Biochemistry and Biophysics Reports 45 (2026) 102448

Contents lists available at ScienceDirect

Biochemistry and Biophysics Reports

journal homepage: www.elsevier.com/locate/bbrep

Single-cell transcriptome sequencing revealing the microenvironment of breast fibroepithelial tumor

Tianrun Shu [a,1], Jing Wang [c,1], Fuying Yang [a,1], Dechun Yang [a,1], Shengnan Ren [a,1], Min Zhao [b,1], Yafeng Lv [a], Ying Song [a], Hanhui Xie [a], Xuan Yang [a], Cong Jiang [a,*], Rong Guo [a,**], Shicong Tang [a,***]

[a] Department of Breast Surgery, Caner Hospital of Yunnan Province, The Third Affiliated Hospital of Kunming Medical University, Kunming, China
[b] Department of Anesthesiology, Caner Hospital of Yunnan Province, The Third Affiliated Hospital of Kunming Medical University, Kunming, China
[c] Department of Gynecology, The Red Cross Center Hospital of Harbin, NO415, Xinyang Road, Daoli District, Harbin, 150000, China

ARTICLE INFO

Keywords:
Fibroepithelial tumor of the breast
Juvenile
Single-cell sequencing
Tumor microenvironments

ABSTRACT

Background: Giant juvenile fibroepithelial tumors of the breast are rare lesions that mainly affect adolescent girls. Their pathogenesis is poorly understood, and current management relies almost entirely on surgery, which may interfere with breast development. To clarify the biological basis of these tumors, we performed single-cell RNA sequencing.

Methods: We profiled 15,295 cells from three juvenile cases (ages 12–16) and 11,442 cells from one adult patient (age 42) using the BD Rhapsody™ platform. This dataset provided a detailed cellular atlas of fibroepithelial lesions and enabled direct comparison of tumor microenvironments between juvenile and adult cases.

Results: Distinct molecular signatures in epithelial cells and fibroblasts: Juvenile FT epithelial cells exhibited significant enrichment of proliferation-associated pathways (e.g., PI3K-Akt signaling, $p = 3.2 \times 10^{-5}$; GnRH signaling, $p = 1.8 \times 10^{-4}$), whereas adult lesions were dominated by estrogen receptor pathways ($p = 2.6 \times 10^{-6}$). **Juvenile-specific antigen-presenting fibroblasts**: Comprising 18.7 % of stromal cells (vs. <1 % in adults), these were immunohistochemically validated to express high levels of HLA-DRA and CD74. **Enhanced intercellular communication networks**: Juvenile lesions showed active stromal-epithelial crosstalk mediated by ligand-receptor interactions such as PTN-SDC4 ($p < 0.001$) and MDK-NCL ($p = 0.003$), suggesting a role in tumor progression.

Conclusions: Our findings reveal molecular features that may underlie the rapid growth of juvenile giant fibroepithelial tumors. The data point to hormone-independent proliferative programs and immune-modulating fibroblast subtypes as key contributors. These insights broaden our understanding of tumor biology and may guide the development of targeted, less invasive treatments for young patients.

1. Introduction

Fibroepithelial tumors (FTs) of the breast include fibroadenomas (FAs) and phyllodes tumors (PTs), both featuring proliferation of epithelial and stromal components [1]. Among them, giant juvenile fibroadenomas are fast-growing benign masses that usually appear during early breast development. Because they compress normal tissue, these lesions can cause both physical discomfort and psychological distress in adolescents [2]. Giant juvenile fibroadenomas are a leading cause of unilateral breast enlargement in adolescent girls. Although

* Corresponding author. Department of Breast Surgery, Caner Hospital of Yunnan Province, The Third Affiliated Hospital of Kunming Medical University, Kunming, 650118, Yunnan, China.
** Corresponding author. Department of Breast Surgery, Caner Hospital of Yunnan Province, The Third Affiliated Hospital of Kunming Medical University, Kunming, 650118, Yunnan, China.
*** Corresponding author. Department of Breast Surgery, Caner Hospital of Yunnan Province, The Third Affiliated Hospital of Kunming Medical University, Kunming, 650118, Yunnan, China.
E-mail addresses: greenm00d@outlook.com (C. Jiang), guorong2320@126.com (R. Guo), tang_shicong@126.com (S. Tang).
[1] Tianrun Shu, Jing Wang, Fuying Yang, Dechun Yang, Shengnan Ren and Min Zhao contribute equally to this work.

https://doi.org/10.1016/j.bbrep.2026.102448
Received 18 April 2025; Received in revised form 29 December 2025; Accepted 8 January 2026

benign, they are associated with a higher long-term risk of breast cancer [3] and frequent recurrence [4]. Safer, non-surgical treatments are therefore needed, but the biological origin of these lesions remains poorly defined.

Phyllodes tumors are less common than fibroadenomas and tend to occur in older women. They are classified as benign, borderline, or malignant according to their histological appearance and behavior [5]. In adults, fibroadenomas are thought to arise from estrogen stimulation [6]. Yet, clinical responses to estrogen receptor modulators such as tamoxifen have been disappointing [7], and surgical excision, though routine, can lead to scarring and tissue loss [8,9]. Understanding their pathogenesis is crucial for the development of safer and more effective therapies.

Previous genomic sequencing studies have identified recurrent mutations, such as those in the MED12 gene, and other genomic alterations in breast fibroepithelial lesions [3,6]. Additionally, pediatric fibroadenomas exhibit morphological differences compared to adult fibroadenomas, and molecular studies have revealed distinct gene mutations between juvenile and conventional fibroadenomas [10]. Single-cell studies have explored epithelial components of adult fibroadenomas and their resistance to estrogen modulators [11]. Recent single-cell transcriptomic analyses of breast cancers have also identified metastasis-related cellular subsets and clarified tumor–microenvironment interactions [12]. Yet, how epithelial and stromal cells differ transcriptionally in fibroepithelial lesions remains unclear, in part due to the lack of suitable human models. Single-cell sequencing has transformed our ability to examine the genome and transcriptome at high resolution. It enables detailed mapping of cellular diversity and interactions, providing new insight into disease mechanisms [13,14]. Since Tang and colleagues first introduced single-cell transcriptome sequencing (scRNA-seq) in 2009 [14], advances such as BD Rhapsody [15], CITE-seq [16], and 10 × Genomics [17] have driven major progress in this field. The single-cell sequencing workflow involves isolating highly active single cells from tissues, lysing the cells to capture RNA, preparing single-cell suspensions, converting mRNA into cDNA for sequencing, and using unique molecular identifiers (UMIs) to quantify gene expression at the single-cell level [18]. This technology makes it possible to profile gene expression in individual cells, offering insight into tumor biology and the cellular complexity of the tumor microenvironment [19]. Despite these advancements, the specific roles of different cell subsets and their interactions in disease pathogenesis remain unresolved.

In this work, we performed single-cell sequencing and bioinformatic analysis of tumor samples from juvenile and adult patients with breast fibroepithelial lesions. To avoid potential harm to breast development in adolescent patients, adjacent normal breast tissue was not sampled. Our goal was to characterize the cellular subsets that contribute to the pathogenesis of juvenile and adult fibroepithelial lesions, identify key genes and pathways in juveniles, and define differences in cellular differentiation between the two groups. By investigating the interactions between different cell subsets, we aim to elucidate the pathogenesis of juvenile breast fibroepithelial lesions.

2. Methods

2.1. Subjects patient number sex age pathological diagnosis mass size (Table S1)

This study enrolled three juvenile patients and one adult patient with breast fibroepithelial lesions. Tissue samples were obtained from the First Department of Breast Surgery, Third Affiliated Hospital of Kunming Medical University. The study was approved by the hospital's Medical Ethics Committee, and informed consent was obtained from all participants or their legal guardians (Approval No. KYCS2024-083, Date: 2024-03-06). Histopathological examination confirmed fibroepithelial lesions in all cases (Table S1).

2.2. Experimental instruments and reagents

Detailed information on reagents and instruments used in this study can be found in Supplementary Material (Tables S2 and S3).

2.3. Single-cell RNA sequencing

2.3.1. Preparation of cell suspension

Fresh tumor specimens were minced, enzymatically digested, and filtered. The resulting cell suspension was processed for RNA capture as detailed in Supplementary Material (Table S4).

2.3.2. Data preprocessing

Prior to analysis, cell identity and viability were verified to ensure accurate sample labeling. Sequencing data were processed using the standard BD Rhapsody™ Whole Transcriptome Analysis pipeline on the Seven Bridges platform (https://www.sevenbridges.com), which performs quality control, filtering, annotation, and generation of gene expression matrices. FASTQ files generated from NovaSeq sequencing were aligned to the human reference genome (hg38) using CellRanger to produce expression matrices for each sample. Quality control was conducted using the Seurat package [20] in R. Low-quality cells were excluded based on two main parameters: potential doublets and mitochondrial gene content. Normally, mitochondrial transcripts remain contained within intact cells; therefore, a high mitochondrial gene percentage indicates cellular damage. Cells with mitochondrial RNA content outside the range of 15–30 % were removed. Doublets were detected and excluded using the DoubletFinder package [21]. Only high-quality single cells were retained for downstream analyses.

2.3.3. Data integration and batch effect removal

To investigate the pathogenesis of fibroepithelial lesions (FELs) in both juvenile and adult patients, multiple samples were analyzed by scRNA-seq to increase data robustness. However, variations in reagents, processing time, and tissue handling can introduce batch effects that obscure biological differences. To mitigate these effects, datasets were integrated and batch-corrected using the Harmony algorithm, which aligns shared cellular features across samples while preserving biological heterogeneity.

2.3.4. Dimensionality reduction clustering

The single-cell gene expression matrix consists of genes as rows and cells as columns. Because of its high dimensionality, clustering directly on raw data is computationally inefficient and prone to noise. To overcome this, dimensionality reduction was performed to project the data into a lower-dimensional space that retains key biological information while minimizing redundancy. This step improves clustering accuracy and computational efficiency.

Principal component analysis (PCA) was first applied to identify major axes of transcriptional variation and highlight genes contributing most to variance in the dataset. Uniform Manifold Approximation and Projection (UMAP) was then used for nonlinear dimensionality reduction and visualization of cellular distributions.

Cells were clustered based on their transcriptional profiles, grouping those with similar gene expression into distinct populations. Cluster identities were determined by comparing juvenile and adult groups, allowing assessment of cell-type composition and transcriptional differences between the two cohorts.

2.3.5. Cell annotation

Cell type annotation is a crucial step in single-cell analysis. After clustering, cells with similar transcriptional profiles were grouped into distinct clusters. Each cluster was annotated based on established marker genes characteristic of known cell types. Reference databases such as CellMarker and SCTPA were used to guide marker selection and improve annotation accuracy.

To reduce subjectivity and improve reproducibility, automated annotation was first performed using the SingleR [22] algorithm, which compares each cell's expression profile to reference datasets and assigns the most correlated cell identity. Annotations were then manually reviewed and cross-validated against known marker genes from published studies to ensure accuracy and consistency.

2.3.6. Difference analysis

Differential expression analysis was performed to characterize transcriptional differences between cell populations across groups. Differentially expressed genes (DEGs) were defined as those with an absolute $\log_2$ fold change greater than 1. The most significant DEGs within each cluster were used to confirm annotation accuracy and to explore pathways activated or suppressed during disease progression. This analysis highlights genes that vary during dynamic processes such as differentiation or phenotypic transition, providing insight into molecular mechanisms underlying disease development.

2.3.7. GO and KEGG enrichment analysis

Gene Ontology (GO) provides a structured vocabulary for describing gene functions across species and includes three domains: Biological Process (BP), Molecular Function (MF), and Cellular Component (CC). The Kyoto Encyclopedia of Genes and Genomes (KEGG) integrates genomic, biochemical, and pathway information to characterize biological systems at the molecular level.

After identifying DEGs between adult and juvenile groups, pathway enrichment analysis was performed using GO and KEGG databases to identify significantly enriched biological functions and pathways, thereby elucidating the functional states of different cell populations.

