gram-positive bacteria like *S. aureus* often causes mild mastitis and the clearance of the pathogens by antibiotics is often ineffective [5]. Currently, antimicrobial treatment is indispensable to keep udder health, animal welfare, and economic aspects in balance. On the contrary, the emergence and spread of antimicrobial resistance is an urgent matter of particular public interest, and as a consequence, there is

an urgent need for future reduction in antimicrobial use in the dairy industry [6]. Efforts at mastitis control include understanding host response mechanisms to infecting pathogens and the development of appropriate control strategies. Cellular response to bacterial infections relies on well-defined networks of molecular interactions based on the regulation of gene expression and protein function. Recent reports suggest that lncRNAs may also play critical roles in the transcriptional regulation of gene expression during innate immune responses. Transcriptomic studies on bovine mammary epithelial cells challenged with mastitic pathogens and/or mammary gland tissues collected after intramammary infection revealed very different mechanisms in host innate immune responses to pathogens [6–8]. Differential cytokines and chemokines (at protein level) and other immune response proteins were also observed in bovine mammary epithelial cells and tissues or milk, in response to *E. coli* and *S. aureus* [9,10].

Long non-coding RNA (lncRNA), is characterized by being over 200

* Corresponding author. Division of Animal Biotechnology, Faculty of Veterinary Sciences and Animal Husbandry, Shuhama, Sher-e- Kashmir University of Agricultural Sciences and Technology, Kashmir, 19006, India.

E-mail address: mudasirbio@skuastkashmir.ac.in (S.M. Ahmad).

bp in length, often polyadenylated, and binding to a large number of chromatin-modifying proteins and guiding them to remodel the structure and/or expression of their neighboring genes [11]. Besides epigenetic control, lncRNAs can regulate transcriptional and post-transcriptional events (alternative splicing), RNA translation, and organize important structures for RNA processing such as nuclear speckles [12–14]. Notably, most studies indicate lncRNAs to play important roles in a certain tissue- and species-specific expression [15], but some multifunctional lncRNAs have displayed broadly and conservative expression profiles [16]. Recent great advances in RNA-Sequencing (RNA-seq) and computational methods for reconstructing transcriptome offer a wonderful opportunity to annotate and characterize lncRNAs, and a large number of lncRNAs have been discovered using RNA-sequencing [17–19]. Therefore, abundant RNA-seq data allow us to comprehensively identify and quantify lncRNAs (also protein-coding genes) and enable us to study the important roles of lncRNAs in various biological processes. Studies on the role of lncRNAs in resistance responses can benefit to understand the mammary gland's immune mechanism. The objectives of the present study were to (1) identify and characterize lncRNAs in response to gram-negative and gram-positive bacteria, and (2) explore the crosstalk network among lncRNAs and their target genes. Based on the RNA-seq data, we performed an analysis of lncRNAs and predicted their regulatory relationships with target genes in pMECs challenged with bacteria. A large amount of lncRNA information in the present study may provide clues to the functional and molecular studies of pMEC-bacteria interactions.

2. Materials and methods

2.1. Cell culture and pathogen challenge

PMEC cultures were established from the milk of early lactating goats at Mountain Research Centre for Sheep and Goat (MRCG) of Sher-e-Kashmir University of Agricultural Sciences and Technology, Kashmir (SKUAST-K). Before milking, the teats were cleaned with 1% iodine solution and the first three strips of milk were discarded. Approximately 30 ml of milk samples was collected into 50 ml sterile centrifuge tubes, kept in iceboxes, and transferred to the laboratory. The isolation process of PMECs was performed within 2 h after milk collection. Milk was defatted by centrifugation at 800g for 15 min. The supernatant and the fatty layer were discarded, and the cell pellet was resuspended with Dulbecco's phosphate-buffered saline (DPBS) without Ca^{2+} and Mg^{2+} (Thermo Fisher Scientific) and then centrifuged at 800 g for 10 min. This washing step was repeated twice, and finally, the cell pellet was resuspended with the culture medium. The cells were cultured in the humidified incubator with 5% CO_2 at 37 °C. After 24 h of culture, the floating cells were removed, the attached cells were washed twice with DPBS without Ca^{2+} and Mg^{2+} , and a fresh culture medium was added to the culture dish. The cells were allowed to grow in the incubator, and the culture medium was changed every 3 days. When cells reached 90% confluence, they were washed with DPBS (Himedia, Mumbai, India) after the culture medium was removed. The cells were then trypsinized with 1X trypsin (0.25% trypsin solution with 1 mM EDTA) and split 1:4 for passaging. The protocol used for the isolation and culture of PMEC was based on that for obtaining MEC from rodents [20]. We used gentamicin (50 µg/ml) in the tissue digestion medium, because gentamicin is known to induce apoptosis (programmed cell death) of fibroblasts [21]. This method helped to eliminate fibroblasts from the PMEC preparation. Primary cells were cryopreserved in freezing media composed of 70% DMEM, 20% FBS, and 10% dimethyl sulfoxide (DMSO, Carl Roth, Karlsruhe, Germany). Before starting the experiments, cells were thawed, seeded into 25 cm² culture flasks and cultured in complete medium for several days at 37 °C and 5% CO_2 in a humidified atmosphere until a confluent monolayer of epithelial cells was formed.

E. coli (ATCC 25922) and *Staphylococcus aureus* subsp. *aureus* *Rosebach* (ATCC 31890) were the infection agents. Bacteria were initially grown overnight on Luria Bertani (LB) agar (*E. coli*) or tryptic soy agar (TSA) (*S. aureus*) aerobically in a humidified incubator at 37 °C. A single colony of *S. aureus* was transferred to a 50 mL conical tube containing 20 ml of tryptic soy broth (TSB) and incubated at 37 °C in an open-air shaker at 225 RPM. Similarly, a single colony of *E. coli* was grown the same way in LB broth. The bacteria were grown up to the logarithmic phase of culture growth (Optical Density at 600 nm [OD 600] 0.5, ~5 × 10⁷/mL). Dilution series were plated to calibrate cell counts from the OD readings. Heat-inactivation of bacteria was performed at 63 °C for 30 min and verified by control plating. Afterward, bacteria were spun down at 3000 rpm for 15 min, washed twice with DME M/F12 medium, and resuspended herein at a density of 1 × 10⁸/mL. Aliquots were stored at -20 °C. Approximately 4.4 × 10⁵ of the isolated PMEC from each individual (three biological replicates) were seeded and cultured in collagen-coated 6-well plates in a complete medium without APS (three technical replicates per individual and treatment condition). On the next day, the medium was changed. Forty-eight hours after seeding, cells reached 90% confluence. PMEC were challenged with 10⁷/mL heat-inactivated *S. aureus* and *E. coli*, respectively for 24 h. Equivalent challenge treatments have been considered as robust cell stimulation based on previously published reports. After incubation periods, pathogen-challenged and unchallenged cells (control) were washed three times with phosphate-buffered saline (PBS) to remove the bacteria. Cells were collected for total RNA isolation.

2.2. RNA extraction, library preparation, and sequencing

The extracted MECs were washed by PBS and were subjected to RNA extraction by Trizol method (Ambion-USA) according to the manufacturer's instructions. RNA was quantified by spectrophotometer (ThermoFisher, USA) and the quality and integrity were assessed by bioanalyzer. IlluminaTruSeq Stranded mRNA Sample Prep kit was used to generate cDNA libraries according to the manufacturer's recommendations. Approximately ~4µg of total RNA was used to prepare the RNA seq library using the TruSeq RNA Sample Prep Kits (Illumina). The poly-A containing mRNA molecules were purified using poly-T oligo-attached magnetic beads. Following purification, the mRNA was fragmented into small pieces using divalent cations under elevated temperatures. The cleaved RNA fragments were used to synthesize first-strand cDNA using reverse transcriptase and random primers followed by the second strand cDNA synthesis using DNA polymerase I and RNase H. After adenylating of 3' ends of DNA fragments, hybridization was initiated by ligating Illumina PE adapter and index. cDNA fragments (200bp) were generated and were selectively enriched to construct the final sequencing library using Illumina PCR Primer Cocktail. An IlluminaHiSeq 2500 platform was used to sequence the libraries.

2.3. Quality control, mapping, and quantification

The raw reads from all samples were cleaned by removing the adapter sequences; low quality reads and reads containing undetermined bases through Trimmomatic software v0.32 [22]. Cleaned reads from all samples underwent quality control analysis using FastQC program version 0.11.9 (www.bioinformatics.babraham.ac.uk/projects/fastqc). To identify lncRNAs, first, we built the reference genome index through Bowtie (v2.2.6) [23] and used the spliced read aligner TopHat [24] version v2.0 (<http://tophat.cbcb.umd.edu/>) to map all sequenced reads to the *Capra hircus* reference genome (ARS1) downloaded from National Centre for Biotechnology Information (NCBI), with the following parameters: min-anchor = 5, min-isoform-fraction = 0 and other parameters set to default. Transcripts were assembled using Cufflinks [25] (version1.0.3) with default parameters (and 'min-frags-per-transfrag = 0'). We allowed up to 1-mismatches in the seed region and all multiple mapped positions were reported. Of all the filtered reads, about 82–87%

of reads in all samples were properly aligned to the reference genome (**Table 1**). After that, the transcripts from each sample were annotated using the ensemble database (v69) [26]. Transcripts (>200 bp) that do not overlap with any known annotation, localized in intronic, antisense, or intergenic region were filtered by CNCI and added to the lncRNA catalog of *Capra hircus*.

