Single-cell transcriptomics reveals dynamic cellular remodeling of porcine mammary gland during pregnancy

Si-Yu Yang[1,*], Tian-Xiong Yao[1,*], Lu-Sheng Huang[1,*]

[1]National Key Laboratory for Swine Genetic Improvement and Germplasm Innovation, Ministry of Science and Technology of China, Jiangxi Agricultural University, Nanchang 330045, China

[*]Corresponding authors: Si-Yu Yang (siyuyang1013@163.com); Tian-Xiong Yao (TianxiongYao@hotmail.com); Lu-Sheng Huang (lushenghuang@hotmail.com)

ABSTRACT

Pregnancy elicits extensive remodeling of the mammary gland to establish the alveolar network required for lactation, yet the cellular programs underlying this transition in pigs remain poorly resolved. Here, we generate a single-cell transcriptomic atlas of the porcine mammary gland across gilt, late pregnancy, and post-lactational stages. Differential expression analyses identify epithelial, fibroblast and endothelial compartments as the most dynamic during pregnancy, motivating higher-resolution analyses of these lineages. Epithelial trajectories toward alveolar fate exhibit induction of fibrillar collagen programs, including *COL1A1*, *COL3A1* and *COL5A1*, implicating epithelial contributions to extracellular matrix assembly. Fibroblast analysis revealed a PGLYRP1^{+} fibroblast subtype enriched for the chemokines *CCL4*, *CCL5* and *CCL21*, suggesting a role in immune cell recruitment. Endothelial cells undergo stage-specific metabolic reprogramming, with late pregnancy characterized by increased expression of glycolysis-associated genes, including *IGF1*, *FLCN* and *P2RX7*, as well as lipid metabolism-related genes such as *RBP7*, *KITLG* and *CD36*. Cross-species comparison reveals conserved cellular identities and key molecular features (*ESR1*, *AR*, *FOXA1*) across human, pig, and mouse mammary glands. Together, this atlas delineates the cellular logic underlying pregnancy-associated remodeling and serves as a resource for elucidating transcriptomic changes in the porcine mammary gland.

Keywords: Porcine; Mammary gland; Pregnancy; Single-cell transcriptomic; Remodeling

INTRODUCTION

Mammary gland development during pregnancy is a tightly regulated physiological process that ensures adequate milk production to support neonatal survival and growth. In mammals, this process involves extensive tissue remodeling, including epithelial expansion, stromal reorganization, and vascular adaptation (Buchholz et al., 2024; Feigman et al., 2020; Twigger et al., 2022). While rodent models have provided valuable insights into the hormonal and cellular drivers of pregnancy-induced mammary changes, the temporal and molecular dynamics of this process in large mammals remain less well understood. The porcine mammary gland closely resembles the human counterpart in both structure and composition, comprising terminal ductal lobular units (TDLUs), glandular epithelium, connective tissue, and adipose stroma (Nagy et al., 2021; Trott et al., 2022). It has therefore emerged as a valuable model for studying human breast development and disease. In the context of pig reproduction, colostrum and milk constitute the exclusive source of neonatal nutrition (Zhang et al., 2018), directly influencing key phenotypes such as immune maturation and early postnatal growth. As modern breeding practices have led to progressively larger litters, elucidating the cellular and molecular mechanisms underlying mammary gland function has become increasingly important for both biomedical and agricultural research.

In both humans and pigs, the mammary gland undergoes profound structural and functional changes across distinct reproductive stages (Hadadi et al., 2022; Hurley, 2019; Seachrist and Keri, 2019), including the gilt, late pregnancy, and post-lactational stages. During this progression, the mammary gland undergoes extensive remodeling, including proliferation, differentiation, and regression of epithelial cells (Gieniec and Davis, 2022), extracellular matrix remodeling mediated by fibroblasts through the deposition of type I, II, and V collagens (Tsutsui et al., 2020; Yamaguchi et al., 2024), and expansion of the vascular network to meet elevated energy demands that are tightly coordinated to support reproduction (Kutys et al., 2020). Rapid advances in single-cell technologies have enabled high-resolution delineation of transcriptional programs at

single-cell resolution, and these approaches are now widely used to interrogate the molecular architecture of mammalian cells (Rao et al., 2023; Yang et al., 2023). Previous studies have profiled transcriptional shifts in specific cell types such as adipocytes across reproductive stages in pigs (Fan et al., 2023). However, the broader transcriptional programs and cell type-specific adaptations in the porcine mammary gland remain poorly characterized, particularly in relation to pregnancy-associated remodeling. Given the physiological similarities between pigs and humans in terms of mammary gland architecture and reproductive endocrinology, elucidating these dynamics in pigs holds both scientific and translational value.

Here, we present a single-cell transcriptomic atlas of the porcine mammary gland across the gilt, late pregnancy and post-lactational stages. Across 14 major cell types in seven pigs, we identify pronounced remodeling of epithelial, fibroblast and endothelial compartments during late pregnancy. Building on this extensive remodeling, we further performed subtype-level analyses within these three lineages. Alveolar epithelial cells upregulated type I, III and V collagen synthesis genes involved in extracellular matrix remodeling in late pregnancy. A fibroblast subtype enriched for chemokine expression was predominantly detected in multiparous glands. Endothelial subtypes exhibited stage-specific metabolic reprogramming, with late pregnancy states showing enhanced vascular energy metabolism. Cross-species comparison revealed conserved mammary gland cell identities and transcriptional features across humans, pigs and mice, supporting the use of pigs as a physiologically relevant complementary model for human mammary gland biology and disease.

MATERIALS AND METHODS

Pig mammary gland tissue collection

Mammary gland tissues were collected from Large White pigs (*Sus scrofa domesticus*) across three reproductive stages. For each stage, biological replicates were obtained from different individuals. Specifically, samples were collected from gilts (days 241–243; *n*=3), late-pregnant sows (gestational day 110; days 485 and 488; *n*=2), and post-lactational sows (day 632; *n*=2). Animals were euthanized with isoflurane, in accordance with protocols approved by the Animal Ethics Committee of Jiangxi Agricultural University (JXAULL-2021-37). To minimize anatomical variation, tissue samples were excised from the mammary gland located near the sixth and seventh ribs. Fresh tissues were either enzymatically dissociated for single-cell transcriptomic analysis and fixed in 4% paraformaldehyde for histological evaluation.

Preparation of single-cell suspensions from mammary gland tissue

Fresh mammary gland tissues (~200 mg per sample) were finely minced on ice and transferred into digestion medium consisting of Dulbecco's Modified Eagle Medium (DMEM) supplemented with 100 mg/mL collagenase type II, 100 mg/mL collagenase type III, 1 mg/mL DNase I, 15uL 1mM $CaCl_2$ and 2% fetal bovine serum (FBS). Samples were incubated at 37°C for 40 min in a shaking water bath to ensure uniform enzymatic dissociation. During digestion, the tissue suspensions were gently agitated every 10 minutes to facilitate mechanical disruption. Enzymatic activity was quenched by adding an equal volume of ice-cold DMEM containing 10% FBS, followed by immediate transfer to ice. The digested material was filtered sequentially through 70-μm and 40-μm strainers to remove undigested fragments and large debris. The cell pellet was resuspended in 1 mL of red blood cell lysis buffer (BD Biosciences) and incubated for 5 min at room temperature to remove residual erythrocytes. The resulting single-cell suspensions were centrifuged and washed twice with Phosphate-buffered saline (PBS) containing 0.04% bovine serum albumin (BSA) to remove residual

enzymes. Cell viability was evaluated using the Cellometer Auto 2 000 Cell Viability Counter, and only cell suspensions with >85% viable cells were processed for scRNA-seq.

Single-cell library preparation and sequencing

Single-cell suspensions were prepared at a concentration of 2×10^5 cells/mL in PBS, with a target recovery of approximately 8 000 cells per sample. Cells were applied to microwells using the Singleron Matrix® Single Cell Processing System. Barcoding beads captured the mRNA from individual cells, which was subsequently reverse-transcribed to generate cDNA and amplified by PCR. The resulting cDNA was fragmented and ligated with sequencing adapters to construct libraries. Libraries for scRNA-seq were constructed using the GEXSCOPE single-cell RNA-seq library preparation kit (Singleron Biotechnologies) according to the manufacturer's protocol. Sequencing was performed on the Illumina NovaSeq 6 000 platform (USA) in paired-end 150 bp (PE150) mode, yielding a minimum of 20 000 reads per cell to ensure sufficient transcriptome coverage and data quality.

Preprocessing and quality control of single-cell RNA-seq data

Raw sequencing reads were aligned to the *Sus scrofa* 11.1 reference genome using the CeleScope pipeline (v.2.4.0), with gene annotations from Ensembl release 113. Initial quality control was performed using Scanpy (v.1.9.2; https://scanpy.readthedocs.io/) (Wolf et al., 2018), retaining cells with 400–5 000 detected genes, fewer than 30 000 unique molecular identifiers (UMIs), and mitochondrial gene content below 15%. Putative doublets were identified and removed using Scrublet (v.0.2.3; https://github.com/AllonKleinLab/scrublet) (Wolock et al., 2019) with a doublet score threshold of <0.25. Gene expression matrices were normalized using sc.pp.normalize_total function, followed by log transformation and identification of the top 2 000 highly variable genes. Principal component analysis (PCA) was then performed for dimensionality reduction. Batch effects across samples were corrected

using BBKNN (v.1.6.0; https://github.com/Teichlab/BBKNN) (Polański et al., 2020) within Scanpy, enabling integration of all cells for downstream clustering and subtype-level analyses.

Computation of gene specificity scores (τ) across cell types

To quantify cell type-specific gene expression, the gene specificity score (τ) was calculated across the 14 major cell types and individual subtypes identified in downstream analyses. As previously described (Yang et al., 2025; Yao et al., 2025), the score was computed using the following formula (Yanai et al., 2005):

$$\tau = \frac{\sum_{i=1}(1 - \hat{x}_i)}{\mathrm{n} - 1} \quad (1)$$

$$\hat{x}_i = \frac{x_i}{\max(x_i)} \quad (2)$$

where $\hat{x}_i$ denotes the average normalized expression of the gene in cell type *i*, and *n* is the total number of cell types. Higher τ values indicate stronger cell type specificity.