2.3.8. CellChat intercellular communication analysis

Cell–cell communication networks were analyzed using the CellChat [23] package in R, which infers intercellular signaling interactions based on known ligand–receptor pairs and expression patterns. This probabilistic framework quantifies the strength and directionality of signaling events between cell types and provides diverse visualization tools to map key signaling pathways and cell-specific roles.

Using CellChat, we examined communication patterns among cell populations to characterize the tumor immune microenvironment and identify signaling pathways potentially relevant to therapeutic intervention.

2.3.9. Analysis of transcription factors in each cell type in FT

Transcription factors (TFs) regulate gene expression by binding specific DNA motifs and modulating chromatin accessibility, transcription initiation, and cofactor interactions [24]. Their tissue-specific expression reflects specialized cellular functions. Analysis of TF activity therefore offers insight into regulatory mechanisms that shape cell identity and environmental responses [25].

Transcription factor regulatory networks were reconstructed using pySCENIC [26], a Python-based implementation of the SCENIC framework [27]. This method efficiently infers gene regulatory modules and TF activity from single-cell RNA-seq data. We applied pySCENIC to quantify TF activity across cell types and groups, enabling identification of key regulators driving cell-type–specific transcriptional programs.

2.3.10. Subpopulation segmentation and quasi-time sequence analysis of epithelial cells and fibroblasts

To characterize heterogeneity among epithelial and fibroblast subtypes—the dominant components of the fibroepithelial tumor microenvironment—we performed additional dimensionality reduction and clustering. Cellular differentiation involves progressive transitions between states driven by gene activation or repression, which can be captured through trajectory inference.

Cell differentiation trajectories were reconstructed using Monocle 2 [28], which orders individual cells along a pseudotime axis based on transcriptional similarity. This analysis enabled visualization of dynamic gene expression changes and inference of lineage relationships among epithelial and fibroblast subpopulations.

2.3.11. Analysis of metabolic characteristics of epithelial cells in different FT groups

To explore cellular metabolism, metabolic activity was analyzed using the scMetabolism [29] package in R. Expression data from epithelial cells in juvenile and adult FT groups were evaluated to assess single-cell metabolic states. The scMetabolism pipeline was used to compute metabolic scores and visualize activity patterns. Heatmaps based on KEGG pathway annotations (https://www.kegg.jp/) were generated to display average expression levels of metabolic genes across groups.

2.4. Immunohistochemistry

Immunohistochemistry (IHC) was performed on tumor tissue sections to validate specific cell types and marker expression. Detailed IHC procedures, including antigen retrieval and antibody staining, are described in Supplementary Material (Table S5).

3. Result

3.1. Single cell data dimensionality reduction clustering results

After quality control, a total of 26,737 high-quality cells were retained from three juvenile fibroepithelial tumor samples and one adult sample. PCA was first applied for dimensionality reduction, and the top 30 components were selected for downstream analysis. UMAP was subsequently used for low-dimensional visualization of cell clusters. At a resolution of 0.2, the data were partitioned into 13 distinct clusters (Fig. 1D), which were further annotated and visualized. DEGs were identified across clusters using the FindAllMarkers function in Seurat. Cluster-specific marker genes were significantly enriched in distinct cell populations and largely absent in others (Table 1), confirming the transcriptional identity of each cluster.

3.2. Identification of cell types

Based on cluster-specific marker expression, the 13 clusters were annotated into six major cell types: (1) neutrophils (FCGR3B$^+$), (2) T/NK cells (CD3E$^+$), (3) macrophages (CD68$^+$), (4) epithelial cells (EPCAM$^+$), (5) fibroblasts (COL3A1$^+$), and (6) endothelial cells (VWF$^+$) (Fig. 1B–E). Marker-gene visualization confirmed the accuracy of these annotations. The relative abundance of these cell types varied markedly among patients (Fig. 1F), underscoring the heterogeneity of breast fibroepithelial tumors. Bar-plot analysis (Fig. 1G) revealed that epithelial and fibroblast populations predominated in both age groups, suggesting a close interplay between these two cell types in the pathogenesis of FT.

3.3. Marker gene expression of each cell type

Bubble-plot visualization (Fig. 1C) demonstrated that signature genes were selectively and strongly expressed in their respective cell populations, confirming the robustness of cell-type annotation and the clear transcriptional separation among populations.

3.4. Results of GSVA enrichment analysis for each cell type in FT

Using Gene Set Variation Analysis (GSVA) package (version 1.46.0) and hallmark gene sets, we compared signaling-pathway activity among clusters (Fig. 1H). Epithelial cells displayed enrichment in early/late estrogen-response and oxidative-phosphorylation pathways, whereas fibroblasts were enriched for EMT and Wnt/β-catenin signaling.

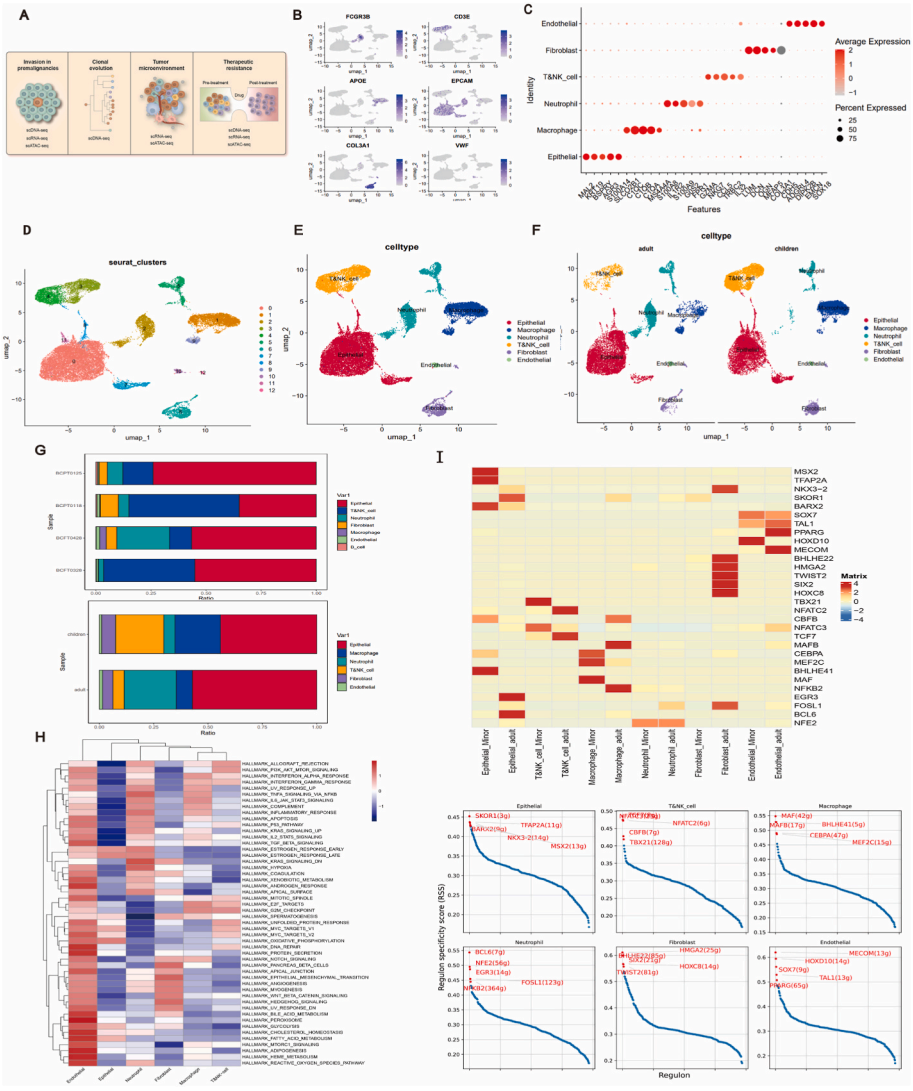

Fig. 1. | **Single-cell data analysis of FT.**
A Single-cell sequencing applications.
B Gene signature map: marker genes and expression levels (darker color = higher expression).
C Bubble plots: six cell type genes in FT (size = enrichment; color = significance).
D UMAP visualization of cell clustering and dimensionality reduction
E UMAP visualization of six cell types in FT
F UMAP of adult and juvenile FEL cells
G Top: proportions of six cell types in four samples; bottom: proportions in different
I Transcription factor (TF) analysis in FEL cell types. (Top) Top 5 TFs per cell type in both groups. (Bottom) TF ranking by regulon specificity score (RSS).
H GSVA enrichment analysis of signaling pathways across FT cell types (red: significantly enriched pathways; blue: downregulated pathways).

Neutrophils were active in TNF–NFκB and KRAS signaling; macrophages in interferon-α and IL6–JAK–STAT3 responses; T cells in PI3K–AKT–mTOR and IFN-γ signaling; and endothelial cells in fatty-acid metabolism. These pathway patterns reveal coordinated metabolic and immune activities within the FT microenvironment.

3.5. FT transcription factor analysis for each cell type

Regulatory-network inference with pySCENIC identified dominant TFs for each cell population (Fig. 1I). TFAP2A and BARX2 were associated with estrogen-receptor regulation and proliferative signaling in epithelial cells. Adult epithelial cells showed strong EGR3 and BCL6 activity, while juvenile fibroblasts selectively expressed SKOR1, and adult fibroblasts expressed HMGA2, TWIST2, and SIX2. These distinct TF profiles suggest age-specific transcriptional control of epithelial and stromal differentiation.

3.6. Epithelial cell annotation

Because epithelial cells were the dominant population in FT, they were isolated for higher-resolution analysis. After reclustering using t-SNE and Seurat functions (FindNeighbors, FindClusters), 11 transcriptionally distinct subclusters were identified. Based on canonical markers, epithelial cells were categorized into three subtypes: basal, ductal type 1, and ductal type 2. Juvenile samples contained a larger proportion of basal cells, whereas adult samples were enriched in ductal type 2 cells (Figs. 2A and 3E). Estrogen-receptor (ER) expression was elevated in basal and ductal 2 subtypes of the juvenile group (Fig. 3A), suggesting age-related differences in hormone responsiveness.

Table 1
Expression of the top 3 marker genes in each cell population after dimensionality reduction clustering of FT cells.

gene	p_val	avg_log2FC	pct.1	pct.2	p_val_adj	cluster
AGR3	0	2.80448	0.756	0.117	0	0
MYB	0	2.839646	0.69	0.1	0	0
PDZK1	0	2.804202	0.608	0.087	0	0
HLA-DQA1	0	5.05931	0.985	0.067	0	1
SLCO2B1	0	5.053447	0.703	0.021	0	1
HLA-DQA2	0	5.357918	0.623	0.021	0	1
S100A8	0	6.497669	0.979	0.06	0	2
IL1R2	0	6.376399	0.899	0.03	0	2
CXCR2	0	6.37961	0.617	0.015	0	2
CCL5	0	3.557566	0.752	0.097	0	3
TRBC2	0	4.565797	0.636	0.055	0	3
CD69	0	4.015615	0.667	0.096	0	3
GNLY	0	6.708889	0.86	0.028	0	4
PRF1	0	6.068262	0.776	0.025	0	4
GZMB	0	5.868239	0.629	0.015	0	4
LYZ	0	5.493938	0.919	0.18	0	5
CTSS	0	3.676096	0.957	0.251	0	5
FCN1	0	6.139432	0.706	0.018	0	5
DCN	0	9.567222	0.999	0.02	0	6
LUM	0	9.843907	0.977	0.007	0	6
OGN	0	9.622056	0.822	0.003	0	6
GABRP	0	5.952134	0.806	0.034	0	7
SFRP1	0	4.929358	0.832	0.081	0	7
NDRG2	0	3.664181	0.808	0.175	0	7
TOP2A	0	6.579666	0.867	0.024	0	8
BIRC5	0	6.611864	0.724	0.011	0	8
PBK	0	7.826611	0.576	0.004	0	8
RNASE1	0	8.1919	0.971	0.034	0	9
FOLR2	0	7.289606	0.882	0.019	0	9
LYVE1	0	8.172953	0.699	0.005	0	9
EMCN	0	9.667542	0.794	0.002	0	10
ADGRL4	0	9.875455	0.73	0.001	0	10
DIPK2B	0	9.715242	0.702	0.001	0	10
FRMD6	4.21E-26	1.306659	0.563	0.242	1.21E-21	11
ERBB3	9.27E-25	1.105554	0.726	0.38	2.66E-20	11
IRX5	2.52E-22	1.105693	0.68	0.358	7.23E-18	11
RGS5	0	9.786813	0.942	0.017	0	12
MYLK	0	7.06784	0.867	0.022	0	12
SYNPO2	0	7.59981	0.78	0.013	0	12

gene: Gene name (marker gene).**p_val**: Raw p-value (statistical significance), where smaller values indicate more significant differential expression. **avg_log2FC**:Average log2 fold change (expression difference), with positive values indicating upregulation and negative values indicating downregulation. **pct.1**:The percentage of cells expressing the gene in the target cluster. **pct.2**:The percentage of cells expressing the gene in all other clusters. **p_val_adj**: Adjusted p-value (e.g., Bonferroni or FDR correction) for multiple hypothesis testing. **Cluster**:fibroepithelial tumor cells subpopulation identifier (0–12), representing distinct cell types or states.

