

Identification of regulatory networks and hub genes controlling mammary gland development and lactation in dairy goats during the late lactation, dry period, and late gestation stages

Rong Xuan

Shandong Agricultural University

Tianle Chao

Shandong Agricultural University

Xiaodong Zhao

Shandong Agricultural University

Aili Wang

Shandong Agricultural University

Yunpeng Chu

Shandong Agricultural University

Qing Li

Shandong Agricultural University

Yilin Zhao

Shandong Agricultural University

Zhibin Ji

Shandong Agricultural University

Jianmin Wang (✉ wangjm@sdau.edu.cn)

Shandong Agricultural University

Research article

Keywords: Dairy goat, Mammary gland, Late lactation, Dry period, Late gestation

Posted Date: August 7th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-46961/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

From the late lactation to late gestation stages, the mammary gland tissue of goats undergoes a process from involution to remodeling and then to high differentiation of mammary gland tissue. From the perspective of lactation, this is a continuous development process of the goat mammary gland from the termination of lactation to the restoration of lactation. We performed transcriptome sequencing on goat mammary gland tissues at three mammary gland developmental stages to screen for differentially expressed genes that affect mammary gland development and the physiological process of lactation and mapped their expression profiles in three stages. The objective of this study is to reveal the expression characteristics of these genes and their potential function during mammary gland development and lactation.

Results

We identified 1,381 differentially expressed genes in the mammary gland during three stages and found that the expression level of genes encoding casein, such as alpha-s1-casein (*CSN1S1*), alpha-s2-casein (*CSN1S2*), beta-casein (*CSN2*), and kappa-casein (*CSN3*), and alpha-lactalbumin (*LALBA*) were higher in mammary gland tissues during the late lactation stage and late gestation stage than those during the dry period. In addition, we constructed six functional networks related to differentially expressed genes and found that these genes are closely related to mammary gland growth and development, apoptosis, immunity, substance transport, biosynthesis, and metabolism. Finally, we identified 35 differentially expressed transcription factors, which were classified into 16 families, and predicted that transcription factors of the basic leucine zipper domain (*bZIP*) family and basic helix-loop-helix (*bHLH*) family regulated the expression levels of genes related to mammary gland development and lactation.

Conclusions

Among the late lactation, dry period, and late gestation stages, there are differences in the expression levels of genes related to mammary gland growth and development, apoptosis, immunity, basic substance transport, biosynthesis, and metabolism in mammary gland tissues. Some genes in the same family or with similar functions have similar expression patterns. These differentially expressed genes or transcription factors work synergistically to participate in the regulation of mammary gland development and the physiological process of lactation.

Background

Mammary gland development and lactogenesis and the factors regulating these two processes have been topics of interest for quite some time. From the beginning, people knew the structure of the

mammary gland through anatomical methods [1–3]. Later, various molecular techniques were used to discover molecules related to mammary gland development and lactation [4, 5]. In recent years, the development and popularization of high-throughput sequencing technologies, such as RNA sequencing (RNA-Seq), have provided more powerful means for mammary gland research [6]. A number of molecules related to mammary gland growth, differentiation, metabolism, lactation, and diseases have been screened and identified. For example, Scheele et al. identified the molecular characteristics of mammary gland stem cells during branching morphogenesis and found that the transcription of mammary gland stem cells in the lumen and basal position might be different [7]. Kosciuczuk et al. discovered the differential expression of the mammary gland immune response-related genes *FOS* and tumor necrosis factor (*TNF*) through transcriptome analysis of mammary parenchyma tissues infected with *Staphylococcus aureus* [8]. Wagner et al. constructed single-cell atlases of mammary gland immunology and tumor ecosystems using single-cell sequencing technology [9]. In addition, in previous studies conducted in our laboratory, Laoshan dairy goats were used as a mammary gland research model to conduct studies on gene and noncoding molecules in mammary gland tissues at the early (20 d postpartum), peak (90 d postpartum) and late stages (210 d postpartum) of lactation. We mapped the expression profiles related to the regulation of mammary gland development and lactation at three stages and constructed a potential regulatory network [10–12]. Although scientists have made some progress in mammary gland research, for the complex physiological processes of mammary gland development and lactation, these findings are just the tip of the iceberg, requiring further exploration.

The physiological structure and function of the mammary gland undergo corresponding changes at different developmental stages [13]. After a long period of lactation, mammary gland tissues begin to involute, and mammary gland cells undergo apoptosis. The number of mammary acinar and duct branches decreases [14]. Moreover, during the late lactation stage, the milk synthesis and secretion capacities and cell viability of epithelial cells significantly decline [15]. Subsequently, the mammary gland of goats stops secreting milk and enters the dry period. Then, the mammary gland tissues undergo remodeling; on the one hand, the lactating epithelial cells in the involution are cleared; on the other hand, the mammary gland cells are regenerated, proliferating and differentiating to form new mammary gland branches and acinar structures, preparing for the next lactation [16]. In late gestation, the structural differentiation of mammary gland tissues is more mature, with more acinus and duct branch numbers. Moreover, mammary epithelial cells already begin to synthesize and secrete milk, and milk is stored in the acinus [17, 18].

Based on the significant changes in the structural and physiological functions of mammary gland tissue during these three stages, we wanted to understand the genes that might play a potential regulatory role in mammary gland development and lactation during these three consecutive mammary gland tissue development stages. Therefore, we performed transcriptome sequencing on the mammary gland tissues of dairy goats at the late lactation (LL), dry period (DP), and late gestation (LG) stages and used bioinformatics analysis to plot the expression profiles of differentially expressed genes (DEGs) at three developmental stages to screen regulatory factors related to mammary gland growth, differentiation,

lactation and immunity. We hope this study will further enrich the theoretical knowledge of lactation and mammary gland development and provide a theoretical basis for the rational breeding of dairy goats.

Results

Basic statistical analysis of the sequencing results

According to the statistical analysis of the sequencing results (Table 1), 147,531,532, 151,341,075, and 161,094,090 clean reads were obtained for mammary gland tissues from the three stages (LL, DP, and LG), respectively; these clean reads were used for downstream and reference genome alignment, and the alignment ratio of the reads was above 95%. Calculation of the expression levels showed 18,472, 19,941, and 19,184 expressed genes during the LL, DP, and LG stages, respectively. We then filtered out low-expression genes and obtained 13,598, 14,357, and 14,221 genes for the LL, DP, and LG stages, respectively, for expression abundance and differential expression analyses.

Table 1 Basic statistical analysis of the sequencing data

ID	Total Reads	Mapped Reads	Expressed Transcripts	Expressed Genes	Genes after Filtering (FPKM≥0.5)
LL	147531532	141040156 (95.6%)	45794	18472	13598
DP	151341075	143960330 (95.1%)	50726	19941	14357
LG	161094090	153729449 (95.4%)	49134	19184	14221

Analysis of gene abundance and identification of DEGs

Analysis of gene expression abundance showed that the LG stage had the largest number of highly abundant genes (129 genes), followed by the DP (104 genes) and LL (53 genes) stages (Table 2). Among the genes with moderate expression levels, the DP stage had the largest number of genes (5104 genes), and the LL stage had the least number of genes (2053 genes). The numbers of genes with low expression levels or ultralow expression levels in the LL stage were 10,122 and 1,370, respectively, much higher than those in the DP and LG stages. Principal component analysis and cluster analysis showed that the samples from the same developmental stage clustered, and correlation heat maps showed that the correlation coefficient for samples from the same developmental stage was greater than or equal to 0.92, higher than the correlation coefficient for samples from different stages (Fig. 1 A-B). This indicates that samples from the same developmental stage have a high correlation and good repeatability. Subsequently, a total of 1,381 DEGs were identified using differential expression analysis among samples (Fig. 2). In the LL vs DP comparison, 627 genes were upregulated, and 158 genes were downregulated (Fig. 2 A-E). Among the upregulated genes, butyrophilin subfamily 1 member A1 (*BTN1A1*), alpha-lactalbumin (*LALBA*), alpha-s1-casein (*CSN1S1*), *LOC102178810*, alpha-s2-casein

(*CSN1S2*), beta-casein (*CSN2*), kappa-casein (*CSN3*), lactoperoxidase (*LPO*), *LOC102169973* and *LOC102177798* had higher fold changes in gene expression, and log2-fold changes in these genes were greater than 10. For the downregulated genes, the absolute values of log2-fold changes in transthyretin (*TTR*), peptidase inhibitor 16 (*PI16*), and transient receptor potential cation channel subfamily v member 6 (*TRPV6*) were greater than four. In the DP vs LG comparison, there were 322 upregulated genes and 495 downregulated genes. In the LL vs LG comparison, there were 413 upregulated genes and 132 downregulated genes. Genes with higher fold changes above two are listed in Figure 2b and 2c. To further understand the expression change patterns of DEGs, we constructed global gene expression heat maps of 1381 genes and identified the top 20 DEGs in the three developmental stages (Fig. 2 F-I). In the LL stage, the expression patterns of *CSN3*, *CSN1S1*, *CSN1S2*, and *LALBA*, which are related to lactoprotein-encoding genes, were similar, and the expression levels were significantly lower in the DP stage than those in the LL and LG stages. In the DP stage, the expression patterns of ribosomal protein L23 (*RPL23*), ribosomal protein S3A (*RPS3A*), *RPL19*, *RPL13A*, *RPLP0*, and *RPS2*, which are related to the encoding of ribosomal protein subunits, were similar, and the expression levels of these genes were lowest in the LL stage and the highest LG stage. In the LG stage, we found the specific high expression of the immunoregulatory genes progesterone-associated endometrial protein (*PAEP*) and lactotransferrin (*LTF*). Please refer to Additional file 1 for specific information of differential expression analysis.

Table 2 Analysis of gene abundance in the three mammary gland developmental stages

Different expression levels	LL	DP	LG
Highly expressed genes (FPKM ≥ 500)	53	104	129
Mediumly expressed genes (500 > FPKM ≥ 10)	2053	5104	4591
Lowly expressed genes (10 > FPKM ≥ 1)	10122	8927	9113
Ultralowly expressed genes (1 > FPKM ≥ 0.5)	1370	222	388

GO annotation and KEGG pathway analysis

GO annotation results showed that apoptosis-related GO terms were identified in the LL vs DP comparison (Fig.3 A-C), in which eight genes were related to intrinsic apoptotic signaling pathways in response to endoplasmic reticulum stress, 20 genes were related to intrinsic apoptotic signaling pathways, and 24 genes were related to regulation of the apoptotic signaling pathway. Among the biosynthesis-related GO terms, nine genes were associated with the monosaccharide biosynthetic process, 12 genes were associated with the cellular amino acid biosynthetic process, and 17 genes were related to fatty acid biosynthetic process. DEGs were also annotated to GO terms associated with growth and development; 14 genes were associated with mammary gland development, and 29 genes were associated with cell growth. In addition, GO terms related to the metabolism of substances (fatty acid metabolic process, regulation of lipid metabolic process, and cellular amino acid metabolic process) and substance transport (protein transmembrane transport, amino acid transmembrane transport, fatty acid

transport, glucose transmembrane transport, and lipid transport) were also significantly enriched. In the DP vs LG comparison, more specific GO terms related to mammary gland growth and development were enriched, including branching involved in mammary gland duct morphogenesis, mammary gland epithelial cell proliferation, mammary gland duct morphogenesis, and mammary gland epithelium development. In the LL vs LG comparison, mammary gland immunity-related GO terms including neutrophil activation involved in immune response, and neutrophil-mediated immunity were also significantly enriched. Please refer to Additional file 2 for specific information of GO analysis. KEGG analysis identified important pathways (Fig. 4 A-C), such as apoptosis, biosynthesis of amino acids, and fatty acid biosynthesis. In addition, the peroxisome proliferator-activated receptor (PPAR) signaling pathway, AMP-activated protein kinase (AMPK) signaling pathway, extracellular matrix (ECM)-receptor interaction and other pathways that are potentially related to mammary gland development were also identified. Please refer to Additional file 3 for specific information of KEGG pathway analysis.

