

# JNK1 downregulation confers fibroblasts with a mammary epithelial cell fate

DOI: <https://dx.doi.org/10.71373/PMKJ7669>

Submitted 14 June 2025

Accepted 22 October 2025

Published 24 October 2025

Dandan Zhang, Liangshan Qin, Quanhui Liu, Shan Deng, Zhigang Lei, Longfei Sun, Hong Pan,

Zhiqiang Wang, Xiang Yuan, Guodong Wang<sup>#</sup>, Ben Huang<sup>#</sup>

The fate of cells is fundamentally determined by their specific transcription patterns, which dictate cellular identity and differentiation pathways. Previously, in our research, we demonstrated that chemical cocktails, including the single small molecule TGF $\beta$ R1 inhibitor RepSox(R), can induce significant changes in transcription patterns and drive the reprogramming of somatic cells into chemically induced mammary epithelial cells (R-iMECs) by effectively downregulating TGF $\beta$ R1 expression. However, the underlying molecular mechanism of chemical induction for such reprogramming remains unclear and poorly defined, limiting its broader application. Here, to address this gap, we employed comprehensive single-cell sequencing (scRNA-seq and scATAC-seq), to systematically uncover RepSox-induced reprogramming events, map detailed molecular trajectories. Importantly, our integrated analysis revealed that JNK1 expression, a critical downstream gene in the TGF $\beta$ R1 signaling pathway, plays an indispensable role in R-iMECs conversion. Thus, the TGF $\beta$ R1-JNK1 signaling cascade may serve as a key regulator of fibroblast-to-mammary epithelial cell fate decision, orchestrating cellular transitions through transcriptional modulation. These findings provide valuable mechanistic insights into the complex processes governing mammary epithelial cell fate decisions and may pave the way for developing novel, efficient strategies to generate large-scale, functional milk-secreting mammary cells in vitro, while also facilitating targeted mammary gland regeneration and repair in vivo for therapeutic advancements.

## Introduction

During fertilization, oocytes fuse with sperm to form zygotes. Subsequently, they undergo division and differentiation to generate an embryo consisting of three germ layers, which eventually develops into the entire organism<sup>[1,2]</sup>. Hence, all specialized cells in an organism originate from a single cell. In most organisms, particularly mammals, embryonic cells irreversibly transform into terminally differentiated cells that ultimately undergo programmed cell death<sup>[1,3]</sup>. Although some stem cells exist in organisms and can differentiate into functional cells to compensate for cellular loss, these stem cells are scarce or become significantly depleted over time<sup>[4]</sup>. The loss of self-regeneration ability may serve as a protective mechanism against oncogenicity since uncontrolled cell division could lead to cancer formation<sup>[5,6,7]</sup>. However, under pathological conditions where specific tissues lose their functionality due to disease or injury, reprogramming terminally differentiated functional cells back into undifferentiated state holds great potential for tissue regeneration by enabling them to differentiate into other normal functional cell types<sup>[8]</sup>.

Previous study has demonstrated that somatic cells can be reprogrammed into induced pluripotent stem cells (iPSCs), which possess the ability to differentiate into various cell types derived from three germ layers<sup>[9]</sup>. However, the utilization of iPSCs for obtaining terminally differentiated functional cells may pose potential oncogenic risks due to the challenge in ensuring complete segregation between differentiated and undifferentiated iPSCs<sup>[10,11]</sup>. To circumvent this issue, researchers have employed transdifferentiation, also known as direct lineage reprogramming, to generate differentiated functional cell types directly from somatic cells<sup>[12]</sup>. Transdifferentiation offers a

lower risk of oncogenicity since it does not involve the induction of pluripotency. Various functional cell types such as myoblasts<sup>[13]</sup>, neurons<sup>[14,15]</sup>, cardiomyocytes<sup>[16]</sup>, neural stem/progenitor cells<sup>[17]</sup>, and hepatocytes<sup>[18]</sup> have been successfully generated through transgene-mediated transdifferentiation in vitro using fibroblasts or other somatic cell sources. Unfortunately, this approach still raises concerns regarding oncogenicity due to the use of transgenes<sup>[19]</sup>. In recent years, novel strategies involving chemical-mediated transdifferentiation have effectively addressed some of these concerns by enabling the successful generation of a wide range of induced cell types from somatic cells including induced neural stem cells<sup>[20,21]</sup>, induced neurons<sup>[22,23]</sup>, and induced cardiomyocytes<sup>[16]</sup> under chemical induction conditions. These chemically-induced cells hold promising clinical applications in tissue regeneration and provide valuable insights into the mechanisms underlying cellular fate conversion from somatic cells.

The mammary gland is a vital organ in mammals, playing a crucial role in milk secretion. Damage to the mammary gland can have severe consequences for mammals. Mammary epithelial cells are capable of restoring the structure and functions of mammary glands both in vivo and in vitro, making them valuable for various applications<sup>[24]</sup>. Therefore, the generation of mammary epithelial cells from nonmammary cells holds great potential for both in vivo and in vitro purposes. Our previous studies have shown that small molecule cocktail induction or TGF $\beta$ R1 inhibitor RepSox alone can induce nonmammary cells to acquire a mammary cell fate<sup>[8]</sup>. In this study, we conducted a comprehensive investigation into the reprogramming events and tracked the conversion of chemically induced mammary epithelial cells (CiMECs) from somatic cells using single-cell sequencing analysis, aiming to elucidate the similarities and differences between iMECs generated through chemical cocktails induction or TGF $\beta$ R1 inhibitor RepSox induction. Additionally, we discovered that TGF $\beta$ R1 downregulation induces this reprogramming event through its regulatory effects on JNK1 downregulation. These findings provide insights into the role of the TGF $\beta$ R1-JNK1 signaling pathway in somatic cell

*I. Guangxi Academy of Medical Sciences, Guangxi Key Laboratory of Eye Health, Nanning, 530021, China.*

*#Corresponding author*

*E-mail addresses: benhuang@gxu.edu.cn/bhuang@gxams.org.cn (B. Huang); guodongwangbio@163.com (GD.Wang).*

*Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Ben Huang.*

reprogramming and decision-making processes related to mammary epithelial cell fate determination. Furthermore, they establish a novel technical platform for generating abundant functional mammary epithelial cells in vitro, which could be applied to various bioapplications and potentially contribute to mammalian gland regeneration in vivo.

## Materials and methods

### Cell culture

Goat ear fibroblasts (GEFs) were cultured in high-glucose DMEM(GIBCO) supplemented with 10% fetal bovine serum (FBS,GIBCO). Once the GEFs approached approximately 80%-90% confluence, they were passaged. GEFs were seeded 24 hours later, and the medium was replaced with R-induction medium (N2B27 and 10 $\mu$ M RepSox). The culture was continued for 8 days, and the induction medium was refreshed every 2 days. The specific experimental methods used were described previously<sup>[8]</sup>.

### Immunofluorescence staining

The procedure was performed according to Zhang's procedure<sup>[8]</sup>. The primary antibodies used were EpCAM(1:100), KRT19 (1:200), ITGA6 (1:100), and KRT8 (1:1,000) at 4°C overnight. The secondary antibodies used were (1:250) Alexa Fluor 488 Donkey Anti-rabbit, Alexa Fluor 555 Donkey Anti-rabbit, and Alexa Fluor 488 Donkey Anti-mouse for 1h in the dark at room temperature. The nuclei were stained with Hoechst 33342 for 15 minutes. The antibody details are listed in the Key Resources Table.

### qRT-PCR

Total RNA was extracted using TRIzol reagent (Life Technologies, 149112), and cDNA synthesis was performed using SuperMix (Vazyme, R223-01). The cDNA samples were subjected to real-time quantitative PCR using a SYBR Green real-time PCR Master Mix Kit (Vazyme, Q711-02). The sequences of primers used for qPCR in this study are listed in the Key Resources Table. The expression of the genes was normalized to that of the internal control gene GAPDH, and the fold changes in the samples were calculated relative to those in the fi-broblasts.

