

Single-cell RNA-Sequencing Reveals Novel Myofibroblasts with Epithelial Cell-Like Features in the Mammary Gland of Dairy Cattle

Fengfei Gu

Zhejiang University

Jiajin Wu

Zhejiang University

Senlin Zhu

Zhejiang University

Teresa G. Valencak

zhejiang university

Jian-Xin Liu

Zhejiang University

Hui-Zeng Sun (✉ huizeng@zju.edu.cn)

Zhejiang University <https://orcid.org/0000-0001-5380-6030>

Research

Keywords: single-cell RNA-sequencing, mammary gland, dairy cows, epithelial cells, myofibroblast

Posted Date: November 13th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-101174/v1>

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

[Read Full License](#)

Abstract

Background: Cow's milk is a highly-nutritious dairy product that is widely consumed worldwide. It is secreted by the developed mammary gland (MG) of dairy cattle. However, a comprehensive understanding of cell-type diversity and cell function within bovine MG is lacking. In the current study, we used single-cell RNA sequencing to investigate the transcriptome of 24,472 high-quality MG cells isolated from newborn and adult cows.

Results: Unbiased clustering analysis revealed the existence of 24 cell types, which could be divided into four categories: 9 immune, 3 epithelial, 9 fibroblast, and 3 endothelial cell types. Other cell subtypes were further identified based on re-clustering and pseudotemporal reconstruction of epithelial cells that included 3 mature luminal epithelial, 1 intermediate, and 2 progenitor cell subtypes. The individual top marker genes of these 3 mature luminal epithelial cell subtypes (L0, L1, and L5) were *APOA1*, *STC2*, and *PTX3*, which were further validated using immunofluorescence. Based on functional analysis, the L0, L1, and L5 cell subtypes were all involved in the upregulation of lipid metabolism, protein and hormone metabolism, and the immune response, respectively. Furthermore, we discovered a novel myofibroblast that expresses *COL1A1* and *CSN3*, has visible epithelial-like characteristics, and shows the potential to differentiate into luminal epithelial cells, especially immune-sensing luminal cells (L5).

Conclusions: We constructed the first single-cell atlas of the dairy cow MG, and our new findings of epithelial-like myofibroblast cells and their differentiation trajectories into luminal cells may provide novel insights into the development and lactogenesis in dairy cattle MGs.

Background

Milk products contain optimal nutrients (e.g., protein, fat, vitamin) and are a vital calcium source in the human diet, especially for children and the elderly. Dairy cattle are important livestock animals since their milk production amounts to 83% of the total dairy products supplied worldwide [1]. Holstein cows have the highest productivity, producing 10-30 tons of milk/year depending on feed nutrition, individual genetics, and management, and are the most widely raised dairy cattle breed, which can largely be attributed to their well-developed mammary gland (MG). Therefore, it is of great significance to investigate the molecular physiology and biological mechanisms of the MG in Holstein cows further to enhance our understanding of the regulatory stages of lactation.

The remarkable dynamic changes of morphogenesis and function of the MG (especially in the epithelium) across the reproductive cycle forms the basis of milk synthesis in the MG. Similar to the developmental processes in mice and humans, there is a rudimentary ductal tree with fat pads at birth that then extends by allometric growth during puberty and reaches full maturity later during pregnancy [2]. Essentially to maintain high and continuous milk production, the MG in dairy cows shows a higher epithelial turnover rate during lactation and incomplete involution in the cessation of lactation [3] to maintain alveolar structure, unlike in mice and humans. Capuco et al. [4] reported that mature mammary

epithelial cells from dairy cows were estimated to undergo a 50% increased turnover rate and established a dynamic equilibrium with mesenchymal cells (e.g., immune cells, fibroblast cells, and adipose cells) during the lactation period. During the incomplete involution that occurs during the “dry” period when females do not produce milk, the number and function of epithelial cells are reduced through a powerful apoptotic mechanism and enhanced through the transition from mesenchymal cells at a later stage [5]. Therefore, the number and diversity of mammary epithelial cells and their association with mesenchymal cells are fundamental to understand the lactation mechanisms in dairy cows fully.

The molecular patterns (e.g., gene expression) of some mammary cell types have been described at the population level in dairy cows [6-8]. However, little is known about the cell lineage specification in the bovine MG, and a comprehensive understanding of the heterogeneity within the different bovine mammary cell populations is lacking. Single-cell RNA Sequencing (scRNA-Seq) allows cellular heterogeneity and functional profiles at the single-cell level to be examined [9]. In recent years, several studies have gradually improved this powerful technique to explore the cell type/subtype profiles, function and fate state of individual cells, communication paths, and developmental trajectories in different cell subtypes in the MG of both mice [10-12] and humans [13, 14]. This approach revealed a series of novel cell subsets, their marker genes, and interactions in the MG. However, to our knowledge, no comparable data are available for dairy cows regarding the MG or other tissues. In addition, although the studies in mouse and human provided essential insights into rare cellular states in MG development, the studies were restricted to epithelial cells, overlooking that mesenchymal cells such as fibroblasts might play an important role in the development of the MG [15]. Using scRNA-Seq to investigate the cell landscape of the MG in dairy cows is an important step and will contribute to a comprehensive understanding of MG development and lactation biology.

In the current study, we used unbiased scRNA-seq to generate a molecular census of cell types/subtypes, function, and states within the MG from newborn and adult dairy cattle. The main objectives of our study were to investigate the MG cell landscape of dairy cows and to reveal the relationships between epithelial and mesenchymal cells.

Materials And Methods

Animal and experimental design

All experimental procedures in this study were approved by the Animal Care Committee of Zhejiang University (Hangzhou, China) and conducted in accordance with the university’s guidelines for animal research. The experimental design is presented in Fig. 1a. A healthy newborn calf (NB; age: 5-days old; body weight: 35 kg) and a lactating adult cow (AD; age: 3-years old; body weight: 730 kg; milk yield: 18 kg/d) were selected.

MG cell sample collection

Animals were euthanized by cervical dislocation using a sharp knife in previously anesthetized individuals within a surgery room to obtain a ~10 g MG sample. Samples then were rinsed in DBPS, transferred into MACS Tissue Storage Solution (NO. 130-100-008, Miltenyi Biotec, Germany), stored in a 37°C vacuum container, and quickly taken to the lab. The dissociation of the MG was done according to published studies [10,11,16] with some modifications. In brief, the MG tissue samples were placed in a sterile RNA-free cell culture dish with 1 × PBS (with 5% FBS) and were cut into 5 mm pieces. The tissues were then washed using 1 × PBS (with 5% FBS) after removing non-targeted tissues such as white adipose tissue. Subsequently, samples were gently placed in 20 mM EDTA (with 5% FBS) that had been incubated at 37°C at 100 rpm for 20 min. After that, samples were digested with 25 U/ml hyaluronidase, 0.5 mg/mL Collagenase \square , 0.5 mg/mL Collagenase IV, 25 U/ml DNase I, and 5% FBS in HBSS for 30 min at 37°C. The mixture was strained through a 70–30 μ m cell strainer and transferred into a new EP tube, and centrifuged at around 25°C at 300 × *g* for 5 min. After removing the supernatant, the cells were washed with 3 mL HBSS (with 5% FBS) and centrifuged at around 25°C at 300 *g* for 5 min. The Countess II Automated Cell Counter (AMQAX1000, Thermo Fisher, USA) was applied to measure cell number and viability. The requirement for cell viability was greater than 90%. If the viability for isolated cells was lower than the threshold, we used a MACS Dead Cell Remove Kit (NO. 130-090-101, Miltenyi Biotec, Germany) to remove dead cells following the manufacturer's recommendations.