GO enrichment and pathway visualization

Gene Ontology (GO) enrichment analysis was performed using the ClueGO plugin (v.2.5.9; https://apps.cytoscape.org/apps/cluego) (Bindea et al., 2009) within Cytoscape (v.3.10.0; https://cytoscape.org/) (Shannon et al., 2003), based on cell type-specific highly expressed genes. Statistical significance was assessed using the Benjamini-Hochberg correction, and pathways with an adjusted $P < 0.05$ were retained. Enriched GO terms were visualized using the ComplexHeatmap package (v.2.16.0; https://bioconductor.org/packages/ComplexHeatmap/) (Gu et al., 2016) in R (v.4.3.0).

Differential gene expression analysis

Differentially expressed genes (DEGs) were identified using the Wilcoxon rank-sum test implemented in Scanpy. Comparisons were performed between late-pregnancy and gilt stages, as well as between late-pregnancy and post-lactational stages. Genes with an adjusted $P < 0.05$ and $\log_2$ fold change >1 were defined as upregulated in late pregnancy, whereas those with an adjusted $P < 0.05$ and $\log_2$ fold change <−1 were

considered downregulated. The intersection of genes upregulated in both comparisons defined a gene set associated with late pregnancy, representing transcripts transiently elevated in late pregnancy and downregulated in the post-lactational stage. Similarly, genes downregulated in both comparisons were defined as transiently suppressed in late pregnancy and restored in the post-lactational stage.

Pseudotime trajectory analysis

Pseudotime trajectory analysis was performed using the Monocle3 (v.1.0.0; https://cole-trapnell-lab.github.io/monocle3/) (Qiu et al., 2017) package in R to reconstruct the transcriptional continuum from luminal progenitor cells to alveolar cells. Cells were ordered along the inferred trajectory using the learn_graph and order_cells functions, with luminal progenitor cells specified as the root node. To identify genes dynamically regulated along pseudotime, graph-autocorrelation statistics were calculated using the graph_test function, and genes with a Q value <0.01 were retained. Smoothing splines were fitted to the expression profiles of significant genes to capture transcriptional dynamics across the trajectory. The top 10 000 pseudotime-associated genes were visualized using the ComplexHeatmap package (v.2.16.0) in R (v.4.3.0).

Cross-species comparative analysis

Cross-species comparative analyses were performed to assess transcriptomic conservation among human, porcine, and murine mammary glands. Publicly available scRNA-seq datasets for human and mouse mammary glands were retrieved (Bhat-Nakshatri et al., 2024; Li et al., 2020). To control for differences in cellular composition and sampling depth, cells were first subsampled based on matched cell identities, with 1 000 cells randomly selected per cell type from each species. Genes were then restricted to one-to-one orthologs shared across the three species to ensure cross-species comparability. Data normalization and feature selection were performed using the Seurat R package (v4.0.5; https://satijalab.org/seurat/). Each dataset was independently normalized using the NormalizeData function, after which the top 3 000 highly variable

genes were identified using FindVariableFeatures function. Cross-species integration anchors were subsequently determined with FindIntegrationAnchors function, and the datasets were integrated using IntegrateData function to enable systematic transcriptome-level comparisons across corresponding cell populations.

Cross-species conservation analysis

As previously described (Xiao et al., 2025), AUROC scores were computed using the MetaNeighbor (v.1.20.0) (Crow et al., 2018) package to quantify transcriptomic similarity across species. Cell type-specific genes were defined per species (τ>0.85; detected in >5% of cells). Conserved cell type-specific genes were obtained by intersecting the three species-specific gene sets for each corresponding cell type.

Cross-species transcriptomic correlation analysis

Highly variable genes were identified for each species using the Seurat function FindVariableFeatures (selection.method="vst", nfeatures=6 000). These 6 000 variable genes from each dataset were retained, and the intersection across the three species was extracted to construct comparable gene expression matrices. Pairwise transcriptomic similarity between species was then quantified using Spearman's correlation with false discovery rate (FDR) adjustment, as implemented in the corr.test function of the psych R package (v. 4.3.3).

Histological staining of mammary gland tissue

Mammary gland tissues from all seven individuals across the gilt, late-pregnancy, and post-lactational stages were subjected to haematoxylin and eosin (H&E) staining. Fresh samples were fixed in 4% paraformaldehyde at 4 °C for 24 hours, dehydrated through a graded ethanol series, cleared in xylene, and embedded in paraffin. Sections were cut at a thickness of 5 μm, mounted on glass slides, and stained with H&E using standard protocols. Whole-slide images were acquired using the Slideview scanning system (OLYMPUS) and visualized for histomorphological assessment across developmental

stages.

Multiplex immunofluorescence (MIF)

To validate the presence of alveolar epithelial cells expressing collagen genes and fibroblasts expressing chemokines, multiplex immunofluorescence was performed on formalin-fixed paraffin-embedded (FFPE) mammary gland tissue sections. Sections were deparaffinized in xylene, rehydrated through graded ethanol, and subjected to antigen retrieval by boiling in Tris-EDTA buffer (95 °C). Endogenous peroxidase activity was quenched by incubation in 3% hydrogen peroxide for 10 min. Sequential rounds of primary and secondary antibody incubation were performed, with each immunofluorescent signal amplified using the Think Color Fluorescent mIHC Kit and a fluorophore-conjugated tyramide signal amplification (TSA) system (Free-Thinking Biosciences). Between rounds, epitope retrieval and protein blocking were repeated to enable multiplex labeling. Nuclei were counterstained with 4’,6-diamidino-2-phenylindole (DAPI). Slides were scanned using a Digital Slice Scanner (3DHISTECH, Hungary) and analyzed with Slide Viewer software (3DHISTECH, Hungary). Primary antibodies and corresponding catalog number, host species, clone number, and dilution ratios are listed in Supplementary Table S1.

RESULTS

Dynamic single-cell transcriptomic landscape of the mammary gland during pregnancy

Mammary gland samples were collected from sows at three physiological stages: gilt, late pregnancy, and post-lactational (Figure 1A; Supplementary Table S2). This comparative design enables the investigation of this complex and dynamic process of the mammary gland remodeling during pregnancy, as it is not possible to obtain healthy human mammary gland samples. Histological analysis via H&E staining revealed distinct morphological changes across stages (Figure 1B–D). Mammary glands from gilts exhibited well-defined ductal networks, with terminal end buds (TEBs) forming proliferative, club-shaped structures embedded in adipose and stromal tissue (Figure 1B). By late pregnancy, alveolar maturation was nearly complete, evidenced by luminal milk accumulation (Figure 1C). In post-lactational stage, the lobuloalveolar structures regressing and the extracellular matrix remodeling (Figure 1D).

Subsequently, we performed scRNA-seq on the same seven samples. After quality control, we obtained a total of 58 940 high-quality cells, with an average of 1 058 genes detected per cell, reflecting robust transcriptome coverage suitable for comprehensive downstream analyses. Clustering analysis using the Leiden algorithm (Traag et al., 2019) identified 36 clusters (Supplementary Figure S1A), which were annotated according to canonical marker gene expression and named following the latest nomenclature standards (Gray et al., 2025) . In total, we defined 14 major cell types (Figure 1E, F; Supplementary Figure S1B; Supplementary Table S3). These included three epithelial subtypes, luminal adaptive secretory precursors (LASP), luminal hormone-sensing (LHS) cells, and basal-myoepithelial (B-Myo) cells. We also identified four lymphoid lineages, including T cells, natural killer (NK) cells, B cells, and plasma cells (PC), as well as two myeloid lineages, macrophages (Macro) and neutrophils (Neu). In addition, five vascular-associated populations were characterized, including fibroblasts (Fibro), pericytes (Peric), smooth muscle cells (SMC), vascular endothelial cells (VEC), and lymphatic endothelial cells (LEC). To assess the accuracy of our annotation, we

compared the identified cell types with reference human mammary gland datasets (Bhat-Nakshatri et al., 2024; Gray et al., 2022). The annotated cell populations exhibited high concordance with their human counterparts (Supplementary Figure S1C, D), confirming the robustness of our classification. Consistent with histological observations, LASP were predominantly derived from late-pregnancy samples (Figure 1G; Supplementary Figure S1E, F), whereas LHS were mainly enriched in the gilt and post-lactational stages. B-Myo cells were most abundant in the gilt stage, suggesting their critical role in early ductal morphogenesis and structural maintenance.

Additionally, we calculated cell-type specificity τ scores and identified 2 474 genes with cell type-specific expression across the major cell populations (Figure 1H). Many of these genes corresponded to canonical cell-type markers. For example, the LASP highly expressed the milk protein genes *CSN3*, *CSN2*, and *CSN1S1* (Figure 1H), which constitute the principal components of milk, indicating that the mammary epithelium had entered a secretory maturation phase. More than half of the identified genes have not been widely reported, providing a valuable resource for future studies. For instance, *PGLYRP1*, specifically enriched in LASP cells (Figure 1H), encodes a peptidoglycan recognition protein previously implicated in antibacterial defense of the mammary gland by sensing bacterial components and activating innate immune responses (Yao et al., 2013). In B-Myo cells, in addition to the classical marker *KRT5*, we found specific expression of *OXTR* (Figure 1H), which encodes the oxytocin receptor that mediates myoepithelial contraction in response to oxytocin, facilitating milk ejection during lactation (Hawley et al., 2018). Furthermore, GO enrichment of cell type-specific genes revealed functions aligned with the expected biological roles of each cell type. LASP-specific genes were enriched for lactation processes, whereas LHS-specific genes were associated with hormone regulation (Supplementary Figure S1G). Together, these findings establish a robust catalogue of cell type-specific genes that not only validates our annotations but also highlights previously unrecognized markers with functional relevance to mammary gland biology.

Dynamic changes in gene expression and pathways during pregnancy

To characterize gene expression dynamics in the mammary gland during pregnancy, we performed differential expression analyses within each main cell type, comparing late pregnancy to both gilt (LP vs GLT) and post-lactational (LP vs PL) stages. These analyses revealed that the number of DEGs significantly downregulated during pregnancy substantially exceeded those upregulated (Figure 2A; Supplementary Table S4). Among these cell types, LHS, Fibro and LASP exhibited the highest number of genes upregulated in late pregnancy, whereas LASP, VEC and Fibro showed the greatest number of downregulated genes (Figure 2A). These findings highlight the critical roles of epithelial, endothelial and fibroblast populations in the mammary gland remodeling associated with pregnancy.