### Western Blot

Add lysis buffer (RIPA) containing 10% protease inhibitor (PMSF) to the cell pellet in proportion. Lyse on ice for 20 minutes, and then centrifuge at 1200 rpm for 10 minutes at 4°C. Use the BCA detection kit to measure the protein concentration. After preparing the separating gel and stacking gel, load the samples and perform electrophoresis. Transfer the proteins to the membrane, and then block with 5% non-fat milk at room temperature for 1 hour (or overnight at 4°C). Dilute the primary antibody in TBST according to the concentration recommended by the manufacturer, and incubate at room temperature for 2 hours. After washing with TBST, dilute the secondary antibody according to the manufacturer's instructions, and incubate on a shaker at room temperature for 40 minutes.

Finally, after washing with TBST, place the membrane in the imager, add the substrate for imaging and recording.

## Animal experiments

In vivo administration of R-CiMEC, followed by hematoxylin-eosin (HE) staining and immunohistochemical staining of paraffin sections (IHC-P), was conducted in accordance with established literature<sup>[8]</sup>.

### Statistical analysis

All the values are presented as the means  $\pm$  SEMs. The values were considered to be significantly different at  $p < 0.05$  according to two-tailed tests assuming equal variance.

## Bulk RNA-seq bioinformatic analysis

For the bulk RNA-seq of RepSox reprogramming samples, a transcriptome library was prepared following the instructions of the TruSeq RNA Sample Preparation Kit from Illumina® (San Diego, CA) using 1  $\mu$ g of total RNA. The bulk RNA-seq data were analyzed on the Majorbio Cloud Platform. The Majorbio Cloud Platform (www.majorbio.com) was used for additional details.

## Single-cell multiome data generation

ScRNA-seq was performed on 0d, 4d and 8d cell samples using a 10 $\times$  Genomics system. Briefly, dissociated cells (~10,000 cells per sample) were loaded into a 10 $\times$  Genomics Chromium Single Cell system using Chromium Single Cell 3' Reagent Kits v3.1 (10 $\times$  Genomics, Pleasanton, CA). ScRNA libraries were generated by following the manufacturer's instructions. The libraries were pooled and sequenced on an Illumina NovaSeq 6000. The sequencing reads were processed through the Cell Ranger 4.0.0 pipeline (10 $\times$  Genomics) using the default parameters.

ScATAC-seq was performed on 0d, 4d and 8d cell using a 10 $\times$  Genomic Single Cell ATAC Reagent v1.1 Kit following the manufacturer's instructions. The libraries were pooled and sequenced on Illumina NovaSeq 6000. The sequencing data were processed through the Cell Ranger ATAC 1.1.0 pipeline (10 $\times$  Genomics) using the default parameters.

## Single-Cell RNA-Sequencing (scRNA-seq) Data Analysis

### Single-Cell RNA-Seq Data Pretreatment

Preprocessing of scRNA-seq data of R-0d, R-4d and R-8d was performed using the Seurat package in R4.4.1. This included removing low-quality cells, normalizing read counts, and selecting highly variable genes.

### Dimensionality Reduction and Clustering

UMAP and t-SNE algorithms from the Seurat package were used for dimensionality reduction of single-cell data to uncover potential cell population structures.

## Automatic Cell-Type Annotation

The SingleR package was used for automatic annotation of cell clusters. By comparing gene expression traits of unknown cell populations with reference datasets, specific cell type labels were assigned to each cell subpopulation.

## Functional Enrichment Analysis

GO/KEGG enrichment analyses of the target gene set were performed using the clusterProfiler package in R4.4.1.

## Pseudotime analysis

Pseudo-time ordering of the cell populations was performed using the monocle 3 package with default parameters.

## Single-CellATAC-Sequencing(scATAC-seq) Data Analysis

For theATAC data, All the analyses (UMAP dimension reduction, cluster identification, and identification of differentially accessible regions) were performed in Signac and the default parameter settings were used to construct cell trajectories with Monocle 3.

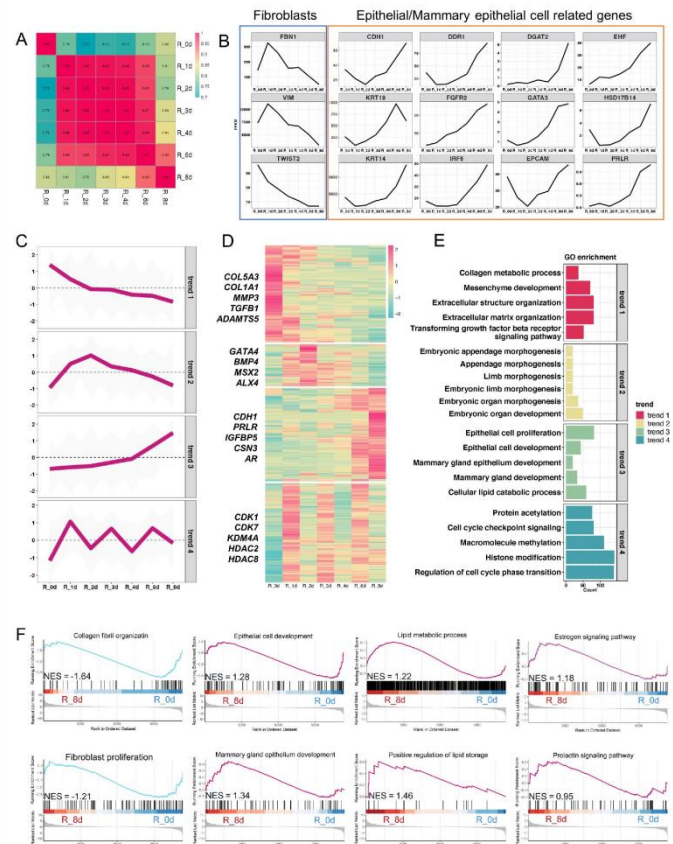
## Results

### The transcriptional dynamics of induced mammary epithelial cells generated by the TGF $\beta$ R1 inhibitor RepSox induction.

Our previous studies<sup>[8]</sup> have demonstrated that BFRTV cocktail(1  $\mu$ M TNPB (B), 10  $\mu$ M forskolin (F), 10  $\mu$ M RepSox (R),10  $\mu$ M tranlycypromine (T) and 500  $\mu$ g/ml VPA (V)) or Rep-Sox alone all can induce the conversion of fibroblasts into induced mammary epithelial cells (referred to as BFRTV-CiMECs or R-CiMECs). However, it remains unclear how the transcriptional profiles of cell fates change under RepSox induction. To address this question, we collected cells at different time points (0, 1, 2, 3, 4, 6, and 8 days post-induction) for bulk RNA sequencing analysis. Firstly,the correlation coefficient graph (Figure 1A) revealed significant differences between day 0 and other time points. Furthermore, during the induction process, fibroblast marker genes (FBN1/VIM/TWIST2) were consistently downregulated while mammary epithelial cell marker genes (CDH1/EPCAM et al.), hormone-related gene PRLR andmilk fat synthesis-related gene DGAT2 showed gradual upregulation over time (Figure 1B).

We subsequently conducted a comprehensive analysis of transcriptional trends during the transdifferentiation process using fuzzy clustering. In total, we identified four distinct gene subsets that exhibited a consistent expression pattern under specific conditions (Figure 1C). These findings indicate a progressive downregulation of somatic programs followed by an upregulation of genes involved in embryonic development as well as activation of programs related to epithelial cell proliferation and mammary epithelial cell development. Notably, genes associated with cell cycle regulation and epigenetics exhibited a fluctuating pattern of change over time (Figure 1D-E). Moreover,GSEA plot analysis also demonstrated that

compared to day-0 samples,the mammary epithelial cell lineage fate was activated in the day-8 sample (Figure 1F). Therefore, this comprehensive transcription profile data provides evidence that R-CiMECs exhibit typical characteristics associated with mammary epithelial cell fate.