Functional network analysis of DEGs

Combining the GO annotation and KEGG analysis results, we identified 40 DEGs and 8 GO terms to construct a development/growth network (Fig. 5 A). Among them, the number of genes related to mammary gland development was the largest, with 25 genes. The expression patterns of the 40 DEGs were divided into three categories: in cluster 1, the expression levels of *PI16*, insulin-like growth factor binding protein 4 (*IGFBP4*), and low density lipoprotein receptor-related protein 1 (*LRP1*) in the LL stage were significantly lower than those in the DP and LG stages. Fibroblast growth factor 7 (*FGF7*) and *FGF2* are members of the fibroblast growth factor family and have extensive mitogenic and cell survival functions [19]. Both *IGFBP4* and *IGFBP3* are members of the insulin-like growth factor binding protein family and play regulatory roles in cell growth by binding to IGF [20]. The highly expressed *IGFBP4* and *IGFBP3* in LG may be closely associated with the proliferation and differentiation of mammary gland cells. In cluster 2, cluster of differentiation 109 (*CD109*), regulator of G protein signaling 4 (*RGS4*), *IGFBP5*, B cell lymphoma 2 like 11 (*BCL2L11*), amphiregulin (*AREG*), and sulfatase 2 (*SULF2*) were highly expressed in the LL and DP stages but were expressed at lower levels in the LG stage. In our previous study, *BCL2L11* was shown to have potential function in promoting the apoptosis of mammary gland epithelial cells [17]. In cluster 3, Wnt family member 5A (*WNT5A*), reticulon 4 (*RTN4*), cysteine dioxygenase type 1 (*CDO1*), and glycerol-3-phosphate acyltransferase 4 (*GPAT4*) were highly expressed in the LL and LG stages and were expressed at low levels in the DP stage. *CSN2* and *CSN3* are important genes encoding casein [21], and the low expression level of these two genes during the DP stage may be closely related to no milk secretion in the DP stage. Similarly, we identified 43 genes and five apoptosis-related pathways to construct an apoptosis network (Fig. 5 B), in which regulation of the apoptotic signaling pathway had the largest number of genes (34 genes). For intrinsic apoptotic signaling pathways, there were 31 genes. For extrinsic apoptotic signaling pathways, there were 13 genes. The expression patterns of the 43 apoptotic genes were also divided into three categories. In cluster 3, the gene expression levels of B cell receptor associated protein 31 (*BCAP31*), *CD24*, tumor protein D52 like 1 (*TPD52L1*), nuclear protein 1 (*NUPR1*), and caspase 8 (*CASP8*) were higher in the LL stage than in the DP and LG stages. *CASP8*, *CASP9*, and *CASP3* encode members of the cysteine-aspartic protease family and

play central roles in the execution of cell apoptosis [22], and their expression patterns may be related to the apoptosis of mammary gland cells during the LL stage. In addition, we also obtained 8 GO terms and 39 genes associated with mammary gland immunity in dairy goats and constructed an immunity network (Fig. 5 C). The DP stage is not only a period for improving the nutrition status of goats and enhancing the constitution of the body but is also a period of susceptibility to mammary gland disease [23]. In cluster 2, we found that complement component 5a receptor 1 (*C5AR1*), S100 calcium-binding protein A9 (*S100A9*), complement component 7 (*C7*), complement component 1s (*C1S*), serpin family G member 1 (*SERPING1*), complement factor B (*CFB*), *C2* and other genes were specifically highly expressed in the DP stage. Among them, complement C1q subcomponent subunit A (*C1QA*), *C1QB*, *C1QC*, *C1S*, *C2*, and *C7* all encode members related to the complement components of the immune system, serve as an important part of the innate immune system [24], and may participate in mammary gland immunity and defense processes in the DP stage. In addition, substance transport, synthesis and metabolism are the most basic physiological processes in the mammary gland and are crucial for mammary gland development and lactation. The substance transport network consists of eight GO terms and 30 genes (Fig. 6 A). In cluster 2, solute carrier family 6 member 14 (*SLC6A14*), *SLC1A5*, *SLC2A10*, *SLC2A9*, and *SLC27A3*, which encode solute carrier family members, had higher expression levels in the LL and LG stages than in the DP stage. The relationship between 41 genes and 7 GO terms was shown in a biosynthetic network (Fig. 6 B). Among them, 12 genes, including glutamate-ammonia ligase (*GLUL*), ornithine aminotransferase (*OAT*), cystathionine gamma-lyase (*CTH*), branched-chain amino acid transaminase 2 (*BCAT2*), cysteine dioxygenase type 1 (*CDO1*), phosphoserine aminotransferase 1 (*PSAT1*), and glutamic-pyruvic transaminase 2 (*GPT2*), are involved in the synthesis of amino acids, 12 genes, including stearoyl-CoA desaturase (*SCD*), fatty acid synthase (*FASN*), lipoprotein lipase (*LPL*), acetyl-CoA synthetase 2 (*ACSS2*), *CD74*, *ACSS1*, and pyruvate dehydrogenase kinase 4 (*PDK4*), are associated with the fatty acids biosynthesis, and nine genes, including transketolase (*TKT*), glucokinase (*GCK*), activating transcription factor 4 (*ATF4*), phosphogluconate dehydrogenase (*PGD*), *SLC25A1*, fructose-bisphosphatase 1 (*FBP1*), L-serine dehydratase (*SDS*), phosphoenolpyruvate carboxykinase 2 (*PCK2*) and sestrin 2 (*SESN2*), are involved in the biosynthetic synthesis of monosaccharides. These genes are involved in the synthesis of basic substances and exhibited significant differential expression characteristics in the three stages. Finally, we also constructed a metabolism network of 40 genes and 9 GO terms (Fig. 6 C). The number of genes associated with cellular amino acid metabolic processes was the largest, with 21 genes. In cluster 1, proteasome 20S subunit beta 8 (*PSMB8*), proteasome activator complex subunit 2 (*PSME2*), *PSMB10*, and *PSME1*, which are related to amino acid metabolism, had the same expression pattern, and their expression levels were higher in the DP and LG stages than in the LL stage. Genes associated with fatty acid metabolism, including fatty-acid-binding protein 3 (*FABP3*), *SCD*, glycerol-3-phosphate acyltransferase 4 (*GPAT4*), fatty acid desaturase 1 (*FADS1*), *SLC27A3*, lysophosphatidylglycerol acyltransferase 1 (*LPGAT1*), *SLC27A5*, *CD36*, *SLC27A4*, acetyl-CoA synthetase 2 (*ACSS2*), *SLC27A1*, and *FASN*, were in cluster 2, and had higher expression levels in the LG and LL stages than that in the DP stage.

Transcription factor analysis

A total of 35 TFs were identified from a total of 1,381 DEGs (Fig. 7 A-D), of which 15 TFs were upregulated in the LL vs DP comparison, 22 TFs were identified in the DP vs LG comparison (nine were upregulated, and 13 were downregulated), and 13 TFs were identified in the LL vs LG comparison (nine were upregulated, and four were downregulated). They were divided into 16 families, of which family information for 4 TFs was not clear. The bZIP family had the largest number of TFs, including 5 TFs, i.e., *ATF4*, CCAAT enhancer binding protein gamma (*CEBPG*), FOS like 2 (*FOSL2*), JunB proto-oncogene, AP-1 transcription factor subunit (*JUNB*), nuclear factor, interleukin 3 regulated (*NFIL3*). The expression patterns of *CEBPG* and *ATF4* were similar according to the expression heat maps, of which the expression levels were higher in the LL stage than in the DP and LG stages, while the expression patterns of *NFIL3*, *JUNB* and *FOSL2* were similar, of which the expression levels were higher in the DP stage than in the LL and LG stages. The bHLH family contains 4 TFs, i.e., aryl hydrocarbon receptor nuclear translocator (*ARNT*), basic helix-loop-helix family member A15 (*BHLHA15*), *BHLHE40*, and microphthalmia-associated transcription factor (*MITF*). The expression patterns of *MITF*, *BHLHE40*, and *ARNT* were similar, of which the expression levels were all higher during the DP stage. In addition, estrogen receptor 1 (*ESR1*), progesterone receptor (*PGR*), and estrogen-related receptor alpha (*ESRRA*) all belong to the *ESR*-like family. They also had similar expression patterns, with high expression in the LG stage and relatively low expression in the LL stage. Similar expression patterns of most TFs in the same family also appear in the homeobox and *zf-C2H2* families. Please refer to Additional file 4 for specific information of transcription factors analysis.

Real-Time Quantitative Reverse Transcription PCR (qRT-PCR)

From the histogram of transcriptome sequencing results (Fig. 7 F), the 15 randomly selected DEGs include genes with relatively high expression levels, such as receptor for activated C kinase 1 (*RACK1*), *GPAT4*, and *CSN2*, as well as genes with relatively low expression levels, such as *PCK2*, *MUC1*, and *SESN2*. qRT-PCR detection showed that the expression patterns of the 15 genes obtained from qRT-PCR were consistent with those obtained from the transcriptome sequencing results (Fig. 7 E-G). Correlation analysis was performed on the RNA-Seq of the 15 genes during three developmental stages and the log2-fold change obtained from qRT-PCR, and the correlation coefficient (cor) = 0.91 (p-value < 2.2e-16), indicating a high correlation of RNA-Seq with qRT-PCR detection, which further verified the accuracy of the sequencing results. All primer sequences are provided in the Additional file 5.

Discussion

In this study, RNA-Seq and bioinformatics analysis were performed on the mammary gland tissues of dairy goats in the LL, DP, and LG stages, and the gene expression profiles related to mammary gland development and lactation were mapped to construct mammary gland growth, proliferation, apoptosis, immunity, substance transport, synthesis, and metabolic-related networks. Prior to this study, studies have reported the transcriptome analysis of mammary gland tissues during the lactating and DP stages in pigs [18], sheep [25], and cattle [26]. The differences between this study and the above studies are that we selected three consecutive periods of mammary gland development. From the perspective of

lactation, this is a process from the gradual termination of milk secretion to the absence of milk secretion and then to the restart of lactation. From the perspective of tissue structure, this is a process from mammary gland tissue involution to tissue remodeling, cell proliferation, and differentiation to form new mammary gland duct branches and acinar structure. Data analysis showed that a total of 1,381 DEGs were identified in the three mammary gland developmental stages, indicating that gene expression was indeed different among the three stages (Fig. 2). Perhaps the differential expression of these genes leads to differences in the morphology, structure and function of the mammary gland during the three stages. We identified the 20 DEGs with the highest expression in the mammary gland during the three stages, and these DEGs included *CSN1S1* and *CSN1S2*, which encode alpha-casein, *CSN2*, which encodes beta-casein, *CSN3*, which encodes kappa-casein, and *LALBA*, which encodes alpha-lactalbumin. The expression levels of these genes were lower in the DP stage than in the LL and LG stages. The expression patterns of these genes found in this study were consistent with those of previous studies. [26, 27], and their high expression levels may be closely related to milk synthesis during the LL and LG stages. *FGF2* and *FGF11*, as members of the fibroblast growth factor family, exhibited similar expression patterns, and their expression levels were lower in the LL stage than in the DP and LG stages. It has been reported that *FGF2* promotes the proliferation of mammary epithelial cells, which is conducive to mammary gland remodeling and maintaining milk secretion [28]. *RPS3A*, *RPS2*, *RPL19*, *RPLP0*, *RPL13A*, *RPL23*, and *RPS8* encode ribosomal proteins. Similarly, they also exhibited similar expression patterns, namely, expression levels during the LL stage were lower than those in the DP and LG stages. Therefore, we hypothesize that genes in the same family or genes with similar functions often have similar expression pattern in the mammary gland during different developmental stages. Other studies reported that in some cases, some gene family members had the same (or almost the same) sequence. These families allow large amounts of genes to be expressed in a short time as needed [29]. However, we used a limited number of goat mammary gland samples, and studies need to be conducted to verify this hypothesis.