Cell capture, cDNA library construction, sequencing, and data pre-processing

The high-quality single-cell suspension was loaded in a 10 × Genomics Chromium machine for the cell capture and cDNA library construction according to the Single-Cell 3' Protocol recommended by the manufacturer. RNA-seq was performed by the NovaSeq 6000 system (2 × 150 bp). The raw reads were demultiplexed and converted to the fastq format using Illumina bcl2fastq software (version 2.20). Cell Ranger (version 3.1.0, <https://support.10xgenomics.com/single-cell-gene-expression/software/overview/welcome>) was used for demultiplexing, barcode processing, and UMI quantification. Seurat (version 3.1.1, <https://satijalab.org/seurat/>) was used for cell filtration, dimensional reduction, clustering, differential gene expression analysis, and marker gene screening of scRNA-Seq data. Overall, cells with 500~4,000 genes, UMI counts less than 50,000, and a mitochondrial gene ratio smaller than 20% were retained for downstream analysis. Using the above pipeline, we obtained 24,472 high-quality single cells (12,432 in NB and 12,040 in AD) from the MG to generate a dairy cow single-cell atlas.

Cell type clusters and marker genes identification

To identify the cell types, we further reduced the dimensionality of the variable genes in all high-quality cells using Seurat (version 3.1.1) and the t-distributed stochastic neighbor embedding (T-SNE) algorithm to project the variables into a two-dimensional space. Briefly, the "Normalization" function in the R package Seurat was used to normalize the gene expression. Then the function "FindVariableGenes" within the R package Seurat was applied to select variable genes. A principal component analysis was performed by using the normalized gene expression value, and finally, Harmony analysis [17] was used to

correct for batch effects. The “FindClusters” function was used to cluster cells with a resolution of 0.80, and cells were visualized using the “RunTSNE” function within Seurat in R. In addition, differential expression analyses for each cluster were performed with a likelihood-ratio test for single-cell gene expression by using the default parameters via the “FindAllMarkers” function again in Seurat. Cluster cell identity was assigned by manual annotation while using known marker genes from already published papers [10-14].

Diffusion maps and pseudotime inference

Diffusion maps were used for inferring the differentiation trajectory between the different cell types[18]. Briefly, all cells from the target clusters were selected, and the highly variable genes were detected. After that, the log-transformed ($\log_2(\text{count} + 1)$) gene counts were used to compute diffusion components using the ‘DiffusionMap’ function [19]. The Monocle2 package was applied to reconstruct cell fate decisions and differentiation trajectories. This method uses a reverse embedded, graph-based user-defined gene list to generate a pseudotime graph explaining both branches and the linear differentiation process [20]. The BEAM function in the Monocle2 package was applied to process the cell dataset that was reduced by DDRTree and classified by the calculated pseudotime before generating the branch heatmap.

Weighted gene co-expression network analysis and Gene Ontology (GO) analysis

A total of 2,827 target genes were selected for weighted gene co-expression networks (WGCNA), which were constructed by the WGCNA package in R (version 4.02) [21]. We set the MEDissThres as 0.25 to merge similar modules, while a minimum module size was comprised of 10 genes. GO analysis of the list of marker genes (supplementary file) per cluster was performed on functional annotation by using the DAVID Bioinformatics Resources 6.8 tool (<https://david.ncifcrf.gov/summary.jsp>).

Ligand-receptor interaction and Pearson correlation analysis

To infer cell-cell communications from scRNA-Seq data, the SingleCellSignalR R package [22] was used to analyze receptor-ligand interactions based on a database including 2,560 previously published ligand-receptor pairs [23]. The iTALK R package was applied to visualize the results with an LRscore > 0.7. Finally, Pearson correlation was performed using OmicStudio tools from <https://www.omicstudio.cn/tool>.

Immunofluorescence analysis

The 5 mm MG samples of both NB and AD were first fixed in 4% paraformaldehyde for 48 h at 4 °C, then dehydrated by gradient concentrations of ethanol, cleared with xylene, and then embedded in melted paraffin. The Leica SM2010 R Sliding Microtome (Leica Biosystems, Wetzlar, Germany) was used to make 3- μm tissue slide sections. Slides were incubated twice for 15-min in xylene to remove paraffin, followed by two 5-min incubations in pure ethanol, and incubated in lower concentrations of ethanol to dehydrate. Subsequently, the slides were washed with distilled water. Next, the slides were immersed in EDTA antigen retrieval buffer (pH 8.0) and maintained at a sub-boiling temperature for 8 min, then rested

for 8 min, and then finally re-exposed to sub-boiling temperatures for 7 min. Tissues were blocked with 3% BAS for 30 min at around 25°C, then incubated with primary antibodies (diluted with PBS appropriately) overnight at 4°C, and washed three times with PBS (pH 7.4). Then the objective tissue was covered with a secondary antibody and incubated at around 25°C for 50 min at dark conditions. Slides were washed three times in PBS (pH 7.4) and then incubated with DAPI solution at around 25°C for 10 min in a dark place. After washing three times with PBS (pH 7.4), the slides were mounted with an anti-fade mounting medium. For the evaluation of the staining (e.g., KRT18 and PTX3 staining), we selected positive cells as a signal around the nuclei (DAPI) and captured them with a laser confocal scanning microscope (IX81-FV1000, Olympus, Japan) in at least three different fields of view using a 40× objective in two different samples.

The primary antibodies we used were anti-cytokeratin 18 (KRT18) mouse mAb diluted 1:100 (NO. M0407-19; HUABIO, Wuhan, China); anti-Ptx3 rabbit pAb diluted 1:100 (NO. ER64704, HUABIO, Wuhan, China); anti-STC2 rabbit pAb diluted 1:100 (NO. ER61963, HUABIO, Wuhan, China); anti-apolipoprotein A1 (APOA1) rabbit mAb diluted 1:100 (NO. ET1702-23, HUABIO, Wuhan, China); anti-collagen I (COL1A1) rabbit mAb diluted 1:50 (NO. ET1609-68, HUABIO, Wuhan, China); and anti-CSN3 rabbit pAb diluted 1:50 (NO. ET7109-60, HUABIO, Wuhan, China).

As secondary antibodies, we used goat anti-mouse IgG conjugated with Alexa Fluor® 488 (ab150113, Abcam, Cambridge, United Kingdom); goat anti-mouse IgG conjugated with Alexa Fluor® 555 (ab150114, Abcam, Cambridge, United Kingdom); goat anti-rabbit IgG conjugated with Alexa Fluor® 555 (ab150078, Abcam, Cambridge, United Kingdom); and goat anti-rabbit IgG conjugated with Alexa Fluor® 488 (ab150077, Abcam, Cambridge, United Kingdom). All the secondary antibodies were diluted 1:1000.

Results

Single-cell transcriptional landscape of MG in dairy cows

A total of 28,393 cells (13,463 in NB and 14,840 in AD) were captured and 24,472 high-quality cells (12,432 in NB and 12,040 in AD) were analyzed after quality control (Fig. 1a). Twenty-four clusters were identified based on the marker genes, with the composition of the cell lineages being very different between NB and AD cows (Fig. 1b and c; Table S1). Based on the gene expression profiles, the 24 clusters were categorized into four major groups as follows: immune cells (clusters 0, 1, 6, 8, 14, 15, 18, 21, and 23); epithelial cells (clusters 16, 18, and 20); fibroblast cells (clusters 2, 3, 5, 7, 9, 10, 11, 13, and 20); and endothelial cells (clusters 4, 12, and 22) (Fig. 1d and e).

All immune cells were determined based on the high expression of *PTPRC* (Fig. 1e). We observed that the immune cells consisted of CD8T, CD4T, NK, Mast, T, macrophages, neutrophils, B, and activated T cells (Table S2). The cells from the MG epithelium were identified based on the high expression of *EPCAM* (Fig. 1e) and co-expression of *KRT19*, *KRT8*, *KRT7*, and *KRT15* (Table S1). We observed clear differences in gene expression profiles between these three epithelial cell types. Specifically, cluster 16 (C16) showed patterns of luminal cells (high expression of *KRT19*, *LTF*), basal cells (high expression of *KRT5* and

KRT15), and their progenitors (*ELF5*) at the same time (Table S1), indicating that C16 may represent a mixed lineage that contains several epithelial cell types. Cluster 17 (C17) was interpreted as mature luminal cells based on the high expression of *KRT8*, *KRT18*, and *AREG* (Table S1). Cluster 19 (C19) expressed not only the marker gene of progenitor cells *ELF5* but also the genes associated with milk protein synthesis, such as *LTF*, *CSN2*, *CSN3*, and *LALBA* (Table S1). Thus, C19 may represent an alveolar luminal progenitor cell or even a mixed lineage. Fibroblast cells expressed genes encoding collagen and extracellular matrix-related molecules, such as *COL1A1*, *DPT*, *COL3A1*, and *DLK1* (Fig. 1e). Endothelial cells were identified due to the high expression of *PECAM1*, *APOLD1*, *CLDN5*, *ECSCR*, and *EGFL7* (Fig. 1d, e).