2.4. Evaluation of coding potential and lncRNA prediction using RNA-Seq data

To identify novel lncRNAs from PMECs with high confidence, we employed a highly stringent pipeline to remove transcripts with evidence of protein-coding potential (**Fig. 1A**). Firstly, single-exon transcripts and transcripts <200 nts long were filtered out. Next, four independent algorithms, CNCI (coding-non-coding index) [27], PLEK (predictor of long non-coding RNAs and messenger RNAs based on an improved k-mer scheme) [28] CPAT (coding potential assessment tool) [29], and Pfam (database providing alignments and hidden Markov models for protein domains) [30] were applied to extract potential non-coding transcripts. CPAT algorithm can rapidly recognize coding and non-coding transcripts from a large pool of candidates using a logistic regression model built with four sequence features: open reading frame size, open reading frame coverage, Fickett TEST CODE statistic, and hexamer usage bias. According to the studies, CPAT coding probability score ranges between 0 and 1, and the optimum cut-off for protein-coding probability varies depending on the species to be analyzed. To extract potential non-coding transcripts with high reliability from our dataset, we selected a very stringent threshold for the CPAT probability with a score <0.02 as ncRNA. The transcripts with a score between the selected thresholds were classified to possess an ambiguous coding potential. For further confidence, we employed Pfam Scan (v1.3) to our dataset to identify the occurrence of any of the known protein family domains documented in the Pfam database.

To further define the clear status of our transcripts we employed CNCI software, a powerful signature tool, which effectively distinguishes between protein-coding and non-coding sequences solely based on their intrinsic sequence composition i.e., by profiling adjoining nucleotide triplets (ANT). We also employed an alignment-free tool called PLEK, which uses a computational pipeline based on an improved k-mer scheme and a support vector machine (SVM) algorithm to distinguish lncRNAs from mRNAs as this tool has >90% accuracy, which is higher than other currently available tools. Transcripts anticipated with coding potential by any of the four tools were filtered out, and the final dataset contained transcripts shared by four tools as potential lncRNAs (**Fig. 1B**).

2.5. Differential expression analysis

For differential gene expression analysis, sequencing reads from each library were mapped to a stringently selected set of reference lncRNAs (31,195) using CuffDiff [31]. Differentially expressed genes (DEGs) between control and bacterial infected samples were screened based on the

Table 1
Read alignment summary.

Sample Name	Total reads	Pre -Processed Read Count	Aligned Read Count	Alignment %
C1	84,225,004	84,095,686	79,442,008	94.47
C2	82,443,136	82,562,531	78,876,132	95.53
C3	79,968,627	79,754,747	73,554,231	92.22
E 1	97,977,762	97,808,150	91,900,218	93.96
E 2	96,234,382	96,104,812	90,832,430	94.51
E 3	99,972,502	99,824,462	94,566,656	94.73
S1	86,904,596	86,744,380	80,534,760	92.84
S2	73,918,958	73,739,902	67,028,121	90.9
S3	78,266,258	78,106,360	71,377,698	91.39

threshold of p-value < 0.05 and absolute log2FC ≥ 1 . The expression value of lncRNAs was calculated in terms of RPKM (reads per kilobase per million). T-test was used to identify significantly differentially expressed lncRNAs and expression differences were considered significant at a fold-change cut-off value of ± 2 and a Benjamini-Hochberg corrected p-value of less than 0.05.

2.6. Gene ontology (GO) and kyoto encyclopedia of genes and genomes analysis (KEGG)

For each DE lncRNA, the nearest upstream and downstream (within 100 kb) protein-coding neighbors were identified as their *cis*-regulatory potential targets. To explore the roles of these target genes, we performed gene ontology (GO) (<http://www.geneontology.org>) and pathway analysis. GO terms are comprised of biological process (BP), cellular component (CC), and molecular function (MF). Pathway analysis is a functional analysis that maps genes to the KEGG (<http://www.genome.jp/kegg/>) pathways. GO terms and pathways with a P-value less than 0.05 were considered significantly enriched. KEGG allowed us to determine the biological pathways in which there is a significant enrichment of differentially expressed mRNAs.

2.7. Quantitative real-time PCR (qRT-PCR) validation of lncRNAs

To validate the repeatability and reproducibility of the RNA-seq data, quantitative real-time PCR was performed on 10 randomly selected lncRNAs including 5 up and 5 downregulated using the total RNA that was used for the sequencing. Primer 3 (<http://bioinfo.ut.ee/primer3-0.4.0/primer3/input.htm>) was used to design specific primers (**Table 2**). The expression levels of the selected lncRNAs were normalized against the housekeeping gene *GAPDH* and *UXT*. The qRT-PCRs were carried out in triplicate with the SYBR Green PCR kit (Roche, Germany) on a Light Cycler 480 (Roche Applied Science, Penzberg, Germany) using the following program: 95 °C for 10 min; 45 cycles of 95 °C for 10 s, 60 °C for 10 s, and 72 °C for 6 min. The relative gene expression was calculated using the $2^{-\Delta\Delta Ct}$ method [32].

3. Results

3.1. Identification and characterization of lncRNAs from RNA-seq data

High-throughput strand-specific RNA-seq was performed in the mammary epithelial cells infected with two types of bacteria *E. coli* and *S. aureus* (E, S), along with control (C), each with three biological replicates. A total of (778,741,030) raw reads were obtained from the nine libraries, with an average of 86.5 million reads per sample. After removing low-quality reads, 728,112,254 clean reads with high quality were retained (**Table 1**).

Clean reads mapped to the *Capra hircus* reference genome (ARS1) (https://www.ncbi.nlm.nih.gov/assembly/GCF_001704415.1/) were then used for transcript assembly and annotation. A total of 68,118 transcripts were initially generated. Then, to identify lncRNAs, transcripts mapped to goat reference genome (ARS1) mRNA, rRNA, etc., were filtered out, further single-exon transcripts and transcripts <200 nt length were filtered out (**Fig. 2C**). The remaining 3,527 transcripts were subsequently used for protein-coding capacity prediction by using the CPAT, CNCI, PFAM, and PLEK (**Fig. 1B**). Finally, 3,130 reliably expressed lncRNAs were obtained among which 1957 were known and 1175 were novel lncRNAs (supplementary 1). Further, the obtained lncRNAs were classified into four categories including ‘u’ (intergenic), ‘i’ (intronic), ‘x’ (anti-sense), and ‘o’ (sense-overlapping) according to their genomic location and referring to the neighboring genes (**Fig. 1C**). Specifically, the ‘u’ category contained transcripts falling in the intergenic regions between two protein-coding loci. The ‘i’ category contained transcripts falling entirely within an intron of a known protein-coding gene. The ‘x’ category contained transcripts that have generic

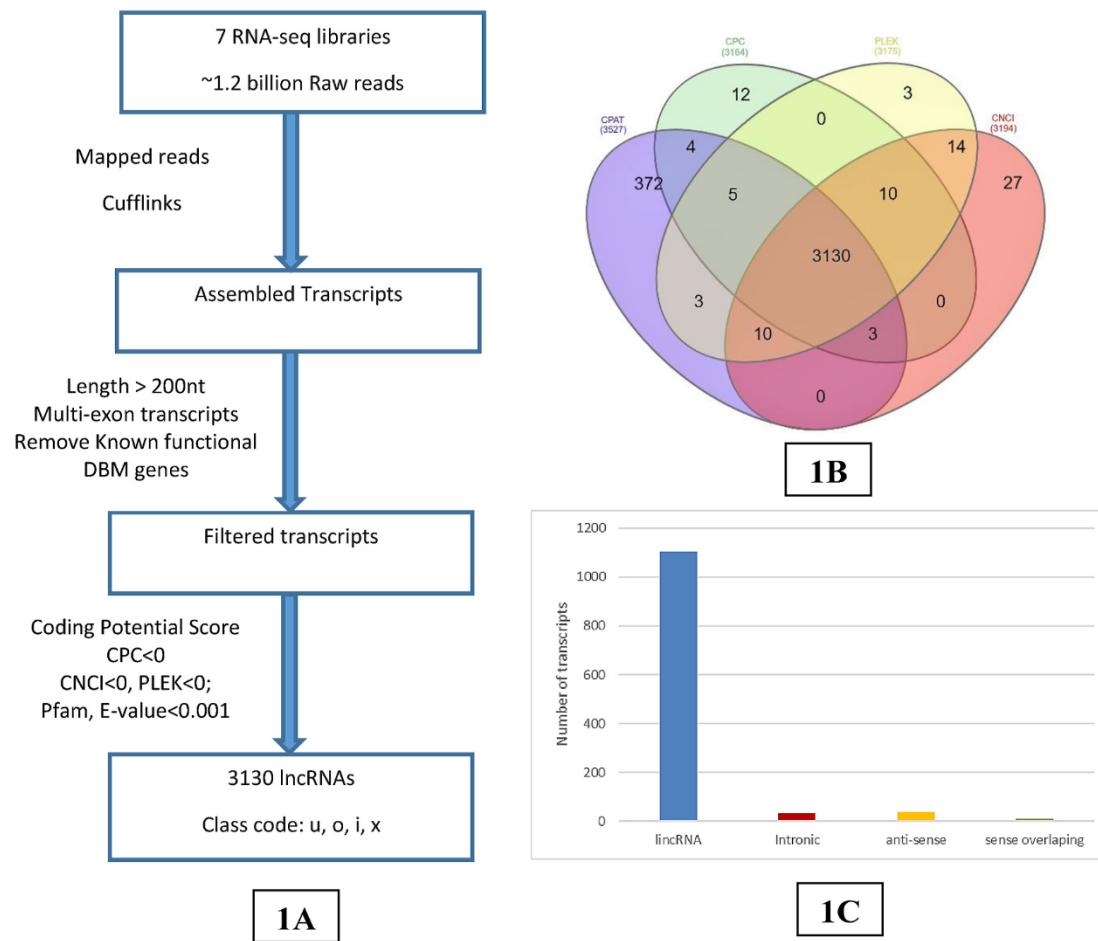


Fig. 1. The computational pipeline for identifying lncRNAs from RNA-seq data and their classification. (A) The lncRNA identification pipeline flowchart; (B) Coding potential analysis using the four methods; (C) The classification of identified lncRNAs, red rectangles or lines represent the exon or intron of a protein-coding gene, respectively; Blue, green, purple, and light blue rectangles or lines represent the exon or intron of lncRNA, respectively.

exonic overlap with a known protein-coding gene on the opposite strand. The 'o' category contained the transcripts partial overlapping with a coding gene on the same genomic strand (Fig. 1C).