The physiological structure of the goat mammary gland changes significantly during the three developmental stages. These changes are closely related to the specific expression of genes during different developmental stages [30]. According to the development/growth network (Fig. 5A), *IGFBP5*, a member of the insulin-like growth factor family, was associated with mammary gland development and morphogenesis and was highly expressed in the LL stage. It has been reported that *IGFBP5* can accelerate the apoptosis of mammary epithelial cells and that the overexpression of *IGFBP5* in vivo can lead to impaired mammary gland development [31]. Different from *IGFBP5*, *IGFBP4* and *IGFBP3* were expressed at low levels in the LL stage and at higher levels in the DP and LG stages. Our findings are consistent with those by Alla et al. in a study of the mammary gland tissues of adolescent, gestational, lactating and degenerative mice [32]. This again confirmed that insulin-like growth factor binding protein family members may have a potential regulatory relationship with mammary gland development. In addition, *WNT5A* and *RTN4* genes involved in mammary epithelial cell proliferation had similar expression patterns, and all were expressed at low levels in the DP stage and high levels in the LG stage. It has been reported that *WNT5A* has an inhibitory effect on mammary gland development [33] and that *RTN4* can inhibit hepatocellular carcinoma cell (HCC) proliferation and promote HCC apoptosis [34],

which may counteract cell proliferation in LG stage. However, mitogen-activated protein kinase 1 (*MAPK1*), *ESR1*, cyclin D1 (*CCND1*), and mediator complex subunit 1 (*MED1*) were expressed at low levels in the LL stage and at high levels in the DP and LG stages. All these genes could promote the proliferation of mammary gland cells [35–38]. Therefore, *RTN4* and *WNT5A* expression in the LG stage may be related to the regulation of mammary gland cell proliferation by other genes and to the differences in roles of these two genes in different tissues and cell types [35]. Transforming growth factor beta 3 (*TGFB3*) and transforming growth factor beta receptor 2 (*TGFBR2*) are both associated with the TGF- β signaling pathway, have become important molecular markers of breast cancer, and play important roles in the TGF- β pathway for mammary gland morphogenesis and lactation [39]. In addition, the elevated levels of prolactin and progesterone during pregnancy are conducive to mammary gland tissue remodeling and the formation of mammary gland lobe and duct branch structures [40]. However, we found that *PGR* and *PRLR*, as receptors for progesterone and prolactin, respectively, were highly expressed in mammary gland tissues during the LG stage, further demonstrating their promoting effects on mammary gland development, differentiation and lactation during the LG stage.

Apoptosis is essential to maintain the stability of the internal environment of the mammary gland, cell renewal and normal cell functions and activities [41]. According to the apoptosis network (Fig. 5B), members of the caspase family, *CASP8*, *CASP9*, and *CASP3*, are associated with extrinsic apoptotic signaling pathways, and their expression levels during the LL stage were much higher than those during the DP and LG stages. This is consistent with previous reports; that is, when the mammary gland degenerates, caspases are activated, and the expression levels of caspases are upregulated [42]. *CASP8* and FADD like apoptosis regulator (*CFLAR*), galectin 3 (*LGALS3*), interferon alpha-inducible protein 6 (*IFI6*), lamin A/C (*LMNA*), janus kinase 2 (*JAK2*), TNF receptor superfamily member 1B (*TNFRSF1B*), secreted frizzled-related protein 1 (*SFRP1*), and protein tyrosine phosphatase receptor type C (*PTPRC*) are also associated with extrinsic apoptotic signaling pathways, and they are highly expressed in the DP stage. Among them, *IFI6* encodes anti-apoptotic proteins that reside on the outer mitochondrial membrane, and interrupting the mitochondrial function of *IFI6* may prevent breast cancer metastasis [43]. The constitutive activation of *JAK2* in mammary epithelial cells can enhance signal transducer and activator of transcription 5 (Stat5) signal transduction, promote the generation of mammary acinus, inhibit apoptosis, and promote the develop of breast tumors [44]. *SFRP1*-encoded protein can block the Wnt signaling pathway, and low expression of *SFRP1* can promote mammary epithelial cell proliferation and the branching of mammary ducts [45]. Based on existing reports and high expression levels in the DP stage, these three genes might be involved in the regulation of cell apoptosis in mammary gland tissue during remodeling in the DP stage. In addition, *CFLAR*-encoded proteins are regulators of cell apoptosis and are structurally similar to caspase-8. *LGALS3* is a member of the galectin family, which encodes carbohydrate binding proteins, and members of this family have affinity for β -galactoside [46]. *LMNA*, *TNFRSF1B* and *PTPRC* are all associated with cell differentiation [47–49]. However, their role in the mammary gland is still unclear; therefore, further verification of their specific function in mammary gland development is needed. *RACK1* is not only associated with the intrinsic apoptotic signaling pathway, but its expression level is in the top 20 during the DP and LG stages. The expression level of *RACK1* in the LL

stage was the lowest, while it was the highest in the LG stage. The scaffold proteins encoded by *RACK1* can participate in the recruitment and assembly of a variety of signaling molecules, including protein kinase C [50], phosphodiesterases [51], Src tyrosine kinase [52], and insulin and insulin-like growth factors [53]. Therefore, *RACK1* may participate in the apoptosis of mammary gland cells by regulating a variety of signaling molecules.

Immunomodulation is essential for mammary gland health and is the basis and guarantee for its normal function. On the one hand, the milk secreted by the mammary gland provides nutrition and immune protection for newborn pups; on the other hand, the mammary gland involution and DP stages are not only periods for mammary gland repair and renewal but are also critical periods for defending against various mammary diseases [54]. According to the immunity network (Fig. 5C), we found that genes associated with the regulation of complement activation, such as *C5AR1*, *C7*, *C1S*, *SERPING1*, *CFB*, *C2*, *C1R*, *C1QB*, *C1QA*, *C1QC*, and *CFI*, had similar expression patterns; they were all highly expressed in the DP stage. These results suggested that these genes might play a potential regulatory role in mammary gland immunity during the DP stage. *C2*, *C1R*, *C1S*, *C1QA*, *C1QB*, and *C1QC* are all associated with the classical immune pathway [55], have antibacterial functions, regulate immune adhesion, mediate inflammation, and play important roles in inherent immunity against microorganisms [56]. In addition, neutrophil-mediated immunity is also associated with innate immunity. lipocalin 2 (*LCN2*) was highly expressed in the LL stage. It is known that protein encoded by *LCN2* is neutrophilic gelatinase-associated lipocalin, which inhibits the growth of bacteria through the chelation of iron-containing carriers, thereby playing a role in innate immunity [57]. In addition, *C5AR1*, *S100A9*, bone marrow stromal cell antigen 1 (*BST1*), fibrinogen like 2 (*FGL2*), *PTPRC*, cathepsin H (*CTSH*), and tubulin beta class I (*TUBB*) were highly expressed in the DP stage. *C5AR1* encodes a receptor for the activated fragment C5a generated after the activation of complement and can be used as a chemokine to induce neutrophil phagocytosis to pathogens [58]. *S100A9* encodes calcium and zinc binding proteins and plays an important role in the regulation of inflammatory processes and the immune response. It can induce neutrophil chemotaxis and adhesion and can enhance the anti-bactericidal activity of neutrophils through promoting phagocytosis by the activation of spleen associated tyrosine kinase (*SYK*), *PI3K/AKT* and extracellular signal-regulated kinase 1/2 (*ERK1/2*) [59]. *BST1*, *FGL2*, *PTPRC*, *CTSH*, and *TUBB* can also induce neutrophils to participate in the fight against pathogens, thus playing a potential protective role in mammary gland health during the DP stage. All three genes, peptidoglycan recognition protein 1 (*PGLYRP1*), beta-2-microglobulin (*B2M*), and *LTF*, were highly expressed in the LG stage. *LTF* (lactoferrin) is a transferrin protein that transfers iron to cells and controls the levels of free iron in blood and external secretions. It is widely present in the milk of humans and other mammals and has antibacterial activity. It is an important part of the nonspecific immune system [60]. In addition, humoral immune responses, B cell-mediated immunity, and T cell-mediated immunity also play important roles in the mammary gland. It is known that the protein encoded by *JCHAIN* can connect two monomeric units of IgM or IgA and induces a larger IgA polymer. This protein also helps to bind these immunoglobulins with secretory components to resist pathogens [61]. Both *CD74* and *CD81* are associated with B cell-mediated immunity, and *CD74* is highly expressed in the DP stage, while *CD81* is highly expressed in the LG stage. In summary, mammary gland immunity is associated

with both the innate immune system and the specific immune system. The collaboration of these genes in the immune system during the three stages allows mammary gland immunity to proceed smoothly, thereby protecting mammary gland health.

The synthesis of proteins, lipids and lactose in milk begins with the transport of basic substances, such as amino acids, fatty acids and glucose, in plasma [62]. It can be seen from the transport network (Fig. 6A) that the members of the solute carrier family that were differentially expressed during the three stages were involved in fatty acid (*SLC27A1*, *SLC27A3*, and *SLC27A4*) amino acid (*SLC1A5*, *SLC6A8*, *SLC6A14*, and *SLC25A29*), and glucose (*SLC2A4*, *SLC2A9*, *SLC2A10*, and *SLC2A8*) transport. The members of these solute carrier families were highly expressed in the LL and LG stages, while low expression was observed in the DP stage. Compared with the gestation stage, the mRNA expression levels of *SLC27A3* and *SLC27A4* were upregulated 37-fold and 1.4-fold, respectively, during the lactation period [63]. In addition, *SLC1A5*, *SLC2A4*, and *SLC2A8* were expressed at higher levels during the lactation stage than during the DP stage [26]. This indicates that the members of the solute carrier family are very important for substance transport to the mammary gland during the lactation stage, and their high expression during lactation is closely related to the higher lactation yield requirement. In addition, *FABP3*, *PLIN2*, *MFSD2A*, *CD36* and other genes are also involved in fatty acid transport in the mammary gland and are highly expressed during the lactation period. By using proteomics, Zheng et al. found that the expression level of *PLIN2* in the LL stage was lower than that in the peak period, which had a potential impact on milk fat synthesis in the mammary gland [64]. *CD36* and *FABP3* were significantly upregulated in the lactation period [65], and *FABP3* is one of the most abundant isoforms in bovine mammary gland tissues; *FABP3* expression is affected by the lactation period [66]. Proteins encoded by *MFSD2A* can mediate lipid transport, participate in the establishment of the blood-brain barrier and are required for brain growth and function [67]. In summary, these four genes might cooperate to participate in fatty acid uptake and transport in the mammary gland. In addition, we also found that ADP ribosylation factor like GTPase 6 interacting protein 1 (*ARL6IP1*) may be involved in amino acid transport; however, its expression pattern is different from that of solute carrier family members. *ARL6IP1* can regulate amino acid transport and negatively regulate cell apoptosis by regulating the activity of caspase-9 (*CASP9*) to protect cell survival [68]. Therefore, *ARL6IP1* is highly expressed in the LG stage and has low expression levels in the LL and DP stages. It may be involved in the transportation of amino acids in the LG stage; on the other hand, it may promote the proliferation and differentiation of mammary gland cells in the LG stage. The above data related to substance transport in the mammary gland indicate that during lactation, at different periods, the mammary gland has different demands for glucose, amino acids and fatty acids, and that the RNA expression of the related three types of transport proteins will also undergo proper adjustment and change, adapting to the lactation requirement of the mammary gland. The entire transport process is complex, which may require the coordination of the proteins encoded by multiple genes to achieve organized substance transport in the mammary gland.