Analyzing these major cell types, we found that the distribution of epithelial and fibroblast cells differed significantly between NB and AD cows (Fig. 1f), which that epithelial cells were mainly observed in the AD, fibroblast cells were mainly identified in the NB. Due to the diversity of epithelial and fibroblast cell types, additional re-clustering was explored.

Reconstruction of mammary epithelium differentiation hierarchy

To better evaluate epithelial cells (C16, 17, and 19), additional re-clustering was performed, and six independent sub-clusters were generated (L0–L5; Fig. 2a and b). All of the sub-clusters featured high expression of the luminal epithelial cell marker genes *KRT18*, *KRT19*, and *KRT8* (Fig. 2c). Moreover, the expression profiles of the top abundant genes were very similar between L2 and L3. Parts of the top abundantly expressed genes were shared between L4 and L0/L1 (Fig. 2b). Both L2 and L3 were the luminal alveolar progenitor cell subtypes because of the high expression of *ELF5*, *CSN2*, and *CSN3* (Table S3). L4 showed characteristics of both progenitor cells (*ALDH1A3*) and mature luminal cells (*KRT18*, *KRT8*, *AREG*; Table S3); thus, it was considered an intermediate luminal cell subtype. The L0, L1, and L5 genes expressed the unique marker gene *APOA1*, *STC2*, and *PTX3*, respectively (Fig. 2c; Table S3).

To further investigate the characteristics of L0, L1, and L5 and understand how these observed cell subtypes and states are related, the pseudotemporal analysis of single cells was performed using Monocle2. We found one tightly connected differentiation trajectory that was separated into two main branches, of which the start points of pseudotime were located within L2 and L3 cells (Fig. 2d). L0, L1, and L5 were present at the end of the two branches (Fig. 2d), indicating these clusters are different subtypes of mature luminal epithelial cells. Except for expressing the luminal cell marker gene *KRT19*, the L0 subtype also expressed *LTF*, as well as *KRT5* and *KRT15* (marker genes of basal cells) (Table S3). Therefore, the L0 subtypes showed the properties of differentiated luminal cells. L1, located at the end of the other branch of the epithelial differentiation hierarchy, showed the typical markers of mature luminal cells such as *KRT18*, *KRT8*, *AREG*, and *PRLR* (Table S3). It was considered a mature luminal cell and was defined as a secretory alveolar cell. L5 expressed *PTX3*, which responds to inflammatory stimuli in epithelial cells and regulates the innate response to pathogens and inflammatory reactions. Thus, we defined this subtype as an immune-sensing luminal cell. Additionally, the immunofluorescence analysis

was used to validate the three subtypes at the protein level, to integrate the newly discovered cell subtypes spatially, and to create the single-cell atlas of the cattle MG (Fig. 2e).

Characteristics and functions of epithelial cell subtypes

Next, we identified 682 genes with branch-specific expression patterns, as well as the key genes and biological processes changing over pseudotime (Fig. 3a). Towards the L1 branch, we found three gene clusters that increased in expression during differentiation. These clusters consisted of key genes that were involved in the development of the MG alveolus and responded to amino acid, mineral, and hormone synthesis (Fig. 3a, Table S4). Towards the L0 branch, one cluster of highly expressed genes was involved in lipid function (e.g., isoprenoid, sterol, steroid, and cholesterol), biosynthetic processes, and regulation of epithelial cell differentiation and proliferation (Fig. 3a, Table S4). We also observed one cluster of genes that gradually decreased expression in both branches, which were involved in response to progesterone and estradiol and the prolactin signaling pathway (Fig. 3a, Table S4).

Five other co-expressed gene modules were generated using WGCNA to evaluate the gene co-expression patterns of L0, L1, and L5. There were 412, 209, 1682, 503, and 21 genes in the blue, brown, grey, turquoise, and yellow modules, respectively (Fig. 3b, Table S5). We found that the turquoise, brown, and yellow modules showed a highly positive correlation with the L0 ($R = 0.81$), L1 ($R = 0.88$), and L5 ($R = 0.44$) subtypes, respectively (Fig. 3b). Next, we performed GO term enrichment analysis to explore the biological processes of the genes in these three key modules (Fig. 3c). The results showed that genes from the yellow module were involved in cytoskeleton organization and immune response. The genes from the brown module were mainly associated with MG epithelial cell differentiation and alveolus development, lactation, and protein, hormone, and mineral metabolism. Finally, the turquoise module genes were mainly involved in epithelium development, epithelial cell differentiation, and the regulation of lipid biosynthetic and metabolic processes, including sterol, steroid, cholesterol, and fatty acid metabolism (Fig. 3c).

Reconstruction of fibroblast cell subtypes

To further reveal the subtypes of fibroblast cells, additional clustering was performed and 14 independent sub-clusters were generated (F0–F13; Fig. 4a and b). Based on the marker gene expression profiles, we defined F0, 3, 6, 7, and 10 as lipofibroblasts and F4 as the matrix fibroblast cell. F5, 8, 13 showed the characteristics of myofibroblasts. However, the other five fibroblast cell subtypes could not be defined (Table S6). Interestingly, we found that F13 showed strong characteristics of epithelial and progenitor cells, which was reflected by the highly expressed marker genes such as *EPCAM*, *KRT18*, *KRT8*, *ELF5*, and *CSN3* (Fig. 4c). Immunofluorescence staining confirmed the existence of F13 with its marker genes *COL1A1* and *CSN3* (Fig. 4d). In addition, F13 was mainly involved in cell differentiation; response to estradiol, progesterone, and estrogen; cell migration; apoptosis; and growth according to the GO analysis (Fig. 4e).

Investigation of luminal and fibroblast cell differentiation hierarchy

For a comprehensive understanding of the relationships between fibroblasts and epithelial cells, a diffusion map was used to computationally reconstruct differentiation stages and the transitions between luminal epithelial and fibroblast cell subtypes. We found that the epithelium and fibroblasts were not completely separated, and ~60% of F13 cells were clustered with epithelial cells (Fig. 5a). The F13 subtype showed a strong relationship with L1 ($R = 0.871$) and L5 ($R = 0.928$) based on Pearson correlation (Fig. 5b). In addition, the epithelial cells (L1–L5) and fibroblasts (F12) interacted with F13 through the ligand-receptor COL1A1-ITGA2B, IGF2-IGF1R/INSR, HMGB1/THBS1-SDC1, and RBP4-STRA6 (Fig. 5c).

Discussion

Defining cell heterogeneity in organs and tissues is critical to understanding its biology, development, and function. However, the current state of knowledge in bovine MG biology is primarily restricted to the population level and mainly focuses on epithelial cell types [6,8,24]. Comprehensive knowledge about expression signatures and cellular identities of mammary cells, therefore, remains inadequate in cows. Our scRNA-seq analysis of the MG from NB and lactating AD cows, to our knowledge, is the first study to consider the unbiased, de novo identification of distinct cell types and states in the MG of cattle. Based on the available bovine MG single-cell atlas, we discovered and verified three mature luminal epithelial cell subtypes with different biological functions. Moreover, we validated a novel myofibroblast with the marker genes *COLIA1* and *CSN3* and revealed its differentiation potential into luminal epithelial cells.