Most of the identified lncRNAs fell into class u, with 1175 lncRNAs (94.04%), whereas 32 (2.72%), 3 (3.23%) and 4 (1.02%) lncRNAs belonged to classes i, x, and o, respectively (Fig. 1C). Previous studies in mammals have shown that the expression of lncRNAs is significantly lower than those of protein-encoding genes [33]. To determine whether mammary epithelial cell lncRNAs have similar features, we measured the expression level (in fragments per kilobase of exon per million fragments mapped (FPKM)) of the identified lncRNAs. We found lncRNAs show a lower level of expression compared to protein-coding mRNAs (Fig. 2A). The length and exon number of the identified lncRNAs were also analyzed. The size distribution of these lncRNAs ranged from 200 nucleotides to 3,000 nucleotides, with approximately 69.6% of lncRNAs shorter than 1500 nucleotides (Fig. 2B). Characterization of the genomic location revealed that the exon number of these lncRNAs ranged from 2 to 13; 1,036 (33.09%) identified lncRNAs had two exons, 386 (12.73%) had three exons, and 1104 (36.4%) lncRNAs had more than five exons (Fig. 2C).

3.2. Analysis of differentially expressed lncRNAs

Hierarchical clustering was used to analyze the expression profiles of lncRNAs in control or bacteria-infected mammary epithelial cells. As it was observed, the lncRNA expression profiles were significantly modified after infection (Fig. 3). In total 27 lncRNAs were identified as

differentially expressed between control vs *E. coli* (C vs E) infected mammary epithelial cells (p -value < 0.05 , \log_2 (fold change) > 1), of which 17 lncRNAs were upregulated and 10 lncRNAs were downregulated (supplementary 2, Fig. 4A). While 85 differentially expressed lncRNAs (68 up-regulated and 17 down-regulated) were identified in the control vs *S. aureus* (C vs S) infected comparison (Fig. 4B). Interestingly, among these differentially expressed lncRNAs, we found 19 lncRNAs (11 up-regulated and 8 down-regulated) that were specifically expressed in control vs *E. coli* infected comparison and 77 lncRNAs (62 up-regulated and 15 down-regulated) that were specifically expressed in control vs *S. aureus* infected comparison whereas 8 lncRNAs (6 up-regulated and 2 down-regulated) were showing differential expression in both the comparisons (Fig. 4C). In addition, 20 lncRNAs were differentially expressed between *E. coli* infected and *S. aureus* infected comparison, of which 6 were down-regulated and 14 were up-regulated (Fig. 4D). The differentially expressed lncRNA was widely scattered in all chromosomes, while the numbers were various in different chromosomes. Chromosomes 3, 5, and 19 had the largest number of altered lncRNAs, while 26, 27, and 28 had the least altered lncRNAs (Fig. 4E).

3.3. Functional analysis of differentially expressed lncRNAs

To explore the functional utility of DE lncRNAs we searched for protein-coding genes located within 100 kb upstream or downstream, as recent studies have shown the *cis*-regulatory role of lncRNAs associated with their neighboring protein-coding genes. we also predicted potential

Table 2
Primer details of lncRNAs used for RT-PCR validation.

lncRNA primer	Primer Sequence	Product Length (bp)
XLOC_009502	Forward: 5'- CTTAGGGACTCCCTCGCA-3' Reverse: 5'- AGGGATAGAACCTGTGCC-3'	70
XLOC_002110	Forward: 5'- CCACCTTCTGCCCTCCTAT -3' Reverse: 5'- TGCTGTCGGTATTCTGTGC-3'	95
XLOC_017603	Forward: 5'- CTGTTGACGACTGTGGGATG -3' Reverse: 5'- CTCTGGGATTGGGTTAGTGG -3'	169
XLOC_005053	Forward: 5'- ATGTTTACGCCCTCCTCAA -3' Reverse: 5'- CCCGTTATCCTGTCCCTCCT -3'	99
XLOC_005292	Forward: 5'- CAAACAAGCCCCACTCTTC -3' Reverse: 5'- GGAAGCAGAGAGCCACAGAG-3'	143
XLOC_007175	Forward: 5'- CTGGGGAAATGAAAAGCAA -3' Reverse: 5'- TGTGAAGAAAATGGGCATCA-3'	97
XLOC_020132	Forward: 5'- CAAAAAATGGGTGAAGTGCT -3' Reverse: 5'- CACTAGGCAACCAGGCAACT-3'	75
XLOC_002802	Forward: 5'- TTCAAAGCCAAAAGGCACT -3' Reverse: 5'- TCCTCCACCGTATTGATTC -3'	113
XLOC_003190	Forward: 5'- CTGGCTGGCTGTGGTCTCTA -3' Reverse: 5'- AACGAAACAAAGATGGCAGAA -3'	107
XLOC_000504	Forward: 5'- TTGCCTTCTCATTGCCCTCT -3' Reverse: 5'- AAGTGTAGTGGGCTCA -3'	96
GAPDH	Forward: 5'- GCAAGTCCACGGCACAG -3' Reverse: 5'- GGTCACGCCATCACAA -3'	249
UXT	Forward: 5'- TGGACCATCGTGACAAGGTA -3' Reverse: 5'- TGAAGTGTCTGGGACCACTG -3'	155

cis and trans targets of identified DE lncRNAs, using in-house developed PERL scripts. A total of 1171 differentially expressed lncRNAs from both the comparisons (*E. coli* infected and *S. aureus* infected PMECs) could cis-regulate 364 and trans-regulate 402 neighboring protein-coding genes respectively. The annotated neighboring genes of DE lncRNAs were further explored for their corresponding biological meaning by GO annotations retrieved from the REVIGO [34] (Fig. 5). Hypergeometric tests were carried out for each enriched GO term, and significant enrichments (adjusted P-value of 0.05 or less) are indicated in (supplementary 3). After that, the terms were mapped against biological pathway information in the KEGG database (version 0.7.2). GO analysis based on these cis- and trans-regulated targets was performed and 684 significant enriched ($P < 0.05$) GO terms were obtained among which RNA processing, Positive regulation of the cellular process, and cell death were highly significant. Significant overrepresented GO terms also included protein binding, Ca^{2+} binding, RNA (poly-A) binding, Enzyme binding, and Lipid Binding. To infer systematic biological behaviors of the neighboring genes of lncRNAs KEGG pathway analysis was conducted. The KEGG results (supplementary 4) revealed that these differentially expressed lncRNAs could cis- or trans-regulate the targets involved in about 37 pathways like apoptotic signaling pathway, regulation of endopeptidase activity, immune response, and inflammatory response (Fig. 6). Most of the identified pathways were related to the regulation of bacterial infection and immune functions. Moreover, some pathways, which have been proved to regulate bacterial infection, have also been detected in the KEGG analyzes of differentially expressed lncRNAs, such as the Bacterial invasion of epithelial cells, Angiogenesis, Apoptosis, B cell receptor signaling pathway, TNF signaling pathway, and Toll-like receptor signaling pathway. These analyzes indicated that the differentially expressed lncRNAs and their annotated pathways should play cis- or trans-regulated roles in the modulation of bacterial infections of

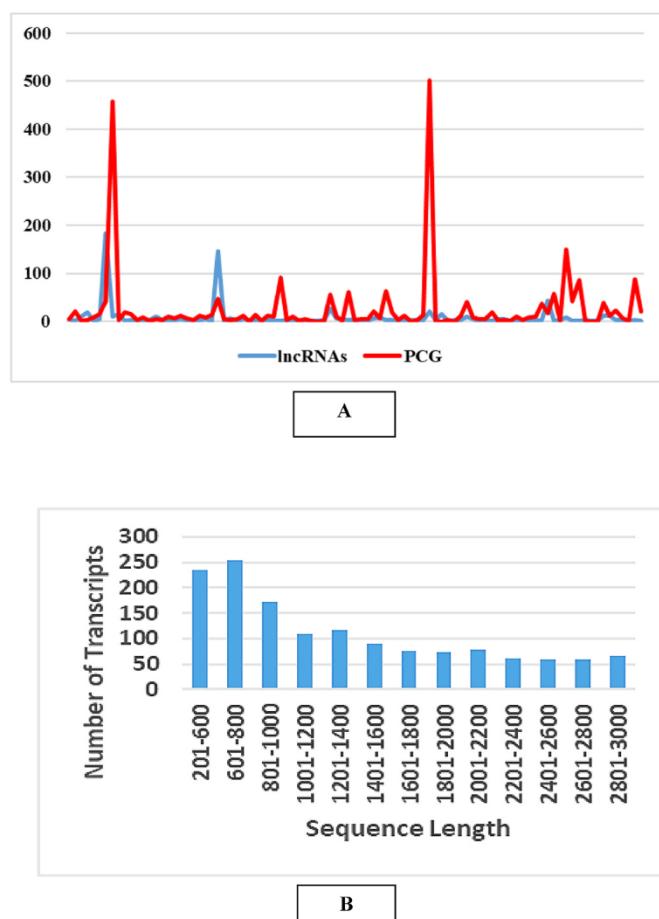


Fig. 2. Characterization of mammary epithelial cell lncRNAs. (A) Comparison of expression values (FPKM) in mammary epithelial cells lncRNAs and protein-coding genes (PCG); (B) Size distribution of mammary epithelial cells lncRNAs; (C) The distribution of exon number of lncRNAs.

mammary epithelial cells.

3.4. Validation through qRT-PCR

To validate the RNA-seq data, 10 lncRNAs (5 upregulated and 5 downregulated) were randomly selected and their relative expression levels were quantified by qRT-PCR. The expression patterns of all selected lncRNAs showed a similar trend between the results of sequencing and qRT-PCR (Fig. 7). However, the expression levels were varying, which may be due to the different cDNA synthesis procedures followed during the two experiments. The Pearson correlation coefficient between RNA-Seq data and qRT-PCR data was 0.970, which indicates that the RNA-Seq data was highly correlated with that of the qRT-PCR.