The biosynthesis and metabolism of substances are the basic physiological processes of mammary gland development and lactation [69]. In addition to the genes encoding the main proteins in milk, such as casein (*CSN1S1*, *CSN1S2*, *CSN2*, and *CSN3*) and lactalbumin (*LALBA*), according to the biosynthesis

network (Fig. 6B), a total of 12 genes (*GLUL*, *OAT*, *CTH*, *BCAT2*, *CDO1*, *PSAT1*, *GPT2*, *PYCR1*, glutaminase (*GLS*), phosphoglycerate dehydrogenase (*PHGDH*), phosphoserine phosphatase (*PSPH*), and asparagine synthetase (*ASNS*)) associated with cellular amino acid biosynthetic processes were identified. *CDO1*, *PYCR1*, *PSAT1*, *BCAT2*, *ASNS* and *PSPH* all had similar expression patterns and were highly expressed in the LL stage. *CDO1* is known to be a key regulator of cysteine concentrations in cells and is identified as a candidate gene affecting milk protein traits [70]. *PYCR1* encodes an enzyme that catalyzes the conversion of NAD(P)H-dependent pyrroline 5-carboxylate to proline [71]. In addition, proteins encoded by *PSAT1* and *PSPH* are involved in serine synthesis [72]. *BCAT2* encodes a branched-chain aminotransferase, and the encoded protein forms a dimer. This dimer catalyzes the first step in the production of branched-chain amino acids, such as leucine, isoleucine and valine. The abnormal expression of *BCAT2* also affects cell proliferation and differentiation [73]. *ASNS*-encoded proteins are involved in the synthesis of asparagine [74]. The expression patterns of *PYCR1*, *PSAT1*, and *ASNS* were consistent with previous reports that the expression levels of these three genes during the lactation period are higher than those during the DP stage [26]. *OAT*, *GPT2*, *PHGDH*, and *CTH* are also involved in amino acid biosynthesis, and their expression levels were higher in the LG stage. These genes play important regulatory roles in amino acid biosynthesis in the mammary gland during the lactation period and provide enough amino acids in the mammary gland to achieve protein synthesis. The mammary gland is not only a place for protein synthesis but also an important place for lipid synthesis. It is known that the mammary gland of a lactating mouse synthesizes and secretes milk fat equivalent to its entire body weight in a 20-day lactation cycle, making the mammary gland one of the most active known lipid synthesis organs [62]. According to the biosynthesis network, we identified 12 genes associated with fatty acid synthesis. These include genes associated with fatty acid uptake in blood (*LPL*), the synthesis of long-chain fatty acids (*TECR* and *LPGAT1*) and short-chain fatty acids (*ACSS1* and *ACPS2*), de novo fatty acid biosynthesis (*ACACA* and *FASN*), fatty acid desaturation (*FADS1* and *SCD*), transcriptional regulation (*INSIG1* and *CD74*), ketone body utilization (*PDK4*), and redox (*GPX4*). It has been reported that *LPL*, *ACSS1*, *ACSS2*, *FASN*, *FADS1*, and *INSIG1* have similar expression patterns, all of which are highly expressed in the peak lactation stage, and the expression pattern of *LPL* in the bovine mammary gland is very similar to the lactation curve pattern, which may indicate that this gene plays an important role in maintaining milk synthesis [75]. *SCD* is a core enzyme in lipid metabolism and is the rate-limiting enzyme in the catalyzed synthesis of monounsaturated fatty acids from saturated fatty acid [76]. These lipid metabolism-related genes show different expression patterns during the three developmental stages, which are mutually adaptive to the fat synthesis function in different developmental stages. In addition, genes involved in carbohydrate synthesis and metabolism were also significantly enriched. *GCK*, *ATF4*, *PGM1*, and *KAT2B* are related to gluconeogenesis and glucose metabolic processes. *GCK* encodes a hexokinase family member, and hexokinase phosphorylates glucose to produce 6-phosphate glucose, which is the first step in most glucose metabolic pathways (Fig. 6C). The enzyme encoded by *PGM1* is involved in the decomposition and synthesis of glucose. *GCK* and *PGM1* are known to have higher expression levels during lactation than during the DP stage [77, 78]. The expression profiles of *ATF4* and *KAT2B* in the mammary gland during different lactation stages were mapped for the first time. In summary, genes related to the biosynthesis and metabolism of basic substances such as glucose, amino

acids and fatty acids in the mammary gland were screened using biosynthesis and metabolism networks, and their expression profiles in the mammary gland during three development stages were revealed. This study provides a reference for studying their roles in mammary gland development and lactation function in the future.

In the mammary gland, TFs interact with specific gene sequences (open or closed), and coordinated gene expression occurs at the correct time and in appropriate numbers. [79]. During mammary gland development and the lactation cycle, many site-specific TFs are essential for the proper growth, branching and involution of the mammary gland [80]. It is known that the overexpression of *ATF4* inhibits the proliferation and differentiation of mammary epithelial cells, leading to lactation dysfunction and accelerated mammary gland involution [81]. Our results showed that *ATF4* was highly expressed during the LL stage, possibly to promote the involution of mammary gland tissue and apoptosis during the LL stage (Fig. 7A-D). In addition, the expression of *JUNB* increases after forced weaning [82]. We also detected high *JUNB* expression in the DP stage, and we speculate that *JUNB* has effects on milk synthesis and secretion. *CEBPG*, *FOSL2*, and *NFIL3* belong to the bZIP family, as do *ATF4* and *JUNB*. bZIP TFs are present in all eukaryotes and are one of the largest families of dimeric TFs. It has been reported that this family of TFs regulates cell growth [83]. We found that this family of TFs had the greatest number of TFs with differential expression in the mammary gland; therefore, these TFs might be involved in mammary gland development and lactation regulation. In addition, four members of the bHLH family (*ARNT*, *BHLHA15*, *BHLHE40*, and *MITF*) were differentially expressed in the three developmental stages. *ARNT* is known to be necessary for mammary gland development in adolescent females [84]. *BHLHE40* is associated with breast cancer [85]. However, the function of *BHLHA15* and *MITF* in the mammary gland has not yet been reported; therefore, further study on their specific role in mammary gland development is warranted. In summary, these 35 TFs are differentially expressed in mammary gland tissues during the three developmental stages and may play potential regulatory roles in gene expression in the mammary gland, thereby affecting mammary gland development and the physiological process of lactation.

Conclusions

We found that the mRNA expression levels of genes associated with mammary gland growth, differentiation, apoptosis, immunity, substance transport, substance synthesis and metabolism were significantly different among the three mammary gland development stages (LL, DP, and LG). The genes in the same family or genes and TFs with similar functions have potential similar expression patterns in mammary gland tissue at different developmental stages. Mammary gland development and the physiological process of lactation involve the synergy of multiple genes involved in regulation. We mapped the expression profiles of DEGs in the mammary gland during the three developmental stages, providing a theoretical reference for further understanding the molecular mechanisms of mammary gland development and lactation.

Methods

Ethics statement

The methods and experimental procedures used in this study were all approved and directed by the animal protection and ethics committee of Shandong Agricultural University (protocol number: SDAUA-2018-048). All participating personnel were required to undergo rigorous training and self-protection and to alleviate the suffering of experimental animals to the greatest extent possible.

Experimental animal treatment and sample collection

In this study, nine healthy and disease-free Laoshan dairy goats were used, all of which came from Aote goat breeding farm (Qingdao, Shandong province, China), and all goats had free access to food and were managed under the same conditions. They were 4 years old, on average (third parity). Mammary gland tissues were collected at the LL (n = 3, 240 d postpartum), DP (n = 3, 300 d postpartum), and LG (n = 3, 140 d after mating) stages. The specific tissue collection methods were as follows. After intravenous injection of pentobarbital sodium (100 mg/kg) for anesthesia, the muscles relaxed, and the heart and respiration were arrested; then, the goats were quickly dissected, and mammary gland tissues were harvested. Tissues were quickly placed in RNase-free cryotubes and stored in liquid nitrogen.

RNA extraction and sequencing analysis

Total RNA was extracted using a Trizol Based RNAiso Plus kit (Code No 9108Q, Takara). RNA quality was analyzed using a NanoDrop 2000 C (Thermo). A sample was used for subsequent experimental analysis when the RNA integrity number (RIN) was > 8. Nine sequencing libraries (LL1, LL2, LL3, DP1, DP2, DP3, LG1, LG2, and LG3) were constructed using TruSeq RNA Library Prep Kit v2 (Illumina), and sequencing was completed using a HiSeq 2500 platform. After obtaining the raw data, sequencing quality was assessed using FastQC software (<http://darlinglab.org/tutorials/fastqc/>), and the sequencing adapters (adapter sequences, primers, poly-A tails and other types of unwanted sequence from high-throughput sequencing reads) [86] were processed through Trimmomatic software. HISAT2 software [87] was used to construct indices of the goat genome (ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/001/704/415/GCF_001704415.1_ARS1), and sequence alignment was performed. The number and percentage of sequence alignment were quantified using SAMtools software [88]. StringTie software was used to assemble the sequence alignments to transcripts [89]. The expression level of each gene was quantified using HTSeq [90]. Statistical analysis of expression levels and differential expression analysis. First, low-expression genes (FPKM < 0.5) were filtered based on fragments per kilobase of transcript per million mapped reads (FPKM). The screening criteria for each gene were as follows: when $FPKM \geq 500$, the gene was defined as highly expressed; when $500 > FPKM \geq 10$, the gene was defined as moderately expressed; when $10 > FPKM \geq 1$, the gene was defined as lowly expressed; and when $1 > FPKM \geq 0.5$, the gene was defined as ultralowly expressed. Prior to differential expression analysis, principal component analysis and intersample correlation analysis were performed on gene expression levels. Then, differential expression analysis was performed using DESeq2 software

[91]. When comparing the same gene between two groups, if the false discovery rate (FDR) was ≤ 0.05 and the log₂-fold change was ≥ 1 , the gene was considered differentially expressed. The DEGs between different groups were analyzed, and differential gene expression was displayed using volcano plots and heat maps. The top 20 DEGs for each developmental stage (LL, DP, and LG) were identified and are displayed using histograms.

Gene Ontology (GO) annotation and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis

The DEGs identified in each group were annotated using the GO database and the KEGG database. The GO terms and KEGG pathways that met $\text{FDR} \leq 0.05$ were selected for downstream analysis. To further clarify and refine the function of genes, we first classified GO terms into multiple categories, including apoptosis-related, biosynthesis-related, growth and development-related, metabolism-related, substance transport-related, and other important function-related terms. Similarly, KEGG pathways were also classified into apoptosis-related, biosynthesis-related, immunity-related, metabolism-related, and other function-related pathways. Finally, the GO terms and KEGG pathways related to mammary gland development and lactation function are presented as bar graphs and bubble plots, respectively.

Functional classification of DEGs and analysis of expression patterns

To further clarify the relationship between specific genes and different important GO terms, we constructed six gene-concept networks, including development/growth, apoptosis, immunity, transport, biosynthesis, and metabolism networks. Subsequently, the patterns of gene expression changes in the network are shown in heat maps. The FPKM values of the gene expression levels were normalized based on Z-score before the heat map was plotted, and then the genes were hierarchically clustered.

Screening and analysis of transcription factors

Transcription factors (TFs) are key regulators of gene expression [92]. To clarify the classification of TFs among the DEGs, we used the AnimalTFDB 3.0 website [93]. DEGs at three developmental stages were annotated. Subsequently, the number of TFs contained in different families was counted, and the expression patterns of these differentially expressed TFs were analyzed.

