



OPEN ACCESS

EDITED BY

Tanvir Alam,
Hamad bin Khalifa University, Qatar

REVIEWED BY

Hui Li,
Guangxi University, China
Mohammad Islam,
Southern Connecticut State University,
United States

*CORRESPONDENCE

Ben Huang,
✉ bhuang@gxams.org.cn
Dandan Zhang,
✉ ddzhang@gxams.org.cn

[†]These authors have contributed equally to this work

RECEIVED 19 March 2023

ACCEPTED 21 August 2023

PUBLISHED 05 September 2023

CITATION

Wei M, Tang W, Lv D, Liu M, Wang G, Liu Q, Qin L, Huang B and Zhang D (2023), Long-chain noncoding RNA sequencing analysis reveals the molecular profiles of chemically induced mammary epithelial cells. *Front. Genet.* 14:1189487. doi: 10.3389/fgene.2023.1189487

COPYRIGHT

© 2023 Wei, Tang, Lv, Liu, Wang, Liu, Qin, Huang and Zhang. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Long-chain noncoding RNA sequencing analysis reveals the molecular profiles of chemically induced mammary epithelial cells

Mengzhen Wei[†], Wenkui Tang[†], Danwei Lv, Mingxing Liu, Guodong Wang, Quanhui Liu, Liangshan Qin, Ben Huang* and Dandan Zhang*

Guangxi Key Laboratory of Eye Health, Guangxi Academy of Medical Sciences, People's Hospital of Guangxi Zhuang Autonomous Region, Nanning, China

Long noncoding RNAs (lncRNAs) were important regulators affecting the cellular reprogramming process. Previous studies from our group have demonstrated that small molecule compounds can induce goat ear fibroblasts to reprogram into mammary epithelial cells with lactation function. In this study, we used lncRNA-Sequencing (lncRNA-seq) to analyze the lncRNA expression profile of cells before and after reprogramming (CK vs. 5i8 d). The results showed that a total of 3,970 candidate differential lncRNAs were detected, 1,170 annotated and 2,800 new lncRNAs. Compared to 0 d cells, 738 lncRNAs were significantly upregulated and 550 were significantly downregulated in 8 d cells. Heat maps of lncRNAs and target genes with significant differences showed that the fate of cell lineages changed. Functional enrichment analysis revealed that these differently expressed (DE) lncRNAs target genes were mainly involved in signaling pathways related to reprogramming and mammary gland development, such as the Wnt signaling pathway, PI3K-Akt signaling pathway, arginine and proline metabolism, ECM-receptor interaction, and MAPK signaling pathway. The accuracy of sequencing was verified by real-time fluorescence quantification (RT-qPCR) of lncRNAs and key candidate genes, and it was also demonstrated that the phenotype and genes of the cells were changed. Therefore, this study offers a foundation for explaining the molecular mechanisms of lncRNAs in chemically induced mammary epithelial cells.

KEYWORDS

long-chain noncoding RNA (lncRNA), reprogramming, mammary epithelial cells, goat, lactation

1 Introduction

1.1 Background

lncRNA is an RNA molecule greater than 200 nucleotides in length that does not have coding potential (Okazaki et al., 2002). Based on the location of lncRNAs on the genome relative to protein-coding genes, lncRNAs were broadly classified as positive, antisense, intragenic, intergenic, bidirectional, and overlapping lncRNAs (Kopp and Mendell, 2018). lncRNAs were initially considered to be genomic transcriptional “noise,” a byproduct of RNA polymerase II transcription. Compared with mRNAs, lncRNAs do not have typical

start codons, promoter conserved regions, stop codons or open reading frames (Ulitsky et al., 2011; Guttman and Rinn, 2012). However, in recent years, an increasing number of lncRNAs with biological functions have been identified, and lncRNAs play important roles in many biological activities, such as participating in cell proliferation (Liu et al., 2018; Qin et al., 2018), differentiation (Song et al., 2016; Touat Todeschini et al., 2017) and apoptosis (Liu et al., 2017; Nan et al., 2017) and promoting myogenic cell differentiation and injury-induced muscle regeneration (Wang et al., 2015), fat deposition (Wang et al., 2017), lactation (Yu et al., 2017), reproduction (Wang et al., 2014) and immunity (Zhou et al., 2017). lncRNAs have also become a hot spot for research in different scientific directions. lncRNAs are important regulatory molecules of gene expression and have highly diverse biological functions. Their modes of action are broadly divided into three categories: one is at the chromatin level, which regulates gene expression through epistatic modification of chromatin; the second is at the transcriptional level, which regulates gene expression by altering the functions of factors and enzymes that mediate the regulation of gene expression; and the third is at the post transcriptional level, which regulates mRNA by binding to coding genes or miRNAs, thereby regulating the expression levels of mRNAs and miRNAs (Lu et al., 2017).

Studies in humans and model animals have shown that lncRNAs were involved in mammary gland development and the regulation of lactation in mammals (Standaert et al., 2014; Yu et al., 2017). Zhang et al. (Zhang et al., 2014) found that overexpression of lncRNAROR increased the self-renewal of mammary stem cells, and further studies of its function showed that lncRNAROR plays a key role in maintaining a normal stem cell subpopulation in mammary epithelial cells. These findings provide a basis for the biological function of lincRNAs in regulating mammary gland development and susceptibility to mastitis in dairy cows (Tong et al., 2017). It has been found that there are specific lncRNAs in the reprogramming process and in mouse and human embryonic stem cells, and these lncRNAs were strongly correlated with the expression of the important reprogramming factors OCT4, NANOG, and SOX2 (Dinger et al., 2008). lncRNAs affect the cell reprogramming process, and lncRNAROR is one of these regulators, and inhibition of their expression decreases the efficiency of cellular reprogramming (Loewer et al., 2010; Guttman et al., 2011). However, less is known about goat lncRNAs, and the regulatory relationship between cell reprogramming and mammary gland development and lncRNA expression is unclear.

1.2 Objective

The aim of this study was to analyze the expression profile of lncRNAs after small molecule compound-induced reprogramming of goat fibroblasts into mammary epithelial cells, to explore the relationship between lncRNAs and mammary gland development and cell reprogramming regulation in goats and to provide a molecular mechanism to explain the role of lncRNAs in mammary gland development and cell reprogramming regulation in goats on a theoretical basis.

2 Materials and methods

2.1 Sample collection and RNA extraction

2.1.1 Sample collection

Goat ear margin fibroblasts were extracted and purified by isolation and pre-digestion of ear tissue from 2 to 3 months old Guanzhong dairy goats, provided by the Guangxi Institute of Animal Science. The test samples used were obtained from laboratory-induced goat fibroblasts. Goat mammary epithelial cells were obtained after 8 days of induction of goat fibroblasts using the 5i system (Induction medium composed of Forskolin, TTNPB, VPA, Repsox, Tranyl-cypromine, five small molecule compounds) (Zhang et al., 2021), and cell precipitates were collected from goat fibroblasts (0 day cells, CK group) and induced goat mammary epithelial cells (8 day cells, 5i8d group), with two biological replicates from each group. Samples were collected and stored at -80°C .

2.1.2 RNA extraction and detection

Total RNA was extracted from goat fibroblasts and goat mammary epithelial cells using the TRIzol method. One percent agarose electrophoresis was used to detect the presence of degradation and impurities in the RNA samples, and the purity of the samples was measured by a Kaio K5500 spectrophotometer. The integrity and purity of the RNA samples were measured by an Agilent 2100 RNA Nano 6000 Assay Kit (Agilent Technologies, CA, United States) to detect the integrity and concentration of the RNA samples.