**Figure1 The transcriptional dynamics of induced mammary epithelial cells generated by the TGF $\beta$ R1 inhibitor RepSox induction.** A.Heatmap illustrating the correlation coefficients of gene expression among R-0d, R-1d, R-2d, R-3d, R-4d, R-6d, andR-8d in the mRNA-seq dataset. B.Line chart depicting the expression patterns of fibroblast and epithelial/mammary epithelial cell-related marker genes obtained from bulk mRNA-seq analysis of samples collected at different time points(R-0d, R-1d, R-2d, R-3d, R-4d, R-6-d and 8-d). C.Fuzzy clustering approach applied to bulk RNA-seq data for investigating gene expression dynamics during specific time periods under defined conditions. D.Heatmap displaying relative gene expression levels across different conditions and time points with each gene group associated with a corresponding trend as described in (C). E.Gene Ontology (GO) analysis revealing enriched terms linked to each identified trend. F.Gene Set Enrichment Analysis (GSEA) demonstrating differential expression of lactation-related signaling pathways between cells collected at day 8 (R-8d) and day 0(R-0d).

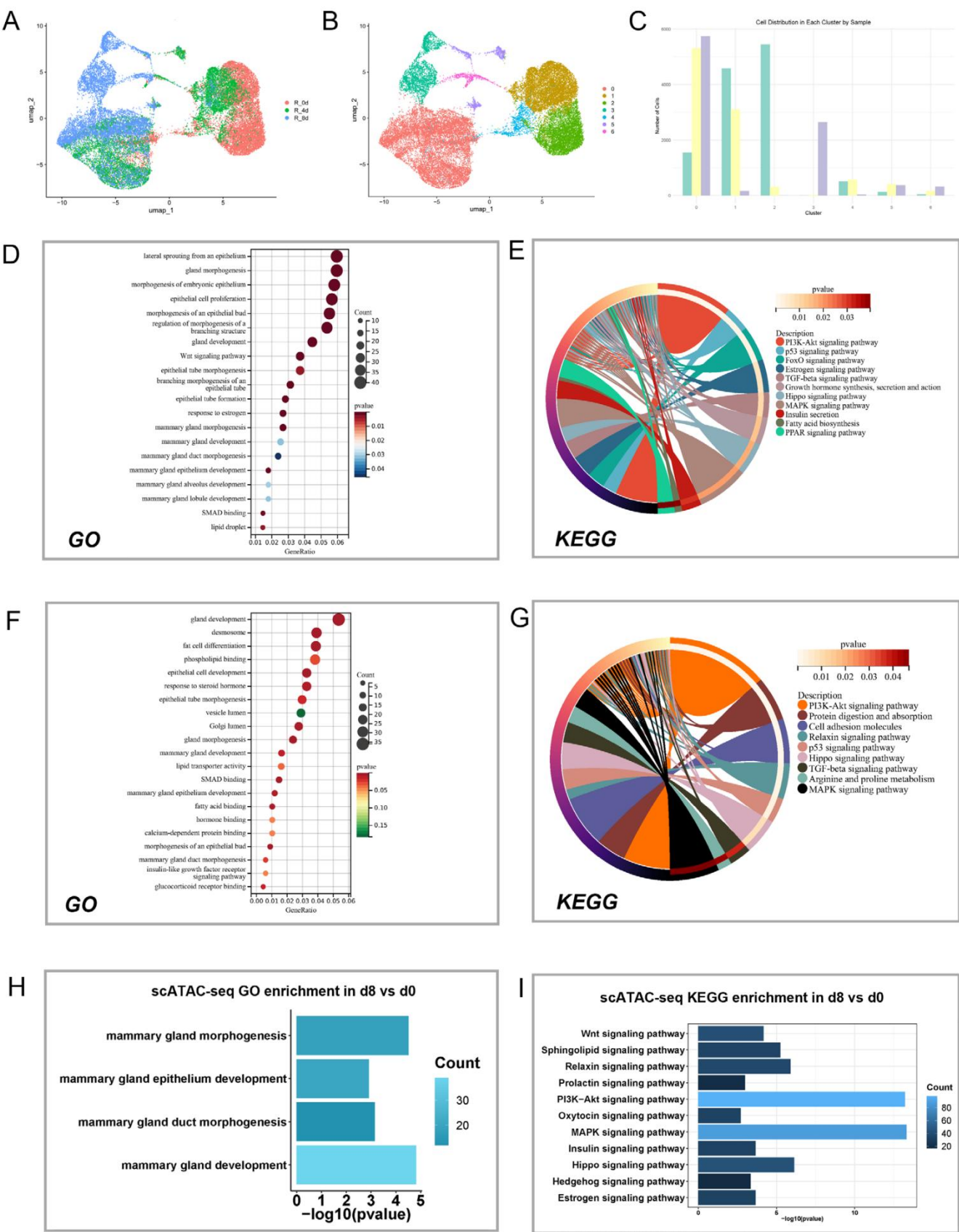
### The progression of R-CiMEC fate acquisition and reprogramming is revealed through single-cell sequencing analyses.

To further investigate the reprogramming trajectories of R-CiMECs, we performed scRNA-seq and scATAC-seq analyses. The analysis results revealed that the R-0d, R-4d, and R-8d cells were segregated into seven distinct clusters (Figure 2A-B) , as well as the cell numbers of each sample and cluster (Figure 2C). The comparison of GO and KEGG between day 4 and day 0 showed that the pathways related to the fate of mammary epithelial cells were activated (Figure 2D-E). The comparison of GO and KEGG between day 8 and day 0 showed that not only the pathways related to the fate of mammary epithelial cells



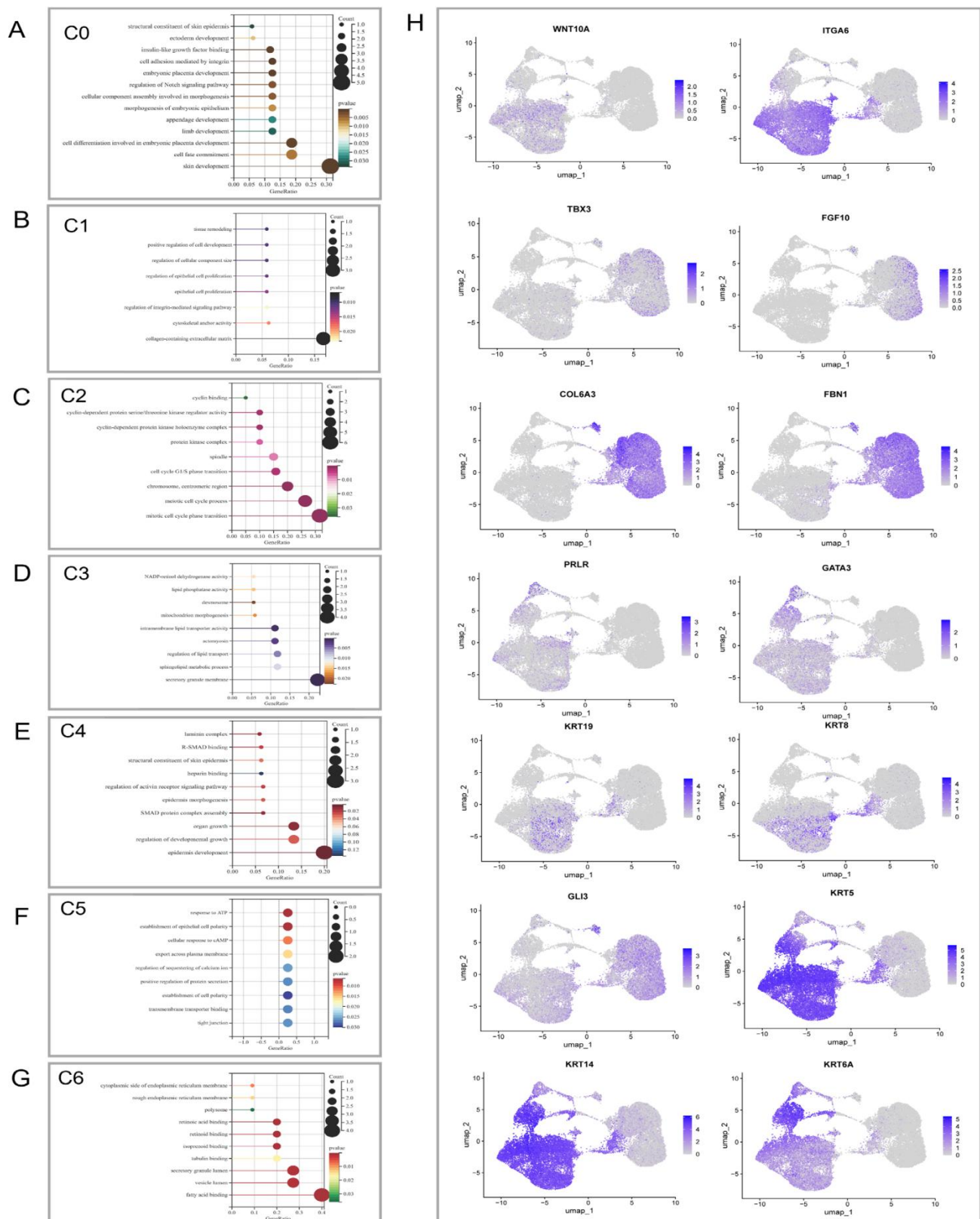
were activated, but also the pathways related to lipid synthesis and protein secretion were activated (Figure 2F-G). Furthermore, the scATAC-seq results further proved that the day 8 cells had obvious characteristics of mammary epithelial cells (Figure