We identified a subset of genes displaying reversible expression patterns by intersecting DEGs across LP vs GLT and LP vs PL comparisons in each cell type. These included genes upregulated in late pregnancy that returned to baseline at the post-lactational stage, and genes downregulated in late pregnancy that were subsequently restored (Figure 2B). Such temporally regulated genes may contribute to the physiological adaptations specific to the late pregnancy stage. This analysis identified 1 686 genes upregulated and 2 429 genes downregulated in late pregnancy, including 671 upregulated and 1 097 downregulated genes shared across multiple cell types (Figure 2B). Among the genes upregulated across multiple cell types were canonical milk proteins (*CSN3*, *CSN2*, *CSN1S1*), extracellular matrix components (*COL4A1*, *MGP*, *POSTN*, *SCRG1*), immune-related secreted proteins (*PGLYRP1*, *JCHAIN*, *MGST1*), and cell fate regulators (*FOXP2*, *GFRA1*) (Figure 2C), reflecting preparatory changes for lactation and extensive tissue remodeling. Conversely, genes downregulated across cell types included inflammatory-response proteins (*LYZ*, *CXCL14*, *BCL2A1*, *SPP1*, *CTSZ*, *CTSD*) and stress-adaptive metabolic proteins (*SQSTM1*, *GPX1*, *APOE*, *ADIRF*). Notably, the concurrent upregulation of antimicrobial defense genes (*PGLYRP1*, *JCHAIN*, *MGST1*) and suppression of inflammatory mediators (*LYZ*, *CXCL14*, *SPP1*, *BCL2A1*, *CTSZ*, *CTSD*) suggest a transcriptional reprogramming of the mammary immune environment during late

pregnancy, reflecting a transition from an inflammation-driven remodeling phase to a tolerant yet protective state that preserves alveolar integrity while establishing antimicrobial readiness before lactation.

To further explore the functional implications of pregnancy-associated transcriptional changes, we conducted GO enrichment analysis of genes upregulated and downregulated in late pregnancy across each cell types. This analysis revealed cell type-shared biological pathways, with genes upregulated in late pregnancy enriched for lactation, collagen fibril organization, oxidative phosphorylation, indicating that the mammary gland undergoes coordinated metabolic activation and extracellular matrix remodeling to support milk production (Figure 2D). Beyond these conserved responses, distinct cell type-specific signatures were also identified. In particular, fibroblasts showed selective activation of the antibacterial humoral response pathway (Figure 2D), indicating that the mammary stroma acquires heightened antimicrobial capacity during late pregnancy. However, genes downregulated in Fibro, T cells, NK cells, Neu, B cell, PC, and LEC in late pregnancy were commonly enriched in the antigen processing and presentation of peptide antigen pathway (Figure 2E), indicating a shift toward reduced antigen presentation, and enhanced immune tolerance within the mammary microenvironment.

Transcriptional dynamics of alveolar differentiation reveal coordinated extracellular matrix remodeling

Significant transcriptional changes were observed in mammary epithelial cells during pregnancy (Feigman et al., 2020). To investigate the molecular transitions associated with these stages, we analyzed 9 631 epithelial cells consisting of LASP and LHS. This analysis identified six transcriptionally distinct epithelial subtypes (Figure 3A). Among these, alveolar cells (ALV), predominantly derived from late-pregnant sows, were characterized by upregulation expression of milk biosynthesis-related genes such as *MFGE8*, *CSN3*, *ELF5*, and *LTF* (Figure 3B). A similar gene expression signature was observed in pre-alveolar luminal progenitors (pre-ALV), primarily originating from gilt and post-lactational stages (Figure 3C). We further identified luminal hormone-sensing-like cells (LHS-like), characterized by high expression of secretory proteins such as *SCGB1D1*, *CRISP3*, and *CHIA*, and hormone-sensing luminal cells (LHS), enriched for estrogen receptor *ESR1*. In addition, two progenitor-like subpopulations were delineated: progenitor-like luminal epithelial cells (LPC-like), marked by elevated expression of progenitor-like luminal epithelial markers, including *EGFR*, *TCF7L2*, and *KIT*; and luminal progenitor cells (LPC), defined by high expression levels of canonical luminal progenitor markers such as *KIT*, *ALDH1A3*, and *CD14*.

Six transcriptionally distinct epithelial subtypes exhibited diverse functional specializations. ALV showed enrichment for genes involved in acyl-CoA metabolism, amino acid response, and collagen fibril organization (Supplementary Figure S2A), consistent with their established roles in milk biosynthesis and extracellular matrix remodeling in late pregnancy. Pre-ALV upregulated genes associated with cytoplasmic translation and intrinsic apoptotic signaling pathways (Supplementary Figure S2A), suggesting a transitional state poised between proliferative expansion and lineage specification. LHS were enriched for transcripts linked to mammary gland development and estrogen receptor signaling (Supplementary Figure S2A), reflecting their role in mediating endocrine cues. The two progenitor-like populations exhibited distinct signaling profiles. LPC-like cells showed elevated expression of genes related to phosphatase regulatory activity and proteoglycan metabolism (Supplementary Figure

S2A). In contrast, LPC were enriched for genes involved in inflammatory and tissue-repair programs (Figure 3D, E; Supplementary Figure S2A), including response to wounding (*CCL2*, *IL34*, *KLF5*, *THBS1*), macrophage chemotaxis (*ANXA1*, *CD34*, *MYLK*, *SERPINE2*, *THBS1*, *TNFRSF12A*) and TGF-β signaling (*CD200*, *CD34*, *THBS1*) (Figure 3E), suggesting a potential role in epithelial regeneration and immune modulation.

To investigate the transcriptional dynamics that underlie the regenerative capacity of luminal progenitor cells and their differentiation toward alveolar epithelial fate, we performed pseudotime trajectory analysis using Monocle3 on LPC, pre-ALV, and ALV to reconstruct the LPC-to-ALV axis (Figure 3F). This analysis revealed 11 444 genes (Q<0.01) associated with this transition (Supplementary Table S5), encompassing 1 487 transcription factors, 578 secreted proteins, and 430 membrane-associated proteins (Supplementary Table S6). These genes were organized into ten distinct functional modules (Figure 3G), each exhibiting marked transcriptional shifts along pseudotime (Figure 3H). Cells from LPC displayed heightened metabolic activity and evidence of mitochondrial remodeling (Figure 3H), consistent with a progenitor state primed for bioenergetic demands. As cells transitioned into the pre-ALV phase, expression of antigen presentation genes such as *CD74*, *CTSD*, *CTSS* and *HLA-DRA* became prominent (Supplementary Figure S2B, C), suggesting a transient phase of immune communication and hormonal sensitivity. Fully differentiated ALV cells were characterized by the upregulation of genes associated with lipid storage, including *LPL*, *PLIN2*, and *SREBF1* (Supplementary Figure S2D, E), reflecting their functional specialization in milk production and the establishment of alveolar architecture in response to increased energetic demand. In addition, these cells exhibited elevated expression of extracellular matrix-related genes such as *COL1A1*, *COL1A2*, *COL3A1*, *COL5A2*, *COL5A3*, and *LUM* (Figure 3I, J). Immunofluorescence analysis further confirmed the colocalization of COL1A1, COL3A1, and COL5A1 with ALV cells (Figure 3K–M), underscoring their structural role in alveolar formation.

Stage-specific fibroblast heterogeneity drives immune remodeling of the mammary gland

Fibroblasts play an active role in tissue remodeling during pregnancy, contributing to collagen fiber synthesis and organization within the mammary gland (Yamaguchi et al., 2024). To characterize their cellular heterogeneity, we performed subclustering analysis on 10 172 fibroblasts and identified five transcriptionally distinct subtypes (Figure 4A), each defined by the dominant subtype-specific genes (Figure 4B). These subtypes displayed pronounced heterogeneity across three physiological stages (Figure 4C), with MDK^+ fibroblasts (Fibro), $LTBP2^+$ Fibro, and $RPGRIP1^+$ Fibro were primarily derived from gilt mammary glands, whereas $GREM^+$ Fibro were enriched in post-lactational samples and $PGLYRP1^+$ Fibro were predominantly present in late pregnancy.

We next performed GO enrichment analysis based on subtype-specific expression gene (Figure 4D). Beyond their shared roles in extracellular matrix organization and collagen biosynthesis (Supplementary Figure S3A), fibroblasts enriched in the gilt mammary gland exhibited distinct transcriptional programs (Figure 4D). The MDK^+ Fibro, marked by elevated expression of *CXCL10*, *IFIT1* and *MX1*, were associated with antiviral innate immune responses (Figure 4E), while co-expression of *ACTA2*, *EMILIN1*, *F3*, *HBEGF*, *MDK*, *RGMA* and *SERPINE2* suggested additional involvement in response to wounding (Figure 4E). These findings indicate that this population may contribute to immune surveillance and regenerative processes. The $LTBP2^+$ Fibro, characterized by high expression of *ASPN*, *CDKN1C*, *FMOD*, *FOLR2*, *ID1*, *LTBP2*, *PMEPA1* and *TGFB3*, were enriched for TGF-β signaling components (Supplementary Figure S3B), implicating a role in stromal regulatory signaling. The $RPGRIP1^+$ Fibro expressed genes such as *ABL1*, *DAAM1*, *PRICKLE1*, *ROR1*, *ROR2* and *RPGRIP1L*, which are involved in epithelial morphogenesis (Supplementary Figure S3C). This finding points to a potential function in structural support and epithelial remodeling.

In contrast, fibroblasts enriched in late pregnancy and post-lactational stages exhibited transcriptional features indicative of dual roles in structural remodeling and immune regulation. The $GREM1^+$ Fibro upregulated genes involved in branching

morphogenesis (Supplementary Figure S3D), including *AGT*, *FGF7* and *NTN4*, as well as genes regulating fatty acid metabolism such as *ACADL*, *FMO2*, *GHSR* and *PDK4* (Supplementary Figure S3D), suggesting a role in ductal patterning and adipose remodeling. PGLYRP1^{+} fibroblasts were enriched for genes associated with antimicrobial humoral responses (Figure 4F), such as *GNLY*, *JCHAIN*, *LYZ*, *PGLYRP1* and *S100A12*, supporting a specialized role in shaping the local immune microenvironment in late pregnancy and mammary gland regression. In addition, this fibroblast subset exhibited upregulation of genes involved in granulocyte chemotaxis and robust expression of pro-inflammatory chemokines (Figure 4D, G), such as *CCL4*, *CCL5*, *CXCL2*, *CCL21*, *CXCL8* and *CXCL14*, suggesting active participation in immune signaling. Immunofluorescence co-localization further confirmed the presence of fibroblast subpopulations expressing CCL5 (Figure 4H).