3.7. DEGs and GSEA enrichment analysis

DEGs were identified using FindAllMarkers in Seurat (v4.0.6) with adjusted $p < 0.05$ (Table 2). The analysis revealed distinct gene expression patterns: genes such as KLK7, KLK6, LCN2, KRT81, and MALL were highly expressed in the Luminal1 subpopulation; MKI67, TOP2A, TPX2, CENPF, and TYMS showed elevated expression in the Basal subpopulation; and SOCS3, ERBB4, PGR, INPP4B, and BTG2 were prominently expressed in the Luminal2 subpopulation (Fig. 2B) (Table 3 and Table 4).

In order to explore the role of up-regulated DEGs in epithelial cells in the occurrence and development of FT, we used GO enrichment analysis of the obtained differential genes (Fig. 2D), which showed that "response to hormones", "proliferation of epithelial cells" and other pathways were significantly enriched in the BP. In the aspect of CC, "proton transport", "extracellular matrix organization" and other pathways were significantly enriched. In the aspect of MF, "DNA binding transcription activator activity", "extracellular matrix structural component" and other pathways were significantly enriched. Table 5 shows the TOP5 pathways of BP, CC and MF sorted by pvalue by GO enrichment analysis of all differential genes obtained from FT. We found that these up-regulated genes were mainly enriched in the response to steroid hormones, epithelial cell proliferation, gland development and other pathways. This suggests that the response of the mammary gland to hormones is likely to be involved in the pathogenesis of breast fibroepithelial cell lesions.KEGG analysis (Fig. 3C) identified MAPK signaling as the most significant pathway, linking epithelial proliferation with tumor expansion [30].

GSEA (Fig. 2C) indicated that both groups shared enrichment in estrogen signaling, but juvenile epithelial cells showed stronger activation of PI3K–AKT, MAPK, and GnRH signaling. The presence of endocrine-resistance pathways in juveniles implies reduced responsiveness to anti-estrogen therapy, aligning with clinical observations that endocrine drugs are ineffective in these lesions. These results highlight age-specific proliferative programs rather than generic growth activity. Moreover, it is highly correlated with the recurrence of fibronenoma [11], which further explores the role of epithelial cells in the pathogenesis of FT.

3.8. Pseudo-time analysis of epithelial cells

Epithelial-cell developmental trajectories were reconstructed with Monocle 2 using variable genes identified in Seurat. Cells from both age groups were ordered along a pseudotime axis (Fig. 3D), revealing three differentiation states originating from basal cells and progressing toward ductal 1 and ductal 2 lineages. These findings align with previous studies [31].

These genes into clusters based on their expression patterns. GO enrichment analysis was conducted using the ClusterGVis package (version 0.1.1). The results (Fig. 2E) demonstrated distinct pathway enrichments: in the

Juvenile group, the early developmental stage (Pre-branch) of epithelial cells was significantly enriched in pathways such as "regulation of chromosome organization" and "DNA repair." One terminal stage (Fate 1) showed enrichment in "cytoplasmic translation" and "ribosome biogenesis," while the other terminal stage (Fate 2) was enriched in "cytoskeleton-dependent extracellular transport," "response to steroid hormones," and "gland morphogenesis." In the adult group, the early developmental stage was enriched in pathways like "negative regulation of the cell cycle" and "sister chromatid separation." Fate 1 was associated with "epidermis development," "skin development," and "cell chemotaxis," while Fate 2 was enriched in "response to endoplasmic reticulum stress." These differences suggest distinct maturation trajectories and hormonal sensitivity between juvenile and adult epithelial populations.

3.9. Metabolic heterogeneity of epithelial cells in different groups of FT

Metabolic-pathway scoring at the single-cell level (Fig. 3B and C) revealed striking metabolic divergence between age groups. Juvenile epithelial cells exhibited enhanced steroid-biosynthesis and amino-acid (lysine, sulfur) metabolism, reflecting active anabolic and proliferative states, whereas adult epithelial cells showed higher tricarboxylic-acid-cycle and fatty-acid-oxidation activity, consistent with more oxidative and differentiated metabolism. These findings emphasize metabolic reprogramming as a key age-dependent feature of FT epithelial cells.

3.10. Fibroblast subpopulation analysis

Given the central role of fibroblasts in FT pathogenesis, we further dissected fibroblast heterogeneity. PCA and UMAP were applied to recluster fibroblasts at a resolution of 0.6, identifying 10 distinct clusters. Differential gene expression (FindAllMarkers) revealed six fibroblast subtypes: Myofibroblasts (mCAFs), inflammatory-associated fibroblasts (iCAFs), vascular fibroblasts (vCAFs), tumor-associated fibroblasts (tCAFs), antigen-presenting fibroblasts (apCAFs), and dividing

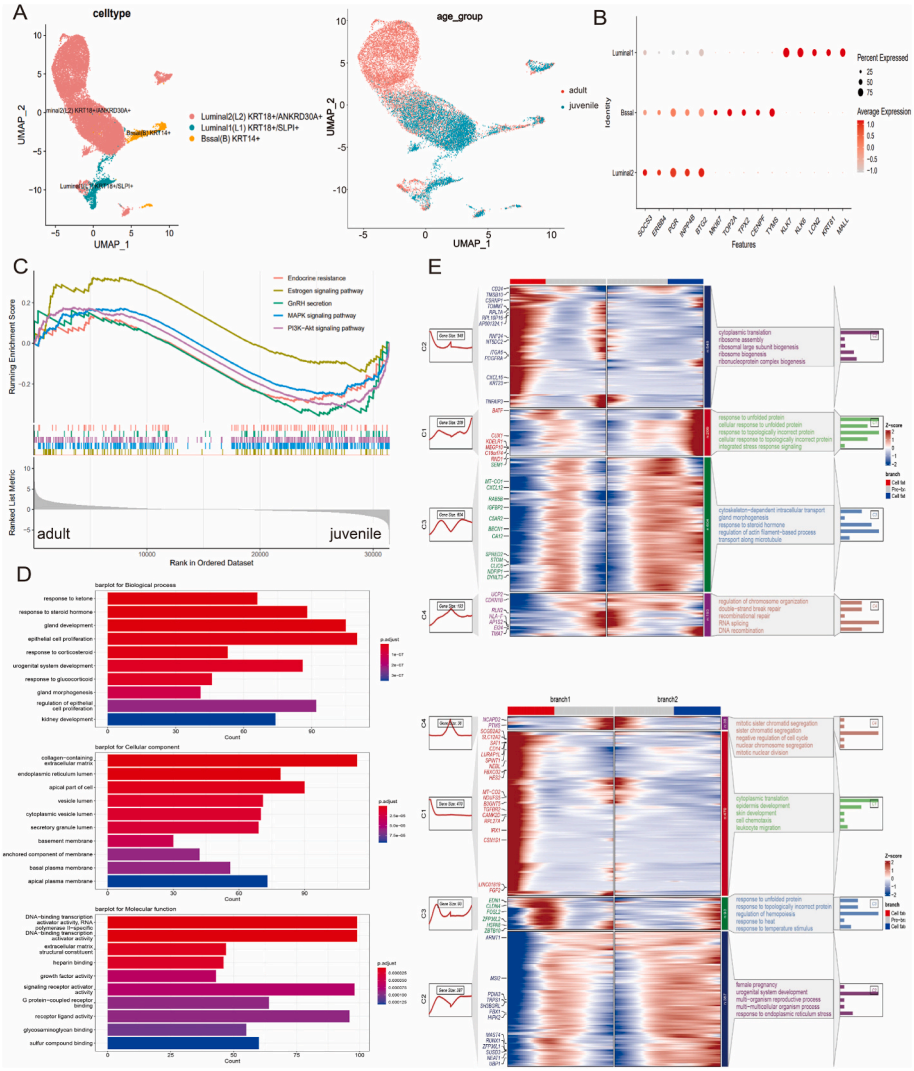

Fig. 2. | **Analysis of epithelial cells.(1)**
A UMAP visualization of epithelial cell subsets (left: annotation; right: distribution in adult and juvenile groups).
B Bubble plot of differential gene analysis in epithelial cells of adult and juvenile FT (bubble size: gene enrichment; color: significance).
C GSEA enrichment analysis (left: adult; right: juvenile).
D GO enrichment analysis of differentially expressed genes in epithelial cells (red: significance).
E Functional changes in epithelial cells during development (top: juvenile group; bottom: adult group).

fibroblasts (dCAFs) (Fig. 4A). This analysis highlights the broad functional diversity of fibroblasts within the FT microenvironment.

Heatmap analysis demonstrated that marker genes were specifically enriched in their respective cell types. For example, POSTN was predominantly expressed in mCAFs, PLA2G2A in iCAFs, RGS5 in vCAFs, MKI67 in dCAFs, and HLA-DRA and CD74 in apCAFs, while HSPH1 was mainly expressed in tCAFs (Fig. 4B). These expression patterns are consistent with previous studies on breast fibroblast heterogeneity [32]. Additionally, the distribution of fibroblast subsets revealed distinct differences between the juvenile and adult groups (Fig. 4C). In the juvenile group, mCAFs were the predominant fibroblast subtype, whereas they were least abundant in adults. apCAFs, characterized by immune-related gene expression (HLA-DRA, CD74), were also more prevalent in the juvenile group, a finding corroborated by immunohistochemical validation (Fig. 5A). In contrast, adult fibroblasts exhibited a higher proportion of iCAFs, followed by tCAFs and vCAFs, suggesting age-dependent shifts in fibroblast composition and immune functionality within the FT microenvironment.

3.11. *Analysis of fibroblast differential genes*

DEGs in fibroblasts were analyzed, revealing 857 upregulated and 118 downregulated genes (Fig. 4D). Angiogenesis-related genes (*MMP16, PTN*) were downregulated [33,34]. while genes linked to cell proliferation, stress response, and hypoxia (*KLF4, ATF3, HIF3A*) were upregulated [35–37], indicating a shift toward proliferative and hypoxic states in FT fibroblasts.