2H-I). At the same time, GO analysis was performed on 7 clusters to define the identity of each cluster (Figure 3A-G), as well as the UMAP display of marker genes of each cluster (Figure 3H).



**Figure 2** Reprogramming of fibroblasts into mammary epithelial cells examined by single cell sequencing analysis. A. UMAP visualization illustrating the distribution of samples among groups (R-0d, R-4d, R-8d) in scRNA-seq data. B. The UMAP plot shows that the samples can be divided into seven different clusters. C. Bar chart shows the cell distribution in each cluster by sample. D. GO enrichment analysis of differentially expressed genes (DEGs) between 0d and 4d in scRNA-seq data. E. KEGG enrichment analysis

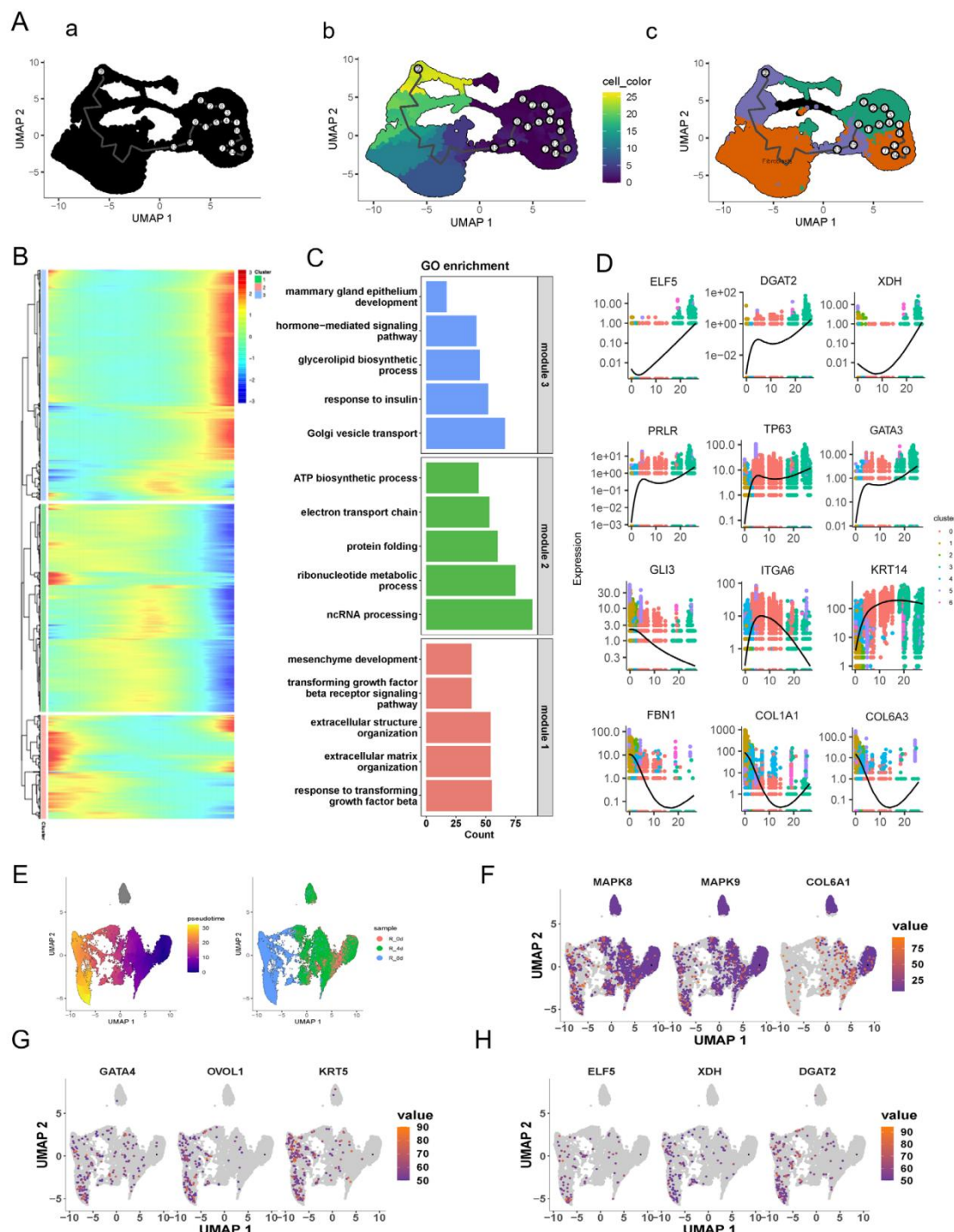
of differentially expressed genes between 0d and 4d in scRNA-seq data. F. GO enrichment analysis of differentially expressed genes between 0d and 8d in scRNA-seq data. G. KEGG enrichment analysis of differentially expressed genes between 0d and 8d in scRNA-seq data. H. GO enrichment analysis of differentially expressed genes between 0d and 8d in scATAC-seq data. I. KEGG enrichment analysis of differentially expressed genes between 0d and 8d in scATAC-seq data.



**Figure 3 Identification of the identities of each cluster in scRNA-seq data and presentation of the expression of marker genes. A-G: GO enrichment of 7 different clusters (C0, C1, C2, C3, C4, C5, C6) in scRNA-seq. H: UMAP plot depicting expression patterns of fibroblasts, development-related genes, and mammary genes across various clusters.**

Furthermore, we reconstructed the transdifferentiation trajectory of R-CiMECs. The cells were ordered and color-coded based on their predicted pseudotime (Figure 4A). Additionally, the genes were further categorized and clustered into three distinct modules according to their coexpression levels (Figure 4B). Module 1 (in red) exhibited enrichment in fibroblast gene sets, including FBN1, COL1A1, and COL6A3, with significant involvement in "extracellular structure organization" and "extracellular matrix organization" GO terms (Figure 4C-D). In Module 2 (in green), there was an enrichment of GO terms related to "ATP biosynthetic process," "ncRNA processing," and "protein folding." This suggests that the ongoing transdifferentiation process may be actively regulated at the transcriptional level (Figure 4C). Furthermore, expression levels of early embryonic development and

epithelial-related genes such as *GLI3*, *KRT14* and *ITGA6* significantly increased over time (Figure 4D). In Module 3 (in blue), GO terms associated with "hormone-mediated signaling pathway" and "mammary gland epithelium development" were enriched. Notably, representative genes like *ELF5*, *PRLR*, and *DGAT2* showed increasing expression along the pseudotime trajectory (Figure 4C-D), indicating a transition towards mammary cell lineage differentiation. Moreover, scATAC data revealed a decrease in activity for fibroblast marker genes *COL6A1*, *MAPK8*, and *MAPK9* with increasing induction days (Figure 4E-F), while markers for epithelial development (*GATA4*, *OVOL1*, *KRT5*) and lactation-related genes (*ELF5*, *XDH*, *DGAT2*) gradually became more active (Figure 4G-H). Collectively, these findings provide insights into the sequential molecular progression underlying successful reprogramming of fibroblasts into mammary epithelial cells.