Lactation is a complex biological process that involves many hormones and synthesizes nutrients such as milk fat and protein [25]. Mammary epithelial cells are primarily responsible for the synthesis and secretion of milk. In a scRNA-seq of mouse MG, alveolar differentiated cells, hormone sensing luminal cells, and some progenitor cells were observed [10,14]. Similar results were obtained by us in the study of cows, with the L0 showing both luminal and basal characteristics and therefore being considered a differentiated cell; L1 was identified as a secretory alveolar cell and marked by the estrogen-regulated secreted glycoprotein *STC2*. Thus, we considered it a hormone sensing luminal cell type. The results of pseudotemporal, WGCNA, and GO analysis of epithelial marker genes further supported our above interpretation. In addition, both epithelial cell types (L0 and L1) showed significant differences in the biological processes of lactation. L0 was mainly involved in lipid metabolism, while L1 was more sensitively responsive to milk protein synthesis. To date, the common textbook knowledge on dairy cows states that all epithelial cells carry out the same function, such as milk fat and protein synthesis [25]. However, our new results suggest that different types of mammary epithelial cells indeed have different effects on the synthesis of different milk constituents. Our data may be helpful for potentially successful experimental manipulations and future targeted investigations on the synthesis and metabolic mechanisms of milk fat or protein.

The mouse cell atlas on the MG was mapped using Microwell-seq in 2018, showing diverse immune cell types in the MG [13]. Consistent with the mouse results, we identified many immune cell types in the bovine MG, especially for AD cows. As a result of continuous lactation, the MG of AD cows is more

susceptible to external stimuli and pathogens, easily leading to inflammatory diseases such as mastitis [26]. It is suggested that abundant immune cells reside in the MG of AD cows as a self-defense mechanism and would be beneficial to health and lactation in the bovine MG. Interestingly, we identified and validated a luminal epithelial cell subtype (L5) featuring *PTX3*, which responds to inflammatory stimuli in epithelial cells [27] and regulates the innate response to pathogens and inflammatory reactions [28]. This immune-sensing luminal epithelial cell subtype has not been reported in the previous scRNA-Seq studies of mouse or human MGs. Recently, Krausgruber et al. [29] reported that epithelial cells were also involved in the immune response through analyzing the relationship between structural cells (epithelial cells, endothelial cells, and stroma) and immune cells in 12 organs of mouse. They tentatively proposed the term 'structural immunity' to emphasize the importance of structural cells for the mammalian immune system. Therefore, our study provides a new perspective that different epithelial cell subtypes may be involved in different primary functions, which not only relate to milk production but also play essential roles in the immunity of the MG during lactation.

Capuco et al. reported that mature mammary epithelial cells were estimated to show a 50% turnover rate during the lactation period in dairy cows [4], which largely contributes to continuous lactation over a long period. Pal et al. [10] identified some luminal progenitors that can differentiate into mature luminal epithelial cells in the mouse MG. Two luminal progenitors and one intermediate cell were identified in our current study and might help explain the high turnover rate of lactating dairy cows. In addition, the mesenchymal microenvironment surrounding epithelial cells also has a significant effect on epithelialization [30]. For example, the mammary fibroblasts, which maintain extracellular matrix homeostasis, are also involved in mammary morphogenesis when interacting with epithelium [31, 32]. The novel epithelial-like myofibroblast subtype F13 identified in this study was mainly involved in cell differentiation and response to estradiol, progesterone, and estrogen. Progesterone and estradiol are significant triggers of MG epithelial cell development [33]. The diffusion map results visually confirmed the relationship between F13 and epithelial cells. Specifically, we further observed that F13 had the strongest correlation with the immune-sensing luminal epithelial cell L5 that positively responded to lactation and was involved in the immune response. Therefore, the combined information suggested that F13 has the potential to differentiate into immune-sensing luminal epithelial cells. Thus, it might play an important role in the remodeling and health of bovine MG.

Cellular transition is regulated by many factors, such as other surrounding cells. We found that L1, L2, L3, and L5 may promote the differentiation of F13 cells through a paracrine mechanism with the ligand insulin growth factor 2 (IGF2) interacting with insulin growth factor receptor (IGF1R) or insulin receptor (INSR). The IGF system comprises the two ligands IGF1 and IGF2, which target tyrosine kinase receptors, including IGF1R and INSR, as well as the IGF2 receptor (IGF2R). These receptors play important roles in regulating critical cellular processes such as cell proliferation, development, differentiation, and survival [34]. Chen et al. reported that the IGF1R-related signaling activation of fibroblasts could induce cellular stemness in humans [35]. In addition, Xu et al. reported that IGF2 secreted by cancer cells instigates fibroblasts to promote cancer progression [36]. The above information indicates that the IGF system and the activation of F13 may be involved in the differentiation of immune-sensing luminal epithelial cells

and that the epithelial cell subtypes of L1, L2, L3, and L5 may enhance the process. However, the underlying molecular mechanisms require further exploration.

Conclusions

In summary, our study is the first one to investigate the heterogeneity of bovine MG and to construct a bovine MG cell atlas using scRNA-Seq. Our findings provide novel insights into the lactation biology of bovine MG. We identified and validated three mature luminal epithelial cells that are involved in different lactation processes as well as in the immune response. In addition, a novel myofibroblast showing potential differentiation into luminal epithelial cells was identified. The IGF pathway system signal may play an important role in the F13-L5 transition, and the hormone-sensing luminal cells (L1), luminal progenitor cells (L2, L3), and immune-sensing luminal cells are correlated with the process. Nevertheless, further studies are warranted to clarify the molecular mechanisms involved in these processes.

Abbreviations

AD, adult cow; GO, Gene Ontology; MG, mammary gland; NB, newborn cow; scRNA-Seq, Single-cell RNA Sequencing; WGCNA, weighted gene co-expression networks.

Declarations

Ethics approval and consent to participate

All experimental procedures involving animals were approved by the Animal Care and Use Committee of Zhejiang University (Hangzhou, China) and were followed the university's guidelines for animal research.

Consent for publication

Not applicable.

Availability of data and materials

The raw file and processed data of scRNA-seq can be accessed at National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) database under the accession number GSE158117.

Competing interests

The authors declare that they have no competing interests.

Funding

This study was financially supported by the National Natural Science Foundation of China (31930107, 32002207, Beijing) and the "Hundred Talents Program" Research Professor Start-up Fund of Zhejiang University.

Authors' contributions

FFG, JXL, and HZS designed the research. FFG, JJW, SLZ, and HZS collected the samples. FFG, SLZ, and JJW performed bioinformatics analysis. FFG drafted the manuscript. TGV, JXL, and HZS supervised the studies, lead the scientific discussion, and revised the manuscript. All authors have read and approved the final manuscript.

Author details

Institute of Dairy Science, College of Animal Sciences, Zhejiang University, Hangzhou 310058, P. R. China.

Acknowledgments

The authors would like to thank the staff of the Yi-Jing Ecological Dairy Farm (Shaoxing, China) and Zheng-Xing Dairy Farm (Hangzhou, China) for providing the experimental animals. We also acknowledge all members of the Institute of Dairy Science Zhejiang University (Hangzhou, China) for their assistance with sampling and sample analysis.

References

1. Burgess K. Milk and dairy products in human nutrition. In: Muehlhoff E, Bennett A, McMahon D, editors. Food and Agriculture Organization of the United Nations (FAO), Rome. E-ISBN: 978-92-5-107864-8. *Int J Dairy Technol.* 2013; vol. 67. p. 303-4.
2. Macias H, Hinck L. Mammary gland development. *Wiley Interdiscip Rev Dev Biol.* 2012; 1: 533-57.
3. Watson CJ, Khaled WT. Mammary development in the embryo and adult: a journey of morphogenesis and commitment. *Development.* 2008; 135: 995-1003.
4. Capuco AV, Wood DL, Baldwin R, McLeod K, Paape MJ. Mammary cell number, proliferation, and apoptosis during a bovine lactation: relation to milk production and effect of bST. *J Dairy Sci.* 2001; 84: 2177-87.
5. Watson CJ, Kreuzaler PA. Remodeling mechanisms of the mammary gland during involution. *Int J Dev Biol.* 2011; 55: 757-62.
6. Finot L, Chanat E, Dessauge F. Mammary epithelial cell lineage changes during cow's life. *J Mammary Gland Biol Neoplasia.* 2019; 24: 185-97.
7. Akers RM. Triennial Lactation Symposium/Bolfa: Plasticity of mammary development in the prepubertal bovine mammary gland. *J Anim Sci.* 2017; 95: 5653-63.
8. Rauner G, Barash I. Enrichment for repopulating cells and identification of differentiation markers in the bovine mammary gland. *J Mammary Gland Biol Neoplasia.* 2016; 21: 41-49.
9. Xue R, Li R, Bai F. Single cell sequencing: technique, application, and future development. *Sci. Bull.* 2015; 60: 33-42