2.2 Library construction, sequencing and transcript assembly

Goat fibroblasts and induced goat mammary epithelial cells were each sampled with $3\ \mu\text{g}$ of total RNA as a starting amount to construct the lncRNA libraries. Ribo-Zero™ GoldKits were used to remove rRNA from the samples, and different index tags were selected for library construction according to the operating instructions of the NEB Next Ultra Directional RNA LibraryPrep Kit for Illumina (NEB, Ipswich, United States).

The specific steps of library construction were as follows: first, a kit was used to remove ribosomal rRNA, fragmentation buffer was added to the reaction system to fragment the RNA into short fragments, the fragmented RNA was used as a template to synthesize the first strand of cDNA with six-base random primers (random hexamers), and buffer, dNTPs, RNase H and DNA polymerase I were added to synthesize the second strand of cDNA. The cDNA second strand was purified by a QiaQuick PCR kit and eluted with EB buffer, end-repair, base A, sequencing junction, agarose gel electrophoresis to recover the target size fragment, digestion of cDNA second strand with UNG enzyme, PCR amplification, and finally agarose gel electrophoresis to recover the target size fragment. The whole library preparation was completed by agarose gel electrophoresis. Finally, the constructed libraries were used for Illumina sequencing.

2.3 Screening of differentially expressed lncRNAs and target gene prediction

lncRNAs were classified into intergenic RNAs (lincRNAs), intronic lncRNAs, anti-sense lncRNAs, sense lncRNAs, bidirectional lncRNAs and other types according to their position in relation to the coding sequence. Among them, the proportion of lincRNAs was the highest, and the screening of novel lncRNAs was mainly for lincRNAs, intronic lncRNAs and antisense lncRNAs. A series of screening conditions were set, and the resulting lncRNAs were used as the final candidate novel lncRNAs for subsequent analysis. The basic screening conditions were as follows: the transcript length was greater than or equal to 200 bp, the number of exons was greater than or equal to 2, the read coverage of each transcript was calculated, transcripts with less than 5 in all samples were screened out, and the known mRNAs and other noncoding RNAs (rRNA, tRNA, snoRNA, snRNA, etc.) in the species were screened out by comparison with the annotation file of the species using gffcompare. Whether the new transcripts have coding potential is the key condition to determine whether the new transcripts were lncRNAs. By using the lncRNA information from the previous step of the initial screening, a variety of coding potential analysis software programs were integrated for screening, mainly CNCI analysis, CPC analysis, PFAM protein structural domain analysis, and CPAT analysis (animal only). Several analysis methods discriminated noncoding transcripts as the final novel lncRNA data set. Differential expression analysis was performed from candidate lncRNAs using DESeq to compare the treatment group with the reference group, and genes with $|\log_2\text{Ratio}| \geq 2$ and $p < 0.05$ were selected as significant differential expression screening conditions to obtain the number of up and downregulated genes. For differentially expressed lncRNA target genes, cis and trans target analyses were performed to indirectly predict their functions by target genes.

2.4 GO (Gene ontology) enrichment and KEGG (Kyoto encyclopedia of genes and genomes) pathway analysis

DAVID was used online for GO analysis and KEGG pathway analysis of the predicted target genes (Sherman et al., 2022).

2.5 RT-qPCR validation of lncRNA and target genes

Six lncRNAs were randomly selected and subjected to RT-qPCR. U6 was selected as the internal reference gene, and the primers were all synthesized by Janus Biological Engineering Co. The total fluorescence PCR system was 20 μL : 0.5 μL of each upstream and downstream primer (20 $\mu\text{mol/L}$), 1 μL of cDNA, 8 μL of ddH₂O, and 10 μL of 2X lncRNA PreMix. The PCR conditions were as follows: 95°C for 3 min; 40 cycles of 95°C for 5 s and 60°C for 15 s; and 3 replicates. The relative expression of lncRNAs in induced goat mammary epithelial cells was calculated by the $2^{-\Delta\Delta\text{CT}}$ method and verified by comparison with the transcriptome sequencing results. The target gene qRT-PCR method was consistent. The lncRNA and target gene primer sequence informations were shown in Table 1.

TABLE 1 lncRNA primer sequence information.

lncRNAs	Primer sequence (5'-3')	Product size/bp
MSTRG12675	F:AGGCAAAGAACAGTCAGGCA	136
	R:TGCTGGTAATTGAGGGTCGG	
LOC102172108	F:AGGTGTGTGTACTGCGAGG	113
	R:CTTCCACTAACTTGCCGGGT	
LOC106503513	F:GACCTTGACTGTGAAGCGGA	121
	R:TTCTCAAATCACCGGGGTG	
MSTRG25656	F:CATGGCAAGCCGCTATTGAC	127
	R:AGAACCAGCCACCATTCTG	
MSTRG29368	F:CCTGAGGCCTCCGTGAAAT	127
	R:GAAACTCTGTGCCGGACTGA	
MSTRG15634	F:GAGACCCAGTGAACCAAGA	140
	R:TGCCCTCTGCCCTGAATTT	
U6	F:CGCACAGACATACGTCCCC	156
	R:TGGTCGGCAGTAAAGCAGAAT	
FGF9	F:AACTGGTACAACAGTACTCC	123
	R:TTTCTGGTGCCGCTTAGTCC	
PDGFRA	F:TCACGGAGATCACCACTGACA	86
	R:GTCTTCTTCTTTGCTCGGAT	
HGF	F:TGCCATTCAAATCGTCCTG	92
	R:ATTGTGGTGCCTTATACGTT	
MAPK10	F:ATGAGCCTCCATTATTGTAC	148
	R:TTTGAGAACCCTAAACGTTG	

Note: F, forward primer; R, reverse primer.

TABLE 2 Quality of data output.

Group	Raw reads	Clean	Clean	Q30/%
Group		Reads/G	Bases/G	
CK1	128,700,068	121,003,698	94.02	92.24
CK2	130,917,490	124,177,272	94.85	92.58
5i8d1	154,055,668	144,290,922	93.66	92.48
5i8d2	125,726,024	122,785,186	97.66	92.14

CK and 5i8d represent goat ear fibroblasts and reprogrammed mammary epithelial cells.

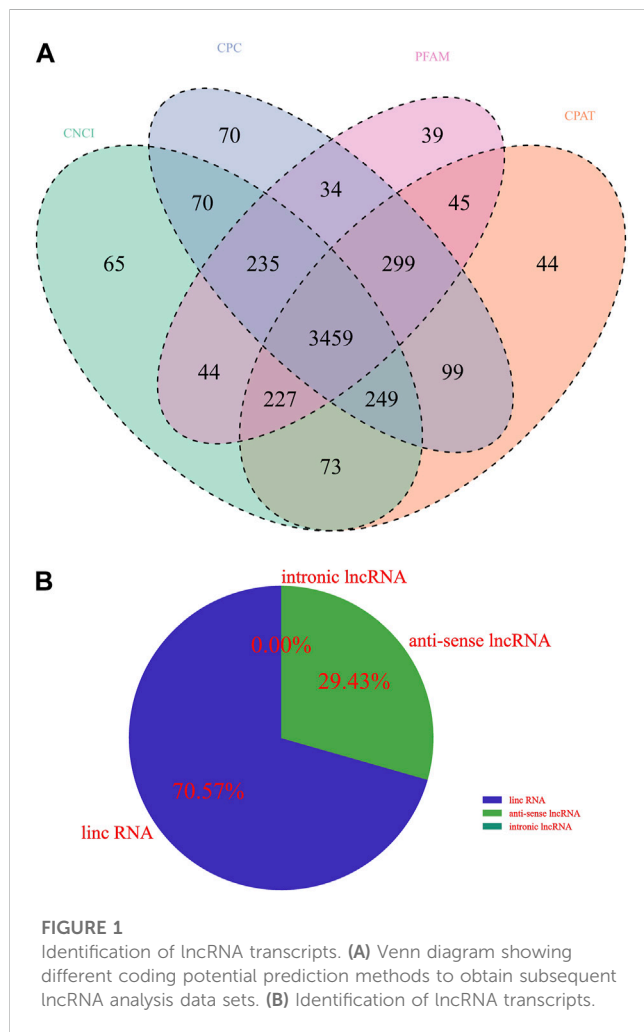
3 Results

3.1 Raw data analysis and quality control

The original downstream sequences were filtered to obtain high-quality clean reads, and then the subsequent analysis was performed. As shown in Table 2, the sizes of the total clean reads of the four groups of samples were 94.02 G, 94.85 G, 93.66 G and 97.66 G. The Q30 ratios were all higher than 90%, indicating that the sequencing quality was reliable.