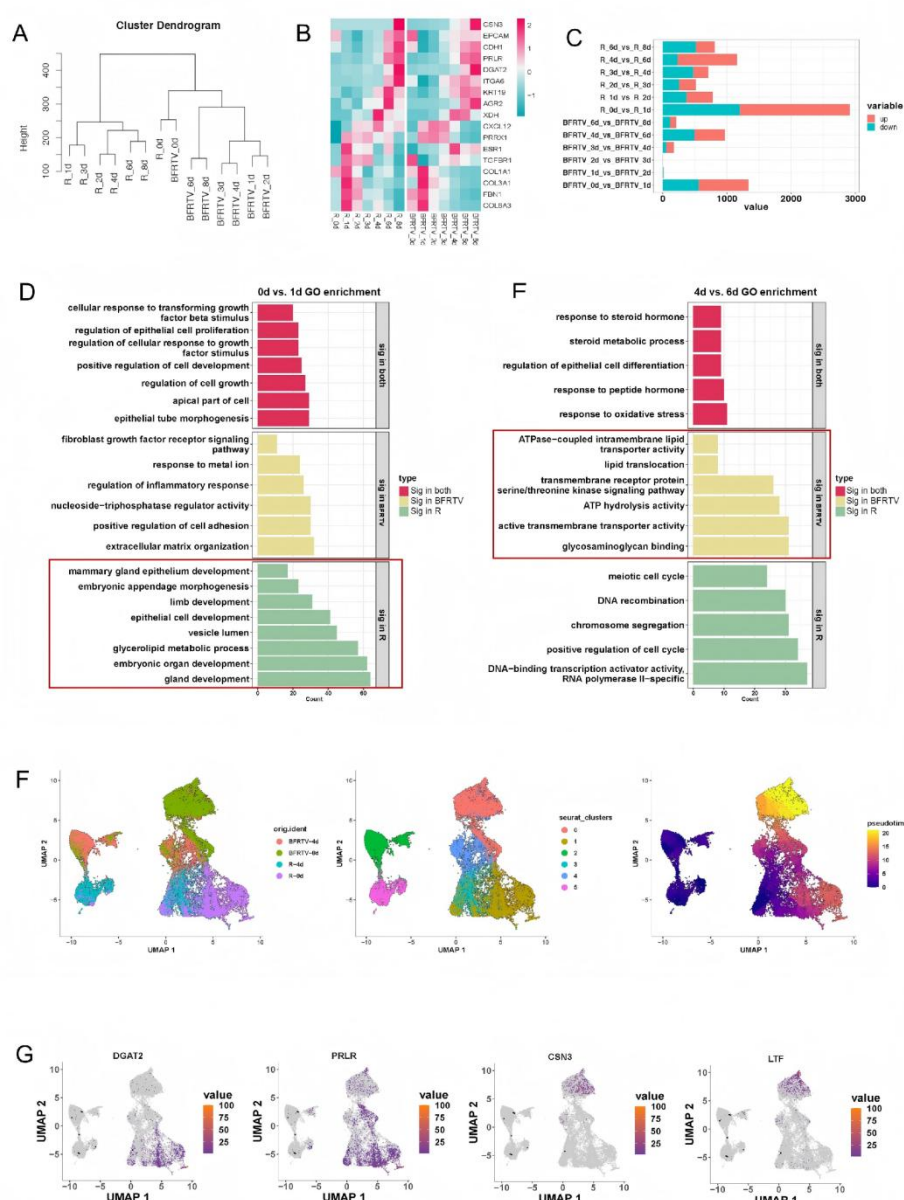
**Figure 4** Reconstruction of the RepSox-induced reprogramming trajectory in a pseudotime manner was performed to elucidate the temporal progression of cellular transformations. A. scRNA-seq data reveals trajectory reconstruction at three time points in a sample of single cells. a, Trajectory plot colored by Seurat cluster (starting point: cluster 1), and the gray lines represent the cell differentiation trajectories constructed by Monocle3. b, Trajectory plot colored by pseudotime. c, Trajectory plot colored by manually annotated cell types. B. A heatmap illustrates the expression patterns of pseudotime-dependent genes along the inferred pseudotime a-axis, which have been hierarchically categorized into three modules. C. The gene modules show enrichment for representative Gene Ontology terms. D. The dynamic expression trends of fibroblast, development and differentiation, and mammary epithelial marker genes during the cell differentiation process (pseudotime). The horizontal axis (pseudotime) represents the time course of cell differentiation (calculated by Monocle3), with larger values indicating that cells are closer to the end of differentiation. The vertical axis represents gene expression levels. E. Pseudotime analysis of scATAC-seq data (R-0d, R-4d, and R-8d samples) throughout RepSox-induced reprogramming is presented on the left side, while a Uniform Manifold Approximation and Projection (UMAP) plot shows sequential changes during this process on the right side. F. UMAP visualization demonstrates fibroblast gene activity in scATAC-seq data during sequential changes. G. UMAP visualization depicts epithelial cell gene activity in scATAC-seq data during sequential changes. H. UMAP visualization showcases mammary gene activity in scATAC-seq data during sequential changes.



# The regenerative function and reprogramming trend of R-CiMECs exhibit similarities to those of BFRTV-CiMECs.

To further elucidate the similarities and discrepancies between R-CiMECs and our previously reported BFRTV-CiMECs<sup>[8]</sup>, we conducted additional investigations to address these pertinent issues. First, we confirmed the regenerative capacity of R-CiMECs *in vivo*. The results demonstrated that hematoxylin and eosin staining clearly revealed well-defined acinar structures (Figure S1A). Staining with goat-specific casein antibodies (CSN2 and CSN3) revealed positive results in the structures of mammary glands derived from transplanted R-CiMECs; however, negative staining was observed in paraffin sections from untreated nude mouse mammary glands which confirmed the specificity of these antibodies for detecting goat antigens (Figure S1B). These findings strongly suggest that similar to BFRTV-CiMECs', R-CiMECs are capable of forming functional mammary gland structures.

Subsequently, we conducted a comparative analysis of gene expression patterns revealed by mRNA-seq data between R-CiMECs and BFRTV-CiMECs. The hierarchical clustering divided RepSox- and BFRTV-treated cells into two distinct groups, both of which exhibited transcriptional differences from the parent R-0d/BFRTV-0d (GEFs) (Figure 5A). However, both induction environments demonstrated that as the induction time increased, the expression of fibroblast marker genes gradually decreased while genes promoting mammary epithelial cell formation and lactation gradually increased (Figure 5B and S2A). The fuzzy clustering analysis of gene transcription trends during transdifferentiation also revealed that R-CiMECs (Figure S2B) exhibited similar transcriptional profiles to BFRTV-CiMECs (Figure S2C), with the distinction being that R-induced reprogramming processes, cell cycles, and epigenetic signaling were more actively engaged.