10. Pal B, Chen Y, Vaillant F, Jamieson P, Gordon L, Rios AC, et al. Construction of developmental lineage relationships in the mouse mammary gland by single-cell RNA profiling. *Nat Commun.* 2017; 8: 1627.
11. Bach K, Pensa S, Grzelak M, Hadfield J, Adams DJ, Marioni JC, et al. Differentiation dynamics of mammary epithelial cells revealed by single-cell RNA sequencing. *Nat Commun.* 2017; 8: 2128.
12. Chung CY, Ma Z, Dravis C, Preissl S, Poirion O, Luna G, et al. Single-Cell Chromatin analysis of mammary gland development reveals cell-state transcriptional regulators and lineage relationships. *Cell Rep.* 2019; 29: 495-510.
13. Han X, Wang R, Zhou Y, Fei L, Sun H, Lai S, et al. Mapping the mouse cell atlas by Microwell-Seq. *Cell.* 2018; 172: 1091-107
14. Nguyen QH, Pervolarakis N, Blake K, Ma D, Davis RT, James N, et al. Profiling human breast epithelial cells using single cell RNA sequencing identifies cell diversity. *Nat Commun.* 2018; 9: 2028.
15. Avagliano A, Fiume G, Ruocco MR, Martucci N, Vecchio E, Insabato L, et al. Influence of fibroblasts on mammary gland development, breast cancer microenvironment remodeling, and cancer cell dissemination. *Cancers (Basel).* 2020; 12: 1697.
16. Kanaya N, Chang G, Wu X, Saeki K, Bernal L, Shim HJ, et al. Single-cell RNA-sequencing analysis of estrogen- and endocrine-disrupting chemical-induced reorganization of mouse mammary gland. *Commun Biol.* 2019; 2: 406.
17. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods.* 2019; 16: 1289-96.
18. Haghverdi L, Büttner F, Theis FJ. Diffusion maps for high-dimensional single-cell analysis of differentiation data. *Bioinformatics.* 2015; 31: 2989-98.
19. Angerer P, Haghverdi L, Büttner M, Theis FJ, Marr C, Büttner F. destiny: diffusion maps for large-scale single-cell data in R. *Bioinformatics.* 2016; 32: 1241-43.
20. Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol.* 2014; 32: 381-86.
21. Yi G, Zhao Y, Xie F, Zhu FX, Li B. Single-cell RNA-seq unveils critical regulators of human FOXP3+ regulatory T cell stability. *Sci. Bull.* 2020; 65: 1114-24.
22. Cabello-Aguilar S, Alame M, Kon-Sun-Tack F, Fau C, Lacroix M, Colinge J. SingleCellSignalR: inference of intercellular networks from single-cell transcriptomics. *Nucleic Acids Res.* 2020; 48: e55.
23. Ramilowski JA, Goldberg T, Harshbarger J, Kloppmann E, Lizio M, Satagopam VP, et al. A draft network of ligand-receptor-mediated multicellular signaling in human. *Nat Commun.* 2015; 6: 7866.
24. Lawson DA, Bhakta NR, Kessenbrock K, Prummel KD, Yu Y, Takai K, et al. Single-cell analysis reveals a stem-cell program in human metastatic breast cancer cells. *Nature.* 2015; 526: 131-5.
25. Krausgruber T, Fortelny N, Fife-Gernedl V, Senekowitsch M, Schuster LC, Lercher A, et al. Structural cells are key regulators of organ-specific immune responses. *Nature.* 2020; 583: 296-302.

26. Angelopoulou A, Warda AK, Hill C, Ross RP Non-antibiotic microbial solutions for bovine mastitis-live biotherapeutics, bacteriophage, and phage lysins. *Crit Rev Microbiol*. 2019; 45: 564-80.
27. Han B, Mura M, Andrade CF, Okutani D, Lodyga M, dos Santos CC, et al. TNF α -induced long pentraxin PTX3 expression in human lung epithelial cells via JNK. *J Immunol*. 2005; 175: 8303-11.
28. Garlanda C, Hirsch E, Bozza S, Salustri A, De Acetis M, Nota R, et al. Non-redundant role of the long pentraxin PTX3 in anti-fungal innate immune response. *Nature*. 2002; 420: 182-6.
29. Truchet S, Honvo-Houéto E. Physiology of milk secretion. *Best Pract Res Clin Endocrinol Metab*. 2017; 31: 367-84.
30. Wolfson B, Zhang Y, Gernapudi R, Duru N, Yao Y, Lo PK, et al. A High-Fat Diet Promotes mammary gland myofibroblast differentiation through microRNA 140 downregulation. *Mol Cell Biol*. 2017; 37: e00461-16.
31. Koledova Z, Zhang X, Streuli C, Clarke RB, Klein OD, Werb Z, et al. SPRY1 regulates mammary epithelial morphogenesis by modulating EGFR-dependent stromal paracrine signaling and ECM remodeling. *Proc Natl Acad Sci USA*. 2016; 113: E5731-40.
32. Inman JL, Robertson C, Mott JD, Bissell MJ. Mammary gland development: cell fate specification, stem cells and the microenvironment. *Development*. 2015; 142: 1028-42.
33. Macias H, Hinck L. Mammary gland development. *Wiley Interdiscip Rev Dev Biol*. 2012; 1: 533-57.
34. Pollak M. The insulin and insulin-like growth factor receptor family in neoplasia: an update. *Nat Rev Cancer*. 2012; 12: 159-69.
35. Chen WJ, Ho CC, Chang YL, Chen HY, Lin CA, Ling TY, et al. Cancer-associated fibroblasts regulate the plasticity of lung cancer stemness via paracrine signalling. *Nat Commun*. 2014; 5: 3472.
36. Xu WW, Li B, Guan XY, Chung SK, Wang Y, Yip YL, et al. Cancer cell-secreted IGF2 instigates fibroblasts and bone marrow-derived vascular progenitor cells to promote cancer progression. *Nat Commun*. 2017; 8: 14399.