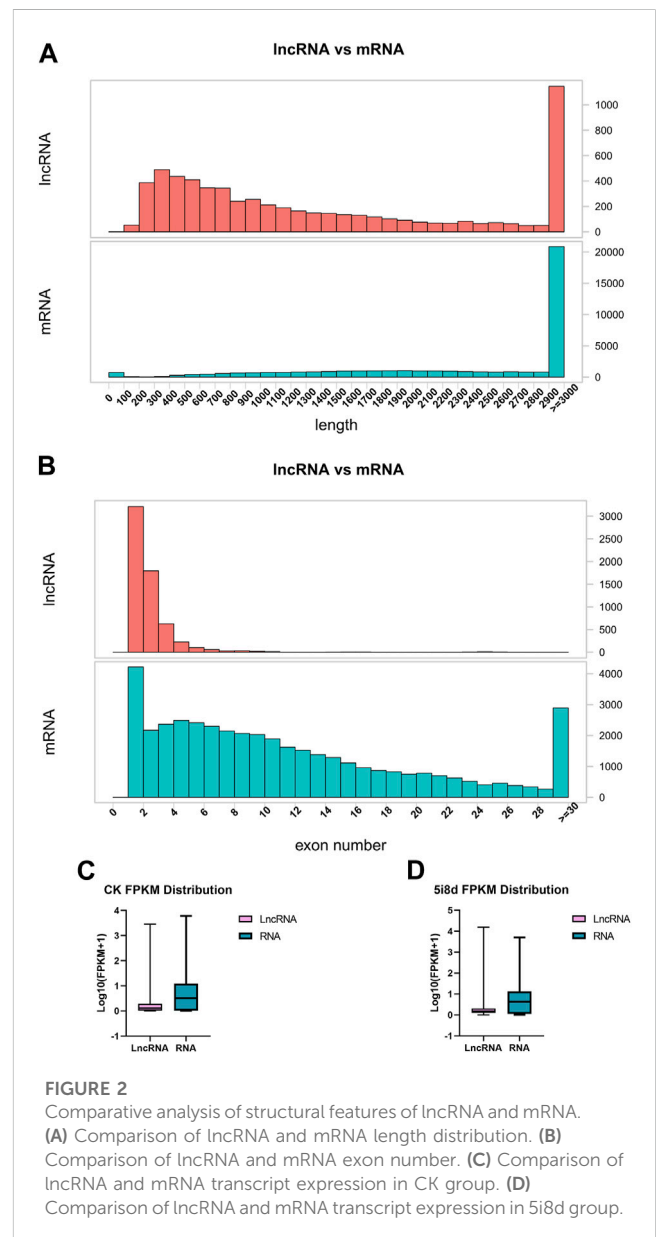
TABLE 3 Comparison of reads and reference genome.

Group	Total number of filtered reads	Number of reads that can be localized to the genome	Number of reads with multiple comparison positions on the reference sequence	Number of reads that have a unique comparison position on the reference sequence
Group	Total reads	Total mapped reads	Multiple mapped reads	Uniquely mapped reads
CK1	121,003,698	116,779,050 (96.51%)	2,631,155 (2.17%)	114,147,895 (94.34%)
CK2	124,177,272	120,090,486 (96.71%)	2,587,299 (2.08%)	117,503,187 (94.63%)
5i8d1	144,290,922	138,150,865 (95.74%)	4,248,247 (2.94%)	133,902,618 (92.80%)
5i8d2	122,785,186	118,813,067 (96.76%)	2,742,768 (2.23%)	116,070,299 (94.53%)

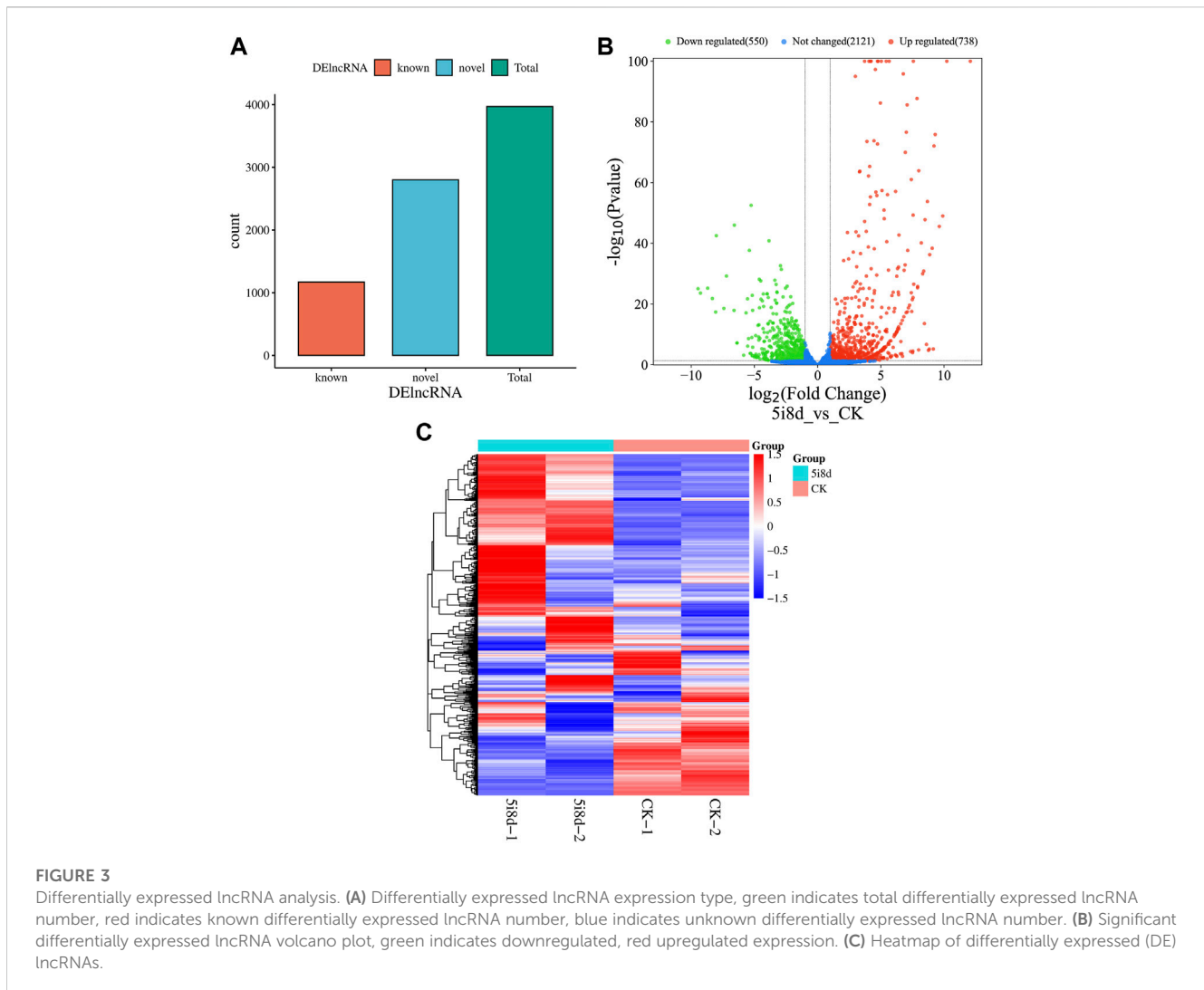


3.2 Sequence matching and splicing

The results of the quality-controlled clean reads compared to the goat reference genome were shown in Table 3. The results show that more than 96% of the clean reads in the control samples could locate sequenced sequences on the genome, 2.08%–2.17% had multiple alignment positions on the reference sequence, and the number of reads with unique alignment positions on the reference sequence accounted for more than 94% of the reads. The percentages of clean reads in



the 5i8d samples that were compared to the genome were all greater than 95%, the percentage of reads that were compared to the reference sequence with multiple comparison positions was



2.23%–2.94%, and the percentage of reads with unique comparison positions was 92.80%–94.53%.