4. Discussion

The infection of the mammary gland in animals leads to decreased milk production and premature culling and production costs. The invading pathogens and their products cause severe damage to the tissue. Certain bacteria produce toxins that destroy cell membranes and damage milk-producing tissue whereas other bacteria can invade and multiply within the bovine mammary epithelial cells before causing cell death. In addition, mammary gland infections are characterized by an influx of somatic cells, primarily polymorphonuclear neutrophils, into the mammary gland. With more immune cells migrating into the mammary gland and the breakdown of the blood-milk barrier, damage

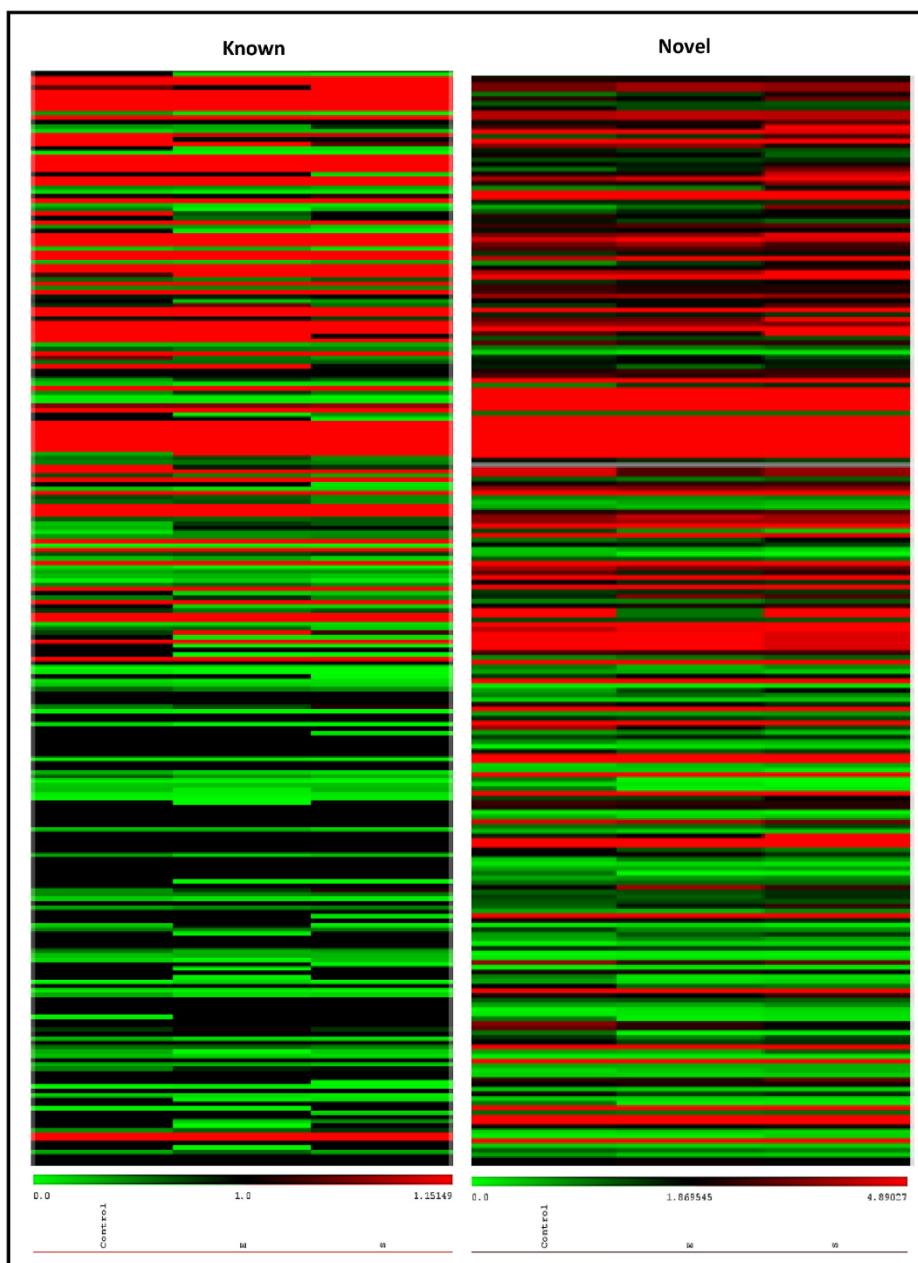


Fig. 3. Cluster analysis of differentially expressed lncRNAs.

to the mammary epithelium worsens [35]. Knowledge of goat mammary cells and their manipulation is applicable to study mammary development and lactation biology, and pathology, including cell damage [36]. This study cataloged the lncRNA repertoire of the primary goat mammary epithelial cells in response to heat-inactivated *E. coli* and *S. aureus* challenge for 24 h under in vitro conditions. Recent research revealed that lncRNAs play a critical role in modulating the immune response to many pathogens [37] and are reported to be influenced by bacterial infection [38]. While the primary cells resemble the physiological environment of the *in-vivo* cells more precisely, the infection intensity from in-vitro to in-vivo conditions is challenging to evaluate. Therefore, to have a standardized experimental condition, we choose heat-inactivated bacteria. So that the timeline of the pathogen-specific immune response is precisely defined in the PMEC model, and bacterial overgrowth and nutrient degradation are prevented during the experiment.

LncRNAs' recognition and functional annotation in livestock are

scarce [39], so one of the study's key challenges was to define lncRNAs along with their biological significance. For this, we adopted a computational pipeline centered upon earlier studies [40–44] that enabled candidate lncRNAs to be identified with higher confidence. In accordance with these reports, we also employed Tophat2/Cufflinks for primary alignment and assembly steps. This helped us to acquire results comparable with previous research on cattle [45,46]. We also got very decent alignment statistics (90% concordant alignment for all samples), with Tophat2, that displayed comparable results relative to its succeeding program [47] with faster search algorithms [48]. In comparison, information from each level in the pipeline synergistically enabled the identification of potentially significant lncRNA transcripts. As recently demonstrated in various studies, the choice of analysis approach (pipeline) has a considerable impact on study outcomes, with a more modest impact from the read aligner and expression model [49,50]. We utilized CuffDiff to perform differential testing and assess the significant changes in the expression profile of lncRNAs [51]. Thus, we

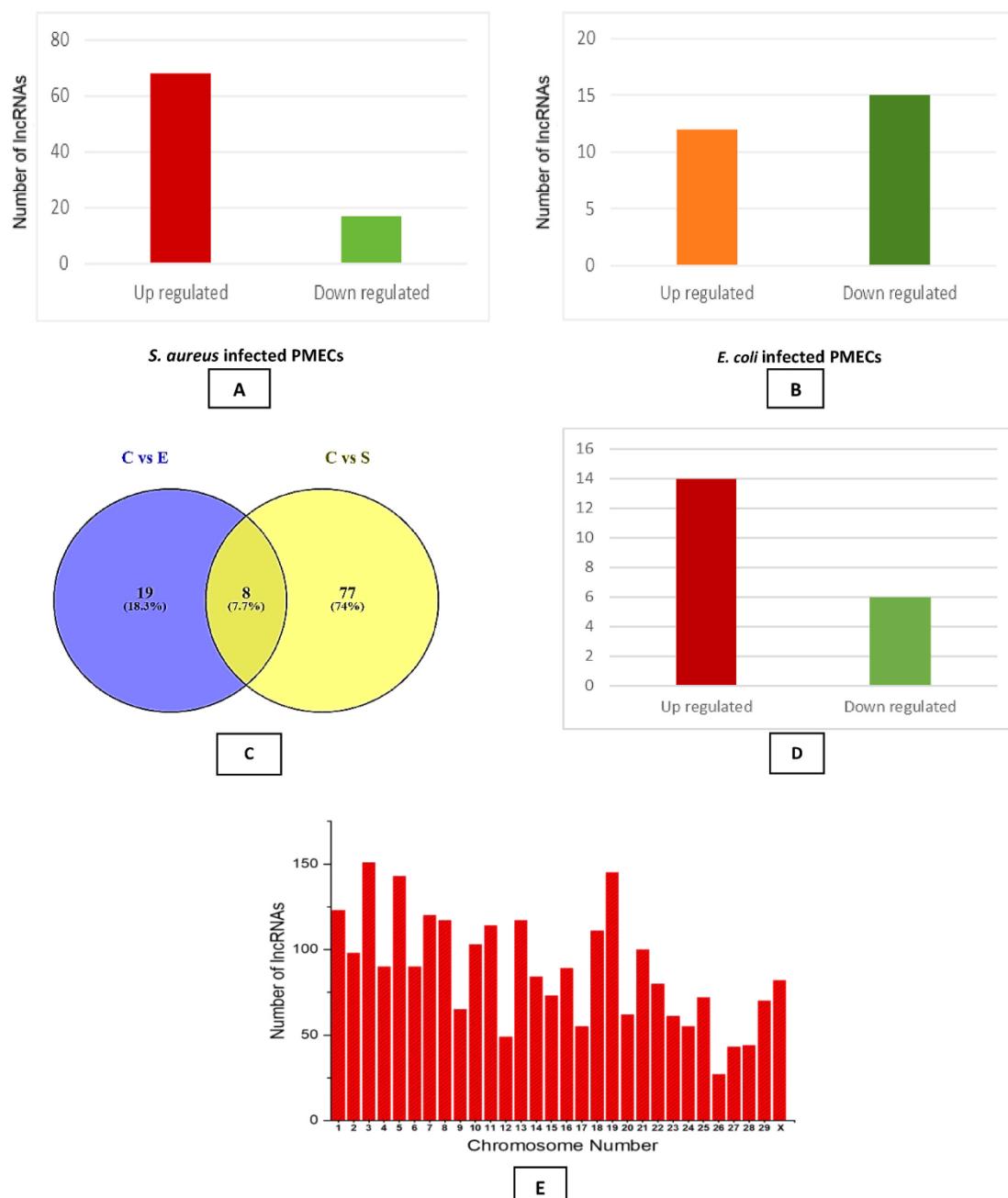


Fig. 4. Differentially expressed lncRNAs identified among control vs *S. aureus* infected PMECs and control vs *E. coli* infected PMECs (A) Hierarchical clustering of the differentially expressed known as well as novel lncRNAs between control, *S. aureus* infected and *E. coli* infected PMECs (A) Analysis of differentially expressed lncRNAs between Control vs *S. aureus* infected PMECs (B) Analysis of differentially expressed lncRNAs between Control vs *E. coli* infected PMECs (C) Venn diagram showing the Specific and common number of lncRNAs between C vs *E. coli* infected PMECs and C vs *S. aureus* infected PMECs comparison (D) Analysis of differentially expressed lncRNAs between *E. coli* infected PMECs vs *S. aureus* infected PMECs comparisons (E) The numbers of lncRNAs localized on each chromosome.