3.12. *GO enrichment analysis of differentially expressed genes in fibroblasts*

GO enrichment of differential genes identified significant pathway activation across BP, CC and MF (Fig. 5E). BP pathways involved extracellular-matrix organization and oxygen-response, consistent with tumor-induced hypoxia. CC pathways showed collagen-rich matrix components, while MF pathways enriched for growth-factor binding. These findings suggest that fibroblasts are adapting to both stress and tumorigenic stimuli, contributing to FT progression. Additionally, Table 5 summarizes the top five enriched pathways for BP, CC, and MF,

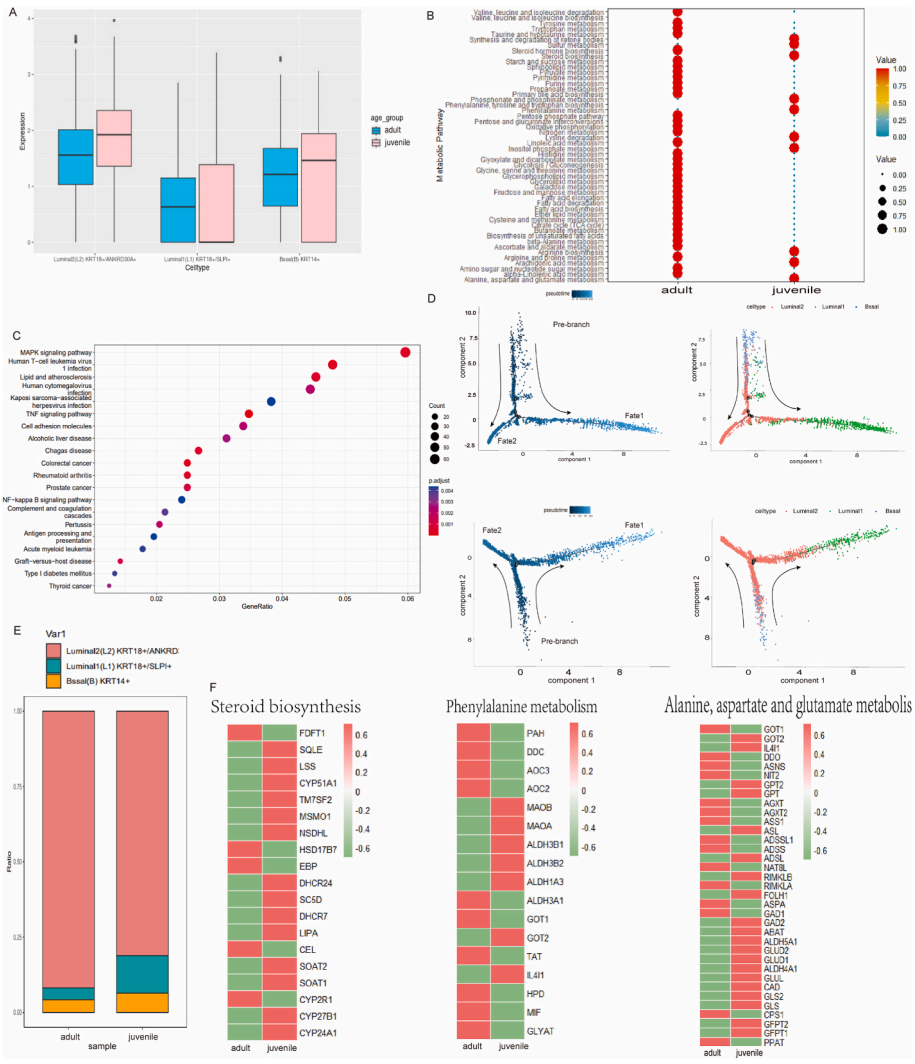

Fig. 3. |**Analysis of epithelial cells.(2)**
A ER expression in epithelial subsets (pink: juvenile; blue: adult).
B Heatmap of metabolic pathway scores in adult and juvenile groups.
C KEGG enrichment analysis of epithelial cells.
D Single-cell pseudotime reconstruction of epithelial cell states. Each point represents an individual cell, colored by: (Left panels) pseudotime progression from dark (Pre-branch, initial state) to light (terminal Fate 1/Fate 2 states); (Right panels) epithelial subpopulations. Top/bottom rows show juvenile and adult groups respectively, demonstrating distinct differentiation trajectories.
E Bargraph of epithelial subpopulation proportions.
F Heatmap of metabolic gene expression in adult and juvenile groups.

ranked by p-value, based on the GO enrichment analysis of fibroblasts in FT (Table 2).

3.13. GSEA pathway enrichment analysis was performed for each subgroup of fibroblasts

GSVA analysis (Fig. 5C) revealed distinct pathway enrichments across fibroblast subtypes. mCAFs were enriched in Hedgehog signaling, tCAFs in EMT and KRAS signaling, and inflammatory iCAFs in pancreatic β-cell pathways. ApCAFs were immune-active, with IL2–STAT5 and IFN-γ pathways enriched, while vCAFs exhibited Wnt/β-catenin and oxidative phosphorylation activation.Additionally, dCAFs were significantly enriched in pathways such as E2F targets and the G2M checkpoint. These findings demonstrate that fibroblasts adopt diverse roles, ranging from immune regulation to tumor progression.

3.14. GSEA enrichment analysis of fibroblasts

GSEA (Fig. 5B) revealed distinct pathway profiles between juvenile and adult fibroblasts. Estrogen signaling was enriched in both groups but more pronounced in adults, while juvenile fibroblasts showed stronger activation of MAPK, HIF-1, and GnRH signaling, along with endocrine resistance pathways. These findings suggest that fibroblasts may play a crucial role in the pathogenesis of FT, potentially through mechanisms involving MAPK signaling and hormonal stimulation.

3.15. Pseudo-time-series analysis of fibroblasts

Monocle2 trajectory analysis revealed distinct developmental trajectories between juvenile and adult fibroblasts. In the juvenile group, stromal mCAFs were at the initial stage of fibroblast development, while iCAFs and tCAFs represented transitional states. apCAFs and vCAFs were identified at the final stage (Fig. 5D). In contrast, in the adult group,

Table 2
Top 3 differentially expressed genes for each epithelial cell subtype.

gene	p_val	avg_log2FC	pct.1	pct.2	p_val_adj	cluster
XIST	9.84E-157	0.73013	0.875	0.827	4.88E-152	0
TMEM47	4.64E-116	0.732107	0.822	0.757	2.30E-111	0
PRXL2A	6.01E-89	0.729864	0.738	0.709	2.98E-84	0
RND1	0	2.947303	0.838	0.23	0	1
BCL3	0	2.582463	0.872	0.313	0	1
SLC25A25	0	2.676772	0.725	0.209	0	1
AREG	7.78E-250	0.933887	0.979	0.885	3.86E-245	2
EREG	4.31E-168	1.707695	0.552	0.294	2.14E-163	2
TCIM	9.92E-154	1.128252	0.761	0.529	4.92E-149	2
PIP	9.94E-293	1.828936	0.765	0.383	4.94E-288	3
SERPINA1	4.00E-256	1.253538	0.878	0.546	1.99E-251	3
TFF1	3.90E-100	1.046417	0.824	0.646	1.93E-95	3
SFRP1	0	5.296426	0.899	0.086	0	4
PTN	0	4.048139	0.736	0.171	0	4
LTF	0	4.706923	0.631	0.093	0	4
TYMS	0	4.456346	0.776	0.063	0	5
TOP2A	0	5.637313	0.637	0.029	0	5
TPX2	0	5.133132	0.567	0.036	0	5
KLK7	0	8.003895	0.902	0.019	0	6
KLK6	0	7.849246	0.872	0.02	0	6
MALL	0	7.413164	0.779	0.016	0	6
CCL5	0	6.064194	0.725	0.011	0	7
PTPRC	0	5.003387	0.699	0.024	0	7
NKG7	0	5.622473	0.542	0.008	0	7
NFKBIA	3.95E-63	1.603749	0.906	0.589	1.96E-58	8
CD74	1.68E-53	2.120964	0.852	0.538	8.33E-49	8
CXCL2	1.38E-48	1.887654	0.614	0.256	6.87E-44	8
MX1	0	4.722183	0.846	0.107	0	9
IFI44L	0	5.330947	0.658	0.032	0	9
IFIT1	0	5.940239	0.625	0.027	0	9
SPN	0	8.188666	0.792	0.009	0	10
GNG2	0	8.410256	0.75	0.012	0	10
GZMB	0	7.905533	0.562	0.007	0	10

gene: Gene name (marker gene). **p_val**: Raw p-value (statistical significance), where smaller values indicate more significant differential expression. **avg_log2FC**:Average log2 fold change (expression difference), with positive values indicating upregulation and negative values indicating downregulation. **pct.1**:The percentage of cells expressing the gene in the target cluster. **pct.2**:The percentage of cells expressing the gene in all other clusters. **p_val_adj**: Adjusted p-value (e.g., Bonferroni or FDR correction) for multiple hypothesis testing. **cluster**,:fibroepithelial tumor Cells subpopulation identifier (0–10), representing distinct cell types or states.

iCAFs occupied the initial state, with mCAFs and apCAFs in transitional states, and vCAFs and mCAFs at the final stages (Fig. 5D). These results underscore significant differences in fibroblast development and differentiation between the two groups.

The pseudo-temporal analysis further revealed that juvenile fibroblast subtypes could be categorized into three developmental states, while adult fibroblast subtypes were divided into five (Fig. 4E). Genes specifically expressed in apCAFs, such as HLA-DRA, CD74, and HLA-DRB1, showed a gradual increase in expression from state 1 to state 3 in the juvenile group, indicating an enhancement of antigen-presenting function and anti-tumor properties during development. In contrast, in the adult group, these genes exhibited a slight increase in expression at state 4, followed by a gradual decline, suggesting a loss of sustained antigen-presenting functionality (Fig. 4E). This contrast highlights the divergent differentiation trajectories of antigen-presenting genes between the juvenile and adult groups.

Additionally, GO enrichment analysis of genes based on their developmental trajectories (Fig. 5F) provided further insights. In the juvenile group, the Pre-branch stage was primarily enriched in pathways related to hormone biosynthesis and glycolysis. Branch 1, representing the differentiation of apCAFs, was enriched in pathways such as antigen processing and presentation via MHC class II, emphasizing their role in immune responses. Branch 2, associated with the differentiation into vascular fibroblasts, was enriched in stress response pathways. In the adult group, the Pre-branch stage was marked by pathways responding

Table 3
GO enrichment analysis of epithelial cells in FT (sorted by pvalue).

ONTOLOGY	ID	Description	pvalue	p. adjust	Count
BP	GO:1901654	response to ketone	2.33E-15	1.43E-11	66
BP	GO:0048545	response to steroid hormone	1.12E-13	3.42E-10	88
BP	GO:0048732	gland development	2.10E-13	4.28E-10	105
BP	GO:0050673	epithelial cell proliferation	9.22E-13	1.41E-09	110
BP	GO:0031960	response to corticosteroid	1.44E-12	1.77E-09	53
CC	GO:0062023	collagen-containing extracellular matrix	2.86E-18	1.97E-15	114
CC	GO:0005788	endoplasmic reticulum lumen	4.04E-12	1.39E-09	79
CC	GO:0045177	apical part of cell	1.14E-08	2.63E-06	90
CC	GO:0031983	vesicle lumen	5.48E-08	9.45E-06	71
CC	GO:0060205	cytoplasmic vesicle lumen	9.46E-08	1.30E-05	70
MF	GO:0001228	DNA-binding transcription activator activity, RNA polymerase II-specific	7.93E-09	7.08E-06	99
MF	GO:0001216	DNA-binding transcription activator activity	1.25E-08	7.08E-06	99
MF	GO:0005201	extracellular matrix structural constituent	3.77E-08	1.42E-05	47
MF	GO:0008201	heparin binding	7.25E-08	2.05E-05	46
MF	GO:0008083	growth factor activity	2.87E-07	6.48E-05	43

Ontology:go classification (BP: biological processes; CC: Cell component; MF: Molecular function).
ID: GO entry unique identifier (Gene Ontology ID). (Unique identifier for the GO term.) **Description**: Description of the GO function. (Functional description of the GO term.) **pvalue**: The original pvalue of the enrichment analysis, the smaller the value, the higher the significance. (Raw p-value of enrichment; smaller values indicate higher significance.) **p.adjust**: Corrected P-value (such as FDR or Bonferroni correction) for multiple hypothesis testing (Adjusted p-value (e.g., FDR or Bonferroni correction). for multiple testing.) **Count**: The number of genes enriched into the GO entry. (Number of genes enriched in the GO term.)

to interleukin-1 and other inflammatory signals. Branch 1 was enriched in extracellular matrix organization and collagen fibril formation, while Branch 2 was associated with vascular development, mammary gland development, and related processes.

Collectively, these findings highlight the differences in gene function dynamics between juvenile and adult fibroblasts, offering preliminary insights into the role of fibroblasts in the pathogenesis of FT.