Figures

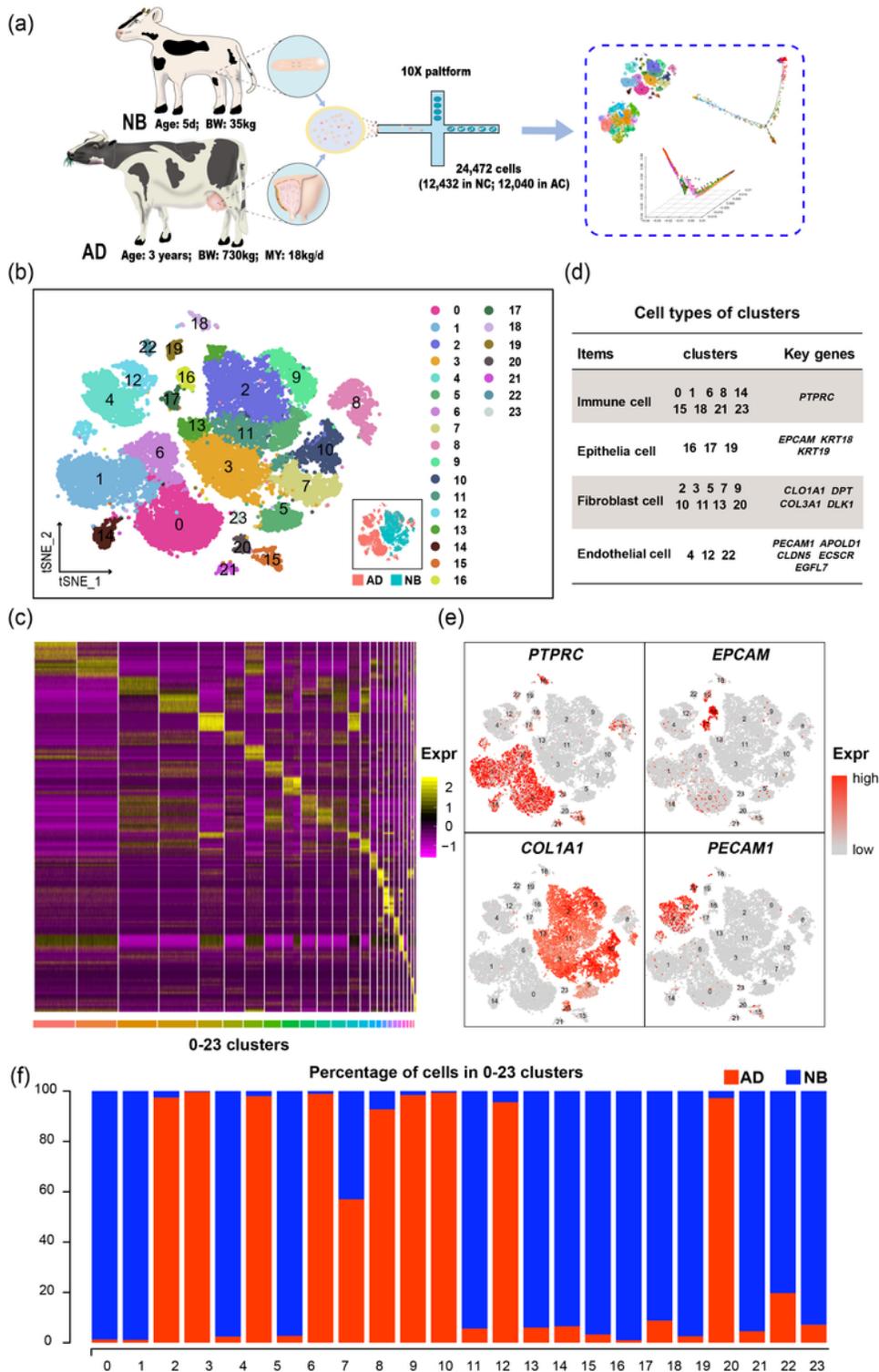


Figure 1

Single-cell atlas of dairy cow MG. (a) The scRNA-seq approach-based workflow of this study. The MG samples from NB and AD were collected and dissociated into a single-cell suspension, which was then subjected to cell capture, sequencing, and bioinformatics analyses using the 10 × chromium system. (b) T-SNE plot map of cell type clustering from the MGs of NB and AD. The cell type clusters are highlighted using different colors. (c) Heatmap displaying the scaled expression patterns of top marker genes of

each cell type cluster; yellow represents high expression of a gene, and purple indicates low expression. (d) The table of major cell types, their corresponding clusters, and marker genes. (e) T-SNE map of different cell type clusters highlighted by the expression of marker genes. PTPRC: immune cells; EPCAM: epithelial cells; COL1A1: fibroblasts; PECAM1: endothelial cells. (f) Percentage distribution of identified 24 clusters in AD and NB. BW, body weight; MY, milk yield; MG: mammary gland; AD: adult cow; NB: newborn cow; scRNA-seq, single cell RNA sequencing; T-SNE, t-distributed stochastic neighbor embedding.

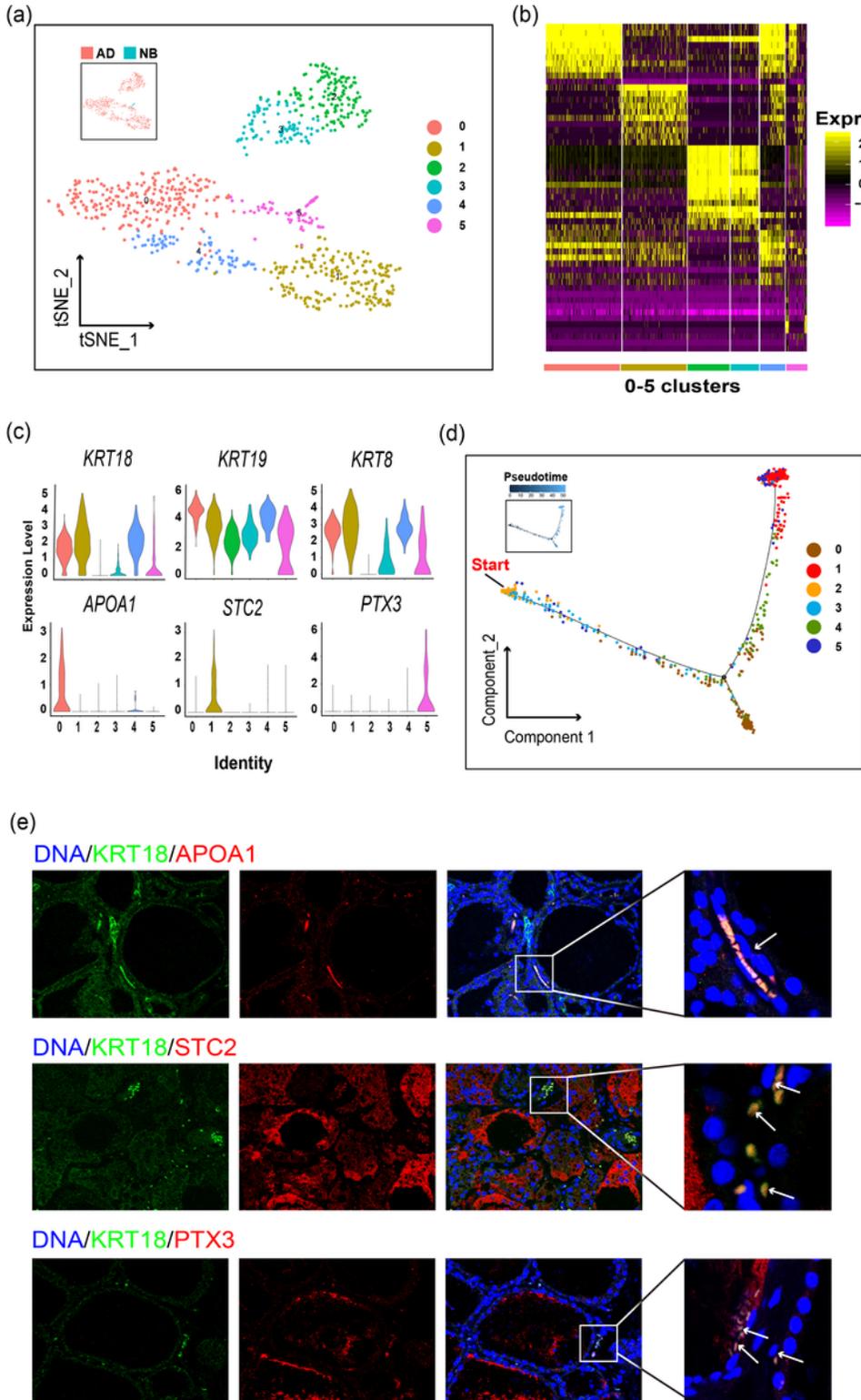


Figure 2

Identification of epithelial cell subtypes. (a) T-SNE plot map of epithelial cell subtypes clustering of NB and AD. (b) Heatmap displaying the scaled expression patterns of top marker genes of each cell subtype cluster; yellow represents high expression of a gene, and purple represents low expression. (c) Violin plots of epithelial marker genes KRT18, KRT19, and KRT8, the L0 marker gene APOA1, L1 marker gene STC2, and L5 marker gene PTX3. (d) The pseudotemporal trajectory of epithelial subtype cells from NB and AD. Color shows different designated cell states. The color in a gradient from dark to light blue indicates the pseudotime. (e) Immunofluorescence analysis of APOA1, STC2, and PTX3 protein expression (red) in combination with luminal epithelial marker KRT18 (green) and DNA stain using DAPI (blue) within tissue sections from AD mammary gland samples. AD, adult cow; NB, newborn cow; T-SNE, t-distributed stochastic neighbor embedding.