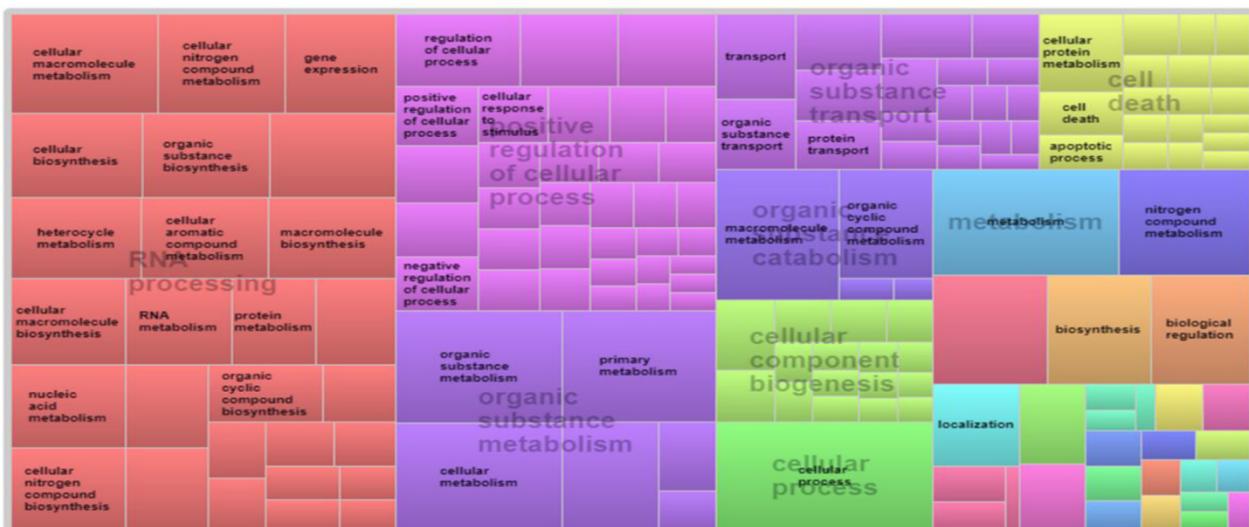
developed a reliable dataset in our study as per the selected filtering parameters.

A set of 1957 known and 1175 novel lncRNAs were reported in this study following strict filtering parameters based on various fundamental genomic features consisting of exon number, coding potential, transcript length, and expression levels. Among which, a total of 112 lncRNAs (*E. coli* and *S. aureus*) were found differentially expressed in bacteria challenged PMECs compared with the control PMECs. As in several differential expression investigations, the repertoire of identified differentially expressed (DE) transcripts greatly depends upon the adopted methodologies (data filtering, read count normalization, and comparing groups) and the methods employed for the correction of

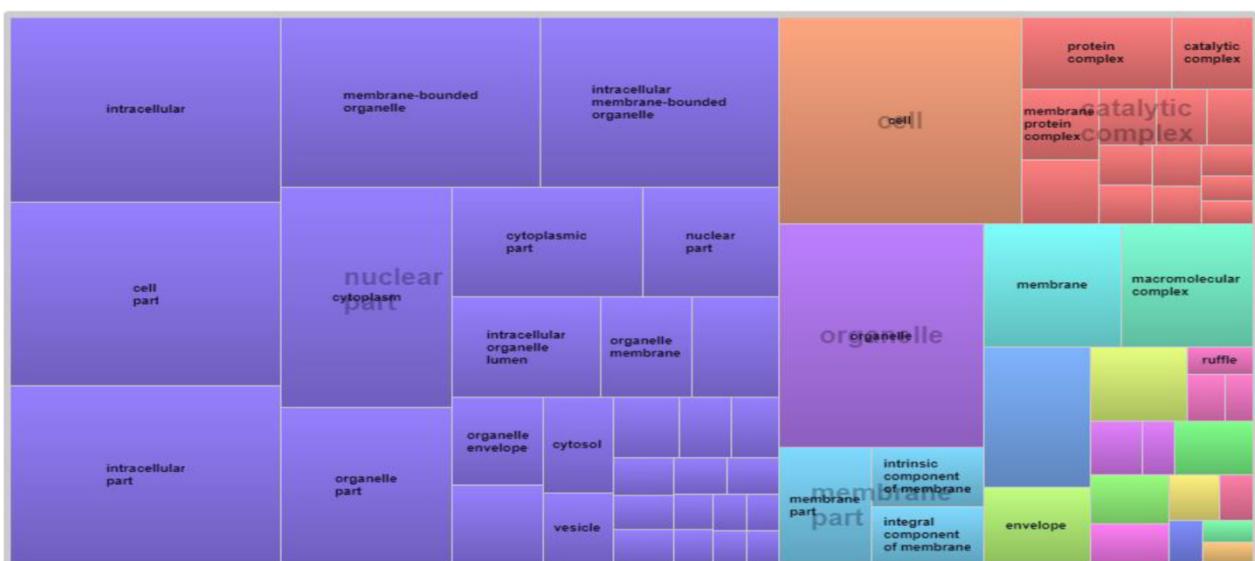
multiple tests and the threshold for the declaration of relevant p-values. In this analysis, we selected the widely used Benjamini and Hochberg method [52] to avoid loss of important transcripts as detected with other traditional methods such as Bonferroni correction. It is well established that the selection of databases for enrichment analysis and the methodology for evaluating enriched terms also affect the study findings [53, 54]. In this study, a widely used hypergeometric test was used for the evaluation of GO enrichment using the ReviGO platform [55–57].

Identified lncRNAs were smaller in size and showed less expression, relative to mRNA transcripts. Earlier studies on the evaluation of both human and bovine lncRNAs also reported fewer exons, lower expression, and smaller size for lncRNAs than mRNAs [58,59] thus supporting our

BIOLOGICAL PROCESS



CELLULAR PROCESS



MOLECULAR FUNCTION

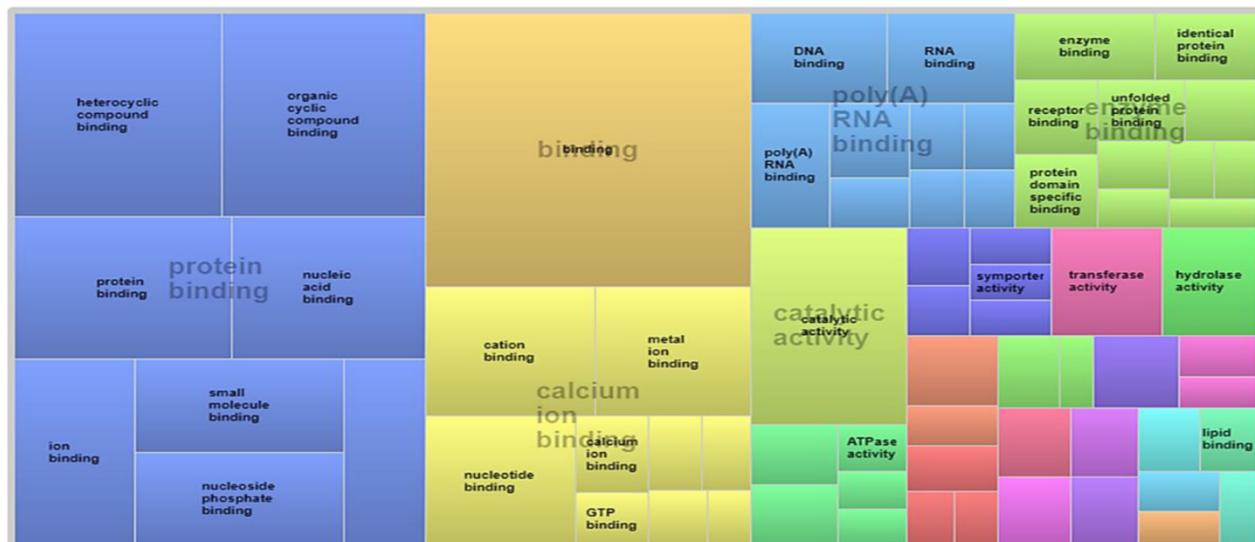


Fig. 5. The “TreeMap” view of REVIGO. Each rectangle is a single cluster representative. The representatives are joined into ‘superclusters’ of loosely related terms, visualized with different colors. The size of the rectangles is adjusted to reflect the p-value of each cluster.