3.16. Communication of different cells in the tissues of the adult and juvenile groups

The tumor microenvironment is highly complex, with distinct cell subpopulations interacting through various signaling pathways, contributing to tumor progression. To assess these interactions, cell-cell communication analysis was conducted using ligand-receptor pairs (Fig. 6A). In the adult group, apCAFs and dCAFs displayed stronger and more frequent interactions with other cell types, suggesting a more immune-immune cell crosstalk. In contrast, the juvenile group showed higher communication frequencies in apCAF, Bssal, and dCAF, with enhanced interaction intensities between stromal and epithelial cells, pointing to more dynamic communication networks in juvenile FT

Table 4
Expression of the first 3 specific genes in each cell cluster of fibroblasts.

gene	p_val	avg_log2FC	pct.1	pct.2	p_val_adj	cluster
TNC	6.16E-77	2.130246	0.826	0.348	3.06E-72	0
AL1	9.90E-73	2.593923	0.648	0.177	4.91E-68	0
LONRF2	3.15E-58	2.591263	0.555	0.158	1.56E-53	0
CCN5	1.43E-180	4.910409	0.781	0.06	7.12E-176	1
ACKR3	1.30E-168	5.086026	0.728	0.052	6.47E-164	1
PLA2G2A	4.78E-145	4.20132	0.669	0.049	2.37E-140	1
FRZB	1.93E-50	1.987753	0.726	0.303	9.59E-46	2
HOPX	5.67E-34	1.677717	0.633	0.282	2.82E-29	2
SBSPON	2.71E-30	1.625831	0.556	0.231	1.34E-25	2
FOXA1	6.81E-169	4.390392	0.866	0.066	3.38E-164	3
AGR2	6.85E-161	4.50396	0.919	0.099	3.40E-156	3
TFF3	3.03E-120	4.645807	0.669	0.054	1.50E-115	3
RGS5	1.29E-201	6.571592	0.992	0.068	6.39E-197	4
NDUFA4L	1.32E-167	6.404404	0.677	0.019	6.53E-163	4
TINAGL1	1.32E-121	5.983164	0.594	0.031	6.54E-117	4
APOD	1.34E-36	2.388996	1	0.83	6.65E-32	5
COL15A1	9.92E-24	2.160343	0.652	0.274	4.93E-19	5
CXCL14	6.12E-12	1.36729	0.636	0.345	3.04E-07	5
PRAME	4.04E-117	4.650225	0.662	0.028	2.00E-112	6
SCG2	3.08E-99	5.413009	0.625	0.034	1.53E-94	6
LRRN1	6.19E-82	4.428568	0.588	0.039	3.07E-77	6
HLA-DQA2	5.21E-187	9.233309	0.583	0.001	2.59E-182	7
CCL3L1	4.30E-112	7.536358	0.55	0.014	2.14E-107	7
HLA-DRA	1.08E-110	7.543371	1	0.096	5.38E-106	7
DHFR	8.85E-21	3.281334	0.562	0.122	4.39E-16	8
CKS2	4.50E-20	3.290285	0.521	0.103	2.23E-15	8
HMGB2	1.12E-19	3.731136	0.667	0.181	5.58E-15	8
CCL5	9.12E-203	8.50334	0.824	0.007	4.53E-198	9
NKG7	3.16E-167	8.591639	0.676	0.005	1.57E-162	9
GZMA	6.46E-135	8.377873	0.588	0.006	3.21E-130	9

gene: Gene name (marker gene).**p_val**: Raw p-value (statistical significance), where smaller values indicate more significant differential expression. **avg_log2FC**:Average log2 fold change (expression difference), with positive values indicating upregulation and negative values indicating downregulation. **pct.1**:The percentage of cells expressing the gene in the target cluster. **pct.2**:The percentage of cells expressing the gene in all other clusters. **p_val_adj**: Adjusted p-value (e.g., Bonferroni or FDR correction) for multiple hypothesis testing. **Cluster**:fibroepithelial tumor cells subpopulation identifier (0–9), representing distinct cell types or states.

tissues.

3.17. Changes in intercellular communication between juvenile and adult groups

A comparative analysis revealed increased interaction frequencies between apCAFs, Bssal, and dCAFs in the juvenile group, while Luminal2 cells exhibited stronger communication strength with other populations (Fig. 6D). These findings indicate distinct functional roles of fibroblasts and epithelial cells in the juvenile tumor microenvironment, which may contribute to more aggressive or immune-sensitive characteristics.

Table 5
Results of GO enrichment analysis of fibroblasts in FT.

ONTOLOGY	ID	Description	pvalue	p.adjust	Count
BP	GO:0030198	extracellular matrix organization	1.34E-27	5.31E-24	98
BP	GO:0043062	extracellular structure organization	1.76E-27	5.31E-24	98
BP	GO:0045229	external encapsulating structure organization	3.04E-27	6.10E-24	98
BP	GO:0001655	urogenital system development	1.30E-18	1.96E-15	90
BP	GO:0072001	renal system development	6.22E-18	7.49E-15	82
CC	GO:0062023	collagen-containing extracellular matrix	8.66E-47	5.92E-44	146
CC	GO:0005775	vacuolar lumen	2.53E-15	8.65E-13	53
CC	GO:0060205	cytoplasmic vesicle lumen	4.30E-15	9.79E-13	77
CC	GO:0031983	vesicle lumen	6.13E-15	1.05E-12	77
CC	GO:0034774	secretory granule lumen	8.06E-15	1.10E-12	76
MF	GO:0005201	extracellular matrix structural constituent	7.09E-21	7.71E-18	62
MF	GO:0005518	collagen binding	1.06E-11	5.77E-09	28
MF	GO:0019838	growth factor binding	2.67E-11	9.68E-09	40
MF	GO:0005178	integrin binding	1.49E-10	4.04E-08	43
MF	GO:0001968	fibronectin binding	3.77E-10	8.19E-08	17

Ontology:go classification (BP: biological processes; CC: Cell component; MF: Molecular function).
ID: GO entry unique identifier (Gene Ontology ID). (Unique identifier for the GO term.) **Description**: Description of the GO function. (Functional description of the GO term.) **pvalue**: The original pvalue of the enrichment analysis, the smaller the value, the higher the significance. (Raw p-value of enrichment; smaller values indicate higher significance.) **p.adjust**: Corrected P-value (such as FDR or Bonferroni correction) for multiple hypothesis testing (Adjusted p-value (e.g., FDR or Bonferroni correction). for multiple testing.) **Count**: The number of genes enriched into the GO entry. (Number of genes enriched in the GO term.)

3.18. Compare primary sources and targets in 2D space

To better understand the directionality of cell signaling, we performed a 2D analysis comparing efferent (signal sending) and afferent (signal receiving) interaction strengths across both groups. Epithelial cells were identified as the primary senders and receivers of signals in both juvenile and adult groups, emphasizing their central role in tumor-stroma communication (Fig. 6B). Furthermore, pathway enrichment analysis of genes from each cell subset in both groups revealed distinct pathway activation patterns. In the juvenile group, significant differences were observed in the RELN, PTN, and HGF pathways, suggesting their potential roles in mediating cellular interactions and functional differences (Fig. 6C).

3.19. Ligand-receptor pairs between cells

We explored ligand-receptor interactions between fibroblast and epithelial subpopulations. In the juvenile group, significant interactions were found between apCAFs-Luminal1, apCAFs-Luminal2, and iCAFs-Luminal2, with key ligand-receptor pairs such as MDK-NCL, MDK-SDC4, and AREG-(EGFR + ERBB2) (Fig. 6E). These interactions suggest that apCAFs may play a prominent role in immune signaling and

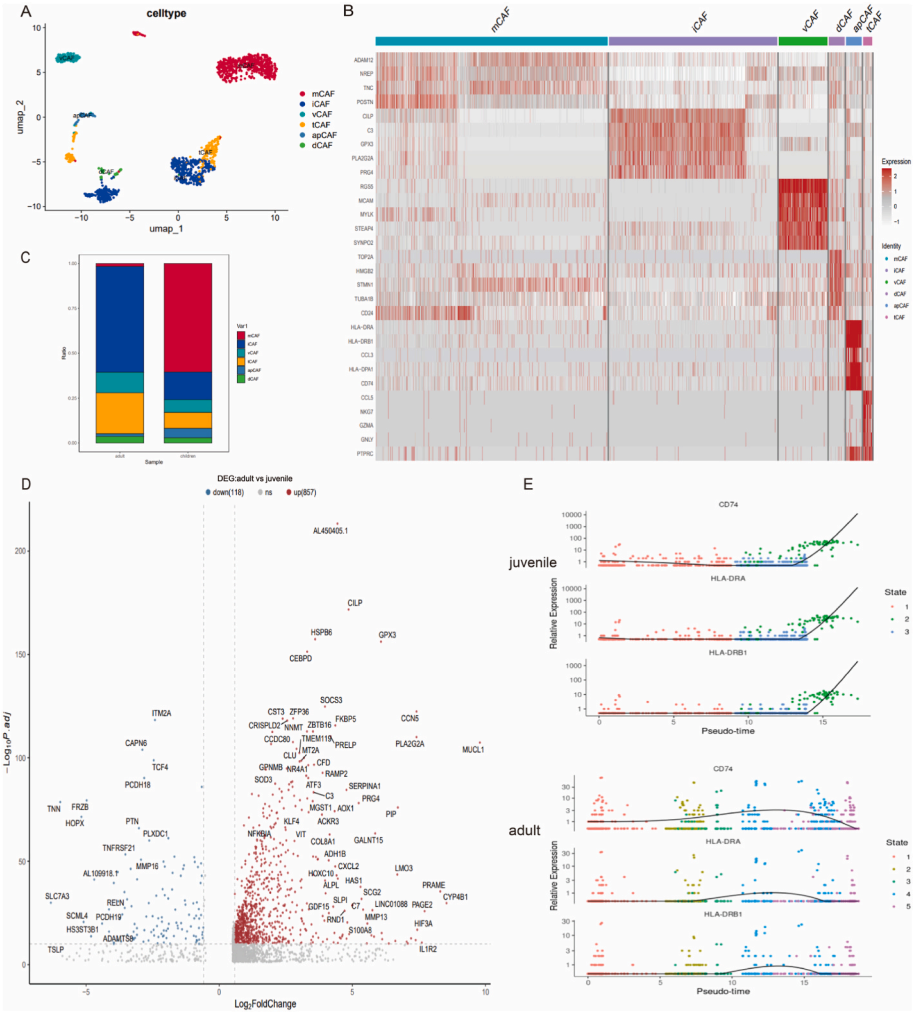

Fig. 4. | **Analysis of fibroblasts.(1)**
A UMAP of fibroblast subsets.
B Heatmap: top 5 marker genes in fibroblasts (red = high expression; gray = low).
C Histogram: fibroblast cluster proportions in adult and juvenile groups.
D Volcano plot: differentially expressed genes in fibroblast subsets (red = up-regulated; blue = down-regulated; x-axis: Log2FoldChange; y-axis: adjusted p-value).
E Expression changes of apCAF marker genes during development (above: juvenile; below: adult).

epithelial-stromal communication in juvenile FT. In contrast, the adult group exhibited downregulated ligand-receptor pairs, such as COL1A2-CD44, COL1A2-SDC1, and COL1A2-SDC4, indicating reduced fibroblast-epithelial crosstalk in adult FT tissues. These findings highlight age-specific differences in cellular signaling and suggest that fibroblast-epithelial interactions may be more immune-active in juvenile tumors.

4. Discussion

In this study, we focused on juvenile giant FT of the breast, investigating the heterogeneity of these lesions through scRNA-seq analysis. Our work revealed the heterogeneity of epithelial cells and fibroblasts, as well as their roles in the pathogenesis of breast fibroepithelial lesions. We also provided a detailed characterization of the tumor microenvironment heterogeneity in juvenile giant FT. By comparing adult and juvenile FEL tissues using scRNA-seq, we uncovered differences in the composition and structural features of epithelial cells and fibroblasts between the two groups. Cell-cell communication analysis of fibroblast and epithelial subpopulations offered new insights into the tumor microenvironment of FT and laid a necessary foundation for future in-depth research.