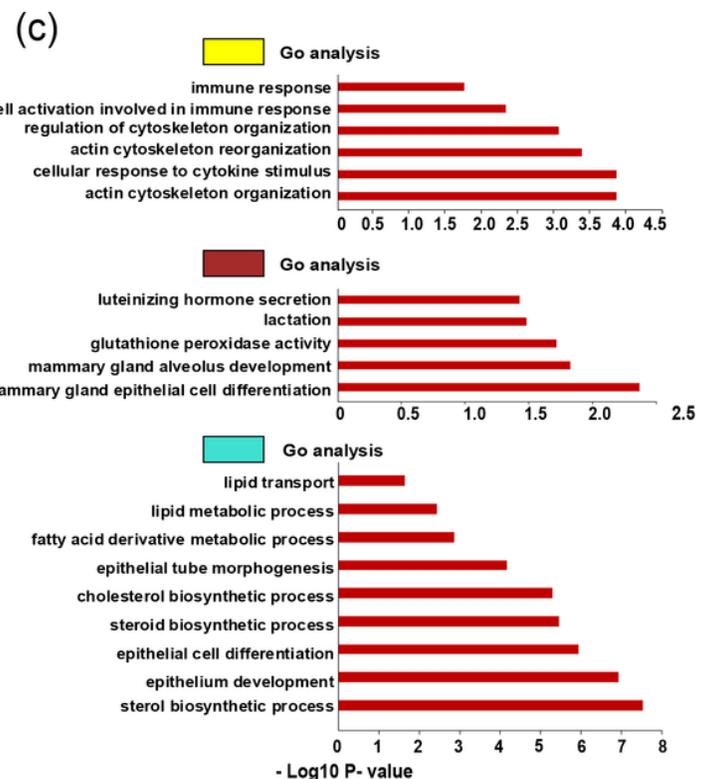
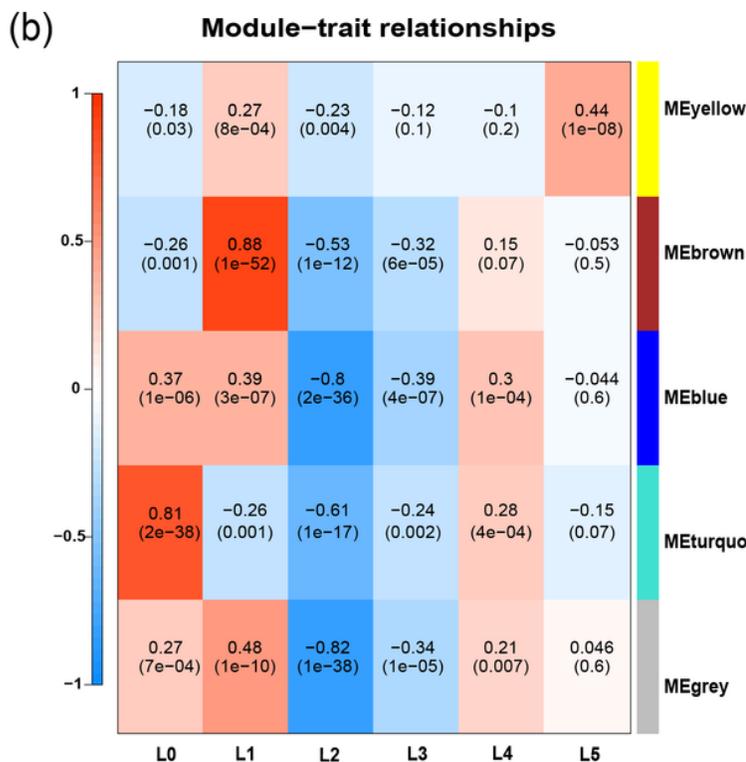
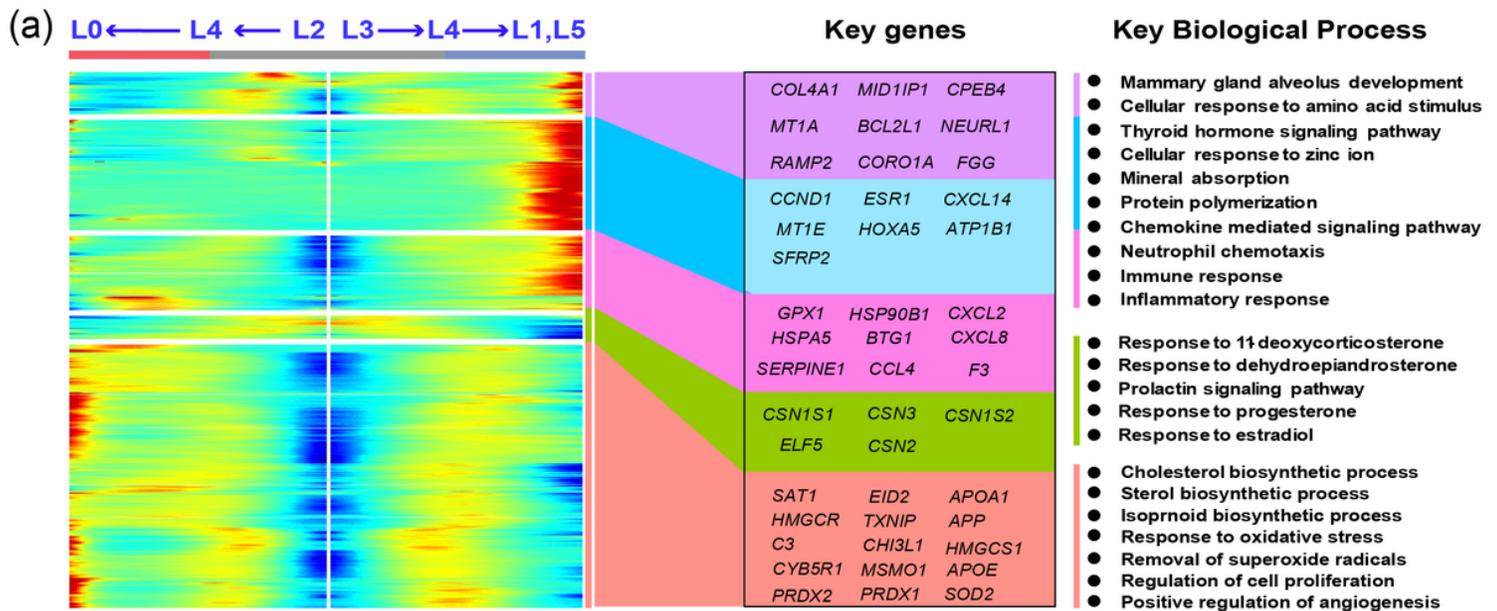


Figure 3

The characteristics of the pseudotime classification to identify genes associated with epithelial cell differentiation. (a) Heatmap of all genes with the branch-specific, pseudotime-dependent expression for the L0 or the L1 and L5 subtypes, key genes of each branch-specific gene cluster, and the enriched biological processes. (b) Heatmap of the correlation between co-expressed gene modules and different epithelial subtypes. The turquoise, brown, and yellow modules were most significantly positively correlated with the L0, L1, and L5 subtypes, respectively. (c) Gene ontology terms of member genes in the turquoise, brown, and yellow modules.