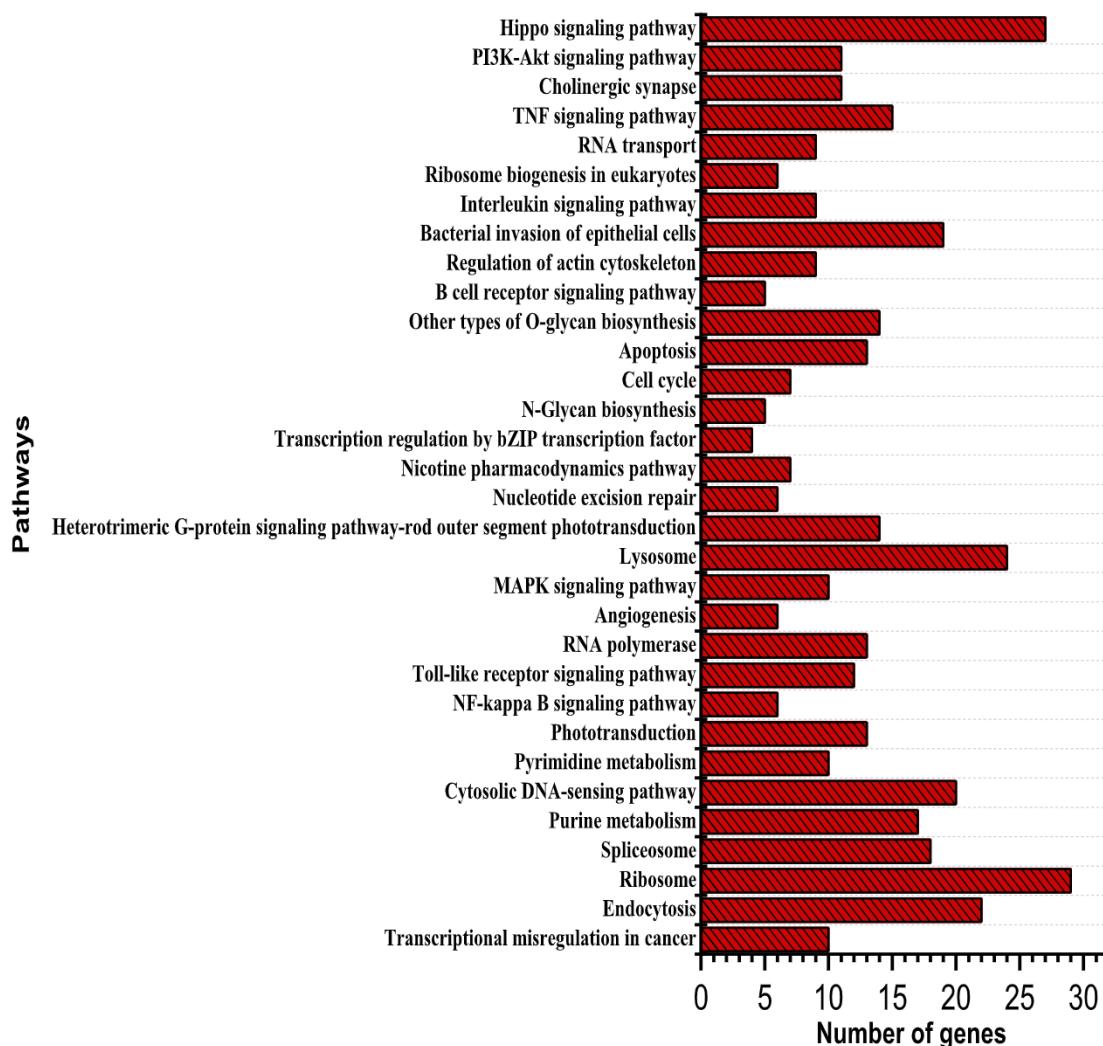


Fig. 6. The KEGG pathway analysis (with $P < 0.05$) based on the *cis*- and *trans*-regulated differentially expressed targets of differentially expressed lncRNAs.

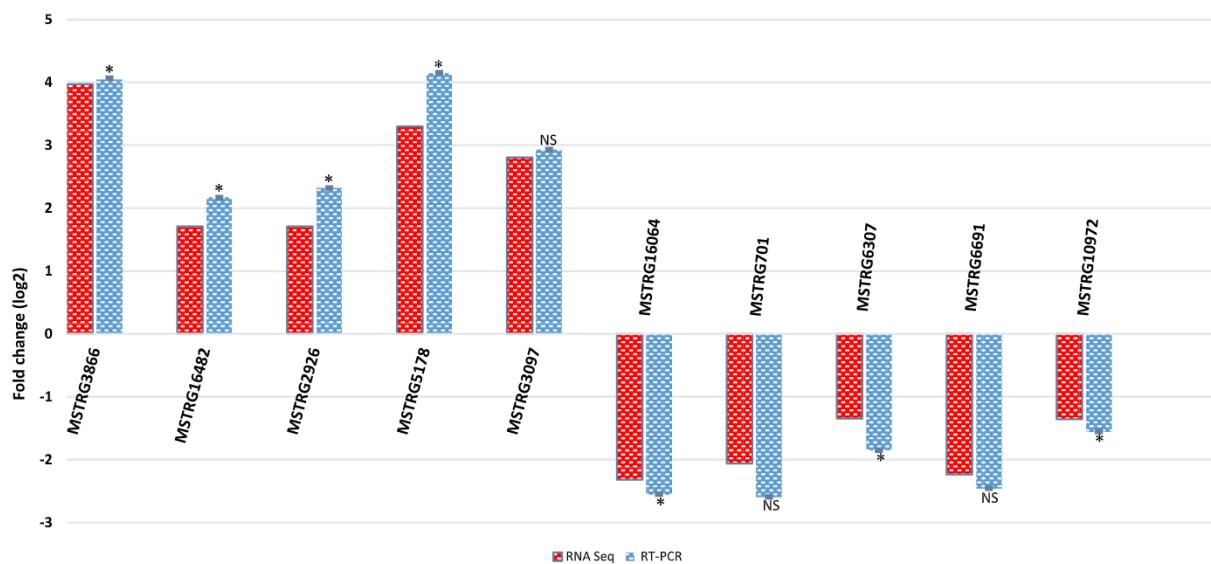


Fig. 7. qRT-PCR validation of significantly differentially expressed lncRNAs among Control, *E. coli* infected and *S. aureus* infected PMECs. The y-axis represents the log₂ fold change of lncRNA expression; the x-axis shows the lncRNA IDs employed for validation. Blue and Red bars depict the RNA-Seq and qPCR results, respectively. Error bars represent Standard Error Mean (SEM). Significance is indicated by * for $p\text{-value} \leq 0.05$ and NS for $p\text{-value} \geq 0.05$.

findings. The count of transcripts per lncRNA in comparison with mRNA in this study came off with the same trend as reported in earlier studies on lncRNAs [60]. These results are consistent with the previous findings that have shown that lncRNAs are localized primarily in the intergenic portion of the genome; however, a smaller proportion overlaps protein-coding genes [58,60,61]. It was observed that more than 67% of intergenic bovine ncRNAs had a neighboring gene within 100 Kb, with a large amount within 5–10 Kb regions of flanking gene [60]. Studies have revealed that lncRNAs could operate in *cis* or *trans* to modulate the expression of neighboring protein-coding genes [62–65]. It was also demonstrated that functional clustering of neighboring protein-coding genes within 5 Kb of intergenic ncRNAs contributed to their over-expression [61]. In addition, the expression of intergenic lncRNAs was strongly associated with neighboring gene expression within 100 Kb [59]. Co-expression of lncRNA and its neighboring mRNAs may be linked to the transcriptional activity of nearby located accessible chromatin [66,67]. Unlike proteins or mRNAs, the roles of ncRNAs cannot be derived from their structural characteristics; therefore, in this study, we predicted the plausible function of the filtered lncRNAs by putting them into service *cis* and *trans* methods. A lncRNA's *cis* nature refers to its capacity to modulate an adjoining gene on a similar allele; it is transcribed by setting up a feedback loop.

During the *cis* prediction, mRNAs located within 100-kb upstream and downstream of all the lncRNAs were listed. Interestingly, GO, and KEGG analysis of these listed coding genes unveiled that the *cis* and *trans* targets of DE lncRNAs were associated with about 15 immune-related pathways like Wnt signaling pathway, VEGF signaling pathway, TGF-beta signaling pathway, Interferon-gamma signaling pathway, FGF signaling pathway, PI3 kinase pathway, indicating that the corresponding lncRNAs are involved in modulation of immunity in the mammary gland. Thus as predicted, lncRNA targets were considerably enriched for the biological process GO terms related to the regulation of RNA processing (GO:0006396). In fact, for their activity lncRNAs may well bind to their target genes [68], so it is quite expected that GO terms controlling nucleic acids were found to be enriched. The pathway of biosynthesis of Aminoacyl-tRNA was a significant KEGG pathway enriched for the bacterial challenge. Recently it was reported that tRNA fragments induce cells to trigger stress-induced tumor suppressor pathway [69]. This indicates that tRNA cleavage induces apoptosis by activating p53, providing a potential correlation between tRNA cleavage and p53-dependent cell death. Another interesting pathway involved "the bacterial invasion of epithelial cells", which is essential for mediating transportation of intracellular vesicles to their destinations through the regulation of actin polymerization. This pathway was also shown to be upregulated during bacterial infection in cholesteatoma [70].

Several recent studies reported *cis*-regulatory activity of lncRNAs in mammary gland immunity [71]; for example, in bovine mastitic tissues and infected MAC-T cells, lncRNA "XIST" expression was abnormally increased. XIST Silencing also intensified the expression of *S. aureus* or *E. coli* induced pro-inflammatory cytokines. XIST may also prevent cellular proliferation, suppress cell viability, and facilitate cell apoptosis during inflammation. XIST also inhibited *E. coli* or *S. aureus*-induced NF- κ B phosphorylation and inflammasome development thereby mediating the inflammation process [72]. Recently, a study revealed that lncRNAs could modulate the expression of the immune gene and significantly come up with disease resistance [73]. Recently a study verified the association of lncRNA H19 with TGF- β 1-induced mesenchymal transformation of bovine epithelial cells and indicated their significance in mammary gland immunity and bovine mastitis [74]. In another study, it was revealed some lncRNAs that expressed differently during bovine diarrhea with a possible role in immunity [75].

In conclusion, we have shown that bacterial challenge triggers the expression of lncRNAs associated with immune response and defense mechanisms in goat mammary epithelial cells. We have cataloged novel DE lncRNAs not yet known to be associated with bacterial infections,

thereby opening new insights into potential mechanisms used by *E. coli* or *S. aureus* or uncertain defensive strategies employed by the host cell. Future studies are required to disclose their true potential as biomarkers or therapeutic agents in mammary gland infections. Additionally, further insight into the fundamental biological questions like; how they regulate gene expression or do, they possess any translational potential.