Verification of the sequencing accuracy using real-time quantitative reverse transcription polymerase chain reaction (qRT-PCR)

To further verify the accuracy of the transcriptome sequencing results, we randomly selected 15 DEGs and designed primers using NCBI Primer-BLAST [94]. Reverse transcription and PCR amplification were performed using a One Step TB Green® PrimeScript™ RT-PCR Kit (Takara). Quantitative PCR was performed using a LightCycler 96 fluorescence quantitative PCR instrument (Roche). The relative expression levels of genes during the three developmental stages were calculated using $2^{-\Delta\Delta\text{CT}}$ [95]. Tukey's honestly significant difference (HSD) test was used for the statistical analysis of gene expression

levels at different developmental stages. Unless otherwise noted, the above plots were all generated using R software (Version 3.6.0).

Declarations

Ethics approval and consent to participate

The methods and experimental procedures used in this study were all approved and directed by the animal protection and ethics committee of Shandong Agricultural University (protocol number: SDAUA-2018-048). All participating personnel were required to undergo rigorous training and self-protection and to alleviate the suffering of experimental animals to the greatest extent possible.

Consent for publication

Not applicable.

Availability of data and materials

All data generated or analysed during this study are included in this published article and its supplementary information files.

Competing interests

The authors declared that they had no competing interests.

Funding

This work was supported by the National Natural Science Foundation of China (31672401), Shandong Provincial Natural Science Foundation of China (ZR2014CM029), a project of the Shandong Province Higher Educational Science and Technology Program of China (J14LF07), the Shandong Provincial Modern Agriculture Industry Technology System (SDAIT-10), and funds of the Shandong “Double Tops” Program (SYL2017YSTD12). The funding body did not exert influence on the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Authors' contributions

WJM and JZB devised the project, the main conceptual ideas and proof outline. XR, CTL, ZXD, WAL, CYP, LQ, ZYL participated in the collection and sequencing of goat mammary gland samples. XR and JZB processed the experimental data, performed the analysis, drafted the manuscript and designed the figures. All authors discussed the results and commented on the manuscript.

Acknowledgements

We are grateful to Prof. Jianmin Wang and Zhibin Ji for designing this experiment. We thank all members of the Animal Protection Laboratory in Shandong Agricultural University for supporting the platform and technology in data analysis.

References

1. Tobon H, Salazar H. Ultrastructure of the human mammary gland. II. Postpartum lactogenesis. *J Clin Endocrinol Metab.* 1975;40:834–44.
2. Silver IA. The anatomy of the mammary gland of the dog and cat. *J Small Anim Pract.* 1966;7:689–96.
3. Sheffield LG. Organization and growth of mammary epithelia in the mammary gland fat pad. *J Dairy Sci.* 1988;71:2855–74.
4. Hens JR, Wysolmerski JJ. Key stages of mammary gland development: molecular mechanisms involved in the formation of the embryonic mammary gland. *Breast Cancer Res.* 2005;7:220–4.
5. Schroeder JA, Masri AA, Adriance MC, Tessier JC, Kotlarczyk KL, Thompson MC, Gendler SJ. MUC1 overexpression results in mammary gland tumorigenesis and prolonged alveolar differentiation. *Oncogene.* 2004;23:5739–47.
6. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet.* 2009;10:57–63.
7. Scheele CL, Hannezo E, Muraro MJ, Zomer A, Langedijk NS, van Oudenaarden A, Simons BD, van Rheenen J. Identity and dynamics of mammary stem cells during branching morphogenesis. *Nature.* 2017;542:313–7.
8. Kosciuczuk EM, Lisowski P, Jarczak J, Majewska A, Rzewuska M, Zwierzchowski L, Bagnicka E. Transcriptome profiling of Staphylococci-infected cow mammary gland parenchyma. *BMC Vet Res.* 2017;13:161.
9. Wagner J, Rapsomaniki MA, Chevrier S, Anzeneder T, Langwieder C, Dykgers A, Rees M, Ramaswamy A, Muenst S, Soysal SD, et al. A Single-Cell Atlas of the Tumor and Immune Ecosystem of Human Breast Cancer. *Cell.* 2019;177:1330–45.e1318.
10. Ji Z, Wang G, Zhang C, Xie Z, Liu Z, Wang J. Identification and Function Prediction of Novel MicroRNAs in Laoshan Dairy Goats. *Asian-Australas J Anim Sci.* 2013;26:309–15.
11. Ji Z, Wang G, Xie Z, Zhang C, Wang J. Identification and characterization of microRNA in the dairy goat (*Capra hircus*) mammary gland by Solexa deep-sequencing technology. *Mol Biol Rep.* 2012;39:9361–71.
12. Ji Z, Liu Z, Chao T, Hou L, Fan R, He R, Wang G, Wang J. Screening of miRNA profiles and construction of regulation networks in early and late lactation of dairy goat mammary glands. *Sci Rep.* 2017;7:11933.

13. Macias H, Hinck L. Mammary gland development. Wiley interdisciplinary reviews Developmental biology. 2012;1:533–57.
14. Orfanou DC, Pourlis A, Ververidis HN, Mavrogianni VS, Taitzoglou IA, Boscors CM, Fthenakis GC. Histological features in the mammary glands of female dogs throughout lactation. Anat Histol Embryol. 2010;39:473–8.
15. Holst BD, Hurley WL, Nelson DR. Involution of the bovine mammary gland: histological and ultrastructural changes. J Dairy Sci. 1987;70:935–44.
16. Safayi S, Theil PK, Hou L, Engbaek M, Nørgaard JV, Sejrsen K, Nielsen MO. Continuous lactation effects on mammary remodeling during late gestation and lactation in dairy goats. J Dairy Sci. 2010;93:203–17.
17. Xuan R, Chao T, Wang A, Zhang F, Sun P, Liu S, Guo M, Wang G, Ji Z, Wang J, et al. Characterization of microRNA profiles in the mammary gland tissue of dairy goats at the late lactation, dry period and late gestation stages. PLoS ONE. 2020;15:e0234427.
18. Zhao W, Shahzad K, Jiang M, Graugnard DE, Rodriguez-Zas SL, Luo J, Looor JJ, Hurley WL. Bioinformatics and Gene Network Analyses of the Swine Mammary Gland Transcriptome during Late Gestation. Bioinform Biol insights. 2013;7:193–216.
19. Ropiquet F, Giri D, Lamb DJ, Ittmann M. FGF7 and FGF2 are increased in benign prostatic hyperplasia and are associated with increased proliferation. J Urol. 1999;162:595–9.
20. Haisan J, Oba M, Ambrose DJ, Steele MA. Short communication: The effects of offering a high or low plane of milk preweaning on insulin-like growth factor and insulin-like growth factor binding proteins in dairy heifer calves. J Dairy Sci. 2018;101:11441–6.
21. Bonfatti V, Di Martino G, Cecchinato A, Degano L, Carnier P. Effects of beta-kappa-casein (CSN2-CSN3) haplotypes, beta-lactoglobulin (BLG) genotypes, and detailed protein composition on coagulation properties of individual milk of Simmental cows. J Dairy Sci. 2010;93:3809–17.
22. Zhou X, Zeng W, Li H, Chen H, Wei G, Yang X, Zhang T, Wang H. Rare mutations in apoptosis related genes APAF1, CASP9, and CASP3 contribute to human neural tube defects. Cell Death Dis. 2018;9:43.
23. Pezeshki A, Capuco AV, De Spiegeleer B, Peelman L, Stevens M, Collier RJ, Burvenich C. An integrated view on how the management of the dry period length of lactating cows could affect mammary biology and defence. J Anim Physiol Anim Nutr (Berl). 2010;94:e7–30.
24. Barnum SR. Complement: A primer for the coming therapeutic revolution. Pharmacol Ther. 2017;172:63–72.
25. Suárez-Vega A, Gutiérrez-Gil B, Klopp C, Tosser-Klopp G, Arranz JJ. Comprehensive RNA-Seq profiling to evaluate lactating sheep mammary gland transcriptome. Scientific data. 2016;3:160051.
26. Dai WT, Zou YX, White RR, Liu JX, Liu HY. Transcriptomic profiles of the bovine mammary gland during lactation and the dry period. Funct Integr Genomics. 2018;18:125–40.
27. Wang J, Zhou H, Hickford JGH, Hao Z, Shen J, Luo Y, Hu J, Liu X, Li S. Comparison of the Transcriptome of the Ovine Mammary Gland in Lactating and Non-lactating Small-Tailed Han Sheep.

Front Genet. 2020;11:472–2.

28. Jeong W, Bae H, Lim W, Bazer FW, Lee H, Song G. The functional effects and mechanisms by which fibroblast growth factor 2 (FGF2) controls bovine mammary epithelial cells: Implications for the development and functionality of the bovine mammary gland. *J Anim Sci.* 2017;95:5365–77.
29. Silver L, Hood L, Hartwell L, Goldberg M, Reynolds AE. *Genetics: From Genes to Genomes.* 4th ed. McGraw-Hill Education; 2010.
30. Clarkson RW, Wayland MT, Lee J, Freeman T, Watson CJ. Gene expression profiling of mammary gland development reveals putative roles for death receptors and immune mediators in post-lactational regression. *Breast Cancer Res.* 2004;6:R92–109.
31. Allan GJ, Beattie J, Flint DJ. The role of IGFBP-5 in mammary gland development and involution. *Domest Anim Endocrinol.* 2004;27:257–66.
32. Allar MA, Wood TL. Expression of the insulin-like growth factor binding proteins during postnatal development of the murine mammary gland. *Endocrinology.* 2004;145:2467–77.
33. Roarty K, Serra R. Wnt5a is required for proper mammary gland development and TGF-beta-mediated inhibition of ductal growth. *Development.* 2007;134:3929–39.
34. Chen YC, Lu DD, Cao XR, Zhang XR. RTN4-C gene expression in hepatocellular carcinoma and its influence on SMMC7721 cell growth and apoptosis. *Yi Chuan Xue Bao.* 2005;32:891–7.
35. Pinkas J, Leder P. MEK1 signaling mediates transformation and metastasis of EpH4 mammary epithelial cells independent of an epithelial to mesenchymal transition. *Cancer Res.* 2002;62:4781–90.
36. Hasegawa N, Sumitomo A, Fujita A, Aritome N, Mizuta S, Matsui K, Ishino R, Inoue K, Urahama N, Nose J, et al. Mediator subunits MED1 and MED24 cooperatively contribute to pubertal mammary gland development and growth of breast carcinoma cells. *Mol Cell Biol.* 2012;32:1483–95.
37. Stevens TA, Meech R. BARX2 and estrogen receptor-alpha (ESR1) coordinately regulate the production of alternatively spliced ESR1 isoforms and control breast cancer cell growth and invasion. *Oncogene.* 2006;25:5426–35.
38. Junk DJ, Cipriano R, Stampfer M, Jackson MW. Constitutive CCND1/CDK2 activity substitutes for p53 loss, or MYC or oncogenic RAS expression in the transformation of human mammary epithelial cells. *PLoS ONE.* 2013;8:e53776.
39. Moses H, Barcellos-Hoff MH. TGF-beta biology in mammary development and breast cancer. *Cold Spring Harb Perspect Biol.* 2011;3:a003277.
40. Oakes SR, Hilton HN, Ormandy CJ. The alveolar switch: coordinating the proliferative cues and cell fate decisions that drive the formation of lobuloalveoli from ductal epithelium. *Breast Cancer Res.* 2006;8:207.
41. Strange R, Li F, Saurer S, Burkhardt A, Friis RR. Apoptotic cell death and tissue remodelling during mouse mammary gland involution. *Development.* 1992;115:49–58.