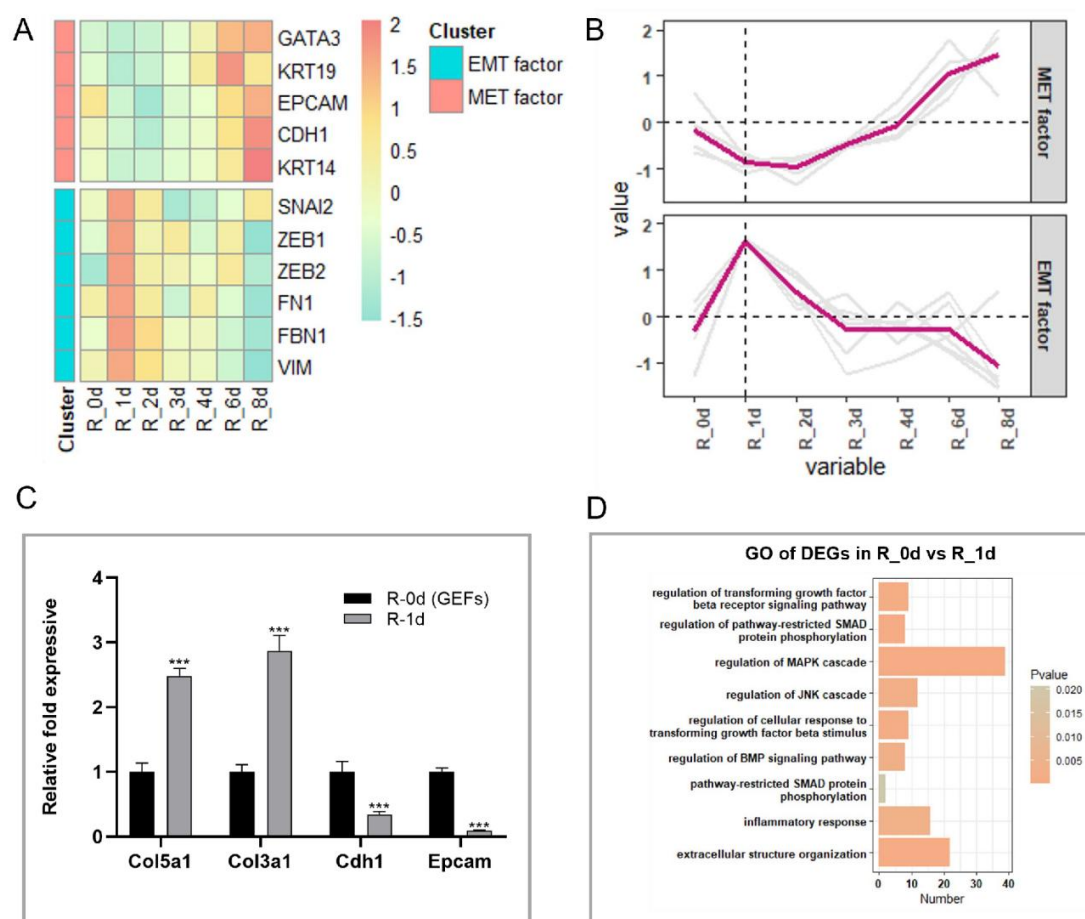


**Figure 5** Discrepancies in the maturation status between R-CiMECs and BFRTV-CiMECs. A.A hierarchical clustering analysis was performed on RepSox- and BFRTV-treated cells using unsupervised methods in bulk mRNA-seq data. B.The expression of fibroblast genes and mammary genes in these cells was visualized through a heatmap in bulk mRNA-seq data. C.Bar plot illustrating the differential expression of genes (DEGs) associated with the R-induced environment and BFRTV-induced environment in bulk mRNA-seq data. D.The GO enrichment map was generated to identify differentially expressed genes (DEGs) between R and BFRTV (5i) on the first day of induction in bulk mRNA-seq data. E.The GO enrichment map was generated to identify differentially expressed genes (DEGs) between R and BFRTV (5i) after 4 days of induction in bulk mRNA-seq data. F.UMAP visualization demonstrates the clustering(left and middle panels)and pseudotime distribution(right panel) of the combined scRNA-seq

Comparing the differentially expressed genes between two consecutive days, it was observed that the up- and down-regulated genes at 0d vs 1d and 4d vs 6d exhibited the most pronounced changes under both induction environments (Figure 5C). Further analysis revealed a significant enrichment of genes and items associated with embryonic development and mammary epithelial cell development during the early stage of transdifferentiation (0d vs 1d) induced by RepSox (R) (Figure 5D and Figure S2D-E). In contrast, in the middle and late stages of transdifferentiation (4d vs 6d), there was a notable enrichment of signals promoting lactation, such as lipid droplet transport and ATP utilization, in the BFRTV-induced environment (Figure 5E). The single cell transcription data of R-CiMECs-4d, R-CiMECs-8d, BFRTV-CiMECs-4d and BFRTV-CiMECs-8d were integrated and analyzed, and it was found that the early lactation genes DGAT2 and PRLR were highly expressed under RepSox induction system. The key genes of milk protein LTF and CSN3 were highly expressed in BFRTV induction system (Figure 5F-G). These findings suggest that RepSox-induced transdifferentiation facilitates rapid epithelialization and entry into the cellular plasticity stage and that BFRTV-CiMECs represent mature mammary epithelial cells that are more closely associated with the lactation stage. The discrepancy may be attributed to the heightened activity of epigenetic genes in R-CiMECs (Figure S2F).

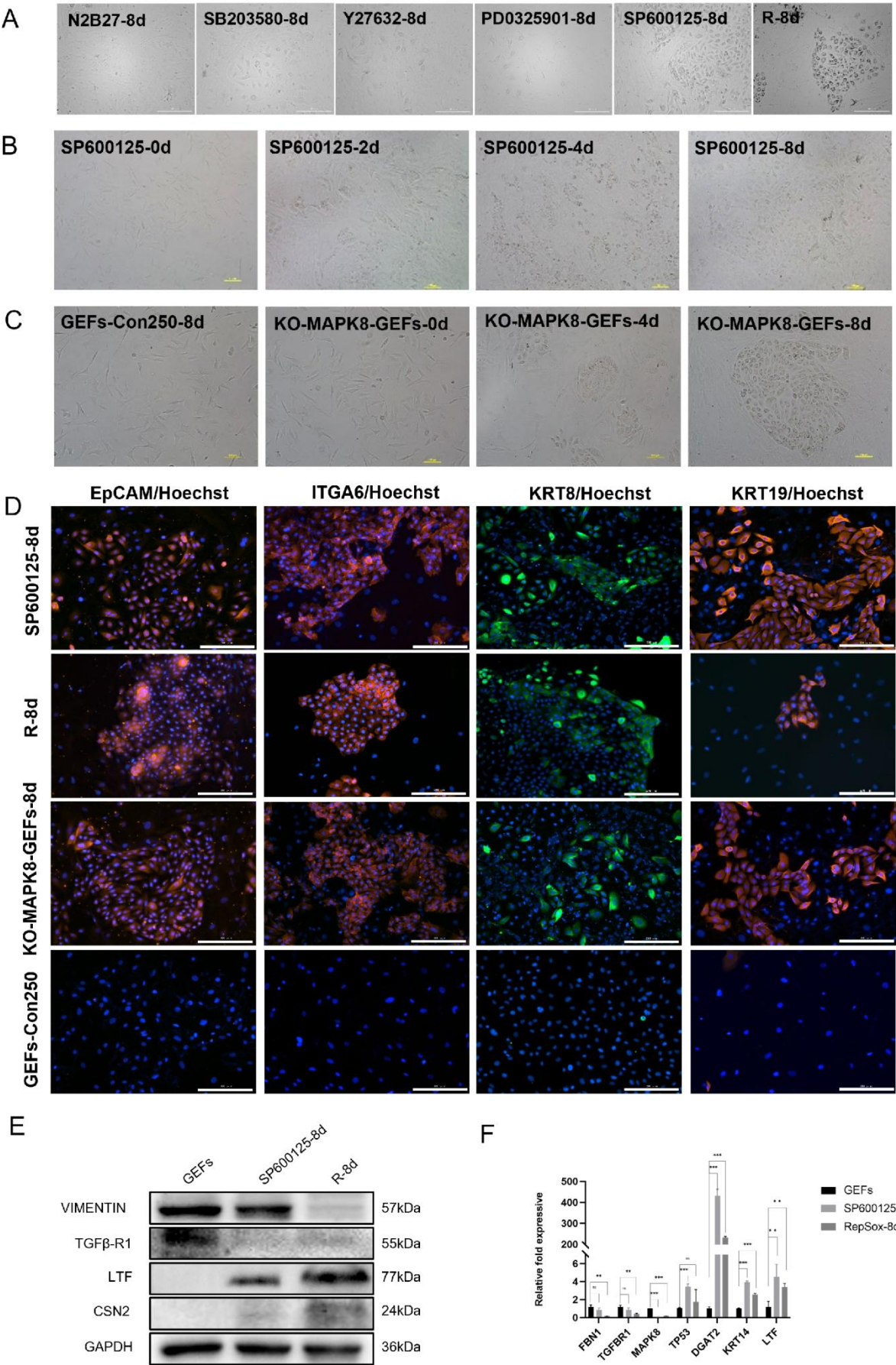
## The JNK signaling pathway governs the determination of mammary epithelial cell fate.