Endothelial metabolic remodeling across mammary developmental stages

The maternal vascular system undergoes extensive remodeling during late gestation to meet the increasing metabolic demands associated with lactation (Hardy et al., 2021). To characterize the endothelial landscape underlying this adaptation, we performed transcriptomic profiling of 3 256 endothelial cells collected across gilt, late-pregnancy, and post-lactational stages. This analysis identified six transcriptionally and functionally distinct endothelial subtypes (Figure 5A), comprising arterial (Art) endothelial cells marked by high expression of *SEMA3G*, three capillary (cap) clusters characterized by elevated expression of *EDNRB* and *SPARCL1*, and two venous (Vein) subtypes defined by *ACKR1* expression (Figure 5B).

GO enrichment analysis of subtype-specific transcriptional profiles revealed functional diversity across endothelial states (Figure 5C). The Cap1 subtype, predominantly observed in the gilt stage (Figure 5D), exhibited upregulated expression of *AGTR1*, *AVPR1A*, *CCN1*, *PDGFRB*, *ANGPTL4*, *CCL19*, *SOCS1*, *SPHK1*, *FADS1*, *FADS3*, *PDGFA* and *SNAI2* (Figure 5E), genes implicated in the regulation of phospholipase activity and lipid signaling, suggesting a preparatory role in maintaining phospholipid homeostasis. The Vein2 subtype exhibited elevated expression of *ABCA1*, *APOE, GPNMB, PLTP* and *SPP1* (Figure 5F), which are involved in lipid transport (Sprenger et al., 2025; Wang et al., 2025), along with *ASAH1*, *SMPDL3A* and *SMPDL3B*, key regulators of sphingolipid metabolism (Jang et al., 2018), consistent with vascular specialization for nutrient delivery during mammary gland repurposing. In contrast, Cap3 cells, which predominated in late pregnancy, exhibited elevated expression of genes such as *FLCN*, *IGF1* and *P2RX7* implicated in the regulation of glycolytic processes (Figure 5G). Analysis of Cap3-specific genes further revealed upregulation of genes involved in both lipid and glycolytic metabolism (Figure 5H), suggesting augmented metabolic regulation to accommodate the heightened energy requirements of late pregnancy. Collectively, these findings delineate stage-specific metabolic specialization among endothelial subtypes, reflecting coordinated energy adaptation during mammary remodeling.

Cross-species conservation of mammary gland transcriptomes among human, pig and mouse

To systematically compare mammary gland cell types across species, we integrated single-cell transcriptomic datasets from humans (Bhat-Nakshatri et al., 2024), pigs, and mice (Li et al., 2020) using an anchor-based mutual nearest neighbor alignment strategy. This cross-species integration revealed that homologous cell types clustered together in the human, pig and mouse mammary gland (Figure 6A, B). Cross-species similarity quantified with MetaNeighbor (Crow et al., 2018) showed uniformly high conservation across corresponding cell types, with AUROC values≥0.91 (Figure 6C; Supplementary Table S7). To further characterize conserved molecular features, we defined cell type-specific genes independently in each species and derived a set of 405 genes shared across all three species (Supplementary Table S8), including canonical markers such as *KRT5* (B-Myo), *ESR1* (LHS), *CSN2*/*CSN3* (LASP), *COL1A1*/*COL3A1* (Fibro), *C1QA*/*C1QB* (Macro), *SKAP1* (T cells), *VWF* (VEC), and *FLT4* (LEC) (Figure 6D). Among these, 34 transcription factors exhibited conserved, cell type-specific expression (Figure 6D). In LHS, *FOXA1*, *ESR1* and *AR* were all conserved across species, with FOXA1 acting as a pioneer factor that facilitates ER (encoded by *ESR1*) and AR recruitment to their regulatory targets (Fu et al., 2023). In LASP, ELF5 and EHF governed epithelial lineage specification and barrier integrity (Nightingale et al., 2024). In endothelial populations, MEOX2 and MEOX1 demarcated vascular and lymphatic compartments (Li et al., 2024; Wang et al., 2022), respectively (Figure 6D). Collectively, these analyses demonstrate conservation of mammary cell identities and regulatory programs across human, pig and mouse.

Given the strong transcriptomic concordance among human, pig, and mouse mammary glands, we next performed pairwise correlation analyses to assess relative similarity across species. Human and pig cell types consistently exhibited higher transcriptomic concordance than their human-mouse counterparts (Figure 6E, F). For instance, B-Myo cells, which mediate milk ejection through contractile activity and play essential roles in mammary development and maintenance (Gieniec and Davis, 2022), showed a higher correlation between pigs and humans (r=0.55) than between

mice and humans (r=0.46). Similarly, fibroblasts, which provide structural support to the gland, facilitate ductal morphogenesis, and contribute to tissue homeostasis and remodeling (Pascual et al., 2025), displayed a stronger pig-human correlation (r=0.62) compared to mouse-human (r=0.53) (Figure 6E, F). To determine whether these cross-species relationships also apply to clinically relevant loci, we next examined the expression patterns of a curated set of breast cancer risk genes (Consortium, 2021) (Figure 6G; Supplementary Table S9). These risk genes were predominantly expressed in human mammary epithelial cells, with most showing no marked cell type specificity. Notably, *CDH1*, a member of the cadherin family, exhibited high expression across epithelial populations in all three species (Figure 6G). Given that germline mutations in *CDH1* confer up to a 55% lifetime risk of developing breast cancer in women (Gamble et al., 2023), this conserved expression underscores its fundamental role in maintaining epithelial integrity and its relevance to breast cancer susceptibility.

DISCUSSION

In this study, we present a single-cell transcriptomic atlas of the porcine mammary gland across three key reproductive stages, providing a comprehensive view of its cellular and molecular remodeling during pregnancy. Transcriptomic profiling revealed extensive transcriptional reprogramming in epithelial, endothelial, and fibroblast compartments. Subtype-level analyses further uncovered pronounced heterogeneity within each lineage. Epithelial cells exhibited active participation in extracellular matrix remodeling during late pregnancy, while endothelial subtypes displayed stage-specific energy metabolic specializations associated with mammary reorganization. A distinct fibroblast population enriched for chemokine expression emerged predominantly in multiparous glands, suggesting an immunomodulatory role. Cross-species comparison revealed conserved transcriptional programs in the mammary gland shared across humans, pigs and mice.

Through systematic differential expression analysis, we established a comprehensive catalogue of pregnancy-associated genes and pathways, offering a global view of the physiological adaptations shaping the mammary gland during pregnancy. Our analysis revealed that epithelial, endothelial, and fibroblast compartments exhibited pronounced transcriptional changes during pregnancy. Notably, we observed a transition of the mammary gland from an inflammation-driven remodeling phase to a tolerant yet protective immune state during late pregnancy. Specifically, this immune functional shift was reflected by the downregulation of inflammatory genes (*LYZ*, *CXCL14*, and *SPP1*) alongside the upregulation of antimicrobial humoral immunity genes (*JCHAIN* and *PGLYRP1*) during late pregnancy. Such transcriptional remodeling likely facilitates tissue preservation and preparation for lactation by minimizing inflammatory damage while maintaining defense against microbial invasion, consistent with the establishment of an immune-tolerant microenvironment characterized by enhanced secretory IgA pathways and dampened inflammatory responses in human and murine mammary glands during late gestation (Donald et al., 2022; Niimi et al., 2018).

Our subtype analysis of epithelial cells delineates the transcriptomic trajectory of alveolar cell differentiation from luminal progenitors, characterizing concomitant shifts in gene expression and pathway activity. This transition was marked by a decline in immune-related transcriptional programs and a concomitant induction of collagen fibril organization genes, including *COL1A1*, *COL1A2*, *COL3A1*, *COL5A2*, and *COL5A3*. These findings are consistent with previous reports identifying type I, III, and V collagens, rather than type VI, VII, XIV, or XV, as the principal fibrillar components of the mammary extracellular matrix (Tsutsui et al., 2020). In rodents and humans, collagen composition and structure in the mammary gland are dynamically regulated across reproductive stages to accommodate shifting functional demands (Guo et al., 2022; Tsutsui et al., 2020). Although extracellular matrix synthesis has long been ascribed to stromal fibroblasts (Boeker and Kalluri, 2025), our data reveal a potential epithelial contribution to collagen deposition during alveolar morphogenesis. Although epithelial involvement in collagen biosynthesis has been described (Luo et al., 2024; Sawhney, 1993), direct evidence for alveolar cells expressing fibrillar collagens I, III, and V to support structural remodeling remains lacking, highlighting an overlooked facet of epithelial-mediated matrix dynamics.

Subtype analysis of endothelial cells revealed enhanced regulation of energy metabolism, indicative of adaptive remodeling to meet the physiological demands of late pregnancy. Both Cap1 and Vein2 subtypes were functionally linked to lipid metabolic pathways, reflecting coordinated endothelial adaptation to dynamic lipid and energy requirements across reproductive stages. In contrast, Cap3 exhibited additional upregulation of glycolysis-associated genes, including *IGF1*, *FLCN*, *P2RX7*, and *SNRK*, suggesting an intensified metabolic shift to support the elevated energy demands of the mammary gland in late pregnancy. This transition parallels observations in human and dairy cows (Armistead et al., 2020; Piantoni and Vandehaar, 2022), in which mammary tissue undergoes pronounced metabolic stress during late gestation. Notably, the mammary capillary endothelium forms the principal semipermeable interface controlling the exchange of oxygen, CO_2, solutes and macromolecules required for cellular energy metabolism (Ryman et al., 2015), implying that metabolic

reprogramming of endothelial subtypes is likely to directly shape substrate availability. Collectively, these findings indicate that endothelial metabolic plasticity may constitute a key regulatory layer in orchestrating nutrient delivery and tissue preparation for milk synthesis prior to parturition.

These findings provide a comprehensive view of the cellular and transcriptional dynamics underlying mammary gland adaptation during pregnancy. By resolving distinct epithelial, fibroblast, and endothelial states, our study highlights stage-specific regulatory programs that orchestrate tissue remodeling, metabolic reprogramming, and immune modulation. The identification of transitional states and functionally specialized subtypes lays a foundation for further mechanistic investigation into hormone-driven reorganization and its implications for porcine pregnancy and reproductive health. Future work integrating spatial context, epigenetic landscapes, and hormonal cues will be critical to fully delineate the regulatory networks guiding mammary gland plasticity.

DATA AVAILABILITY

The raw sequencing data have been deposited in the China National Center for Bioinformation database of GSA (https://ngdc.cncb.ac.cn/gsa/) (CRA028311).

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

L.S.H. conceptualized and designed the study. S.Y.Y. was responsible for data quality control and analysis. S.Y.Y. wrote the original draft. S.Y.Y., T.X.Y. and L.S.H. reviewed and revised the manuscript. L.S.H supervised the project. All authors read and approved the final version of the manuscript.