42. Marti A, Ritter PM, Jäger R, Lazar H, Baltzer A, Schenkel J, Declercq W, Vandenabeele P, Jaggi R. Mouse mammary gland involution is associated with cytochrome c release and caspase activation. *Mech Dev.* 2001;104:89–98.
43. Cheriya V, Kaur J, Davenport A, Khaleel A, Chowdhury N, Gaddipati L. G1P3 (IFI6), a mitochondrial localised antiapoptotic protein, promotes metastatic potential of breast cancer cells through mtROS. *Br J Cancer.* 2018;119:52–64.
44. Caffarel MM, Zaragoza R, Pensa S, Li J, Green AR, Watson CJ. Constitutive activation of JAK2 in mammary epithelium elevates Stat5 signalling, promotes alveologenesis and resistance to cell death, and contributes to tumorigenesis. *Cell Death Differ.* 2012;19:511–22.
45. Gauger KJ, Shimono A, Crisi GM, Schneider SS. Loss of SFRP1 promotes ductal branching in the murine mammary gland. *BMC Dev Biol.* 2012;12:25.
46. Chaudoin TR, Bonasera SJ. Mice lacking galectin-3 (Lgals3) function have decreased home cage movement. *BMC Neurosci.* 2018;19:27.
47. Liu H, Li DM, Zhu LY, Lai LJ, Yan WY, Lu YH, Wei Y, Huang YQ, Fang M, Su YG, et al. Research on the knockout of LMNA gene by CRISPR/Cas9 system in human cell lines. *Yi chuan.* 2019;41:66–75.
48. Sun L, Lian JX, Meng S. MiR-125a-5p promotes osteoclastogenesis by targeting TNFRSF1B. *Cell Mol Biol Lett.* 2019;24:23.
49. Ferreira-Iglesias A, Montes A, Perez-Pampin E, Cañete JD, Raya E, Magro-Checa C, Vasilopoulos Y, Sarafidou T, Caliz R, Ferrer MA, et al. Replication of PTPRC as genetic biomarker of response to TNF inhibitors in patients with rheumatoid arthritis. *Pharmacogenomics J.* 2016;16:137–40.
50. Schechtman D, Mochly-Rosen D. Adaptor proteins in protein kinase C-mediated signal transduction. *Oncogene.* 2001;20:6339–47.
51. Yarwood SJ, Steele MR, Scotland G, Houslay MD, Bolger GB. The RACK1 signaling scaffold protein selectively interacts with the cAMP-specific phosphodiesterase PDE4D5 isoform. *J Biol Chem.* 1999;274:14909–17.
52. Fan Y, Si W, Ji W, Wang Z, Gao Z, Tian R, Song W, Zhang H, Niu R, Zhang F. Rack1 mediates tyrosine phosphorylation of Anxa2 by Src and promotes invasion and metastasis in drug-resistant breast cancer cells. *Breast Cancer Res.* 2019;21:66.
53. Zhang W, Zong CS, Hermanto U, Lopez-Bergami P, Ronai Z, Wang LH. Correction for, Zhang, et al., RACK1 Recruits STAT3 Specifically to Insulin and Insulin-Like Growth Factor 1 Receptors for Activation, Which Is Important for Regulating Anchorage-Independent Growth. *Mol Cell Biol.* 2017;37.
54. Sordillo LM. Mammary Gland Immunobiology and Resistance to Mastitis. *Vet Clin North Am Food Anim Pract.* 2018;34:507–23.
55. Nesargikar PN, Spiller B, Chavez R. The complement system: history, pathways, cascade and inhibitors. *European journal of microbiology immunology.* 2012;2:103–11.
56. Rainard P. The complement in milk and defense of the bovine mammary gland against infections. *Vet Res.* 2003;34:647–70.

57. Wilson BR, Bogdan AR, Miyazawa M, Hashimoto K, Tsuji Y. Siderophores in Iron Metabolism: From Mechanism to Therapy Potential. *Trends Mol Med*. 2016;22:1077–90.
58. Hernandez MX, Jiang S, Cole TA, Chu SH, Fonseca MI, Fang MJ, Hohsfield LA, Torres MD, Green KN, Wetsel RA, et al. Prevention of C5aR1 signaling delays microglial inflammatory polarization, favors clearance pathways and suppresses cognitive loss. *Mol Neurodegener*. 2017;12:66.
59. Shabani F, Farasat A, Mahdavi M, Gheibi N. Calprotectin. (S100A8/S100A9): a key protein between inflammation and cancer. *Inflamm Res*. 2018;67:801–12.
60. Czosnykowska-Lukacka M, Orczyk-Pawitowicz M, Broers B, Królak-Olejnik B. Lactoferrin in Human Milk of Prolonged Lactation. *Nutrients*. 2019;11.
61. Li Y, Wang G, Li N, Wang Y, Zhu Q, Chu H, Wu W, Tan Y, Yu F, Su XD, et al. Structural insights into immunoglobulin M. *Science*. 2020;367:1014–7.
62. Rudolph MC, McManaman JL, Phang T, Russell T, Kominsky DJ, Serkova NJ, Stein T, Anderson SM, Neville MC. Metabolic regulation in the lactating mammary gland: a lipid synthesizing machine. *Physiol Genomics*. 2007;28:323–36.
63. Han LQ, Li HJ, Wang YY, Zhu HS, Wang LF, Guo YJ, Lu WF, Wang YL, Yang GY. mRNA abundance and expression of SLC27A, ACC, SCD, FADS, LPIN, INSIG, and PPARGC1 gene isoforms in mouse mammary glands during the lactation cycle. *Genet Mol Res*. 2010;9:1250–7.
64. Zheng X, Ning C, Dong Y, Zhao P, Li J, Fan Z, Li J, Yu Y, Mrode R, Liu JF. Quantitative proteome analysis of bovine mammary gland reveals protein dynamic changes involved in peak and late lactation stages. *Biochem Biophys Res Commun*. 2017;494:292–7.
65. Lee JN, Wang Y, Xu YO, Li YC, Tian F, Jiang MF. Characterisation of gene expression related to milk fat synthesis in the mammary tissue of lactating yaks. *J Dairy Res*. 2017;84:283–8.
66. Bionaz M, Looor JJ. ACSL1, AGPAT6, FABP3, LPIN1, and SLC27A6 are the most abundant isoforms in bovine mammary tissue and their expression is affected by stage of lactation. *J Nutr*. 2008;138:1019–24.
67. Ben-Zvi A, Lacoste B, Kur E, Andreone BJ, Mayshar Y, Yan H, Gu C. Mfsd2a is critical for the formation and function of the blood-brain barrier. *Nature*. 2014;509:507–11.
68. Lui HM, Chen J, Wang L, Naumovski L. ARMER, apoptotic regulator in the membrane of the endoplasmic reticulum, a novel inhibitor of apoptosis. *Mol Cancer Res*. 2003;1:508–18.
69. Truchet S, Honvo-Houéto E. Physiology of milk secretion. *Best Pract Res Clin Endocrinol Metab*. 2017;31:367–84.
70. Li C, Cai W, Zhou C, Yin H, Zhang Z, Looor JJ, Sun D, Zhang Q, Liu J, Zhang S. RNA-Seq reveals 10 novel promising candidate genes affecting milk protein concentration in the Chinese Holstein population. *Sci Rep*. 2016;6:26813.
71. Reversade B, Escande-Beillard N, Dimopoulou A, Fischer B, Chng SC, Li Y, Shboul M, Tham PY, Kayserili H, Al-Gazali L, et al. Mutations in PYCR1 cause cutis laxa with progeroid features. *Nat Genet*. 2009;41:1016–21.

72. Brearley MC, Li C, Daniel Z, Loughna PT, Parr T, Brameld JM. Changes in expression of serine biosynthesis and integrated stress response genes during myogenic differentiation of C2C12 cells. *Biochemistry biophysics reports*. 2019;20:100694.
73. Dhanani ZN, Mann G, Adegoke OAJ. Depletion of branched-chain aminotransferase 2 (BCAT2) enzyme impairs myoblast survival and myotube formation. *Physiological reports*. 2019;7:e14299.
74. Wrona E, Jakubowska J, Pawlik B, Pastorczak A, Madzio J, Lejman M, Sędek Ł, Kowalczyk J, Szczepański T, Młynarski W. Gene expression of ASNS, LGMN and CTSB is elevated in a subgroup of childhood BCP-ALL with PAX5 deletion. *Oncology letters*. 2019;18:6926–32.
75. Bionaz M, Looor JJ. Gene networks driving bovine milk fat synthesis during the lactation cycle. *BMC Genom*. 2008;9:366.
76. Karlengen IJ, Harstad OM, Taugbøl O, Berget I, Aastveit AH, Våge DI. The effect of excess cobalt on milk fatty acid profiles and transcriptional regulation of SCD, FASN, DGAT1 and DGAT2 in the mammary gland of lactating dairy cows. *J Anim Physiol Anim Nutr (Berl)*. 2012;96:1065–73.
77. Lin J, Bao ZK, Zhang Q, Hu WW, Yu QH, Yang Q. Transcriptome analysis of the mammary gland from GH transgenic goats during involution. *Gene*. 2015;565:228–34.
78. Cantu JM, Ibarra B. Phosphoglucomutase: evidence for a new locus expressed in human milk. *Science*. 1982;216:639–40.
79. Mitchell PJ, Tjian R. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science*. 1989;245:371–8.
80. Murtagh J, Martin F, Gronostajski RM. The Nuclear Factor I (NFI) gene family in mammary gland development and function. *J Mammary Gland Biol Neoplasia*. 2003;8:241–54.
81. Bagheri-Yarmand R, Vadlamudi RK, Kumar R. Activating transcription factor 4 overexpression inhibits proliferation and differentiation of mammary epithelium resulting in impaired lactation and accelerated involution. *J Biol Chem*. 2003;278:17421–9.
82. Stein T, Salomonis N, Gusterson BA. Mammary gland involution as a multi-step process. *J Mammary Gland Biol Neoplasia*. 2007;12:25–35.
83. Szuplewski S, Kottler B, Terracol R. The Drosophila bZIP transcription factor VriI is involved in hair and cell growth. *Development*. 2003;130:3651–62.
84. Le Provost F, Riedlinger G, Hee Yim S, Benedict J, Gonzalez FJ, Flaws J, Hennighausen L. The aryl hydrocarbon receptor (AhR) and its nuclear translocator (Arnt) are dispensable for normal mammary gland development but are required for fertility. *Genesis*. 2002;32:231–9.
85. Sethuraman A, Brown M, Krutilina R, Wu ZH, Seagroves TN, Pfeiffer LM, Fan M. BHLHE40 confers a pro-survival and pro-metastatic phenotype to breast cancer cells by modulating HBEGF secretion. *Breast Cancer Res*. 2018;20:117.
86. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30:2114–20.

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

88. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25:2078–9.

89. Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat Protoc*. 2016;11:1650–67.

90. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;31:166–9.

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

92. Latchman DS. Transcription factors: an overview. *Int J Biochem Cell Biol*. 1997;29:1305–12.

93. Hu H, Miao YR, Jia LH, Yu QY, Zhang Q, Guo AY. AnimalTFDB 3.0: a comprehensive resource for annotation and prediction of animal transcription factors. *Nucleic Acids Res*. 2019;47:D33-d38.

94. Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*. 2012;13:134.

95. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*. 2001;25:402–8.