Contributions

Dr. Mudasir Ahmad: conception of study design, supervised the experiments and data analysis, and provided input on data interpretation. Peerzada Tajamul Mumtaz: conducted the experiment, analyzed the data, and drafted the article. Basharat Bhat: performed bioinformatics analysis of data. Mengqi Wang: performed correlation analysis of data. Dr. Shakil Ahmad Bhat and Dinesh Velayutham: performed mRNA sequencing and acquisition of data. Dr. Eveline M. Ibeagha-Awemu: provided intellectual content, proofreading, and data interpretation. Prof. Dr. Riyaz: provided intellectual content. Qamar Taban, Mashooq Ahmad Dar, and Zulfiqar Ul Haq: provided input on data interpretation. All authors revised and approved the final version of the manuscript.

Funding

The Department of Biotechnology, Ministry of Science and Technology, Government of India supported the study under grant BT/PR114 08/AAQ/1/591/2014. We would like to acknowledge Bioinformatics team of AgriGenome private limited chochin, for their support and providing data timely.

Data availability

The sequencing data is available in NCBI under accession number [GSE118778](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE118778).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.micpath.2021.105367>.

References

- [1] H. Hogeweij, K. Huijps, T.J. Lam, Economic aspects of mastitis: new developments, *N. Z. Vet. J.* 59 (2011) 16–23.
- [2] R.N. Zadoks, J.R. Middleton, S. McDougall, J. Katholm, Y.H. Schukken, Molecular epidemiology of mastitis pathogens of dairy cattle and comparative relevance to humans, *J. Mammary Gland Biol. Neoplasia* 16 (2011) 357–372.
- [3] F. Vangroenweghe, I. Lamote, C. Burvenich, Physiology of the periparturient period and its relation to severity of clinical mastitis, *Domest. Anim. Endocrinol.* 29 (2005) 283–293.
- [4] E.M. Ibeagha-Awemu, A.E. Ibeagha, S. Messier, X. Zhao, Proteomics, genomics, and pathway analyzes of *Escherichia coli* and *Staphylococcus aureus* infected milk whey reveal molecular pathways and networks involved in mastitis, *J. Protozool. Res.* 9 (2010) 4604–4619.
- [5] L. Sutra, B. Poutrel, Virulence factors involved in the pathogenesis of bovine intramammary infections due to *Staphylococcus aureus*, *J. Med. Microbiol.* 40 (1994) 79–89.
- [6] B. Buitenhuis, C.M. Rontved, S.M. Edwards, K.L. Ingvarsson, P. Sorensen, In depth analysis of genes and pathways of the mammary gland involved in the pathogenesis of bovine Escherichia coli-mastitis, *BMC Genom.* 12 (2011) 130.
- [7] J. Gunther, K. Esch, N. Poschadel, W. Petzl, H. Zerbe, S. Mitterhuemer, H. Blum, H.-M. Seyfert, Comparative kinetics of *Escherichia coli*- and *Staphylococcus aureus*-specific activation of key immune pathways in mammary epithelial cells demonstrates that *S. aureus* elicits a delayed response dominated by interleukin-6 (IL-6) but not by IL-1A or tumor necrosis factor alpha, *Infect. Immun.* 79 (2011) 695–707.