ACKNOWLEDGEMENTS

We thank our lab member Yi-Zhong Huang and Jia-Cheng Wei from Jiangxi Agricultural University for collecting the samples.

FUNDING

Supported by Key Project of Jiangxi Provincial Key Research and Development Program (Grant No. 20243BCC31002) and National Modern Agricultural Industry Technology System (CARS-35)

REFERENCES

Armistead B, Johnson E, Vanderkamp R, et al. 2020. Placental regulation of energy homeostasis during human pregnancy. *Endocrinology*, **161**(7): bqaa076.

Bhat-Nakshatri P, Gao H, Khatpe AS, et al. 2024. Single-nucleus chromatin accessibility and transcriptomic map of breast tissues of women of diverse genetic ancestry. *Nature medicine*, **30**(12): 3482–3494.

Bindea G, Mlecnik B, Hackl H, et al. 2009. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics*, **25**(8): 1091–1093.

Boeker V & Kalluri R. 2025. Fibroblast dynamics during mammary oncogenesis: senescence, Wnt9a and beyond. *The EMBO Journal*, **44**(11): 3053–3055.

Buchholz M-B, Scheerman DI, Levato R, et al. 2024. Human breast tissue engineering in health and disease. *EMBO Molecular Medicine*, **16**(10): 2299–2321.

Consortium BCA. 2021. Breast cancer risk genes—association analysis in more than 113,000 women. *New England Journal of Medicine*, **384**(5): 428–439.

Crow M, Paul A, Ballouz S, et al. 2018. Characterizing the replicability of cell types defined by single cell RNA-sequencing data using MetaNeighbor. *Nature communications*, **9**(1): 884.

Donald K, Petersen C, Turvey SE, et al. 2022. Secretory IgA: linking microbes, maternal health, and infant health through human milk. *Cell host & microbe*, **30**(5): 650–659.

Fan YL, Jin L, He ZP, et al. 2023. A cell transcriptomic profile provides insights into adipocytes of porcine mammary gland across development. *Journal of Animal Science and Biotechnology*, **14**(1): 126.

Feigman MJ, Moss MA, Chen C, et al. 2020. Pregnancy reprograms the epigenome of mammary epithelial cells and blocks the development of premalignant lesions. *Nature communications*, **11**(1): 2649.

Fu XY, Pereira R, Liu C-C, et al. 2023. High FOXA1 levels induce ER transcriptional reprogramming, a pro-metastatic secretome, and metastasis in endocrine-resistant breast cancer. *Cell reports*, **42**(8).

Gamble LA, Mcclelland PH, Teke ME, et al. 2023. Defining features of hereditary lobular breast cancer due to CDH1 with magnetic resonance imaging and tumor characteristics. *NPJ Breast Cancer*, **9**(1): 77.

Gieniec KA & Davis FM. 2022. Mammary basal cells: stars of the show. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, **1869**(1): 119159.

Gray GK, Carlson EG, Lev T, et al. 2025. Defining breast epithelial cell types in the single-cell era. *Developmental cell*, **60**(17): 2218–2236.

Gray GK, Li CM-C, Rosenbluth JM, et al. 2022. A human breast atlas integrating single-cell proteomics and transcriptomics. *Developmental cell*, **57**(11): 1400–1420. e1407.

Gu ZG, Eils R, Schlesner M. 2016. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*, **32**(18): 2847–2849.

Guo QC, Sun DC, Barrett AS, et al. 2022. Mammary collagen is under reproductive control with implications for breast cancer. *Matrix Biology*, **105**: 104–126.

Hadadi E, Deschoemaeker S, Venegas GV, et al. 2022. Heterogeneity and function of macrophages in the breast during homeostasis and cancer. *International Review of Cell and Molecular Biology*, **367**: 149–182.

Hardy N, Enger K, Hanson J, et al. 2021. Organization of mammary blood vessels as affected by mammary parenchymal region and estradiol administration in Holstein heifer calves. *Journal of Dairy Science*, **104**(5): 6200–6211.
Hawley D, Tang X, Zyrianova T, et al. 2018. Myoepithelial cell-driven acini contraction in response to oxytocin receptor stimulation is impaired in lacrimal glands of Sjögren's syndrome animal models. *Scientific reports*, **8**(1): 9919.
Hurley W. 2019. Mammary gland development in swine: embryo to early lactation. *Animal*, **13**(S1): s11–s19.
Jang SW, Park W-J, Min H, et al. 2018. Altered mRNA expression levels of the major components of sphingolipid metabolism, ceramide synthases and their clinical implication in colorectal cancer. *Oncology Reports*, **40**(6): 3489–3500.
Kutys ML, Polacheck WJ, Welch MK, et al. 2020. Uncovering mutation-specific morphogenic phenotypes and paracrine-mediated vessel dysfunction in a biomimetic vascularized mammary duct platform. *Nature communications*, **11**(1): 3377.
Li CM-C, Shapiro H, Tsiobikas C, et al. 2020. Aging-associated alterations in mammary epithelia and stroma revealed by single-cell RNA sequencing. *Cell reports*, **33**(13).
Li JJ, Sun YH, Zhi XL, et al. 2024. Unraveling the molecular mechanisms of lymph node metastasis in ovarian cancer: focus on MEOX1. *Journal of Ovarian Research*, **17**(1): 61.
Luo LY, Zhang W, You SY, et al. 2024. The role of epithelial cells in fibrosis: Mechanisms and treatment. *Pharmacological Research*, **202**: 107144.
Nagy D, Gillis CM, Davies K, et al. 2021. Developing ovine mammary terminal duct lobular units have a dynamic mucosal and stromal immune microenvironment. *Communications biology*, **4**(1): 993.
Nightingale R, Reehorst CM, Vukelic N, et al. 2024. Ehf controls mammary alveolar lineage differentiation and is a putative suppressor of breast tumorigenesis. *Developmental cell*, **59**(15): 1988–2004. e1911.
Niimi K, Usami K, Fujita Y, et al. 2018. Development of immune and microbial environments is independently regulated in the mammary gland. *Mucosal immunology*, **11**(3): 643–653.
Pascual R, Cheng J, De Smet AH, et al. 2025. Fibroblast hierarchy dynamics during mammary gland morphogenesis and tumorigenesis. *The EMBO Journal*, **44**(11): 3266–3300.
Piantoni P & Vandehaar M. 2022. The impact of absorbed nutrients on energy partitioning throughout lactation. Pages 119–119 in Proc. Journal of Dairy Science. ELSEVIER SCIENCE INC STE 800, 230 PARK AVE, NEW YORK, NY 10169 USA.
Polański K, Young MD, Miao Z, et al. 2020. BBKNN: fast batch alignment of single cell transcriptomes. *Bioinformatics*, **36**(3): 964–965.
Qiu XJ, Mao Q, Tang Y, et al. 2017. Reversed graph embedding resolves complex single-cell trajectories. *Nature methods*, **14**(10): 979–982.
Rao L, Cai LP, Huang LS. 2023. Single-cell dynamics of liver development in postnatal pigs. *Science Bulletin*, **68**(21): 2583–2597.
Ryman VE, Packiriswamy N, Sordillo LM. 2015. Role of endothelial cells in bovine mammary gland health and disease. *Animal health research reviews*, **16**(2): 135–149.
Sawhney RS. 1993. Expression of type I and type III procollagen by lens epithelial cells. *Investigative ophthalmology & visual science*, **34**(7): 2195–2202.
Seachrist DD & Keri RA. 2019. The activin social network: activin, inhibin, and follistatin in breast

development and cancer. *Endocrinology*, **160**(5): 1097–1110.
Shannon P, Markiel A, Ozier O, et al. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research*, **13**(11): 2498–2504.
Sprenger KG, Lietzke EE, Melchior JT, et al. 2025. Lipid and lipoprotein metabolism in microglia: Alzheimer's disease mechanisms and interventions. *Journal of Lipid Research*, **66**(10): 100872.
Traag VA, Waltman L, Van Eck NJ. 2019. From Louvain to Leiden: guaranteeing well-connected communities. *Scientific reports*, **9**(1): 1–12.
Trott JF, Schennink A, Horigan KC, et al. 2022. Unique transcriptomic changes underlie hormonal interactions during mammary histomorphogenesis in female pigs. *Endocrinology*, **163**(3): bqab256.
Tsutsui S, Wakasa H, Tsugami Y, et al. 2020. Distinct expression patterns of fibrillar collagen types i, iii, and v in association with mammary gland remodeling during pregnancy, lactation and weaning. *Journal of mammary gland biology and neoplasia*, **25**(3): 219–232.
Twigger A-J, Engelbrecht LK, Bach K, et al. 2022. Transcriptional changes in the mammary gland during lactation revealed by single cell sequencing of cells from human milk. *Nature communications*, **13**(1): 562.
Wang J, Chen YM, Wang Q, et al. 2022. MEOX2-mediated regulation of Cathepsin S promotes cell proliferation and motility in glioma. *Cell Death & Disease*, **13**(4): 360.
Wang JQ, Peng LM, Wang H, et al. 2025. Regulation of lipoprotein processing by GPNMB in foamy macrophages: potential therapeutic targets for atherosclerosis. *Nature communications*, **16**(1): 10030.
Wolf FA, Angerer P, Theis FJ. 2018. SCANPY: large-scale single-cell gene expression data analysis. *Genome biology*, **19**(1): 15.
Wolock SL, Lopez R, Klein AM. 2019. Scrublet: computational identification of cell doublets in single-cell transcriptomic data. *Cell systems*, **8**(4): 281–291. e289.
Xiao YY, Zou XX, Yang B, et al. 2025. Single-nucleus transcriptome profiling of wild boar and domestic pig intestines reveals the spatiotemporal dynamics of immunity and nutrient absorption. *Science China Life Sciences*: 1–18.
Yamaguchi K, Nakayama J, Yamamoto T, et al. 2024. Collagen induction of immune cells in the mammary glands during pregnancy. *Physiological Genomics*, **56**(2): 128–135.
Yanai I, Benjamin H, Shmoish M, et al. 2005. Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics*, **21**(5): 650–659.
Yang SY, Chen D, Xie L, et al. 2023. Developmental dynamics of the single nucleus regulatory landscape of pig hippocampus. *Science China Life Sciences*, **66**(11): 2614–2628.
Yang SY, Jing QQ, Duan WX, et al. 2025. Microglial function alterations during healthy ageing and canine cognitive dysfunction revealed by single-nucleus and spatial transcriptomics. *National Science Review*, doi: 10.1093/nsr/nwaf388
Yao TX, Zhang Y, Chen XY, et al. 2026. Dynamic single-nucleus transcriptomic profiling of kidney development in Bamaxiang and Wild Boar. *Zoological Research*, **47**(1): 58–71.
Yao XL, Gao MX, Dai CL, et al. 2013. Peptidoglycan Recognition Protein 1 Promotes House Dust Mite–Induced Airway Inflammation in Mice. *American Journal of Respiratory Cell and Molecular Biology*, **49**(6): 902–911.
Zhang SH, Chen F, Zhang YZ, et al. 2018. Recent progress of porcine milk components and mammary gland function. *J Anim Sci Biotechnol*, **9**: 77.