3.3 Identification of lncRNA transcripts

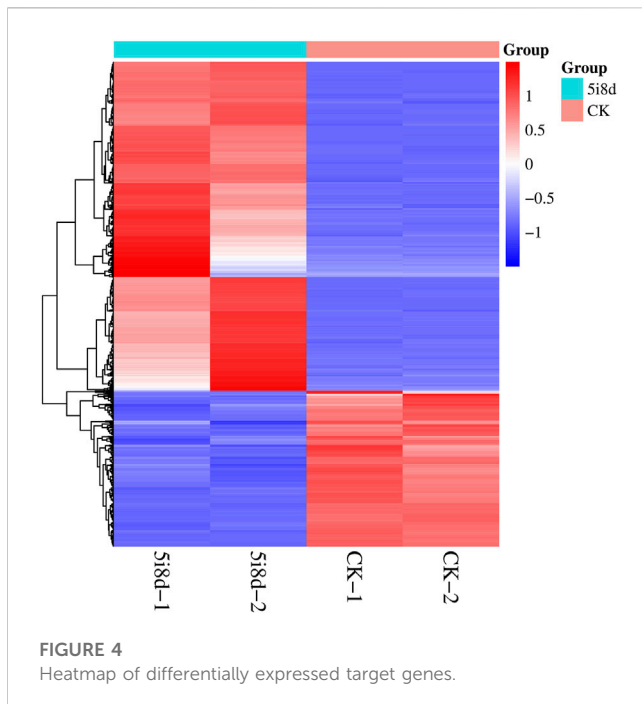
Whether a new transcript has coding potential is the key condition to determine whether it is an lncRNA. The lncRNA information from the previous step was screened by combining various coding potential analysis software, mainly coding-noncoding index (CNCI) analysis, coding potential calculator (CPC) analysis, Pfam-scan protein structural domain analysis, and Coding Potential Assessment Tool (CPAT) analysis (animal only). The noncoding transcripts identified by the above coding potential prediction methods were counted, and the common and unique numbers of each method were displayed as Venn diagrams. The intersection of the predicted results was taken as the subsequent data set for novel lncRNA analysis (Figure 1A). Upon further classification of unknown lncRNAs, as shown in Figure 1B, lincRNAs accounted for 70.57%, intronic lncRNAs accounted for 29.43%, and no antisense lncRNAs were found.

3.4 Structure and characterization of lncRNAs

To investigate the basic characteristics of lncRNAs in fibroblasts and induced mammary epithelial cells of dairy goats in Guanzhong, the transcript length, exon number and transcript expression levels of lncRNAs and mRNAs were compared after bioinformatics analysis. The results showed that compared with mRNA, the lncRNA length interval was concentrated between 200–3,000 bp (Figure 2A), and the number of lncRNA exons was between 1 and 11, mostly between 1 and 3 (Figure 2B). The expression of lncRNAs obtained from 0 d to 8 d sequencing was compared with that of mRNAs, and the overall expression of mRNAs was found to be significantly higher than that of lncRNAs (Figures 2 C, D).

3.5 Differential expression analysis of lncRNAs

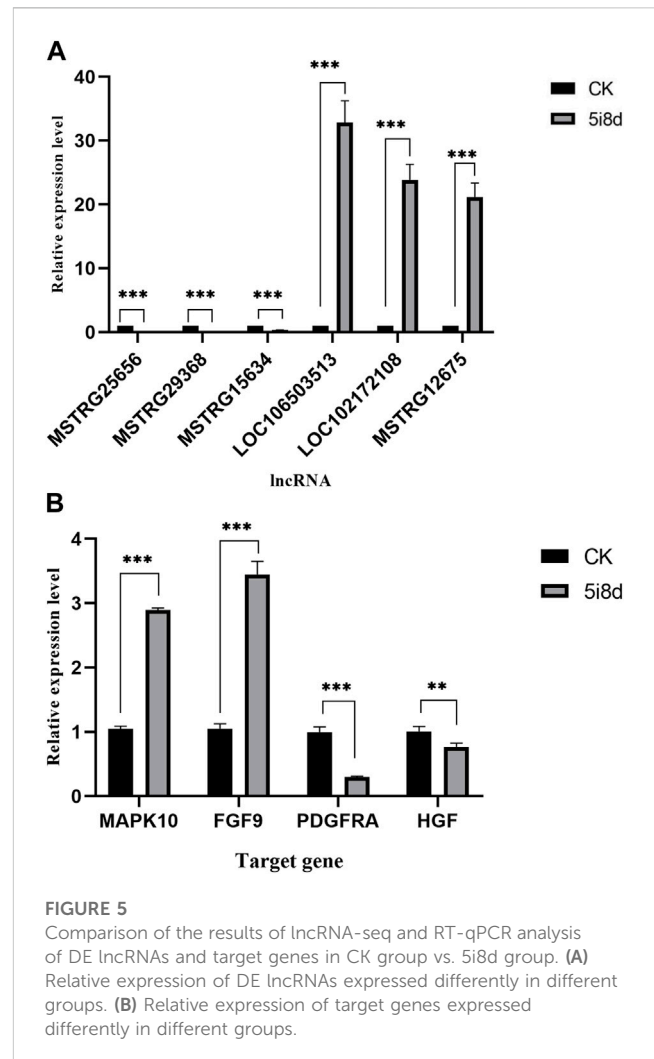
DEseq was used for differential expression analysis, and genes with $|\log_2\text{Fold Change}|\geq 2$ and $p < 0.05$ were selected as significant



differential expression screening conditions. The screening results were shown in Figure 3. The results showed that 3,970 candidate differential lncRNAs were obtained in the 5i8d group and the CK group, which included 1,170 known lncRNAs and 2,800 novel lncRNAs (Figure 3A). Further differential expression analysis of the goat fibroblast group and induced mammary epithelial cell group resulted in 1,288 significantly differentially expressed lncRNAs, of which 738 were upregulated and 550 were downregulated (Figure 3B). The heat maps of DE lncRNAs revealed different expression patterns between the 5i8d group and the CK group (Figure 3C).

3.6 Differentially expressed lncRNA-targeted mRNA prediction

To further investigate the role of DE lncRNAs in the reprogramming process of mammary epithelial cells, the possible target genes of DE lncRNAs were predicted by target analysis through cis and trans methods, and their functions were indirectly predicted through target genes. The target gene prediction results showed that there were 3,970 different lncRNAs, of which 2,800 were unknown and 1,170 were known, corresponding to 5,003 target genes. Among them, there were multiple significantly different lncRNAs with common target genes, such as MSTRG. 100417, MSTRG. 10063, and MSTRG. 10088 all correspond to the target gene MAPK10, and there were also lncRNAs with multiple variable shears targeting multiple coding genes, such as the coding genes FLT1, FGF9, HGF, and PDGFRA, which were potential targets of MSTRG.100336, and these lncRNAs exhibited regulatory effects in different directions. The heat maps of differently expressed target genes revealed different expression patterns between the 5i8d group and the CK group (Figure 4).