Figures

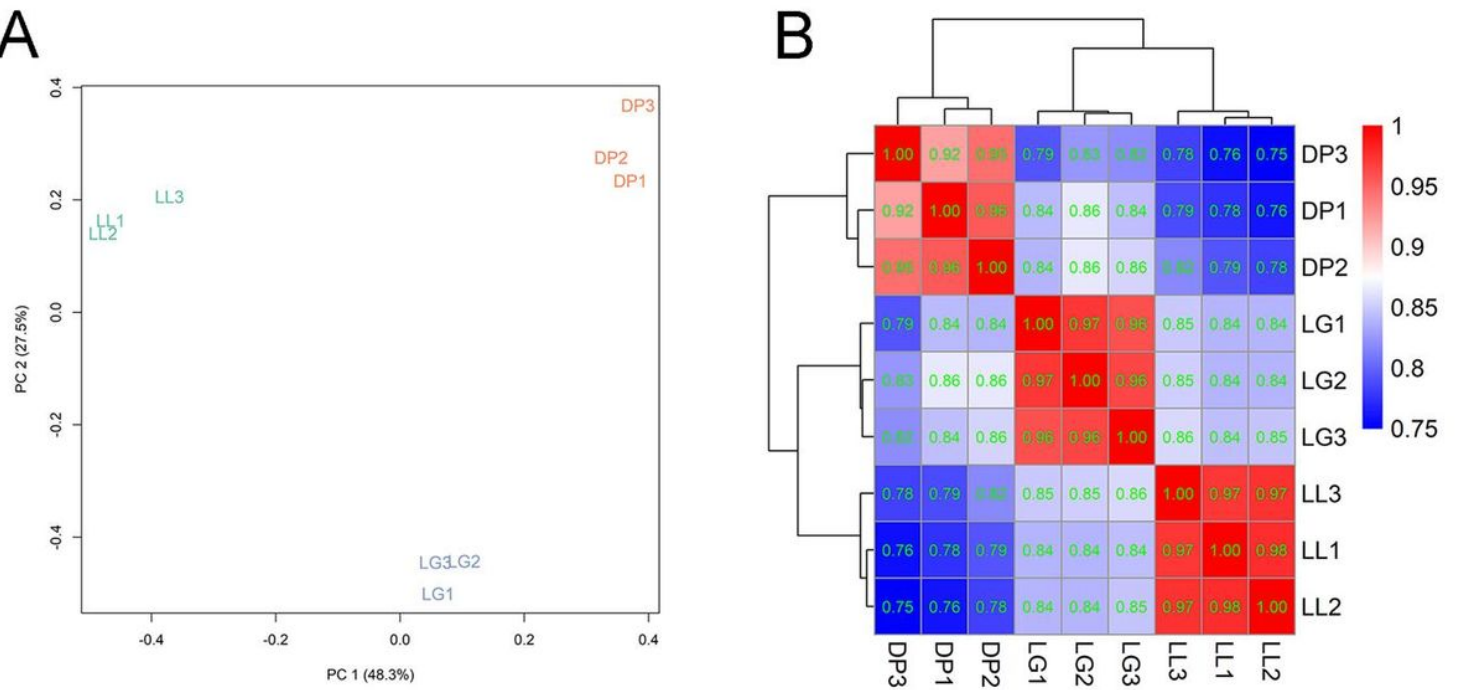


Figure 1

Principal component analysis and correlation analysis between sequencing samples A Principal component analysis of the nine sequencing samples. B Correlation heat map of the nine sequencing samples.

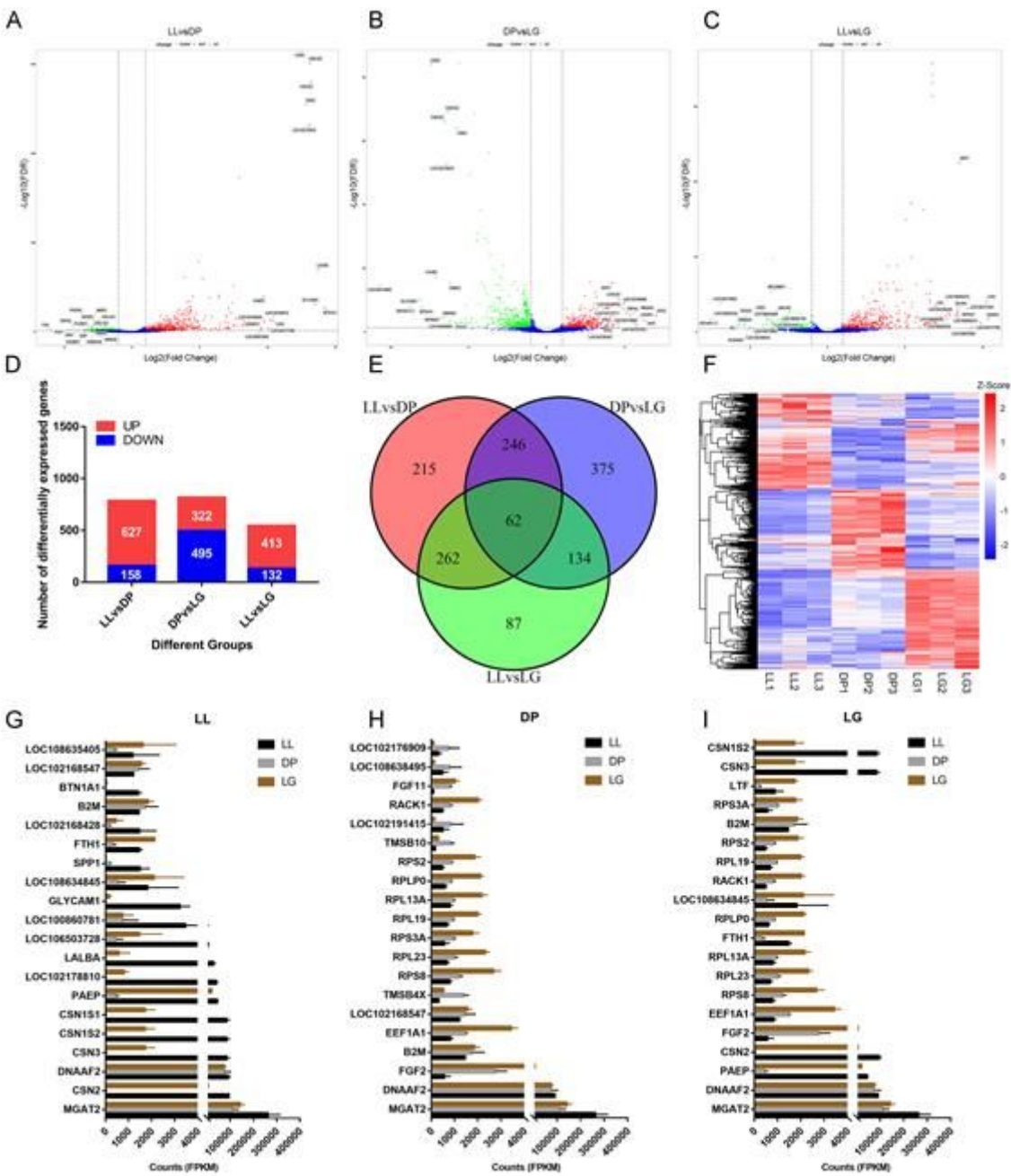


Figure 2

Screening DEGs in mammary gland tissue at different developmental stages A-C Volcano plots of the DEGs in the LL vs DP, DP vs LG, and LL vs LG comparisons. D Bar graphs of the DEGs that were up- and downregulated in the three comparison groups. E Venn diagram of DEGs in the three comparison groups. F Heat map of the expression patterns of all DEGs. G-I Bar graphs of the expression levels of the top 20 DEGs in the LL, DP, and LG stages.

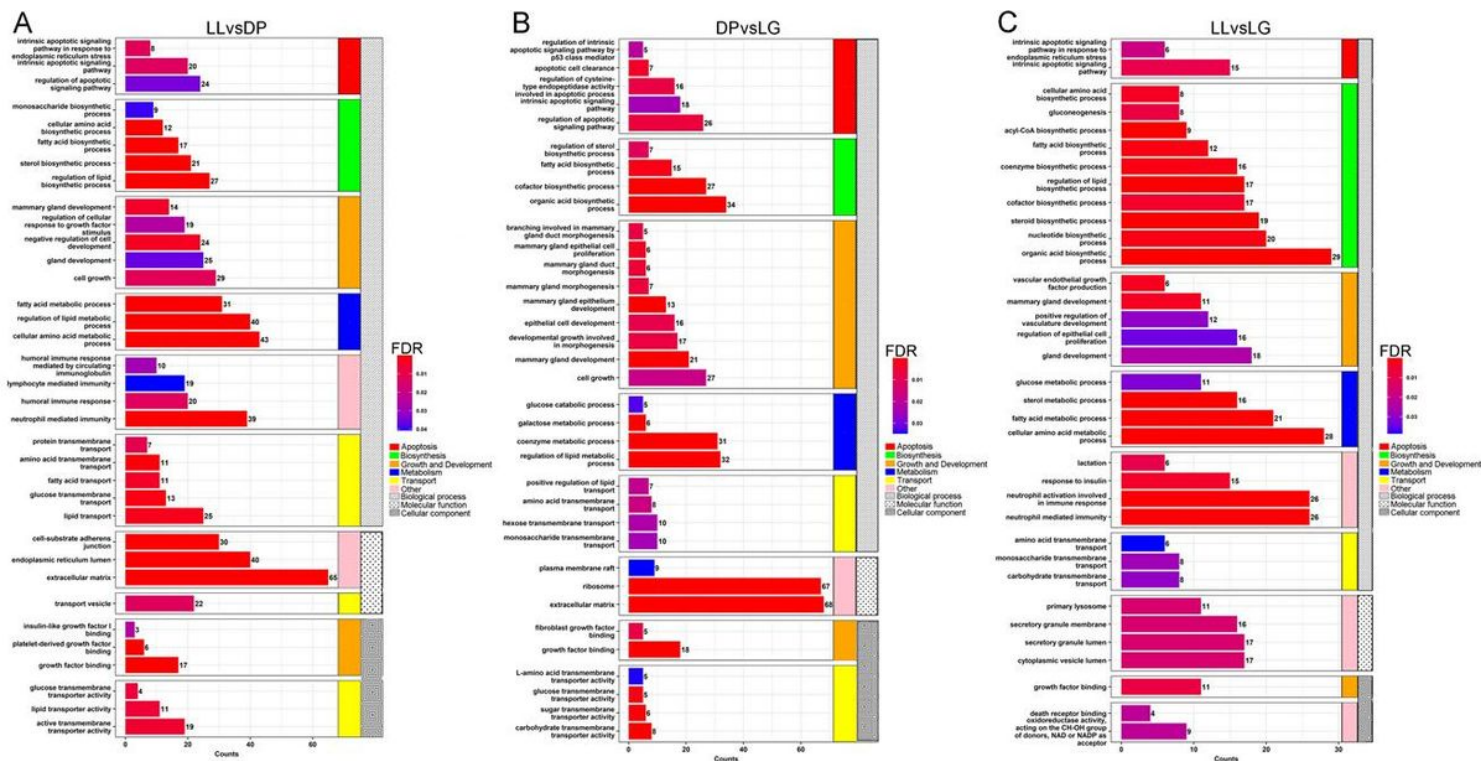


Figure 3

GO annotation of DEGs in the three developmental stages A GO functional annotation bar graphs for the DEGs identified in the LL vs DP comparison group. B GO functional annotation bar graphs for the DEGs identified in the DP vs LG comparison group. C GO functional annotation bar graphs for the DEGs identified in the LL vs LG comparison group.