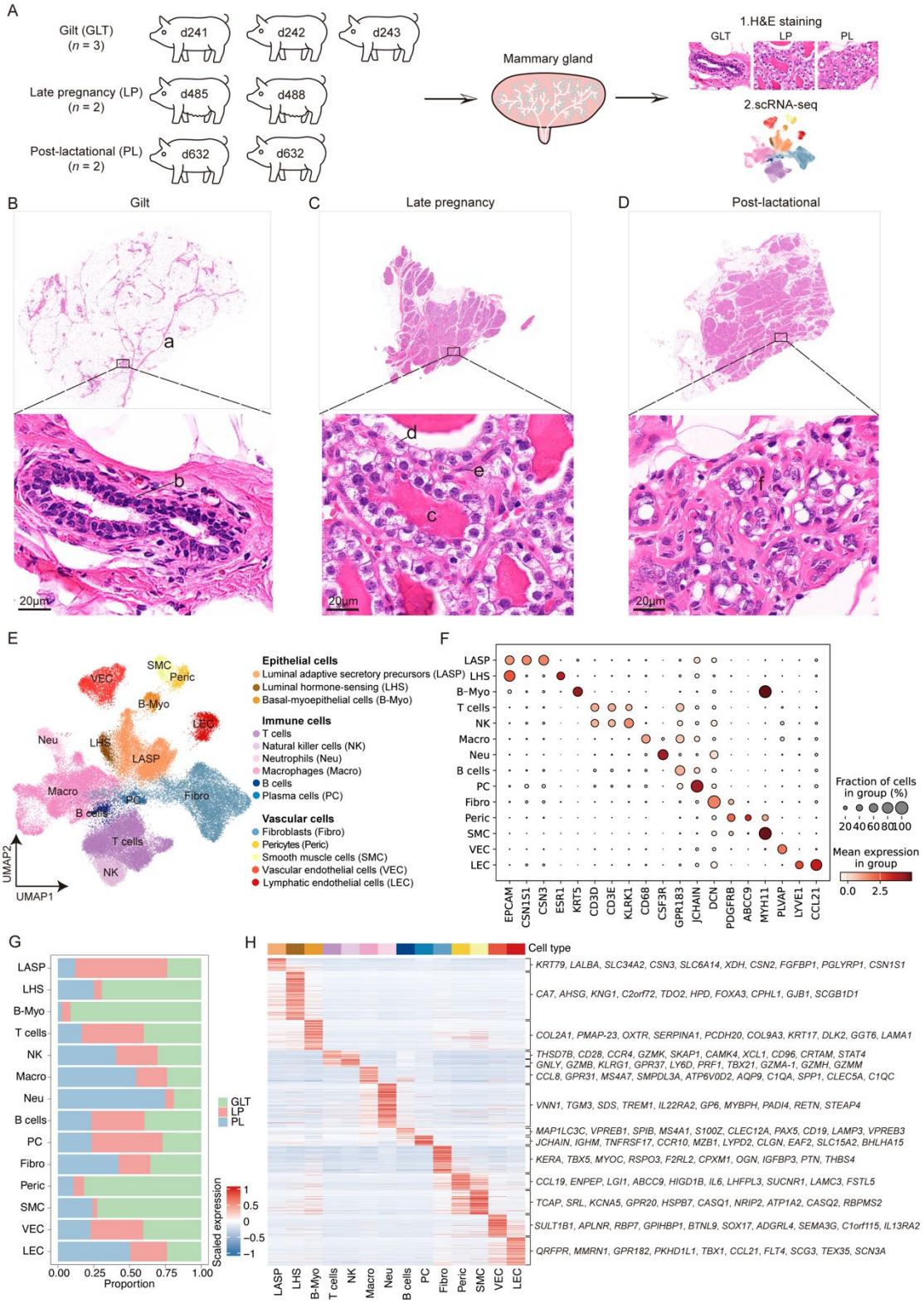

Figure 1 Single-cell transcriptomic atlas of the mammary gland across reproductive stages

A: Experimental design illustrating H&E staining and scRNA-seq of mammary gland tissues collected from gilt, late-pregnancy, and post-lactational stages. *n* indicates the number of biological replicates (individual pigs). d denotes the postnatal day (e.g., d241 represents day 241 after birth). B–D: Representative H&E-stained sections of mammary glands from gilt (B), late pregnancy (C), and post lactational stages (D). Labels a–f indicate the following structures: a, ductal network; b,

immature alveolar structure; c, milk in the alveolar lumen; d, alveolar epithelial cells; e, basal myoepithelial cells; f, regressed alveolar structure. Scale bar, 20 μm. E: UMAP visualization of single-cell transcriptomes from seven mammary samples, colored by cell type. LASP: luminal adaptive secretory precursors, LHS: luminal hormone-sensing cells, B-Myo: basal-myoepithelial cells, NK: natural killer cells, PC: plasma cells, Macro: macrophages, Neu: neutrophil, Fibro: fibroblasts, Peric: pericytes, SMC: smooth muscle cells, VEC: vascular endothelial cells, LEC: lymphatic endothelial cells. F: Dot plot showing the expression patterns of canonical marker genes across major cell types. G: Bar plot showing the proportion of cells from each stage contributing to the 14 major cell types. H: Heatmap showing the expression levels of 2 474 cell type-specific genes, τ>0.85. The top 10 marker genes for each cell type, ranked by τ scores, are highlighted on the right.

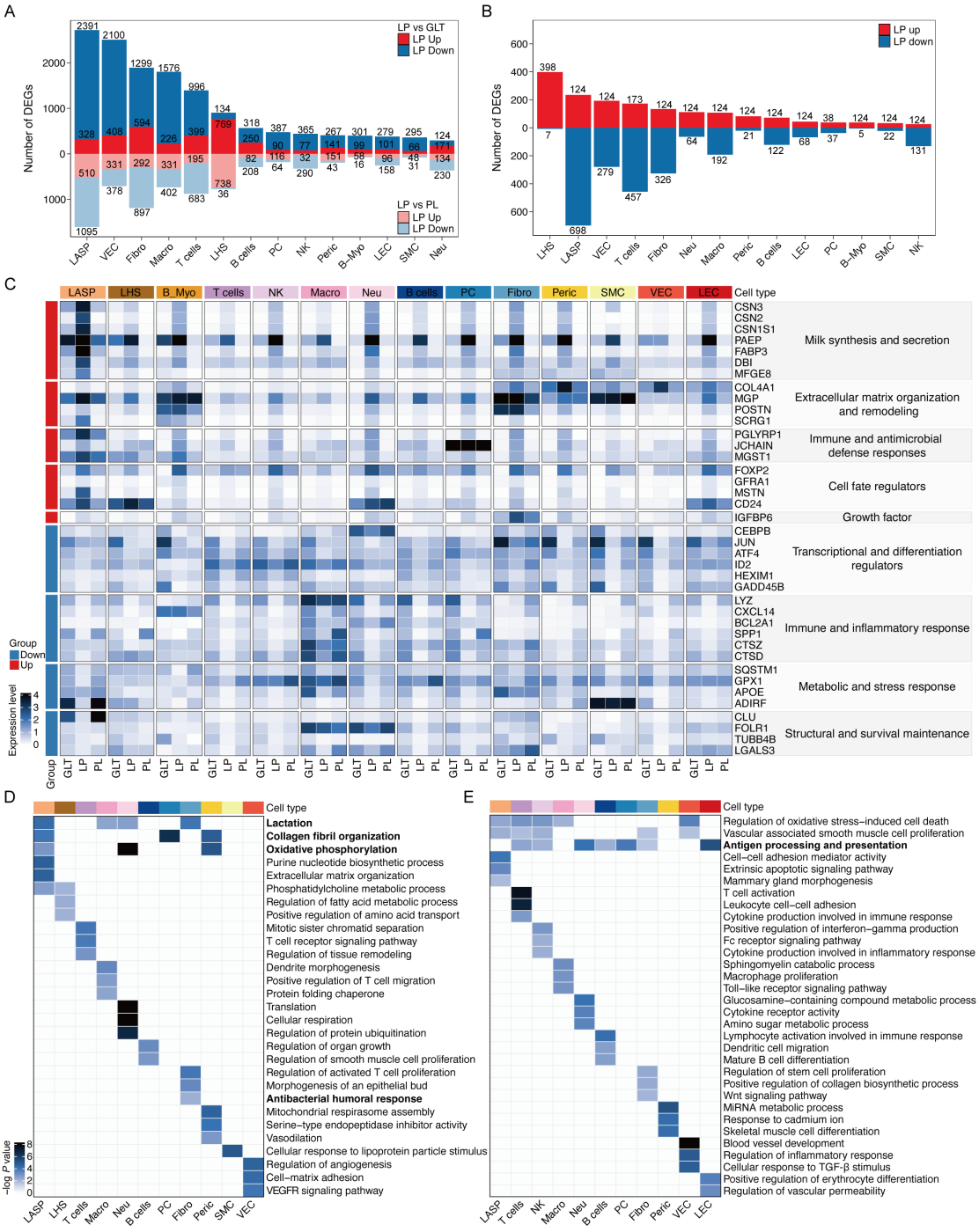

Figure 2 Transcriptional dynamics of mammary gland cell types during pregnancy

A: Bar plot showing the number of DEGs in late pregnancy compared to gilt (top) and post-lactational (bottom) stages (logFC>1, P<0.05). Red indicates genes upregulated in late pregnancy; blue indicates downregulated genes. B: Intersections of upregulated and downregulated DEGs in each cell type for the comparisons of late pregnancy versus gilt and post-lactational versus post-lactational. Red bars represent the number of shared upregulated genes in late pregnancy; blue bars represent shared downregulated genes in late pregnancy. C: Heatmap showing the expression of

genes that are commonly upregulated (red) or downregulated (blue) across multiple cell types in late pregnancy. Cell type-shared genes are defined as those significantly altered in at least two cell types. D–E: Heatmaps showing enriched pathways for upregulated (D) and downregulated (E) genes in each cell type. Cell type abbreviations are as defined in Figure 1E.