3.7 Validation of lncRNAs and target genes expression by RT-PCR

To further validate the sequencing results and detect their expression in induced goat mammary epithelial cells, six differentially expressed lncRNAs were randomly selected, in which MSTRG12675, LOC102172108, and LOC106503513 were upregulated in the transcriptome data, MSTRG25656, MSTRG29368, and MSTRG15634 were downregulated in the transcriptomic data. Four candidate target genes were selected at the same time, among which MAPK10 and FGF9 were upregulated in transcriptome data, and PDGFRA and HGF were downregulated in transcriptome data. Its expression in goat fibroblasts and induced goat mammary epithelial cells was measured by RT-qPCR. As shown, the RT-qPCR results showed that the expression of MSTRG12675, LOC102172108, and LOC106503513 was upregulated and that of MSTRG25656, MSTRG29368, and MSTRG15634 was downregulated (Figure 5A), which were all consistent with the results of lncRNA-seq. The RT-qPCR results showed that the expression of MAPK10 and FGF9 was upregulated, and the expression of PDGFRA and HGF was downregulated (Figure 5B), which were all consistent with the sequencing results. This result demonstrated that the RT-qPCR results were consistent with the expression trend of the transcriptome sequencing results.

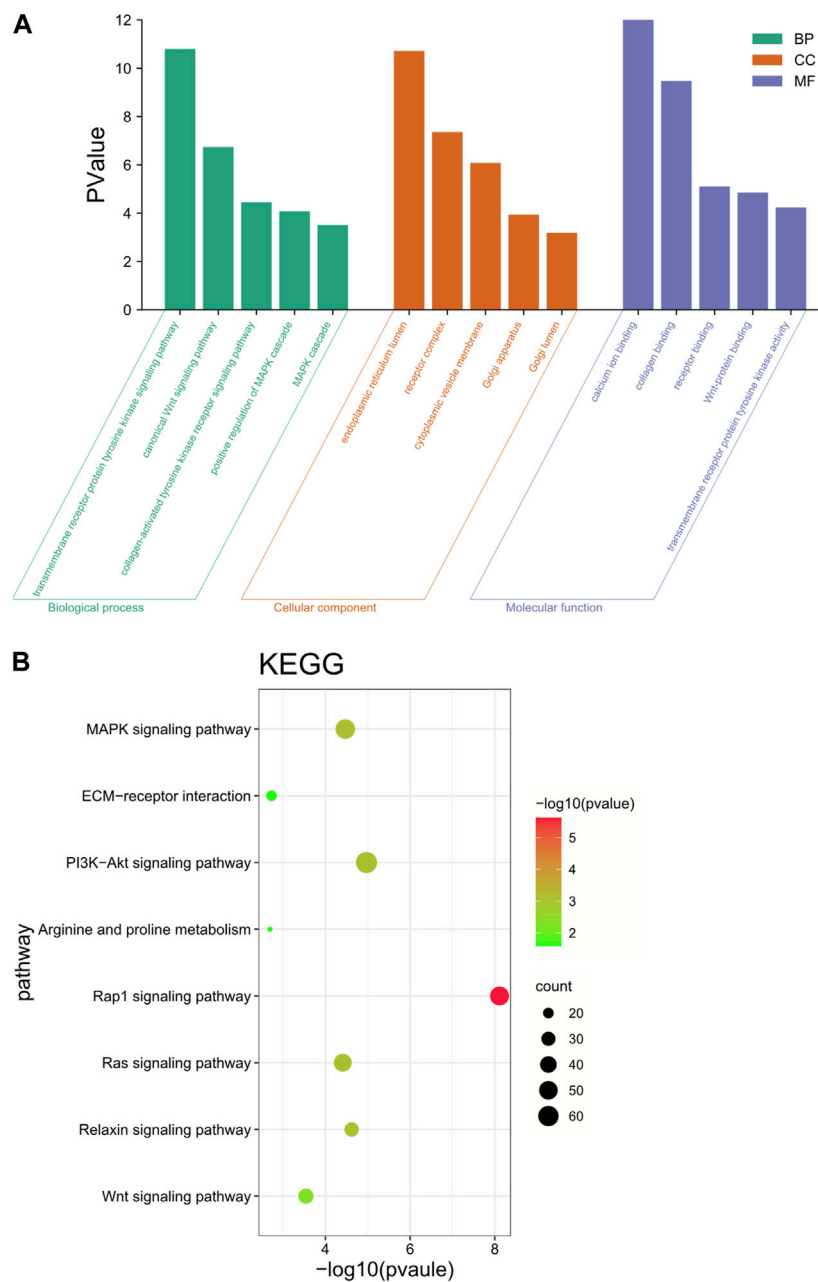


FIGURE 6 Functional annotation analysis of target genes. **(A)** GO enrichment analysis of differential lncRNA target genes. **(B)** KEGG enrichment analysis of differential lncRNA target genes.

3.8 Functional annotation analysis of target genes: GO and KEGG

3.8.1 GO enrichment analysis

These neighboring $|\log_2(\text{Ratio})| \geq 2$ and $p < 0.05$ differentially significant target genes were analyzed for functional annotation, and the GO results were classified into three entries: biological process (BP), cellular component (CC) and molecular function (MF). The GO results were classified by three entries, namely, biological process (BP), cellular component (CC) and molecular function (MF). The results were shown in Figure 6A. These target

genes were involved in the regulation of a variety of biological processes, including the regulation of metabolic processes, protein synthesis, receptor binding and regulation of tyrosine kinase activity.

3.8.2 KEGG enrichment analysis

KEGG enrichment results showed that these target genes were mainly involved in the Wnt signaling pathway, PI3K-Akt signaling pathway, arginine and proline metabolism, ECM-receptor interaction, and MAPK signaling pathway (Figure 6B). These genes were significantly enriched in pathways that are associated with reprogramming and mammary gland development, while

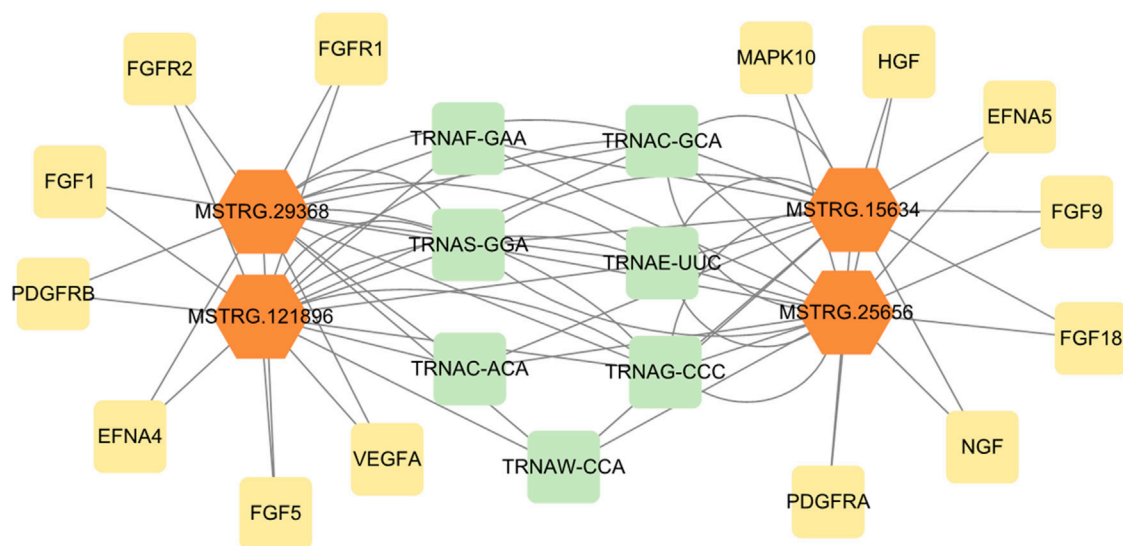


FIGURE 7
The regulatory network of differentially expressed lncRNAs and target genes.

lncRNAs target these genes, thus suggesting that lncRNAs play a key role in reprogramming and mammary gland development.