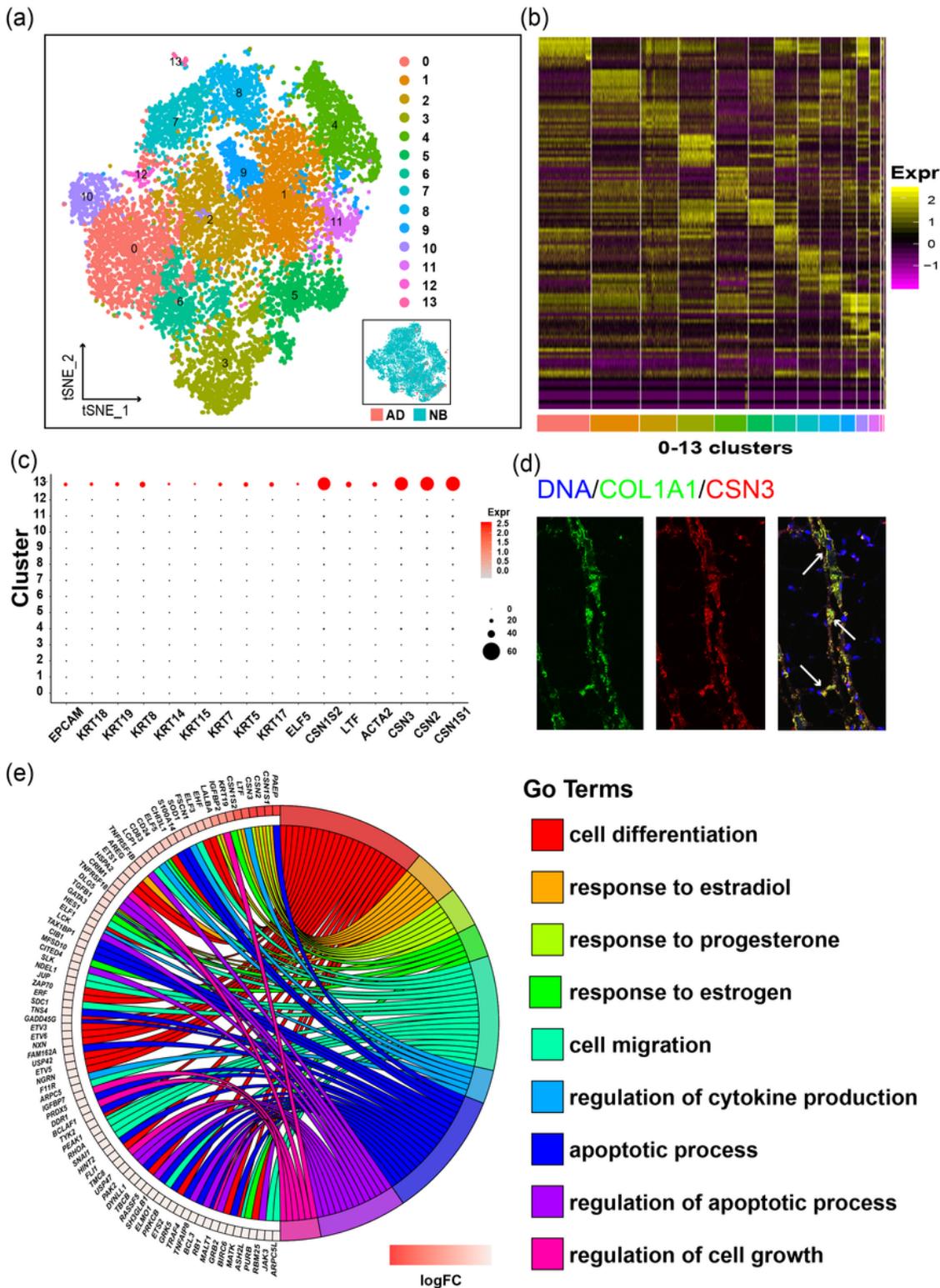


Figure 4

Identification of fibroblast cell subtypes and a novel epithelial-like fibroblast. (a) T-SNE plot map of fibroblast cell subtypes clustering of NB and AD. (b) Heatmap displaying the scaled expression patterns of top marker genes of each cell subtype cluster of fibroblast cells; yellow represents high expression of a gene, and purple represents low expression. (c) Scatterplot of the epithelial cell marker genes expressed in the fibroblast subtypes. (d) Immunofluorescence analysis of CSN3 protein expression (red) in

combination with fibroblast marker COL1A1 (green) and DNA stain using DAPI (blue) within tissue sections from the mammary gland samples of NB. (e) Gene ontology terms of marker genes in the novel epithelial-like fibroblast (F13). AD, adult cow; NB, newborn cow.

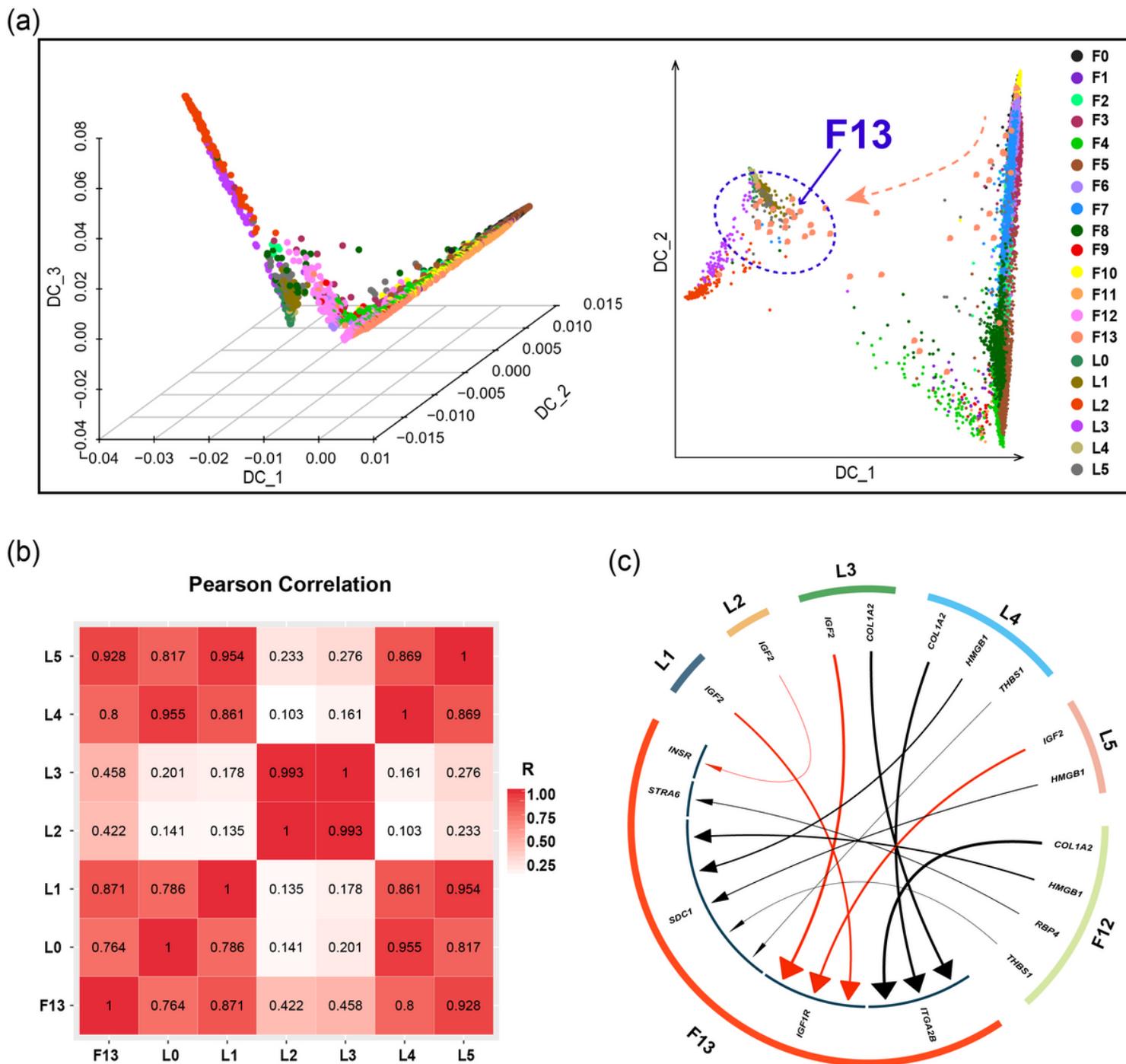


Figure 5

The association between fibroblast and luminal epithelial cells. (a) Diffusion map of epithelial and fibroblast cells from the newborn and adult cows. (b) Heatmap of Pearson correlation between fibroblast 13 (F13) and epithelial cell subtypes (L0–L5). (c) Ligands from fibroblasts and epithelial subtypes cells interact with the receptors of F13.

Supplementary Files

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

- [SupplementaryTables.xlsx](#)