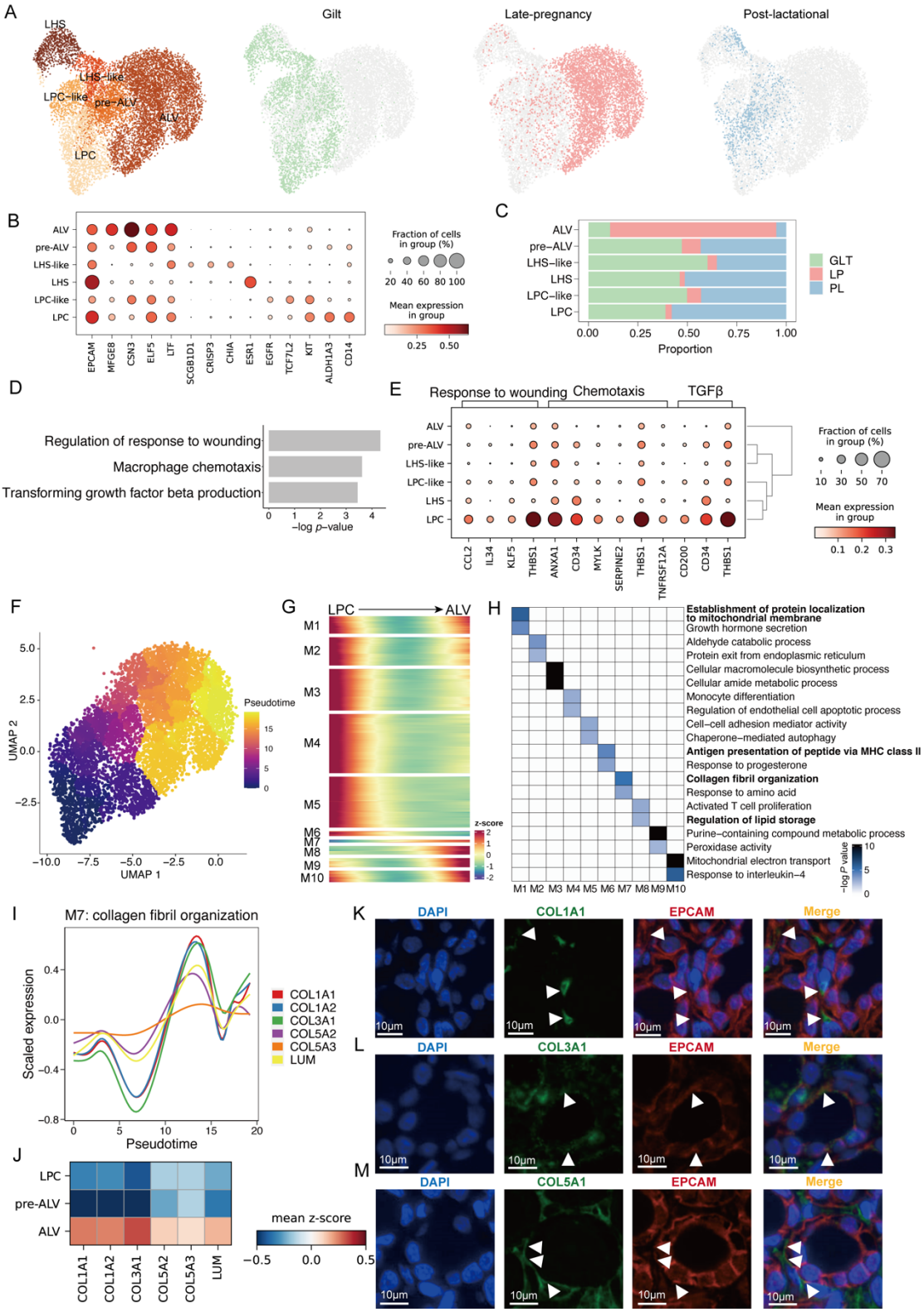

Figure 3 Transcriptional heterogeneity and lineage progression of epithelial subtypes in the pig mammary gland

A: UMAP visualization of mammary epithelial cells colored by subtype (left) and by reproductive stage (right). B: Dot plot showing the expression of canonical marker genes across epithelial subtypes. C: Bar plot showing the stage contribution of cells to each of the six epithelial subtypes. D: GO term enrichment of LPC-specific highly expressed genes. E: Dot plot showing the expression

of LPC-enriched genes associated with selected GO terms. F: Pseudotime trajectory of LPC, pre-ALV, and ALV cells reconstructed using Monocle3, colored by pseudotime value. G: Heatmap showing the expression dynamics of the top 10 000 genes significantly associated with pseudotime (Q<0.01) along the LPC-to-ALV axis. H: GO enrichment heatmap of the ten gene modules derived from pseudotime-associated genes. I: Expression dynamics of genes within Module 7 enriched for collagen fibril organization along pseudotime trajectory. J: Expression level of genes within Module 7 enriched for collagen fibril organization across three cell types. K–M: Immunofluorescence staining showing the colocalization of type I, III, and V collagens with the epithelial marker EPCAM in mammary tissue at the late pregnancy stage. Arrows highlight alveolar epithelial cells that are positive for COL1A1, COL3A1, and COL5A1. Scale bar, 10 μm.

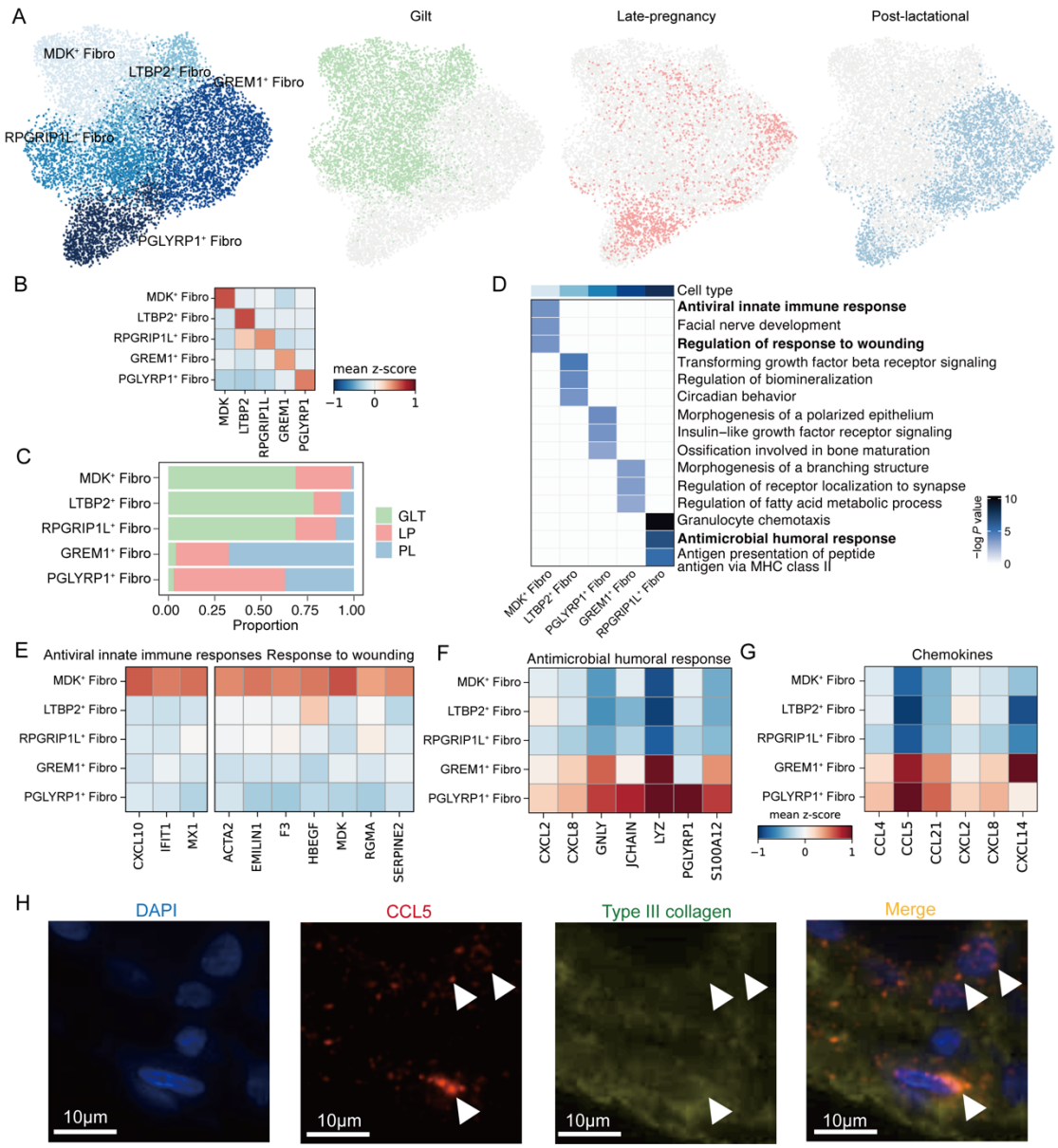

Figure 4 Transcriptional heterogeneity of fibroblast subtypes in the mammary gland during pregnancy

A: UMAP visualization of fibroblast populations colored by subtype (left) and by reproductive stage (right). B: Heatmap showing the expression of subtype-specific marker genes defining fibroblast identities. C: Bar plot showing the contribution of cells from each reproductive stage to the five fibroblast subtypes. D: GO enrichment heatmap of subtype-specific highly expressed genes across fibroblast populations. E: Heatmap showing the expression of genes enriched in MDK$^+$ fibroblasts, associated with antiviral innate immune response and regulation of wound healing. F: Heatmaps showing expression of genes enriched in PGLYRP1$^+$ fibroblasts, associated with antimicrobial humoral response. G: Heatmaps showing expression of genes enriched in PGLYRP1$^+$ fibroblasts, associated with granulocyte chemotaxis. H: Immunofluorescence staining showing co-localization of the fibroblast marker type III collagen (yellow) with CCL5 (red) in the mammary gland from the late-pregnancy stage. Arrows highlight fibroblasts that are positive for CCL5. Scale bar: 10μm.