- [8] F.B. Gilbert, P. Cunha, K. Jensen, E.J. Glass, G. Foucras, C. Robert-Granié, R. Rupp, P. Rainard, Differential response of bovine mammary epithelial cells to *Staphylococcus aureus* or *Escherichia coli* agonists of the innate immune system, *Vet. Res.* 44 (2013) 40.
- [9] D.D. Bannerman, M.J. Paape, J.W. Lee, X. Zhao, J.C. Hope, P. Rainard, *Escherichia coli* and *Staphylococcus aureus* elicit differential innate immune responses following intramammary infection, *Clin. Diagn. Lab. Immunol.* 11 (2004) 463–472.
- [10] H. Lahouassa, E. Moussay, P. Rainard, C. Rioillet, Differential cytokine and chemokine responses of bovine mammary epithelial cells to *Staphylococcus aureus* and *Escherichia coli*, *Cytokine* 38 (2007) 12–21, <https://doi.org/10.1016/j.cyto.2007.04.006>.
- [11] J. Castillo, T.R. Stueve, C.N. Marconett, Intersecting transcriptomic profiling technologies and long non-coding RNA function in lung adenocarcinoma: discovery, mechanisms, and therapeutic applications, *Oncotarget* 8 (46) (2017 Jun 9) 81538–81557.
- [12] P. Grote, et al., The tissue-specific lncRNA Fendrr is an essential regulator of heart and body wall development in the mouse, *Dev. Cell* 24 (2013) 206–214.
- [13] Shakil Ahmad Bhat, Syed Mudasir Ahmad, Peerzada Tajamul Mumtaz, Abrar Ahad Malik, Mashoq Ahmad Dar, Uneeb Urwat, Riaz Ahmad Shah, Nazir Ahmad Ganai, Long non-coding RNAs: mechanism of action and functional utility, *Non-coding RNA Res.* 1 (no. 1) (2016) 43–50.
- [14] C.J. Brown, A. Ballabio, J.L. Rupert, R.G. Lafreniere, M. Grompe, R. Tonlorenzi, H. F. Willard, A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome, *Nature* 349 (1991) 38–44.
- [15] N. Brockdorff, A. Ashworth, G.F. Kay, V.M. McCabe, D.P. Norris, P.J. Cooper, S. Swift, S. Rastan, The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus, *Cell* 71 (1992) 515–526.
- [16] L.B. Herzing, J.T. Romer, J.M. Horn, A. Ashworth, Xist has properties of the X-chromosome inactivation centre, *Nature* 386 (1997) 272–275.
- [17] V. Tripathi, J.D. Ellis, Z. Shen, D.Y. Song, Q. Pan, A.T. Watt, S.M. Freier, C. F. Bennett, A. Sharma, P.A. Bubulya, et al., The nuclear-retained non-coding RNAMALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation, *Mol. Cell.* 39 (2010) 925–938.
- [18] V. Tripathi, D.Y. Song, X. Zong, S.P. Shevtsov, S. Hearn, X.-D. Fu, M. Dundr, K. V. Prasanth, SRSF1 regulates the assembly of pre-mRNA processing factors in nuclear speckles, *Mol. Biol. Cell* 23 (2012) 3694–3706.
- [19] J.-H. Yoon, K. Abdelmohsen, M. Gorospe, Posttranscriptional gene regulation by long non-coding RNA, *J. Mol. Biol.* 425 (2013) 3723–3730.
- [20] S.R. Atkinson, S. Marguerat, J. Bähler, Exploring long non-coding RNAs through sequencing, *Semin. Cell Dev. Biol.* 23 (2012) 200–205.
- [21] H. Servais, P. Van Der Smissen, G. Thirion, G. Van der Essen, F. Van Bambeke, P. M. Tulkens, M.-P. Mingeot-Leclercq, Gentamicin-induced apoptosis in LLC-PK1 cells: involvement of lysosomes and mitochondria, *Toxicol. Appl. Pharmacol.* 206 (2005) 321–333.
- [22] A.M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data, *Bioinformatics* (2014) btu170.
- [23] Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10:R25.
- [24] C. Trapnell, L. Pachter, S. L. TopHat Salzberg, Discovering splice junctions with RNA-Seq, *Bioinformatics* 25 (2009) 1105–1111.
- [25] C. Trapnell, et al., Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks, *Nat. Protoc.* 7 (2012) 562.
- [26] F. Cunningham, P. Achuthan, W. Akanni, J. Allen, M.R. Amode, I.M. Armean, R. Bennett, J. Bhai, K. Billis, S. Boddu, C. Cummins, Ensembl 2019, *Nucleic Acids Res.* 47 (D1) (2019) D745–D751.
- [27] L. Sun, et al., Utilising sequence intrinsic composition to classify protein-coding and long non-coding transcripts, *Nucleic Acids Res.* 41 (2013).
- [28] A. Li, J. Zhang, Z.P.L.E.K. Zhou, A tool for predicting long non-coding RNAs and messenger RNAs based on an improved k-mer scheme, *BMC Bioinf.* 15 (2014).
- [29] L. Wang, H.J. Park, S. Dasari, S. Wang, J.P. Kocher, W.C.P.A.T. Li, Coding-Potential Assessment Tool using an alignment-free logistic regression model, *Nucleic Acids Res.* 41 (2013).
- [30] R.D. Finn, et al., Pfam: the protein families database, *Nucleic Acids Res.* 42 (2013).
- [31] C. Trapnell, et al., Transcript assembly and quantification by RNA-seq reveals unannotated transcripts and isoform switching during cell differentiation, *Nat. Biotechnol.* 28 (2010) 511–515.
- [32] K.J. Livak, T.D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-ΔΔCT} method, *Methods* 25 (4) (2001) 402–408, pmid:11846609.
- [33] L. Statello, C.J. Guo, L.L. Chen, et al., Gene regulation by long non-coding RNAs and its biological functions, *Nat. Rev. Mol. Cell Biol.* 22 (2021) 96–118.
- [34] F. Supek, M. Bošnjak, N. Škunca, T. Šmuc, REVIGO summarises and visualises long lists of gene ontology terms, *PLoS One* 6 (2011) 1800–1801.
- [35] X. Zhao, P. Lacasse, Mammary tissue damage during bovine mastitis: causes and control, *J. Anim. Sci.* 86 (2008) 57–65.
- [36] Jernej Ogorevc, Minja Zorc, Dovč Peter, Development of an in Vitro Goat Mammary Gland Model: Establishment, Characterization, and Applications of Primary Goat Mammary Cell Cultures, 2017, <https://doi.org/10.5772/intechopen.71853>.
- [37] S.I. Miller, R.K. Ernst, M.W. Bader, LPS, TLR4 and infectious disease diversity, *Nat. Rev. Microbiol.* 3 (2005) 36–46.
- [38] J. Zur Bruegghe, R. Einspanier, S. Sharbati, A long journey ahead: long non-coding RNAs in bacterial infections, *Front. Cell. Infect. Microbiol.* 7 (2017) 95–96.
- [39] X. Yang, J. Yang, J. Wang, Q. Wen, H. Wang, J. He, S. Hu, W. He, X. Du, S. Liu, L. Ma, Microarray analysis of long non-coding RNA and mRNA expression profiles in human macrophages infected with *Mycobacterium tuberculosis*, *Sci. Rep.* 6 (2016), 38963.
- [40] R. Weikard, W. Demasius, C. Kuehn, Mining long noncoding RNA in livestock, *Anim. Genet.* 48 (2017) 3–18.
- [41] J.R. Presner, M.K. Iyer, O.A. Balbin, S.M. Dhanasekaran, Q. Cao, J.C. Brenner, B. Laxman, I.A. Asangani, C.S. Grasso, H.D. Kominsky, X. Cao, Transcriptome sequencing across a prostate cancer cohort identifies PCAT-1, an unannotated lncRNA implicated in disease progression, *Nat. Biotechnol.* 29 (2011) 742.
- [42] R. Weikard, F. Hadlich, C. Kuehn, Identification of novel transcripts and non-coding RNAs in bovine skin by deep next generation sequencing, *BMC Genom.* 14 (2013) 789–790.
- [43] C. Billerey, M. Boussaha, D. Esquerré, E. Rebours, A. Djari, C. Meersseman, C. Klopp, D. Gautheret, D. Rocha, Identification of large intergenic non-coding RNAs in bovine muscle using next-generation transcriptomic sequencing, *BMC Genom.* 15 (2014) 499.
- [44] M. Sun, S.S. Gadad, D.S. Kim, W.L. Kraus, Discovery, annotation, and functional analysis of long non-coding RNAs controlling cell-cycle gene expression and proliferation in breast cancer cells, *Mol. Cell* 59 (2015) 698–711.
- [45] L.C. Tsai, M.K. Iyer, P.E. Stuart, W.R. Swindell, J.E. Gudjonsson, T. Tejasvi, M. K. Sarkar, B. Li, J. Ding, J.J. Voorhees, H.M. Kang, Analysis of long non-coding RNAs highlights tissue-specific expression patterns and epigenetic profiles in normal and psoriatic skin, *Genome Biol.* 16 (2015) 24–25.
- [46] L.T. Koufarisiotis, Y.P.P. Chen, A. Chamberlain, C. Vander Jagt, B.J. Hayes, A catalogue of novel bovine long non-coding RNA across 18 tissues, *PLoS One* 10 (2015) 1225–1226.
- [47] C. Kern, Y. Wang, J. Chitwood, I. Korf, M. Delany, H. Cheng, J.F. Medrano, A. L. Van Eenennaam, C. Ernst, P. Ross, H. Zhou, Genome-wide identification of tissue-specific long non-coding RNA in three farm animal species, *BMC Genom.* 19 (2018) 684–685.
- [48] D. Kim, B. Langmead, S.L. Salzberg, HISAT: a fast spliced aligner with low memory requirements, *Nat. Methods* 12 (2015) 357–360.
- [49] G. Baruzzo, K.E. Hayer, E.J. Kim, B. Di Camillo, G.A. FitzGerald, G.R. Grant, Simulation-based comprehensive benchmarking of RNA-seq aligners, *Nat. Methods* 14 (2017) 135.
- [50] M. Teng, M.I. Love, C.A. Davis, S. Djebali, A. Dobin, B.R. Graveley, S. Li, C. E. Mason, S. Olson, D. Pervouchine, C.A. Sloan, A benchmark for RNA-seq quantification pipelines, *Genome Biol.* 17 (2016) 74–75.
- [51] C. Trapnell, D.G. Hendrickson, M. Sauvageau, L. Goff, J.L. Rinn, L. Pachter, Differential analysis of gene regulation at transcript resolution with RNA-seq, *Nat. Biotechnol.* 31 (2013) 46–53.
- [52] Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing, *J. Roy. Stat. Soc. Ser. B* 57 (1995) 289–300.
- [53] D.W. Huang, B.T. Sherman, R.A. Lempicki, Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists, *Nucleic Acids Res.* 37 (2009) 1–13.
- [54] M.D. Young, M.J. Wakefield, G.K. Smyth, A. Oshlack, Gene ontology analysis for RNA-seq: accounting for selection bias, *Genome Biol.* 11 (2010) 14–15.
- [55] S. Sharma, M. Taneja, S. Tyagi, K. Singh, S.K. Upadhyay, Survey of high throughput RNA-Seq data reveals potential roles for lncRNAs during development and stress response in bread wheat, *Front. Plant Sci.* 8 (2017) 1019.
- [56] N. Nohata, M.C. Abba, J.S. Gutkind, Unraveling the oral cancer lncRNAome: identification of novel lncRNAs associated with malignant progression and HPV infection, *Oral Oncol.* 59 (2016) 58–66.
- [57] J.Y. Tan, T. Sirey, F. Honti, B. Graham, A. Piovesan, M. Merkenschlager, C. Webber, C.P. Ponting, A.C. Marques, Extensive microRNA-mediated crosstalk between lncRNAs and mRNAs in mouse embryonic stem cells, *Genome Res.* 25 (2015) 655–666.
- [58] X.Y. Zhao, J.D. Lin, Long non-coding RNAs: a new regulatory code in metabolic control, *Trends Biochem. Sci.* 40 (2015) 586–596.
- [59] L. Yang, P. Li, W. Yang, X. Ruan, K. Kiesewetter, J. Zhu, H. Cao, Integrative transcriptome analyzes of metabolic responses in mice define pivotal lncRNA metabolic regulators, *Cell Metabol.* 24 (2016) 627–639.
- [60] Z. Qu, D.L. Adelson, Bovine ncRNAs are abundant, primarily intergenic, conserved and associated with regulatory genes, *PLoS One* 7 (2012) 638–639.
- [61] J. Xia, L. Xin, W. Zhu, L. Li, C. Li, Y. Wang, Y. Mu, S. Yang, K. Li, Characterization of long non-coding RNA transcriptome in high-energy diet induced nonalcoholic steatohepatitis minipigs, *Sci. Rep.* 6 (2016) 1–11.
- [62] R. Weikard, F. Hadlich, H.M. Hammon, D. Frieten, C. Gerbert, C. Koch, G. Dusel, C. Kuehn, Long non-coding RNAs are associated with metabolic and cellular processes in the jejunum mucosa of pre-weaning calves in response to different diets, *Oncotarget* 9 (2018) 52–53.
- [63] F. Yang, X.S. Huo, S.X. Yuan, L. Zhang, W.P. Zhou, F. Wang, S.H. Sun, Repression of the long non-coding RNA-LET by histone deacetylase 3 contributes to hypoxia-mediated metastasis, *Mol. Cell* 49 (2013) 1083–1096.
- [64] A. Haug, A.T. Höstmark, O.M. Harstad, Bovine milk in human nutrition—a review, *Lipids Health Dis.* 6 (2007) 25–26.
- [65] A.L. Lock, D.E. Bauman, Modifying milk fat composition of dairy cows to enhance fatty acids beneficial to human health, *Lipids* 39 (2004) 1197–1206.
- [66] J.J. Loor, A. Ferlay, A. Ollier, M. Doreau, Y. Chilliard, Relationship among trans and conjugated fatty acids and bovine milk fat yield due to dietary concentrate and linseed oil, *J. Dairy Sci.* 88 (2005) 726–740.
- [67] L. Flintoff, Non-coding RNA: structure and function for lncRNAs, *Nat. Rev. Genet.* 14 (9) (2013) 598.
- [68] M.L. Kelly, J.R. Berry, D.A. Dwyer, J.M. Grinari, P.Y. Chouinard, M.E. Van Amburgh, D.E. Bauman, Dietary fatty acid sources affect conjugated linoleic acid concentrations in milk from lactating dairy cows, *J. Nutr.* 128 (1998) 881–885.

- [69] T. Hanada, S. Weitzer, B. Mair, C. Bernreuther, B.J. Wainger, J. Ichida, R. Hanada, M. Orthofer, S.J. Cronin, V. Komnenovic, A. Minis, CLP1 links tRNA metabolism to progressive motor-neuron loss, *Nature* 495 (2013) 474–480.
- [70] I. Brook, P. Burke, The management of acute, serous and chronic otitis media: the role of anaerobic bacteria, *J. Hosp. Infect.* 22 (1992) 75–87.
- [71] E. Ibeagha-Awemu, R. Li, P.L. Dudemaine, D. Do, N. Bissonnette, Transcriptome analysis of long non-coding RNA in the bovine mammary gland following dietary supplementation with linseed oil and safflower oil, *Int. J. Mol. Sci.* 19 (2018) 3610.
- [72] M. Ma, Y. Pei, X. Wang, J. Feng, Y. Zhang, M.Q. Gao, LncRNA XIST mediates bovine mammary epithelial cell inflammatory response via NF- κ B/NLRP3 inflammasome pathway, *Cell Prolif* 52 (2019), 12525.
- [73] J. Chen, L. Ao, J. Yang, Long non-coding RNAs in diseases related to inflammation and immunity, *Ann. Transl. Med.* 7 (18) (2019) 494.
- [74] W. Yang, X. Li, S. Qi, X. Li, K. Zhou, S. Qing, Y. Zhang, M.Q. Gao, lncRNA H19 is involved in TGF- β 1-induced epithelial to mesenchymal transition in bovine epithelial cells through PI3K/AKT Signaling Pathway, *PeerJ* 5 (2017) 3950–3951.
- [75] Qiman Ma, Liangyuan Li, Yan Tang, Qiang Fu, Sheng Liu, Shengwei Hu, Jun Qiao, Chuangfu Chen, Wei Ni, Analyzes of long non-coding RNAs and mRNA profiling through RNA sequencing of MDBK cells at different stages of bovine viral diarrhea virus infection, *Res. Vet. Sci.* 115 (2017) 508–516.