To further explore the downstream molecular mechanisms following repsox treatment, we analyzed the bulk RNA-seq data from the early induction stage. The results revealed a transient EMT-MET process on the first day of induction (Figure 6 A-B). Quantitative PCR results also confirmed this finding (Figure 6C). The differentially expressed genes between day 0 and day 1 were enriched in the MAPK and JNK signaling pathways (Figure 6D). We hypothesized that the downregulation of TGF  $\beta$  R1 may modulate the JNK/MAPK signaling pathway to regulate the reprogramming process. To investigate this regulatory mechanism, we conducted further experiments based on our previous single-cell sequencing analyses. Surprisingly, screening for small molecule compounds targeting the JNK/MAPK signaling pathway revealed that SP600125 (SP), a specific inhibitor of JNK1/MAPK8, induced goat ear fibroblasts (GEFs) to form epithelial cell colonies similar to R-CiMECs (Figure 7A-B). Additionally, we performed knockout experiments on JNK1/MAPK8 to determine its effects on reprogramming. The results demonstrated that downregulation of MAPK8 can induce iMEC generation (Figure 7C). The immunofluorescence staining results revealed that SP600125-induced mammary epithelial cells (SP600125-8d cells) and KO-MAPK8-GEFs-8d cells exhibited a similar expression pattern to R-8d cells (positive control), demonstrating the presence of epithelium-specific marker antigens EpCAM, ITGA6, KRT8, and KRT19 (Figure 7D). The WB (Figure 7E) and qRT-PCR (Figure 7F) results further confirmed that the inhibition of TGF  $\beta$  R1 or JNK1/MAPK8 significantly increased the expression levels of genes related to mammary gland development, differentiation and lactation (TP53, LTF, CSN2, DGAT2). These findings suggest that inhibiting JNK1 is crucial for converting nonmammary-derived stromal fibroblasts into mammary epithelial cells and indicate a potential role for the TGF  $\beta$  R1-JNK1 signaling pathway in mediating this reprogramming event.





**Figure 6** Identification of key nodes in the RepSox-induced reprogramming process through bulkRNA-seq data analysis. A.Heat map of representative genes showing a sequential EMT-MET during the chemical conversion process. B.The corresponding expression dynamic of EMT- and MET-related genes at the indicated time points during chemical reprogramming. Red lines indicate the average gene expression patterns in each cluster. C.qRT-PCR results show the expression of stromal and epithelial cell marker genes. (mean  $\pm$  SEM, n=3 biological replicates, \*p < 0.05, \*\*p < 0.01, \*\*\*p < 0.001, one-way ANOVA). D.GO analyses of DEGs in d1.



**Figure 7** The downregulation of JNK1/MAPK8 gene expression plays a crucial role in the formation of induced mammary epithelial cells.A.Reprogramming cellular morphology through the application of diverse small molecule compounds to regulate JNK/MAPK8 signaling.Scale bars, 200  $\mu$ m. B.The cellular morphology of mammary epithelial cell formation induced by SP600125 at different time points is illustrated.Scale bar, 100  $\mu$ m. C.The morphological landscape of induced mammary epithelial cell formation triggered by MAPK8 knockout was examined. GEF-Con250-8d cells were utilized as negative controls.Scale bar, 100  $\mu$ m. D.Immunofluorescence was performed to assess the expression of MEC marker proteins (EpCAM,ITGA6,KRT8,and KRT19) in SP600125-8d cells and KO-MAPK8-GEFs-8d cells.GEFs-Con250-8d cells were utilized as

negative controls, while R-8d cells served as positive controls. Scale bar, 200  $\mu$ m. E. Western blot results showed the expression changes of fibroblast marker proteins and lactation-related proteins in cell samples of cells treated with GEFs, SP600125-8d and R-8d. F. qRT-PCR analysis was conducted to investigate alterations in the expression of fibroblast-related marker genes and MEC marker genes in samples from GEFs, SP600125-8d cells, and RepSox-8d cells. Three biological replicates were included for each group. The data are presented as means  $\pm$  SEMs with statistical significance denoted by \*\* $p < 0.01$ , \*\*\* $p < 0.001$  based on one-way ANOVA.

## Discussion

Transdifferentiation offers an alternative approach to derive various functional cell types from somatic cells and holds great potential in clinical applications for tissue regeneration<sup>[25]</sup>. Previously, we demonstrated that the downregulation of TGF $\beta$  R1 can induce the conversion of fibroblasts into mammary epithelial cells<sup>[8]</sup>. In this study, we further elucidate the mechanism by which goat ear fibroblasts (GEFs) can be reprogrammed into milk-secreting mammary epithelial cells through the suppression of TGF $\beta$  R1 expression. Additionally, we establish an efficient platform for generating a large number of induced mammary epithelial cells in vitro. Given that goats possess two breasts similar to humans, they serve as suitable model animals for investigating human mammary gland development and regeneration. Moreover, goats are commonly utilized as mammary bioreactors for producing valuable recombinant human proteins used in disease treatment. Therefore, our findings may pave the way for a novel strategy to generate abundant functional mammary epithelial cells and mammary bioreactors capable of producing "culture dish milk" and recombinant proteins at a potentially lower cost compared to traditional methods.

Moreover, based on the results of single-cell sequencing analysis, we investigated reprogramming events and tracks. Our findings revealed that RepSox induction showed a trend similar to our previously reported small-molecule cocktail induction method. The results indicated that the small molecule cocktail (BFRTV) erased somatic cell epigenetic memory and subsequently induced cells to enter the mammary epithelial cell fate. However, R induction directly induced somatic cells to enter the mammary epithelial cell fate without significant signals for erasing epigenetic memory. Our previous studies<sup>[8]</sup> have shown that BFTV functions include epigenetic modification and promotion of mammary development. Therefore, under R induction, epigenetic barriers may force somatic cell reprogramming to resemble a "wavy" transcriptional wave which requires additional time for induced cells under R induction to reach lactation functionality. These findings indicate that TGF $\beta$  R1 inhibitor RepSox plays an indispensable role in cocktails and further emphasize downregulating TGF $\beta$  R1 as key in mammary epithelial cell fate decisions while other small molecules in cocktails (BFTV) may facilitate smoother reprogramming progress leading to faster secretion stages.

Importantly, our findings demonstrate that the downregulation of TGF $\beta$  R1 may exert regulatory effects by suppressing JNK1/MAP-K8 signaling, suggesting that the TGF $\beta$  R1-JNK1 pathway mediates these reprogramming events. JNKs play crucial role in processes such as proliferation, migration, apoptosis, and embryonic development<sup>[26]</sup>. Previous studies have shown that JNK negatively influences somatic cell reprogramming and its downregulation can enhance reprogramming efficiency<sup>[27]</sup>. In this study, we discovered that the downregulation of JNK1 following TGF $\beta$  R1 downregulation directly initiates somatic cell reprogramming rather than merely promoting it. Recent research indicates<sup>[28]</sup> that the JNK pathway acts as a major obstacle to chemical reprogramming. Suppression of JNK is essential for inducing cellular plasticity and a regeneration-like program by inhibiting proinflammatory pathways. During the induction of human cells into pluripotent stem

cells using combination small molecule compounds, researchers<sup>[29]</sup> found that the c-Jun N-terminal kinase inhibitor (JNKIN8) is a critical small molecule capable of inducing plasticity characteristics and activating SALL4 expression along with genes related to embryonic development. Furthermore, addition of proinflammatory factors TNF- $\alpha$  and IL-1 $\beta$  impairs the generation of plasticity features, further confirming that the JNK pathway serves as a significant barrier in recreating plasticity in human somatic cells through regulation of inflammation-related pathways.