3.9 Interaction network of lncRNA and mRNA

The regulatory network analysis of differential lncRNAs with target genes was plotted based on the relationship of the identified differentially expressed lncRNA genes with mRNA genes and genes predicted by the cis and trans targets of lncRNAs (Figure 7). Yellow and green were target genes, and orange were lncRNAs. The results show that MSTRG.15634, MSTRG.25656, MSTRG.121896, and MSTRG.29368 were the differential lncRNAs that simultaneously target the most genes related to signaling pathways of mammary gland development and reprogramming, with MSTRG.15634 and MSTRG.25656 in MSTRG.15634 and MSTRG.25656 being the most upregulated differential lncRNAs that simultaneously targeted the above genes, and MSTRG.121896 and MSTRG.29368 being the most downregulated differential lncRNAs that targeted the above genes. The network plot results showed that MSTRG.15634 and MSTRG.25656 targeted MAPK10, FGF9, HGF, FGF18, EFNA5, PDGFRA, NGF, and MSTRG.121896 and MSTRG.29368 targeted FGF1, FGFR1, FGF5, FGFR2, PDGFRB, VEGFA, and EFNA4, while the above four lncRNAs targeted 7 transporter RNAs.

4 Discussion

lncRNAs play important roles in many biological activities, such as the dosage compensation effect, epigenetic regulation, cell cycle regulation, and cell differentiation regulation (Xiao et al., 2009), which has become a research hotspot in different scientific directions. In this study, transcriptome sequencing analysis of lncRNAs in Guanzhong dairy goat fibroblasts and induced mammary epithelial cells was

performed, and then the target genes of lncRNAs were further predicted to reveal their functions. This study provides a theoretical basis for exploring the role of lncRNAs in reprogramming and lactation.

A total of 4,317 lncRNAs and 3,970 DELncRNAs were detected before and after the induction of the small molecule compounds. A total of 1,288 of these lncRNAs were significantly differentially expressed. In the present study, transcriptome sequencing analysis of lncRNAs in fibroblasts and induced mammary epithelial cells of dairy goats in Guanzhong was performed, and then the target genes of lncRNAs were further predicted to reveal their functions. The sequencing analysis results were consistent with those of (Li et al., 2019). These similarities suggest that the lncRNAs analyzed in this study were more reliable.

The expression level was verified by qRT-PCR, and the trend was consistent with the lncRNA-Seq data. We enriched for differentially expressed lncRNA target genes, and GO enrichment revealed that the tyrosine kinase activity and WNT pathway involved in significantly enriched GO term entries were associated with mammary gland development and reprogramming; for example, tyrosine kinase influences mammary gland differentiation through prolactin receptors (Xie et al., 2002). Analysis and quantification of pluripotent pluripotency reprogramming showed that the WNT pathway undergoes extensive reprogramming and regulates various cellular and developmental processes (Hu et al., 2020). Therefore, these lncRNAs may play an important regulatory role in the process of mammary gland development.

KEGG enrichment significantly contains the MAPK signaling pathway, PI3K-Akt signaling pathway, Wnt signaling pathway, and relaxin signaling pathway related to mammary gland development; arginine and proline metabolism is a related pathway affecting milk protein synthesis by influencing the amino acid metabolism of milk proteins (Manni et al., 1998; Xie et al., 2002; HSU et al., 2005; Dinger et al., 2008; Xiao et al., 2009; Loewer et al., 2010; Guttman et al., 2011; Jardé and Dale, 2011; Meng et al., 2018; Li et al., 2019; Hu et al., 2020; Zhang et al., 2021; Sherman et al., 2022). There were also signaling pathways related to reprogramming ECM-receptor interactions, the Rap1 signaling pathway, the MAPK pathway, and the Ras signaling

pathway (Mezl and Knox, 1977; Zwartkruis and Bos, 1999; Dickinson et al., 2010; Lopez-Vicente et al., 2011; Zhang et al., 2016; Xu et al., 2019; Dillon et al., 2021). Several genes exist in these pathways that were simultaneously involved in subprocesses, including MAPK10, FGF9, HGF, FGF18, EFNA5, PDGFRA, NGF, FGF1, FGFR1, FGF5, FGFR2, PDGFRB, VEGFA, and EFNA4. MAPK10 is a member of the MAPK signaling pathway, which is an important pathway for cell differentiation and proliferation. It has been shown that Fgf signaling affects the induction and development of the embryonic mammary gland by binding to a class of cell surface enzymes belonging to the family of receptor tyrosine kinases (FgfRs) (Dillon et al., 2004) that Fgf signaling has a role in normal mammary gland lobular alveolar development (Jackson et al., 1997a) and that Fgf signaling is also required for pregnancy-dependent lobular alveolar development in the mammary gland (Jackson et al., 1997b). HGF has mitogenic effects on human and mouse mammary epithelial cells, as well as kinetic and morphological effects, including the induction of extensive ductal branching and the formation of good lumens (Niranjan et al., 1995). HGF plays a key role in the estrogen-induced proliferation of mammary epithelial cells *in vivo* (Zhang et al., 2002). EfnA5, a member of the ephrin gene family, has receptors whose spatiotemporal expression patterns are important for the morphogenesis and function of mammary epithelial tissue, and the Eph receptor acts as a signaling factor mediating hormones that are critical for the process of mammary gland expansion and regression (Perez White and Getsios, 2014). PDGF-BB expression upregulation affects mammary tissue fibrosis after *Staphylococcus aureus* infection, and PDGF-BB was found to be associated with breast fibrosis (Bi et al., 2020). Stromal PDGFRA signaling disrupts ECM homeostasis during mammary gland development, leading to increased mammary gland stiffness and increased tumor growth potential (Hammer et al., 2017). NGF and its high-affinity receptors are expressed during mammary gland development and lactation in sheep, and this expression is dependent on the stage of mammary gland development and lactation. NGF and its cognate receptors are expressed during mammary gland development, lactation, and degeneration in sheep through phosphatidylinositol 3-kinase (PI3K), and protein kinase B (AKT) plays a central role in promoting mammary cell survival during mammary gland development, lactation and degeneration (Colitti, 2015). Under the influence of hormonal and epithelial-stromal interactions, VEGF has multiple roles in the development and function of mammary glands, such as contributing to vascular development and supporting fat pads within the mammary gland (Hovey et al., 2001). Vascular endothelial growth factor affects mammary gland development by regulating mammary angiogenesis (Dangat et al., 2020).

In conclusion, all of the above target genes showed their role in reprogramming and mammary gland development relationships in previous studies, while the functional enrichment analysis of the present sequencing results showed that these target genes were also involved in reprogramming and mammary gland development-related pathways, among which two lncRNAs, MSTRG.15634 and MSTRG.25656, target both MAPK10, FGF9, HGF, MSTRG.25656, MSTRG.29368, and MSTRG.121896. The four lncRNAs involved in the above mentioned signaling pathways, through the regulation of their target genes, allowed fibroblasts to undergo fate transformation and obtain functional mammary epithelial cells.

5 Conclusion

We used lncRNA-Seq data of the transdifferentiation process to understand the expression profile of lncRNAs in the reprogramming of Guanzhong dairy goat fibroblasts into mammary epithelial cells, while analysis of the data revealed that four lncRNAs, MSTRG.15634, MSTRG.25656, MSTRG.29368, and MSTRG.121896, may have a key role in mammary gland development and lactation. Our findings can provide new ideas for further study of lncRNAs and mRNAs associated with mammary gland development and reprogramming and provide a theoretical basis for further elaboration of the reprogramming process from dairy goat fibroblasts to mammary epithelial cells in Guanzhong and the molecular mechanism of mammary gland development.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://www.ncbi.nlm.nih.gov/geo/GSE142551>.