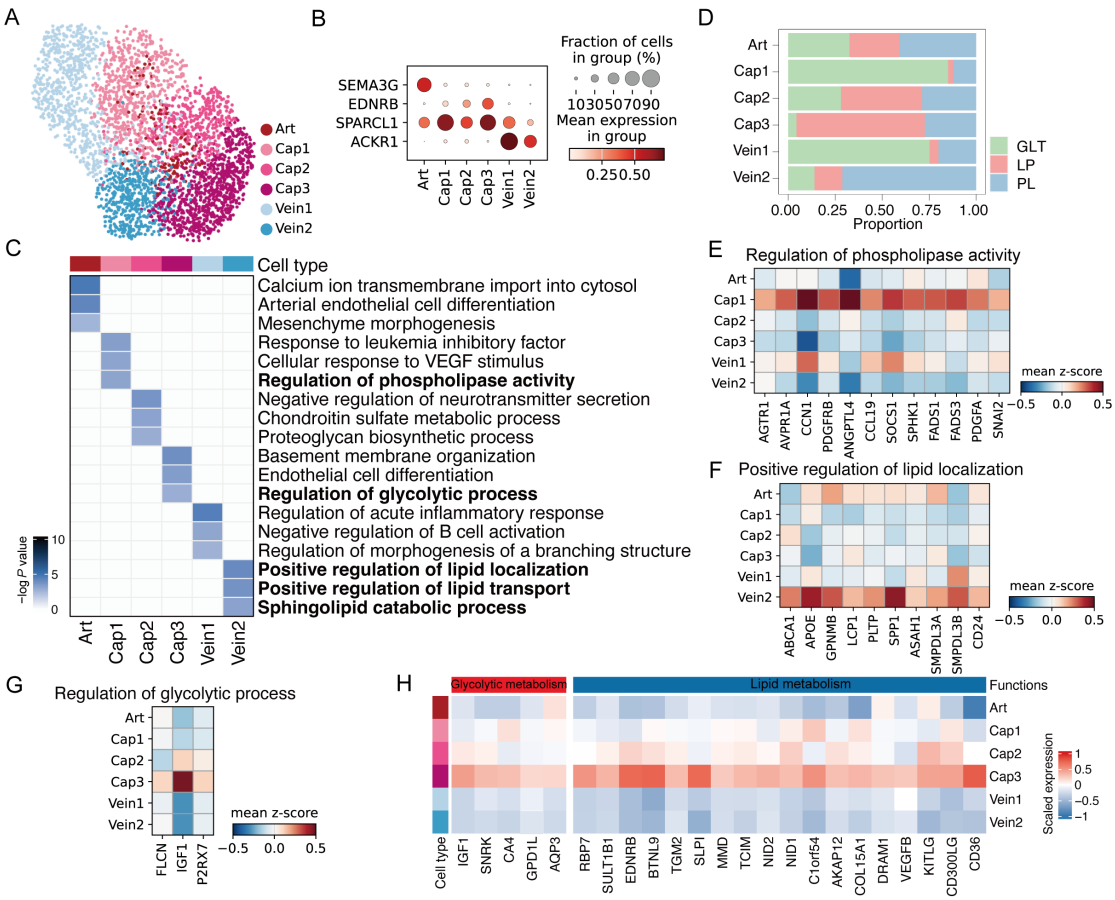

Figure 5 Transcriptional heterogeneity of endothelial subtypes in the mammary gland during pregnancy

A: UMAP visualization of endothelial cells colored by subtype (left) and by reproductive stage (right). B: Dot plot showing the expression of marker genes distinguishing arterial, capillary and venous endothelial identities. C: GO enrichment heatmap of subtype-specific highly expressed genes across endothelial subtypes. D: Bar plot showing the contribution of cells from each reproductive stage to the six endothelial subtypes. E: Heatmap showing the expression of genes enriched in Cap1, associated with the regulation of phospholipase activity. F: Heatmap showing the expression of genes enriched in Vein2, linked to the positive regulation of lipid localization. G: Heatmap showing the expression of genes enriched in Cap3, involved in the regulation of glycolytic processes. H: Heatmap illustrating the expression patterns of Cap3-specific genes involved in glycolytic and lipid metabolism.

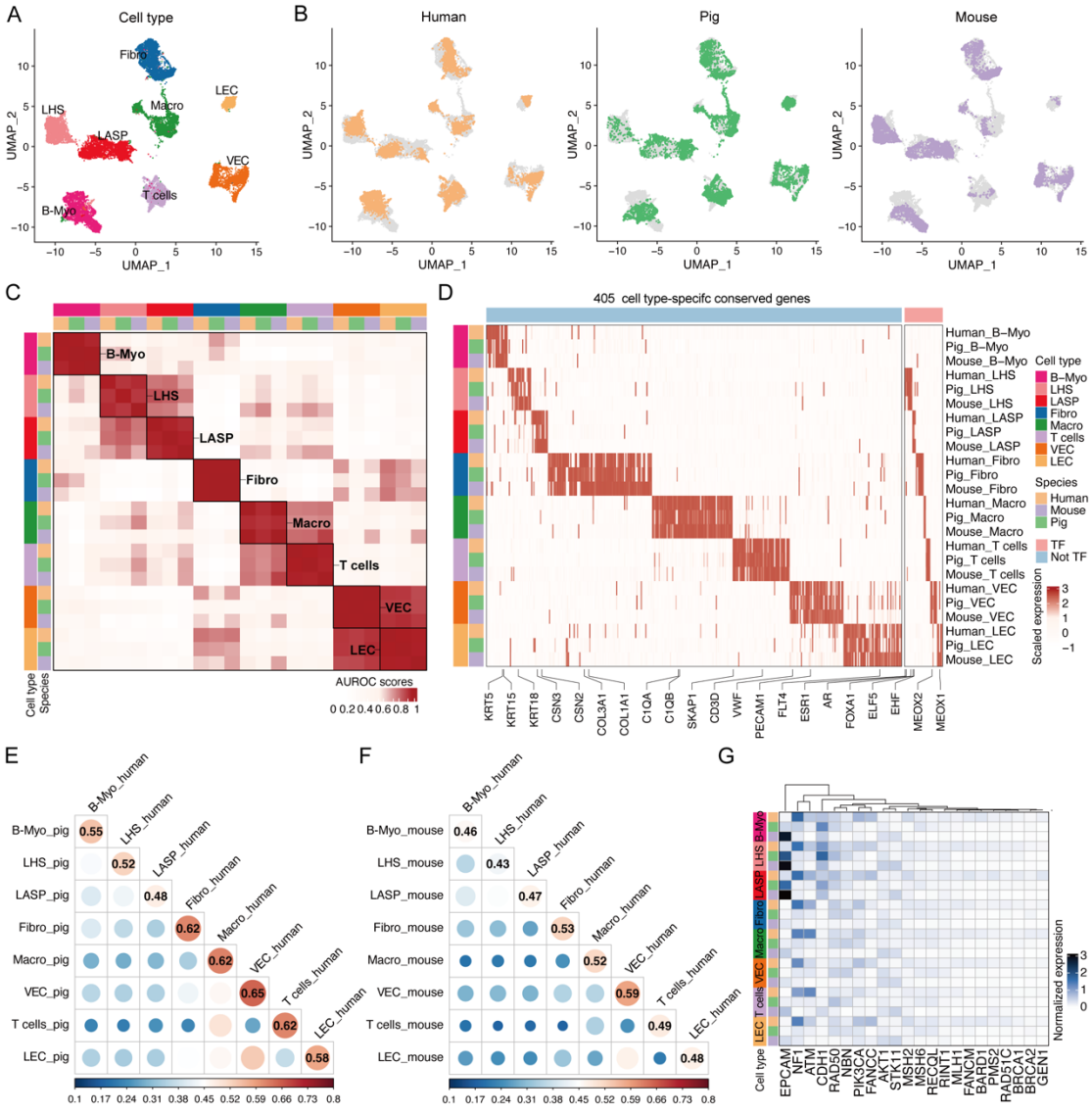

Figure 6 Cross-species single-cell transcriptomic comparison of mammary gland cell types in humans, pigs, and mice.

(A) UMAP visualization of integrated single-cell transcriptomes from human, pig, and mouse mammary glands, colored by cell type. Cell type abbreviations are as defined in Figure 1E. (B) UMAP plots of individual species, colored by species identity. (C) Heatmap showing AUROC scores quantifying cross-species similarity among human, pig, and mouse mammary gland cell types. All corresponding cell types exhibit high concordance, with a minimum AUROC of 0.91. (D) Heatmap showing the expression patterns of 405 cross-species conserved cell type-specific genes, defined by τ>0.85 and detection in more than 5% of cells within the corresponding cell type. (E–F) Heatmaps showing pairwise correlation coefficients between cell types from humans and pigs (E), humans and mice (F). (G) Heatmap showing the expression patterns of breast cancer risk genes across different cell types in the human, pig, and mouse mammary glands.

单细胞转录组学揭示妊娠期猪乳腺的动态细胞重塑

杨思雨[1,*]，姚天雄[1,*]，黄路生[1,*]

[1]江西农业大学猪遗传改良与种质创新全国重点实验室，南昌 330045

[*]通讯作者：杨思雨 (siyuyang1013@163.com)；姚天雄 (TianxiongYao@hotmail.com)；黄路生 (lushenghuang@hotmail.com)

摘要：妊娠会诱导乳腺发生广泛的重塑，以建立哺乳所需的腺泡网络，然而这一转变在猪中的细胞程序仍缺乏系统解析。在本研究中，我们构建了覆盖后备母猪、妊娠晚期及泌乳后阶段的猪乳腺单细胞转录组图谱。差异表达分析显示，上皮细胞、成纤维细胞和内皮细胞在妊娠过程中表现出最为显著的动态变化，促使我们对这些细胞谱系进行更高分辨率的分析。上皮细胞向腺泡命运分化的轨迹中诱导了以 COL1A1、COL3A1 和 COL5A1 为代表的纤维型胶原相关程序，提示上皮细胞在细胞外基质组装中的潜在贡献。成纤维细胞分析揭示了一类 $PGLYRP1^{+}$ 成纤维细胞亚型，该亚型富集表达趋化因子 CCL4、CCL5 和 CCL21，提示其可能参与免疫细胞的募集。内皮细胞在不同阶段经历特异性的代谢重编程，其中妊娠晚期以内皮细胞中糖酵解相关基因（包括 IGF1、FLCN 和 P2RX7）以及脂质代谢相关基因（如 RBP7、KITLG 和 CD36）表达上调为特征。跨物种比较分析表明，人、猪和小鼠乳腺在细胞身份及关键分子特征（ESR1、AR 和 FOXA1）方面具有高度保守性。总体而言，该图谱系统刻画了妊娠相关乳腺重塑的细胞学逻辑，并为解析猪乳腺转录组变化提供了重要资源。

关键词：猪；乳腺；妊娠；单细胞转录组；重塑