In summary, our findings demonstrate that the suppression of the TGF $\beta$  R1-JNK1 signaling plays a crucial role in facilitating the entire reprogramming process. Chemical reprogramming leads to the transformation from fibroblast characteristics to mammary epithelial cell characteristics, providing valuable insights into fate determination of mammary epithelial cells. Therefore, our study not only sheds light on the novel regulatory effects of the TGF $\beta$  R1-JNK1 signaling pathway in somatic cell reprogramming but also establishes an alternative research platform for investigating mammary gland development and regeneration.

## Reference

1. Daley GQ (2015) Stem cells and the evolving notion of cellular identity. *Philos Trans R Soc Lond B Biol Sci* 370: 20140376.
2. Valcourt JR, Huang RE, Kundu S, Venkatasubramanian D, Kingston RE, et al. (2021) Modulating mesendoderm competence during human germ layer differentiation. *Cell Rep* 37:109990.
3. Morris SA, Daley GQ (2013) A blueprint for engineering cell fate: current technologies to reprogram cell identity. *Cell Res* 23:33-48.
4. Bacakova L, Zarubova J, Travnickova M, Musilkova J, Pajorova J, et al. (2018) Stem cells: their source, potency and use in regenerative therapies with focus on adipose-derived stem cells-a review. *Biotechnol Adv* 36:1111-1126.
5. Ahmad A, Ahsan H (2020) Ras-Mediated Activation of NF-kappaB and DNA Damage Response in Carcinogenesis. *Cancer Invest* 38:185-208.
6. Qayum A, Shah SM, Singh SK (2022) Divergent Signaling Pathways May Lead to Convergence in Cancer Therapy-A Review. *Cell Physiol Biochem* 56:180-208.
7. Zhang C, Cao S, Toole BP, Xu Y (2015) Cancer may be a pathway to cell survival under persistent hypoxia and elevated ROS: a model for solid-cancer initiation and early development. *Int J Cancer* 136:2001-2011.
8. Zhang D, Wang G, Qin L, Liu Q, Zhu S, et al. (2021) Restoring mammary gland structures and functions with autogenous cell therapy. *Biomaterials* 277:121075.
9. Malik N, Rao MS (2013) A review of the methods for human iPSC derivation. *Methods Mol Biol* 997:23-33.



10. Scesa G, Adami R, Bottai D (2021) iPSC Preparation and Epigenetic Memory: Does the Tissue Origin Matter? *Cells* 10.
11. Tan Y, Ooi S, Wang L (2014) Immunogenicity and tumorigenicity of pluripotent stem cells and their derivatives: genetic and epigenetic perspectives. *Curr Stem Cell Res Ther* 9: 63-72.
12. Chambers SM, Studer L (2011) Cell fate plug and play: direct reprogramming and induced pluripotency. *Cell* 145: 827-830.
13. Abdel-Raouf K, Rezgui R, Stefanini C, Teo J, Christoforou N (2021) Transdifferentiation of Human Fibroblasts into Skeletal Muscle Cells: Optimization and Assembly into Engineered Tissue Constructs through Biological Ligands. *Biology (Basel)* 10.
14. Han F, Liu Y, Huang J, Zhang X, Wei C (2021) Current Approaches and Molecular Mechanisms for Directly Reprogramming Fibroblasts Into Neurons and Dopamine Neurons. *Front Aging Neurosci* 13:738529.
15. Wapinski OL, Vierbuchen T, Qu K, Lee QY, Chanda S, et al. (2013) Hierarchical mechanisms for direct reprogramming of fibroblasts to neurons. *Cell* 155:621-635.
16. Cao N, Huang Y, Zheng J, Spencer CI, Zhang Y, et al. (2016) Conversion of human fibroblasts into functional cardiomyocytes by small molecules. *Science* 352:1216-1220.
17. Kim J, Efe JA, Zhu S, Talantova M, Yuan X, et al. (2011) Direct reprogramming of mouse fibroblasts to neural progenitors. *Proc Natl Acad Sci U S A* 108:7838-7843.
18. Chen YF, Tseng CY, Wang HW, Kuo HC, Yang VW, et al. (2012) Rapid generation of mature hepatocyte-like cells from human induced pluripotent stem cells by an efficient three-step protocol. *Hepatology* 55:1193-1203.
19. Scesa G, Adami R, Bottai D (2021) iPSC Preparation and Epigenetic Memory: Does the Tissue Origin Matter? *Cells* 10.
20. Han YC, Lim Y, Duffield MD, Li H, Liu J, et al. (2016) Direct Reprogramming of Mouse Fibroblasts to Neural Stem Cells by Small Molecules. *Stem Cells Int* 2016:4304916.
21. Wang YY, Sun TT, Yang P, Xu JJ, Liang Y, et al. (2022) [VCR, a Small Molecule Compound, Induces Reprogramming of Rat Fibroblasts into Neural Progenitor Cells under Hypoxic Condition]. *Sichuan Da Xue Xue Bao Yi Xue Ban* 53:790-797.
22. Qin H, Zhao AD, Sun ML, Ma K, Fu XB (2020) Direct conversion of human fibroblasts into dopaminergic neuron-like cells using small molecules and protein factors. *Mil Med Res* 7:52.
23. Zhao AD, Qin H, Sun ML, Ma K, Fu XB (2020) Efficient and rapid conversion of human astrocytes and ALS mouse model spinal cord astrocytes into motor neuron-like cells by defined small molecules. *Mil Med Res* 7:42.
24. Macias H, Hinck L (2012) Mammary gland development. *Wiley Interdiscip Rev Dev Biol* 1:533-557.
25. Xie X, Fu Y, Liu J (2017) Chemical reprogramming and transdifferentiation. *Curr Opin Genet Dev* 46:104-113.
26. Semba T, Sammons R, Wang X, Xie X, Dalby KN, et al. (2020) JNK Signaling in Stem Cell Self-Renewal and Differentiation. *Int J Mol Sci* 21.
27. Yao K, Ki MO, Chen H, Cho YY, Kim SH, et al. (2014) JNK1 and 2 play a negative role in reprogramming to pluripotent stem cells by suppressing Klf4 activity. *Stem Cell Res* 12:139-152.
28. Guan J, Wang G, Wang J, Zhang Z, Fu Y, et al. (2022) Chemical reprogramming of human somatic cells to pluripotent stem cells. *Nature* 605:325-331.
29. Guan J, Wang G, Wang J, Zhang Z, Fu Y, et al. (2022) Chemical reprogramming of human somatic cells to pluripotent stem cells. *Nature* 605:325-331.

#### Acknowledgments:

We thank Personal Biotechnology (Shanghai) Co., Ltd., for performing the scRNA-seq service and Majorbio Biomedical Science and Technology (Shanghai) Co., Ltd., for enabling the bulk mRNA-seq. This study was supported by grants from the National Natural Science Foundation of China (Grant No. 32160171), Natural Science Foundation of Guangxi (Grant No. 2023GXNSFBA026023), and Startup funds for high-level talents of Guangxi Academy of Medical Science.

#### Authors' contributions

D. D.Z. developed the chemical induction protocol, performed the experiments, analyzed the data, and wrote and revised the manuscript. L.S.Q. contributed to the bulk RNA-seq sequencing data analysis. Q.H.L. contributed to animal experiment. S.D. and L.F.S. contributed to the cell culture. Z.G.L. and H.P. contributed to molecular biology experiments. Z.Q.W. and X.Y. revised the manuscript. G.D.W. analyzed single-cell data and revised the manuscript. B.H. conceptualized the study, supervised the entire project, and revised the manuscript.

#### Ethical issues

The animal experiments were approved and monitored by the animal experiments ethical review committee of Guangxi University, Nanning, China.

#### Conflict of interest statement

The authors declare no competing interests.

#### SUPPORTING INFORMATION

Additional supplementary information is available for download and review in the supplementary information section located on the right-hand side of this article's HTML page.

**Figure S1.**

**Figure S2.**