Breast fibroepithelial lesions, such as fibroadenomas and phyllodes tumors, are typically benign, but their pathogenesis remains poorly understood, especially in juvenile cases. Juvenile giant fibroadenomas, though rare, are characterized by rapid growth and significant size, often occurring during prepuberty or early puberty. These tumors are thought to be influenced by endocrine hormone stimulation during a time of hormonal fluctuations. Our findings suggest that juvenile FT may be more sensitive to estrogen signaling, as indicated by higher levels of estrogen receptors in juvenile epithelial cells. This is consistent with previous research, which has linked endocrine factors to the development of fibroadenomas in reproductive-aged women [9].

The development of giant juvenile fibroadenomas is generally believed to be associated with endocrine hormone stimulation of local breast tissue. Since the disease occurs in children during prepuberty or early puberty, unstable ovarian function leads to disturbances in estrogen and progesterone levels. Alternatively, some breast glandular tissues may exhibit heightened sensitivity to estrogen stimulation, resulting in hyperplasia [38], irregular ductal growth, ductal dilation, cyst formation, and eventual fibroadenoma development. Our data showed that juvenile epithelial cells, including type 2 ductal cells and basal cells, expressed higher levels of estrogen receptors compared to adult tissues. Chen [11] used scRNA-seq to investigate why adult fibroadenomas are insensitive to the estrogen-modulating drug tamoxifen and the

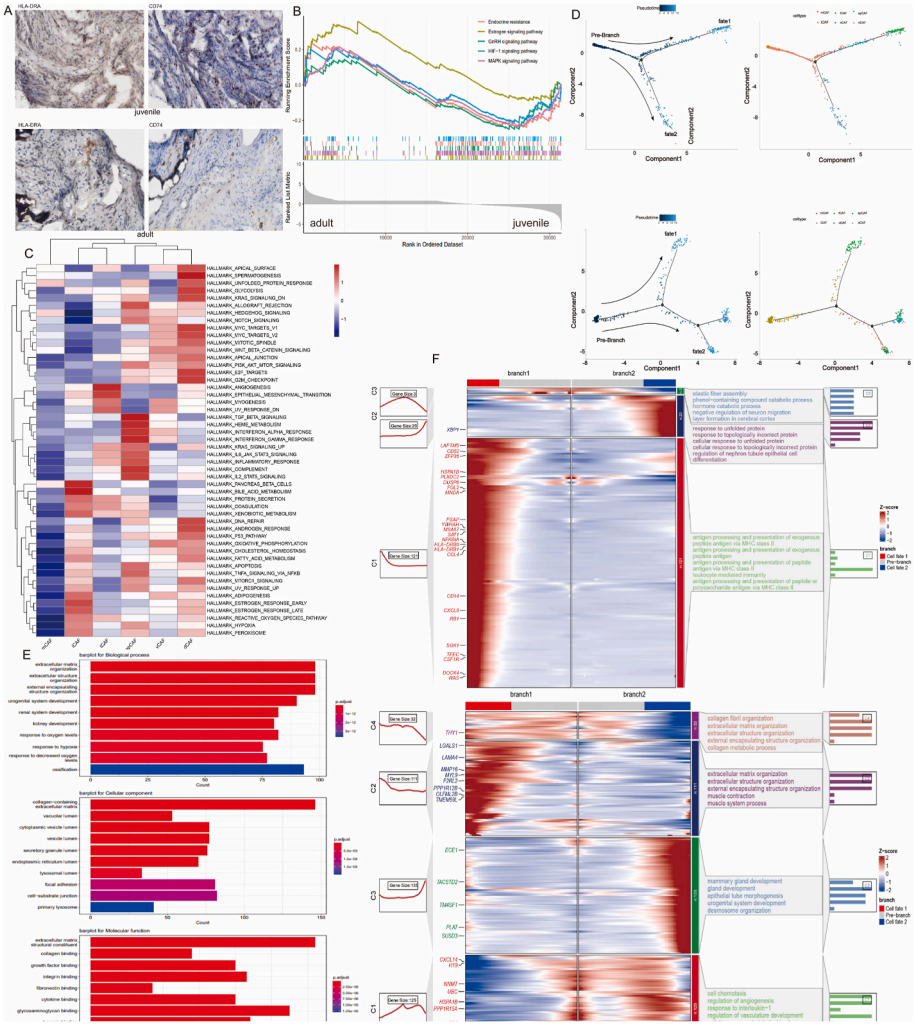

Fig. 5. Analysis of fibroblasts.(2)

A Immunohistochemistry of apCAF marker genes (top left and right: juvenile group HLA-DRA, CD74; bottom left and right: adult group HLA-DRA, CD74; Original magnification, x20; Scale bar, 100 μm)

B GSEA enrichment: fibroblasts (left: adult; right: juvenile).

C GSVA enrichment: fibroblast subsets (red = increase; blue = decrease).

D Single-cell pseudotime reconstruction of fibroblast differentiation. Each point represents an individual cell colored by: (Top row) Juvenile group trajectories; (Bottom row) Adult group trajectories. Left panels show temporal progression from dark (Pre-branch initial state) to light (terminal Fate 1/Fate 2) along pseudotime; Right panels display fibroblast subpopulation trajectories. Distinct developmental paths are evident between age groups.

E GO enrichment: differentially expressed genes in fibroblasts (red = significance).

F Functional changes during fibroblast development (top: juvenile group; bottom: adult group).

mechanisms underlying drug resistance. Previous research has primarily focused on adult fibroadenomas, with genomic studies indicating that mutations in MED12 and TERT promoter hotspots are common in adult breast FT [39]. However, the pathogenesis of juvenile giant fibroepithelial lesions requires further investigation.

Single-cell RNA sequencing is a key technology for studying physiological and pathological processes, enabling the discovery of complex tumor microenvironments and high cellular heterogeneity. It provides insights into disease mechanisms and potential new treatment strategies. Our scRNA-seq analysis explored the heterogeneity of breast fibroepithelial lesions, categorizing cells into different types and revealing their roles in various subgroups of FT, with a focus on epithelial cells and fibroblasts. We further investigated the heterogeneity of these two cell types and their contributions to the pathogenesis of FT, providing a detailed characterization of the tumor microenvironment heterogeneity in juvenile giant FT.

To elucidate the disease mechanism, we compared adult and juvenile FEL tissues using scRNA-seq, generating single-cell transcriptomic profiles for both groups. This revealed differences in the composition and structural features of epithelial cells and fibroblasts between adult and juvenile FT, which differ from normal breast tissue. Previous studies have shown that breast fibroepithelial lesions are biphasic tumors, but their cellular composition remains unclear [40]. We analyzed tissues from three cases of juvenile giant FT and one adult fibroadenoma, identifying changes in cellular composition and structural alterations in epithelial cells and fibroblasts between the two groups. Cell annotation identified six cell types in FT, including epithelial cells, endothelial cells, fibroblasts, macrophages, neutrophils, and T & NK cells. The proliferation of epithelial and stromal cells is a major factor in FEL pathogenesis.

Although only one adult fibroadenoma was included as a control sample, it was carefully selected according to strict histopathological criteria and RNA quality standards. Given the rarity of juvenile giant FT and the ethical limitations of obtaining matched adult controls, expanding the sample number was challenging. However, consistent transcriptional patterns observed across all three juvenile cases support the robustness of our observations. Future multicenter studies with

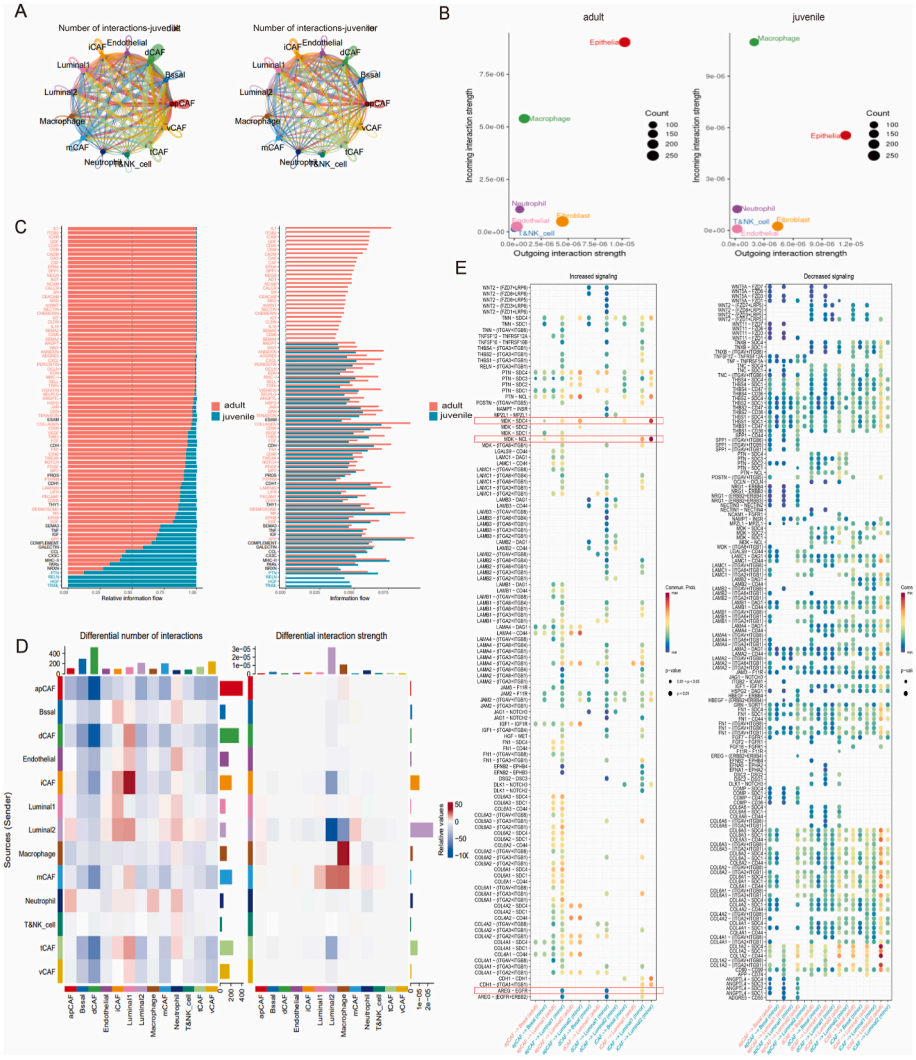

Fig. 6. Communication between different cells.
A Cell communication among 13 populations in adult and juvenile groups.
B Scatter plots: communication sources and targets per subpopulation.
C Pathway proportions in adult vs. juvenile groups.
D Interaction changes (number/strength) between cell types in juvenile vs. adult groups.
E Ligand-receptor pairs between fibroblast and epithelial subsets (left: upregulated; right: downregulated; red = up, blue = down, blank = none).

larger and age-balanced cohorts are needed to validate and expand these findings.

GSVA of nine cell types revealed characteristic pathways, such as "early estrogen response," "late estrogen response," "oxidative phosphorylation," and "MYC targets v2" in epithelial cells. Estrogen mediates biological homeostasis, but as a target organ, excessive estrogen stimulation can lead to breast tumorigenesis. Endocrine resistance pathways present in both groups are highly associated with fibroadenoma recurrence. MYC, a transcription factor that promotes cell growth and proliferation, is involved in the pathogenesis of various tumors [41]. Based on differentially expressed genes, epithelial cells were classified into three subtypes. Although the developmental states of epithelial cells were similar between adult and juvenile groups, gene regulation varied at different stages, highlighting potential differences in the role of epithelial cells in juvenile FT. EMT and Wnt/β-catenin signaling pathways were enriched in fibroblasts.