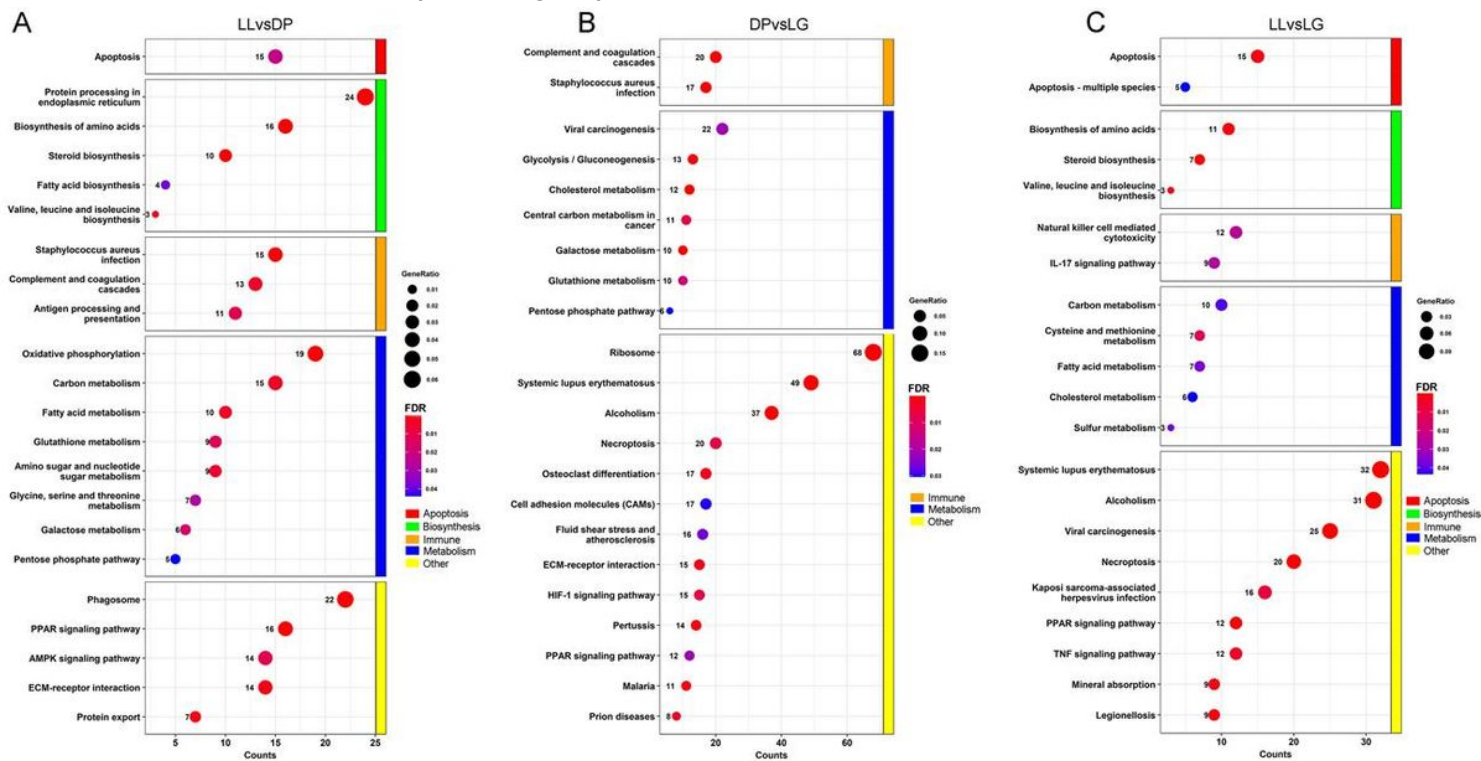


Figure 4

Page 26/30

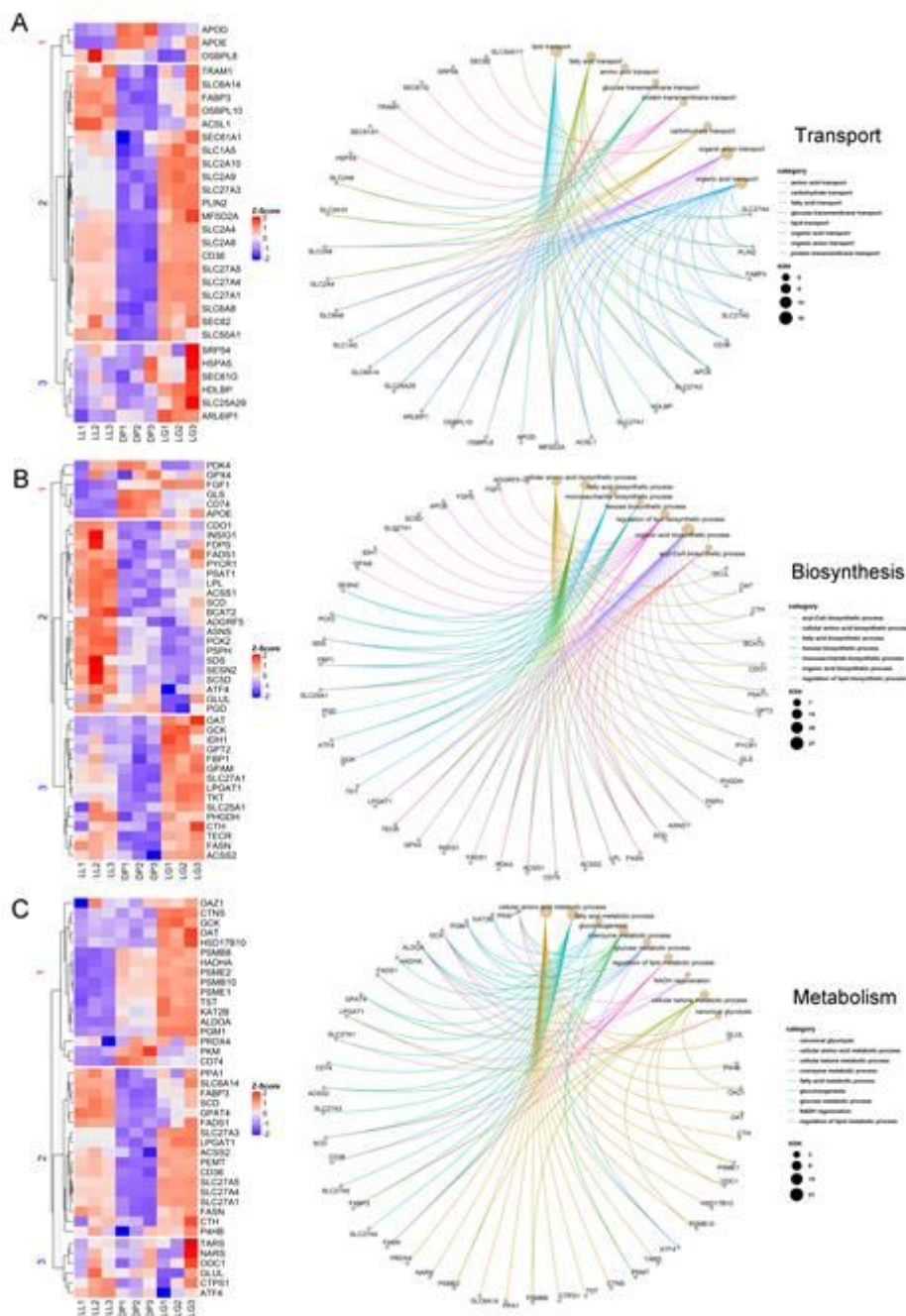


Figure 6

Functional networks of substance transport, biosynthesis and metabolism in the mammary gland A Functional network of genes related to substance transport in the mammary gland and heat maps of expression patterns of DEGs. B Functional network of genes related to biosynthesis in the mammary gland and heat maps of expression patterns of DEGs. C Functional network of genes related to substance metabolism in the mammary gland and heat maps of expression patterns of DEGs.

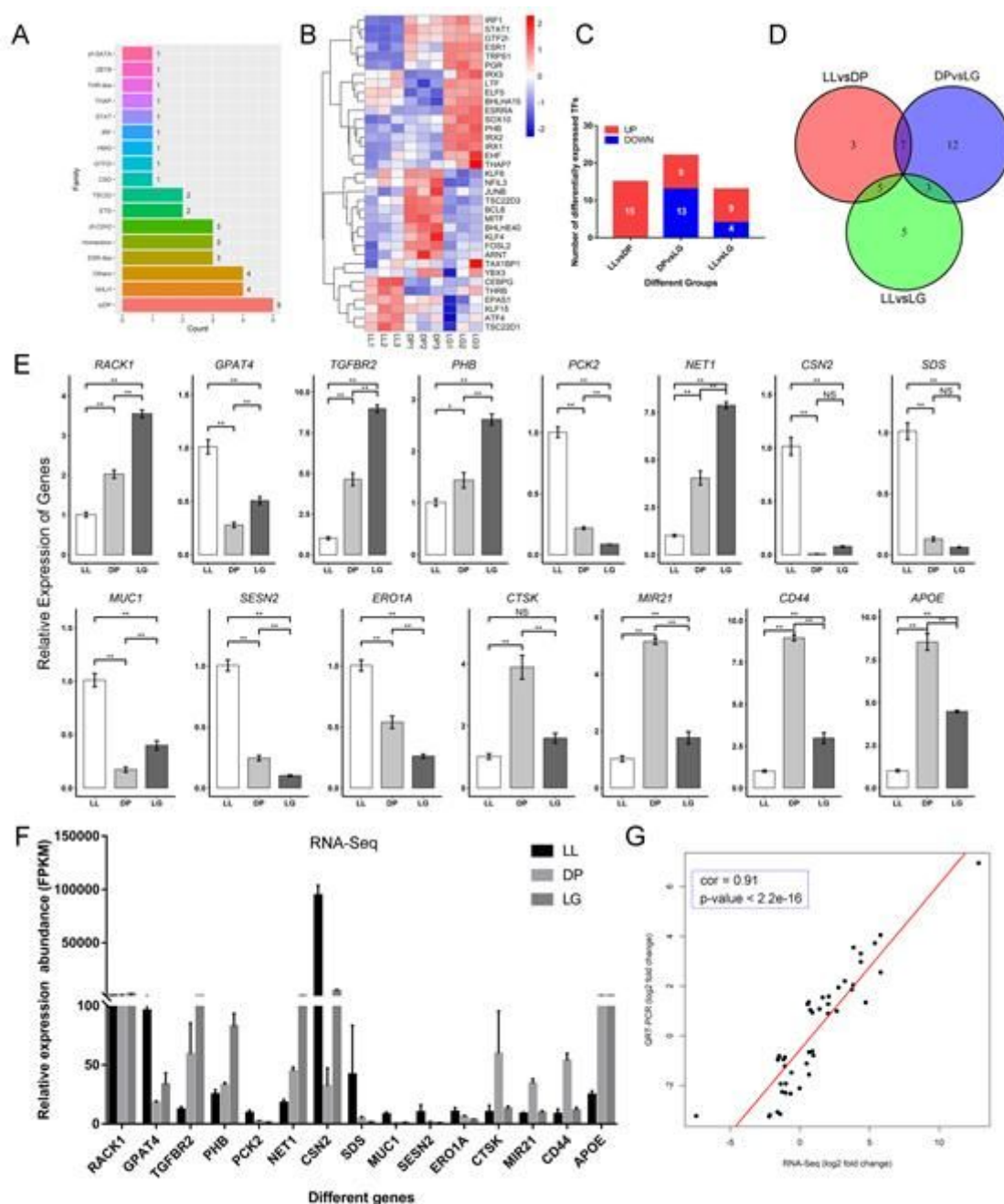


Figure 7

Identification of TFs and validation of DEGs by Real-Time Quantitative Reverse Transcription PCR A Classification of differentially expressed TF families in the three developmental stages. B Heat map of the expression of 35 differentially expressed TFs. C Statistical analysis of the upregulated and downregulated TFs in the three groups. D Venn diagram of the differentially expressed TFs in the three groups. E Histogram of the expression levels of the 15 DEGs detected by Real-Time Quantitative Reverse Transcription PCR. F Histogram of the expression levels of the 15 DEGs using transcriptome sequencing. G Scatter plot of the correlation between Real-Time Quantitative Reverse Transcription PCR and fold changes in transcriptome sequencing (log 2-fold change) of the 15 genes.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile7.pdf](#)
- [Additionalfile6.tif](#)
- [Additionalfile5.xlsx](#)
- [Additionalfile4.xlsx](#)
- [Additionalfile3.xlsx](#)
- [Additionalfile2.xlsx](#)
- [Additionalfile1.xlsx](#)