Ethics statement

The animal studies were approved by the Experimental Committee of Animal Ethics Review of Guangxi University, Guangxi University. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent was obtained from the owners for the participation of their animals in this study.

Author contributions

BH designed the research. DL collected the data. LQ and GW cultured cells. ML and QL collected literature-related information. MW and WT analyzed the data and drafted the manuscript. BH and DZ revised the final manuscript. All authors contributed to the article and approved the submitted version.

Funding

This research was partially supported by the National Natural Science Foundation of China grants: No. 32160171 and No. 31960160.

Acknowledgments

We thank Anno Youda Gene Technology Co., Ltd. for technical support.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Bi, Y., Ding, Y., Wu, J., Miao, Z., Wang, J., and Wang, F. (2020). *Staphylococcus aureus* induces mammary gland fibrosis through activating the TLR/NF- κ B and TLR/AP-1 signaling pathways in mice. *Microb. Pathog.* 148, 104427. doi:10.1016/j.micpath.2020.104427
- Colitti, M. (2015). Expression of NGF, BDNF and their high-affinity receptors in ovine mammary glands during development and lactation. *Histochem. Cell Biol.* 144 (6), 559–570. doi:10.1007/s00418-015-1360-0
- Dangat, K., Khaire, A., and Joshi, S. (2020). Cross talk of vascular endothelial growth factor and neurotrophins in mammary gland development. *Growth factors.* 38 (1), 16–24. doi:10.1080/08977194.2020.1792469
- Dickinson, L. E., Kusuma, S., and Gerecht, S. (2010). Reconstructing the differentiation niche of embryonic stem cells using biomaterials. *Macromol. Biosci.* 11 (1), 36–49. doi:10.1002/mabi.201000245
- Dillon, C., Spencer-Dene, B., and Dickson, C. (2004). A crucial role for fibroblast growth factor signaling in embryonic mammary gland development. *J. Mammary Gl. Biol. Neoplasia* 9 (2), 207–215. doi:10.1023/b:jomg.0000037163.56461.1e
- Dillon, M., Lopez, A., Lin, E., Sales, D., Perets, R., and Jain, P. (2021). Progress on Ras/MAPK signaling research and targeting in blood and solid cancers. *Cancers* 13 (20), 5059. doi:10.3390/cancers13205059
- Dinger, M. E., Amaral, P. P., Mercer, T. R., Pang, K. C., Bruce, S. J., Gardiner, B. B., et al. (2008). Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. *Genome Res.* 18 (9), 1433–1445. doi:10.1101/gr.078378.108
- Guttman, M., Donaghy, J., Carey, B. W., Garber, M., Grenier, J. K., Munson, G., et al. (2011). lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature* 477 (7364), 295–300. doi:10.1038/nature10398
- Guttman, M., and Rinn, J. L. (2012). Modular regulatory principles of large noncoding RNAs. *Nature* 482 (7385), 339–346. doi:10.1038/nature10887
- Hammer, A. M., Sizemore, G. M., Shukla, V. C., Avendano, A., Sizemore, S. T., Chang, J. J., et al. (2017). Stromal PDGFR- α activation enhances matrix stiffness, impedes mammary ductal development, and accelerates tumor growth. *Neoplasia* 19 (6), 496–508. doi:10.1016/j.neo.2017.04.004
- Hovey, R. C., Goldhar, A. S., Baffi, J., and Vonderhaar, B. K. (2001). Transcriptional regulation of vascular endothelial growth factor expression in epithelial and stromal cells during mouse mammary gland development. *Mol. Endocrinol.* 15 (5), 819–831. doi:10.1210/mend.15.5.0635
- Hsu, S. Y. T., Semyonov, J., Park, J.-I., and Chang, C. L. (2005). Evolution of the signaling system in relaxin-family peptides. *Ann. N. Y. Acad. Sci.* 1041 (1), 520–529. doi:10.1196/annals.1282.078
- Hu, K., Ianov, L., and Crossman, D. (2020). Profiling and quantification of pluripotency reprogramming reveal that WNT pathways and cell morphology have to be reprogrammed extensively. *Heliyon* 6 (5), e04035. doi:10.1016/j.heliyon.2020.e04035
- Jackson, D., Bresnick, J., and Dickson, C. (1997a). A role for fibroblast growth factor signaling in the lobuloalveolar development of the mammary gland. *J. Mammary Gl. Biol. Neoplasia* 2 (4), 385–392. doi:10.1023/a:1026351414004
- Jackson, D., Bresnick, J., Rosewell, I., Crafton, T., Poulosom, R., Stamp, G., et al. (1997b). Fibroblast growth factor receptor signalling has a role in lobuloalveolar development of the mammary gland. *J. Cell Sci.* 110 (11), 1261–1268. doi:10.1242/jcs.110.11.1261
- Jardé, T., and Dale, T. (2011). Wnt signalling in murine postnatal mammary gland development. *Acta Physiol.* 204 (1), 118–127. doi:10.1111/j.1748-1716.2011.02283.x
- Kopp, F., and Mendell, J. T. (2018). Functional classification and experimental dissection of long noncoding RNAs. *Cell* 172 (3), 393–407. doi:10.1016/j.cell.2018.01.011
- Li, C., Li, X., Liu, Z., Ni, W., Zhang, X., Hazi, W., et al. (2019). Identification and characterization of long noncoding RNA in prenatal and postnatal skeletal muscle of sheep. *Genomics* 111 (2), 133–141. doi:10.1016/j.ygeno.2018.01.009
- Liu, W., Liu, X., Luo, M., Liu, X., Luo, Q., Tao, H., et al. (2017). dNK derived IFN- γ mediates VSMC migration and apoptosis via the induction of lincRNA MEG3: A role in uterovascular transformation. *Placenta* 50, 32–39. doi:10.1016/j.placenta.2016.12.023
- Liu, Y., Sun, Z., Zhu, J., Xiao, B., Dong, J., and Li, X. (2018). lincRNA-TCONS_00034812 in cell proliferation and apoptosis of pulmonary artery smooth muscle cells and its mechanism. *J. Cell. Physiology* 233 (6), 4801–4814. doi:10.1002/jcp.26279
- Loewer, S., Cabili, M. N., Guttman, M., Loh, Y.-H., Thomas, K., Park, I. H., et al. (2010). Large intergenic noncoding RNA-RoR modulates reprogramming of human induced pluripotent stem cells. *Nat. Genet.* 42 (12), 1113–1117. doi:10.1038/ng.710
- Lopez-Vicente, L., Pons, B., Coch, L., Teixido, C., Hernandez-Losa, J., Armengol, G., et al. (2011). RSK4 inhibition results in bypass of stress-induced and oncogene-induced senescence. *Carcinogenesis* 32 (4), 470–476. doi:10.1093/carcin/bgr003
- Lu, Y., Liu, X., Xie, M., Liu, M., Ye, M., Li, M., et al. (2017). The NF- κ B-Responsive long noncoding RNA FIRRE regulates posttranscriptional regulation of inflammatory gene expression through interacting with hnRNP U. *J. Immunol.* 199 (10), 3571–3582. doi:10.4049/jimmunol.1700091
- Manni, A., Wechter, R., Verderame, M. F., and Mauger, D. (1998). Cooperativity between the polyamine pathway and HER-2neu in transformation of human mammary epithelial cells in culture: role of the MAPK pathway. *Int. J. Cancer* 76 (4), 563–570. doi:10.1002/(sici)1097-0215(19980518)76:4<563::aid-ijc20>3.0.co;2-7
- Meng, Y., Zhang, J., Yuan, C., Zhang, F., Fu, Q., Su, H., et al. (2018). Oleic acid stimulates HC11 mammary epithelial cells proliferation and mammary gland development in peripubertal mice through activation of CD36-Ca2+ and PI3K/Akt signaling pathway. *Oncotarget* 9 (16), 12982–12994. doi:10.18632/oncotarget.24204
- Mezl, V. A., and Knox, W. E. (1977). Metabolism of arginine in lactating rat mammary gland. *Biochem. J.* 166 (1), 105–113. doi:10.1042/bj1660105
- Nan, A., Jia, Y., Li, X., Liu, M., Zhang, N., Chen, L., et al. (2017). Editor's highlight: lincRNAL20992 regulates apoptotic proteins to promote lead-induced neuronal apoptosis. *Toxicol. Sci.* 161 (1), 115–124. doi:10.1093/toxsci/kfx203
- Niranjana, B., Buluwela, L., Yant, J., Perusinghe, N., Atherton, A., Phippard, D., et al. (1995). HGF/SF: A potent cytokine for mammary growth, morphogenesis and development. *Development* 121 (9), 2897–2908. doi:10.1242/dev.121.9.2897
- Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., et al. (2002). Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature* 420 (6915), 563–573. doi:10.1038/nature01266
- Perez White, B. E., and Getsios, S. (2014). Eph receptor and ephrin function in breast, gut, and skin epithelia. *Cell Adhesion Migr.* 8 (4), 327–338. doi:10.4161/19336918.2014.970012
- Qin, Y., Sun, W., Zhang, H., Zhang, P., Wang, Z., Dong, W., et al. (2018). lincRNA GAS8-AS1 inhibits cell proliferation through ATG5-mediated autophagy in papillary thyroid cancer. *Endocrine* 59 (3), 555–564. doi:10.1007/s12020-017-1520-1
- Sherman, B. T., Hao, M., Qiu, J., Jiao, X., Baseler, M. W., Lane, H. C., et al. (2022). David: A web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res.* 50 (W1), W216–W221. doi:10.1093/nar/gkac194
- Song, G., Shen, Y., Ruan, Z., Li, X., Chen, Y., Yuan, W., et al. (2016). lincRNA-uc.167 influences cell proliferation, apoptosis and differentiation of P19 cells by regulating Mef2c. *Gene* 590 (1), 97–108. doi:10.1016/j.gene.2016.06.006
- Standaert, L., Adriaens, C., Radaelli, E., Van Keymeulen, A., Blanpain, C., Hirose, T., et al. (2014). The long noncoding RNA Neat1 is required for mammary gland development and lactation. *RNA* 20 (12), 1844–1849. doi:10.1261/rna.047332.114
- Tong, C., Chen, Q., Zhao, L., Ma, J., Ibeagha-Awemu, E. M., and Zhao, X. (2017). Identification and characterization of long intergenic noncoding rnas in bovine mammary glands. *BMC Genomics* 18 (1), 468. doi:10.1186/s12864-017-3858-4
- Touat Todeschini, L., Shichino, Y., Dangin, M., Thierry Mieg, N., Gilquin, B., Hiriart, E., et al. (2017). Selective termination of lincRNA transcription promotes heterochromatin silencing and cell differentiation. *EMBO J.* 36 (17), 2626–2641. doi:10.15252/embj.201796571
- Ulitsky, I., Shkumatava, A., Jan, Calvin H., Sive, H., and Bartel, David P. (2011). Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell* 147 (7), 1537–1550. doi:10.1016/j.cell.2011.11.055
- Wang, H., Cao, Q., Ge, J., Liu, C., Ma, Y., Meng, Y., et al. (2014). lincRNA-regulated infection and inflammation pathways associated with pregnancy loss: genome wide differential expression of lincRNAs in early spontaneous abortion. *Am. J. Reproductive Immunol.* 72 (4), 359–375. doi:10.1111/aji.12275
- Wang, J., Hua, L., Chen, J., Zhang, J., Bai, X., Gao, B., et al. (2017). Identification and characterization of long noncoding RNAs in subcutaneous adipose tissue from castrated and intact full-SIB pair Huainan Male Pigs. *BMC Genomics* 18 (1), 542. doi:10.1186/s12864-017-3907-z
- Wang, L., Zhao, Y., Bao, X., Zhu, X., Kwok, Y. K., Sun, K., et al. (2015). lincRNA Dum interacts with Dnmts to regulate Dppa2 expression during myogenic differentiation and muscle regeneration. *Cell Res.* 25 (3), 335–350. doi:10.1038/cr.2015.21
- Xiao, B., Zhang, X., Li, Y., Tang, Z., Yang, S., Mu, Y., et al. (2009). Identification, bioinformatic analysis and expression profiling of candidate mRNA-like noncoding RNAs in *Sus scrofa*. *J. Genet. Genomics* 36 (12), 695–702. doi:10.1016/s1673-8527(08)60162-9