In parallel, epithelial transcriptomic analysis revealed the activation of the PI3K–AKT signaling pathway, typically associated with malignant tumors, representing a distinctive feature in benign fibroadenomas. While such proliferative signaling is expected in neoplastic epithelium, our analysis identified additional juvenile-specific mechanisms. Notably, co-activation of the GnRH, MAPK, and endocrine resistance pathways suggests a hormone-driven proliferative program unique to juvenile tissues. Furthermore, pseudotime analysis indicated that these pathways are activated earlier during epithelial differentiation, potentially accounting for the rapid growth and enhanced proliferative potential of juvenile fibroadenomas. Beyond these canonical pathways, we also identified distinct activation of RELN, PTN, and HGF signaling in juvenile epithelial cells, providing novel mechanistic insights into the pathogenesis of juvenile fibroadenomas. Together, these findings highlight coordinated endocrine and immune signaling between fibroblasts and epithelial cells, suggesting that stromal–epithelial crosstalk plays a pivotal role in shaping the growth dynamics and immune landscape of juvenile fibroadenomas.

Fibroblasts in FT were divided into six subpopulations, with juvenile tissues showing a higher proportion of antigen-presenting cancer-associated fibroblasts (apCAFs), characterized by the expression of immune-related genes such as HLA-DRA and CD74, as confirmed by immunohistochemistry. HLA-DRA and CD74 are classical MHC class II–related molecules typically restricted to professional antigen-presenting

immune cells. Their high expression in fibroblasts indicates the acquisition of antigen-presenting properties, defining a subset with immune-like characteristics. In malignant tumors, apCAFs expressing these markers are usually associated with immune modulation or immune evasion; however, their presence in benign fibroadenomas is uncommon. This benign-specific activation suggests an immunoregulatory rather than immunosuppressive microenvironment, potentially supporting immune surveillance and limiting malignant transformation. Pseudotime analysis revealed three developmental states of fibroblasts, in which early-stage cells were enriched in metabolic pathways including "glycolysis," resembling normal fibroblasts, whereas mid-to late-stage fibroblasts were enriched in immune-related pathways such as "antigen processing and presentation via MHC class II," indicating their potential roles in immune regulation [42]. The enrichment of these immune-active fibroblasts in juvenile FT highlights a distinct, immunomodulatory tumor microenvironment compared with that of adult FT, which may contribute to tumor growth regulation and immune homeostasis.

Cell-cell communication analysis of fibroblast and epithelial subpopulations identified ligand-receptor interactions in different groups. The MDK-NCL and MDK-SDC4 pathways were key communication signals between epithelial cells and fibroblasts in juvenile tissues. MDK is a heparin-binding growth factor that promotes cell growth, survival, migration, angiogenesis, cytokine expression, and differentiation, facilitating interactions with surrounding cells [43]. Other ligand-receptor pairs, such as AREG-(EGFR + ERBB2) and AREG-EGFR, are essential for adolescent breast development [44]. Our findings provide new insights into the tumor microenvironment of FT and establish a foundation for future research. These ligand–receptor pairs suggest that fibroblast-derived signals may directly influence epithelial cell proliferation and maturation, offering potential therapeutic targets to modulate fibroblast–epithelial communication in FT. Recent single-cell studies in breast cancer have also characterized immunosuppressive microenvironments that facilitate tumor progression and therapeutic resistance, particularly in HER2-positive inflammatory breast cancer [45]. These findings underscore the importance of dissecting the immune and stromal components of the tumor microenvironment, supporting our observation that immune-active fibroblasts and stromal–epithelial interactions play central roles in juvenile fibroepithelial tumor pathogenesis.

This study has some limitations. First, due to the low incidence of juvenile giant FT, the sample size was small. This limited sample size inevitably restricts the statistical power and generalizability of our conclusions. While the single-cell resolution provides detailed mechanistic insights, the small cohort may not capture the full biological diversity among patients. Therefore, the interpretations should be considered exploratory rather than definitive, and future studies involving larger, independent cohorts are needed to validate these findings. In future work, we plan to collect additional specimens and generate more comprehensive datasets to further investigate the pathogenic mechanisms and clinical characteristics of these tumors. Additionally, because patients were adolescents, we minimized damage to normal breast tissue during sampling to avoid affecting future breast development, resulting in a lack of adjacent normal tissue as a control. Consequently, some subtle but important cell populations or genetic loci may have been overlooked. Furthermore, it remains unclear whether differences in gene expression between groups stem from developmental stages or disease-related changes. Our study focused on mRNA-level transcription without validation at other levels. Future research should include larger sample sizes, morphological analysis, whole-exome sequencing, and characteristic molecular markers to comprehensively investigate the pathogenesis of juvenile giant FT. Such efforts could advance personalized treatment and improve outcomes for FT patients, particularly those with high recurrence rates or increased breast cancer risk. Future research should expand sample sizes, incorporate matched adult and normal controls, and apply integrated multi-omics approaches—including spatial transcriptomics and proteomics—to validate these findings and clarify the developmental versus pathological origins of observed gene expression differences.

5. Conclusion

In conclusion, this study employs single-cell transcriptomic analysis to uncover the cellular heterogeneity and molecular characteristics of juvenile giant mammary fibroepithelial lesions. It identifies elevated estrogen receptor expression in the juvenile group and significant upregulation of immune-related genes in fibroblasts, highlighting the critical roles of estrogen signaling and the immune microenvironment in disease pathogenesis. These findings provide novel insights into the mechanisms underlying juvenile fibroadenomas and establish a foundation for personalized therapeutic strategies, particularly offering innovative approaches for managing patients at high risk of recurrence.

Ethics approval and consent to participate

The study involving human participants was reviewed and approved by the Ethics Committee of the Third Affiliated Hospital of Kunming Medical University (Approval No. KYCS2024-083, Date: 2024-03-06). Written informed consent was obtained from all participants or their legal guardians.

Consent for publication

Not applicable.

Clinical trial number

Not applicable.

Human ethics and consent to participate declarations

Written informed consent was obtained from all patients prior to study enrollment.

Clinical trial number

Not applicable.

Statement

AI-assisted technologies (ChatGPT, OpenAI) were used to improve the clarity and readability of the text. All scientific content and interpretations were created solely by the authors.

Funding declaration

The authors acknowledge the financial support received from the Program: the First-Class Discipline Team of Kunming Medical University (Kunming Medical University Breast Cancer Precision & Translational Medicine Research Team 2024XKTDYS08). National Natural Science Foundation of China (grant no. 82360614), the Science and Technology Project of Yunnan Provincial Science and Technology Department (grant no. 202401AT070005), Yunnan Health Training Project of High Level Talents (grant number: H-2024087), National Natural Science Foundation of China (grant no. 82560481), Kunming Medical University two thousand twenty-five Master's Degree Student Education Innovation Fund Project(grant no. 2025S276).

CRediT authorship contribution statement

Tianrun Shu: Validation, Writing – review & editing. **Jing Wang:** Investigation. **Fuying Yang:** Validation, Writing – original draft.

Dechun Yang: Data curation, Funding acquisition. **Shengnan Ren:** Project administration, Software. **Min Zhao:** Formal analysis, Investigation. **Yafeng Lv:** Methodology, Software. **Ying Song:** Data curation, Validation. **Hanhui Xie:** Methodology, Resources. **Xuan Yang:** Project administration, Resources. **Cong Jiang:** Data curation, Visualization. **Rong Guo:** Formal analysis, Methodology, Software. **Shicong Tang:** Conceptualization, Funding acquisition, Methodology, Resources, Visualization.

Declaration of competing interest

The authors declare that they have no competing interests.

Acknowledgements

Not applicable.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bbrep.2026.102448.

Data availability

The original data applied in this research are accessible from the corresponding author.

References

[1] K.S. Mon, P. Tang, Fibroepithelial lesions of the breast: update on molecular profile with focus on pediatric population, Arch. Pathol. Lab Med. 147 (1) (2023) 38–45.
[2] A. Eleftheriades, E. Tsarna, K. Toutoudaki, E. Paschalidou, N. Christopoulos, I. Georgopoulos, G. Mitropoulou, P. Christopoulos, Giant Juvenile fibroadenoma: case report and review of the literature, J. Clin. Med. 12 (5) (2023) 1855.
[3] K.S. Mon, P. Tang, Fibroepithelial lesions of the breast: update on molecular profile with focus on pediatric population, Arch. Pathol. Lab Med. 147 (1) (2023 Jan 1) 38–45, https://doi.org/10.5858/arpa.2022-0011-RAIF:3.2.Q2. PMID: 35776911.
[4] Murtala Abubakar, Balarabe Kabir, Almustapha Aliyu Liman, Sani Mohammed Shehu, Saad Aliyu Ahmed, Peter Pase Abur 1, Mohammed Bashir 1, Zainab A. Adamu, Kamarudeen Owolabi Sani, Osita Ufere, Ibiam Murphy 2, Shehu Abdullahi, Aliyu Usman Bappa, Histopathological spectrum of benign lesions of the breast from a university teaching hospital in Northern Nigeria, Ann. Tropical Pathol. 13 (1) (Jan–Jun 2022) 12–17, https://doi.org/10.4103/atp.atp_19_21.
[5] S. Yılmaz, M.R. Aykota, Y. Arman Karakaya, U. Özgen, E. Erdem, Phyllodes tumors of the breast: a single-center experience, Eur. J. Breast Health 17 (1) (2021) 36–41.
[6] Z. Chen, Y. Zhang, W. Li, et al., Single cell profiling of female breast fibroadenoma reveals distinct epithelial cell compositions and therapeutic targets, Nat. Commun. 14 (2023) 3469, https://doi.org/10.1038/s41467-023-39059-3.
[7] S. Alipour, M. Abedi, A. Saberi, A. Maleki-Hajiagha, F. Faiz, S. Shahsavari, B. Eslami, Metformin as a new option in the medical management of breast fibroadenoma; a randomized clinical trial, BMC Endocr. Disord. 21 (1) (2021 Aug 20) 169, https://doi.org/10.1186/s12902-021-00824-4. PMID: 34416879; PMCID: PMC8377455.
[8] J. Limberg, K. Barker, S. Hoda, R. Simmons, A. Michaels, J.L. Marti, Fibroepithelial lesions (FELs) of the breast: is routine excision always necessary? World J. Surg. 44 (5) (2020) 1552–1558.
[9] L.M. Kopicky, D.A. Pratt, Cryoablation and breast disease, Curr. Breast Cancer Rep. 16 (2024) 279–287, https://doi.org/10.1007/s12609-024-00561-2.
[10] T.K.Y. Tay, P. Guan, B.N. Loke, N.D.M. Nasir, V. Rajasegaran, A.A. Thike, D. Lian, K.T.E. Chang, B.T. Teh, C.C.Y. Ng, P. Tan, Molecular insights into paediatric breast fibroepithelial tumours, Histopathology 73 (5) (2018) 809–818.
[11] Z. Chen, Y. Zhang, W. Li, C. Gao, F. Huang, L. Cheng, M. Jin, X. Xu, J. Huang, Single cell profiling of female breast fibroadenoma reveals distinct epithelial cell compositions and therapeutic targets, Nat. Commun. 14 (1) (2023) 3469.
[12] S. Zhu, M. Zhang, X. Liu, et al., Single-cell transcriptomics provide insight into metastasis-related subsets of breast cancer, Breast Cancer Res. 25 (2023) 126, https://doi.org/10.1186/s13058-023-01728-y.
[13] L. Wen, F. Tang, Recent advances in single-cell sequencing technologies, Precis. Clin. Med. 5 (1) (2022) pbac002.
[14] F. Tang, C. Barbacioru, Y. Wang, E. Nordman, C. Lee, N. Xu, X. Wang, J. Bodeau, B. B. Tuch, A. Siddiqui, K. Lao, M.A. Surani, mRNA-Seq whole-transcriptome analysis of a single cell, Nat. Methods 6 (5) (2009) 377–382.
[15] J. Ulbrich, V. Lopez-Salmeron, I. Gerrard, BD Rhapsody™ single-cell analysis system workflow: from sample to multimodal single-cell sequencing data[G], in: R. A. Calogero, V. Benes (Eds.), Single Cell Transcriptomics: Methods and Protocols, Springer US, New York, NY, 2023, pp. 29–56.
[16] A.M. Leader, J.A. Grout, B.B. Maier, B.Y. Nabet, M.D. Park, A. Tabachnikova, C. Chang, L. Walker, A. Lansky, J. Le Berichel, L. Troncoso, N. Malissen, M. Davila, J.C. Martin, G. Magri, K. Tuballes, Z. Zhao, F. Petralia, R. Samstein, N.R. D'Amore, G. Thurston, A.O. Kamphorst, A. Wolf, R. Flores, P. Wang, S. Müller, I. Mellman, M. B. Beasley, H. Salmon, A.H. Rahman, T.U. Marron, E. Kenigsberg, M. Merad, Single-cell analysis of human non-small cell lung cancer lesions refines tumor classification and patient stratification, Cancer Cell 39 (12) (2021) 1594–1609.e12.
[17] P.V. Kharchenko, The triumphs and limitations of computational methods for scRNA-seq, Nat. Methods 18 (7) (2021) 723–732.
[18] A. Lafzi, C. Moutinho, S. Picelli, H. Heyn, Tutorial: guidelines for the experimental design of single-cell RNA sequencing studies, Nat. Protoc. 13 (12) (2018) 2742–2757.
[19] Y. Chao, Z. Liu, Biomaterials tools to modulate the tumour microenvironment in immunotherapy, Nat. Rev. Bioeng. 1 (2) (2023) 125–138.
[20] Y. Hao, S. Hao, E. Andersen-Nissen, W.M. Mauck 3rd, S. Zheng, A. Butler, M.J. Lee, A.J. Wilk, C. Darby, M. Zager, P. Hoffman, M. Stoeckius, E. Papalexi, E.P. Mimitou, J. Jain, A. Srivastava, T. Stuart, L.M. Fleming, B. Yeung, A.J. Rogers, J. M. McElrath, C.A. Blish, R. Gottardo, P. Smibert, R. Satija, Integrated analysis of multimodal single-cell data, Cell 184 (13) (2021 Jun 24) 3573–3587.e29, https://doi.org/10.1016/j.cell.2021.04.048. Epub 2021 May 31. PMID: 34062119; PMCID: PMC8238499.
[21] C.S. McGinnis, L.M. Murrow, Z.J. Gartner, DoubletFinder: doublet detection in single-cell RNA sequencing data using artificial nearest neighbors, Cell Syst. 8 (4) (2019) 329–337.e4.
[22] D. Aran, A.P. Looney, L. Liu, E. Wu, V. Fong, A. Hsu, S. Chak, R.P. Naikawadi, P. J. Wolters, A.R. Abate, A.J. Butte, M. Bhattacharya, Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage, Nat. Immunol. 20 (2) (2019 Feb) 163–172, https://doi.org/10.1038/s41590-018-0276-yIF:27.6.Q1. Epub 2019 Jan 14. PMID: 30643263; PMCID: PMC6340744.
[23] S. Jin, C.F. Guerrero-Juarez, L. Zhang, I. Chang, R. Ramos, C.-H. Kuan, P. Myung, M.V. Plikus, Q. Nie, Inference and analysis of cell-cell communication using CellChat, Nat. Commun. 12 (1) (2021) 1088.
[24] S.A. Lambert, A. Jolma, L.F. Campitelli, P.K. Das, Y. Yin, M. Albu, X. Chen, J. Taipale, T.R. Hughes, M.T. Weirauch, The human transcription factors, Cell 172 (4) (2018) 650–665.
[25] P. Weidemüller, M. Kholmatov, E. Petsalaki, J.B. Zaugg, Transcription factors: bridge between cell signaling and gene regulation, Proteomics 21 (23–24) (2021 Dec) e2000034, https://doi.org/10.1002/pmic.202000034. Epub 2021 Aug 9. PMID: 34314098.
[26] B. Van de Sande, C. Flerin, K. Davie, M. De Waegeneer, G. Hulselmans, S. Aibar, R. Seurinck, W. Saelens, R. Cannoodt, Q. Rouchon, T. Verbeiren, D. De Maeyer, J. Reumers, Y. Saeys, S. Aerts, A scalable SCENIC workflow for single-cell gene regulatory network analysis, Nat. Protoc. 15 (7) (2020) 2247–2276.
[27] B. Van de Sande, C. Flerin, K. Davie, M. De Waegeneer, G. Hulselmans, S. Aibar, R. Seurinck, W. Saelens, R. Cannoodt, Q. Rouchon, T. Verbeiren, D. De Maeyer, J. Reumers, Y. Saeys, S. Aerts, A scalable SCENIC workflow for single-cell gene regulatory network analysis, Nat. Protoc. 15 (7) (2020 Jul) 2247–2276, https://doi.org/10.1038/s41596-020-0336-2IF:16.0.Q1. Epub 2020 Jun 19. PMID: 32561888.
[28] W. Hou, Z. Ji, Z. Chen, E.J. Wherry, S.C. Hicks, H. Ji, A statistical framework for differential pseudotime analysis with multiple single-cell RNA-seq samples, bioRxiv [Preprint]. 2021 Jul 12:2021.07.10.451910. doi: 10.1101/2021.07.10.451910. Update, Nat. Commun. 14 (1) (2023 Nov 10) 7286, https://doi.org/10.1038/s41467-023-42841-y. PMID: 34282418; PMCID: PMC8288148.
[29] Y. Wu, S. Yang, J. Ma, Z. Chen, G. Song, D. Rao, Y. Cheng, S. Huang, Y. Liu, S. Jiang, J. Liu, X. Huang, X. Wang, S. Qiu, J. Xu, R. Xi, F. Bai, J. Zhou, J. Fan, X. Zhang, Q. Gao, Spatiotemporal immune landscape of colorectal cancer liver metastasis at single-cell level, Cancer Discov. 12 (1) (2022) 134–153.
[30] Y. Guo, W. Pan, S. Liu, Z. Shen, Y. Xu, L. Hu, ERK/MAPK Signalling Pathway and Tumorigenesis (Review)[J], Experimental and Therapeutic Medicine, 2020.
[31] B. Pal, Y. Chen, M.J.G. Milevskiy, F. Vaillant, L. Prokopuk, C.A. Dawson, B. D. Capaldo, X. Song, F. Jackling, P. Timpson, G.J. Lindeman, G.K. Smyth, J. E. Visvader, Single cell transcriptome atlas of mouse mammary epithelial cells across development, Breast Cancer Res. 23 (1) (2021 Jun 29) 69, https://doi.org/10.1186/s13058-021-01445-4. PMID: 34187545 ; PMCID: PMC8243869.
[32] L. Cords, S. Tietscher, T. Anzeneder, C. Langwieder, M. Rees, N. De Souza, B. Bodenmiller, Cancer-associated fibroblast classification in single-cell and spatial proteomics data, Nat. Commun. 14 (1) (2023) 4294.
[33] Z. Zhang, J. Wang, Z. Zheng, X. Chen, G. Xu, S. Chen, F. Liu, L. Chen, M. Ding, L. Yuan, Y. Li, J. Qian, X. Xie, B. Deng, W. Lu, A protective polymorphism in *MMP16* , improved blood gas levels, and chronic obstructive pulmonary diseases: family and two population-based studies, Hum. Mutat. 41 (7) (2020) 1280–1297.
[34] S. Liu, M. Shen, E.-C. Hsu, C.A. Zhang, F. Garcia-Marques, R. Nolley, K. Koul, M. A. Rice, M. Aslan, S.J. Pitteri, C. Massie, A. George, J.D. Brooks, V. J. Gnanapragasam, T. Stoyanova, Discovery of PTN as a serum-based biomarker of pro-metastatic prostate cancer, Br. J. Cancer 124 (5) (2021) 896–900.
[35] Z. He, J. He, K. Xie, KLF4 transcription factor in tumorigenesis, Cell Death Discov. 9 (1) (2023) 118.
[36] Y. Inaba, E. Hashiuchi, H. Watanabe, K. Kimura, Y. Oshima, K. Tsuchiya, S. Murai, C. Takahashi, M. Matsumoto, S. Kitajima, Y. Yamamoto, M. Honda, S. Asahara, K. Ravnskjaer, S. Horike, S. Kaneko, M. Kasuga, H. Nakano, K. Harada, H. Inoue, The transcription factor ATF3 switches cell death from apoptosis to necroptosis in hepatic steatosis in male mice, Nat. Commun. 14 (1) (2023) 167.
[37] I.K. Townley, C.H. Babin, T.E. Murphy, C.M. Summa, B.B. Rees, Genomic analysis of hypoxia inducible factor alpha in ray-finned fishes reveals missing Ohnologs and evidence of widespread positive selection, Sci. Rep. 12 (1) (2022) 22312.

[38] S. Baral, M. Gyawali, N. Thapa, R.K. Chhetri, P. Dahal, Giant juvenile fibroadenoma in an adolescent female: a case report, Clin. Case Rep. 8 (12) (2020) 3488–3492.
[39] Y. Zhang, F. Zhong, L. Liu, Single-cell transcriptional atlas of tumor-associated macrophages in breast cancer, Breast Cancer Res. 26 (2024) 129, https://doi.org/10.1186/s13058-024-01887-6.
[40] M.F. Lerwill, A.H.S. Lee, P.H. Tan, Fibroepithelial tumours of the breast-a review, Virchows Arch. 480 (1) (2022 Jan) 45–63, https://doi.org/10.1007/s00428-021-03175-6. Epub 2021 Sep 10. PMID: 34505197.
[41] C. Lourenco, D. Resetca, C. Redel, P. Lin, A.S. MacDonald, R. Ciaccio, T.M. G. Kenney, Y. Wei, D.W. Andrews, M. Sunnerhagen, C.H. Arrowsmith, B. Raught, L. Z. Penn, MYC protein interactors in gene transcription and cancer, Nat. Rev. Cancer 21 (9) (2021) 579–591.
[42] N. Thapa, M. Chen, H.T. Horn, S. Choi, T. Wen, R.A. Anderson, Phosphatidylinositol 3-kinase signalling is spatially organized at endosomal compartments by microtubule-associated protein 4, Nat. Cell Biol. 22 (11) (2020) 1357–1370.
[43] D.W. McKellar, L.D. Walter, L.T. Song, M. Mantri, M.F.Z. Wang, I. De Vlaminck, B. D. Cosgrove, Large-scale integration of single-cell transcriptomic data captures transitional progenitor states in mouse skeletal muscle regeneration, Commun. Biol. 4 (1) (2021) 1280.
[44] A. Morato, P. Accornero, R.C. Hovey, ERBB receptors and their ligands in the developing mammary glands of different species: fifteen characters in search of an author, J. Mammary Gland Biol. Neoplasia 28 (2023) 10, https://doi.org/10.1007/s10911-023-09538-w.
[45] J. Huang, Y. Zhu, W. Zeng, et al., Single-cell transcriptome analysis reveals the malignant characteristics of tumour cells and the immunosuppressive landscape in HER2-positive inflammatory breast cancer, J. Exp. Clin. Cancer Res. 44 (2025) 196, https://doi.org/10.1186/s13046-025-03454-z.

List of abbreviations

Complete title: abbreviation
Fibroepithelial tumor and: FT
fibroadenomas: FA
phyllodes tumors: PT
single-cell transcriptome sequencing: scRNA-seq
unique molecular identifiers: UMIs
Hanks' balanced salt solution: HBSS
Principal component analysis: PCA
Uniform Manifold Approximation and Projection: UMAP
Differentially expressed genes: DEGs
Gene Ontology (GO): GO
Cellular Component: CC
Molecular Function: MF
Biological Process: BP
the Kyoto Encyclopedia of Genes and Genomes: KEGG
Transcription factors: TFs
tricarboxylic acid: TCA
Myofibroblasts: mCAFs
inflammatory associated fibroblasts: iCAFs
vascular fibroblasts: vCAFs
tumor-associated fibroblastsantigen-presenting: tCAFs
associated fibroblasts: apCAFs
fission fibroblasts: dCAFs
mediator complex subunit 12: MED12
telomerase reverse transcriptase: TERT
Gene Set Variation Analysis: GSVA
Epithelial-mesenchymal transition: EMT