- Xie, J., LeBaron, M. J., Nevalainen, M. T., and Rui, H. (2002). Role of tyrosine kinase Jak2 in prolactin-induced differentiation and growth of mammary epithelial cells. *J. Biol. Chem.* 277 (16), 14020–14030. doi:10.1074/jbc.m112399200
- Xu, W., Chen, X., Wang, Y., Fan, B., Guo, K., Yang, C., et al. (2019). Chitooligosaccharide inhibits RANKL-induced osteoclastogenesis and ligation-induced periodontitis by suppressing MAPK/c-fos/NFATC1 signaling. *J. Cell. Physiology* 235 (3), 3022–3032. doi:10.1002/jcp.29207
- Yu, S., Zhao, Y., Lai, F., Chu, M., Hao, Y., Feng, Y., et al. (2017). LncRNA as ceRNAs may be involved in lactation process. *Oncotarget* 8 (58), 98014–98028. doi:10.18632/oncotarget.20439
- Zhang, D., Ren, Y., Qin, L., Liu, Q., Wang, G., Sun, L., et al. (2021). Transdifferentiation of goat ear fibroblasts into lactating mammary epithelial cells induced by small molecule compounds. *Biochem. Biophys. Res. Commun.* 573, 55–61. Epub 2021 Aug 5. PMID: 34388455. doi:10.1016/j.bbrc.2021.07.087
- Zhang, H.-Z., Bennett, J. M., Smith, K. T., Sunil, N., and Haslam, S. Z. (2002). Estrogen mediates mammary epithelial cell proliferation in serum-free culture indirectly via mammary stroma-derived hepatocyte growth factor. *Endocrinology* 143 (9), 3427–3434. doi:10.1210/en.2002-220007
- Zhang, H., Tao, J., Sheng, L., Hu, X., Rong, R., Xu, M., et al. (2016). Retracted: twist2 promotes kidney cancer cell proliferation and invasion via regulating ITGA6 and CD44 expression in the ECM-Receptor-Interaction pathway. *Biomed. Pharmacother.* 81, 453–459. doi:10.1016/j.biopha.2016.02.042
- Zhang, Y., Xia, J., Li, Q., Yao, Y., Eades, G., Gernapudi, R., et al. (2014). NRF2/Long noncoding RNA ROR signaling regulates mammary stem cell expansion and protects against estrogen genotoxicity. *J. Biol. Chem.* 289 (45), 31310–31318. doi:10.1074/jbc.m114.604868
- Zhou, M., Zhang, Z., Zhao, H., Bao, S., Cheng, L., and Sun, J. (2017). An immune-related six-lncRNA signature to improve prognosis prediction of glioblastoma multiforme. *Mol. Neurobiol.* 55 (5), 3684–3697. doi:10.1007/s12035-017-0572-9
- Zwartkruis, F. J., and Bos, J. L. (1999). Ras and Rap1: two highly related small GTPases with distinct function. *Exp. Cell Res.* 253 (1), 157–165. doi:10.1006/excr.1999.4695