



Unraveling the metastatic niche in breast cancer bone metastasis through single-cell RNA sequencing

Xiangyu Li¹ · Ziyu Gao¹ · Meiling Yang^{2,3} · Ciqiu Yang⁴ · Dongyang Yang⁴ · Wenhui Cui¹ · Dandan Wu⁵ · Jie Zhou⁵

Received: 23 February 2024 / Accepted: 1 July 2024

© The Author(s), under exclusive licence to Federación de Sociedades Españolas de Oncología (FESEO) 2024

Abstract

Purpose Breast cancer (BRCA) is characterized by a unique metastatic pattern, often presenting with bone metastasis (BoM), posing significant clinical challenges. Through the study of the immune microenvironment in BRCA BoM offer perspectives for therapeutic interventions targeting this specific metastatic manifestation of BRCA.

Methods This study employs single-cell RNA sequencing and TCGA data analysis to comprehensively compare primary-tumors (PT), lymph node metastasis (LN), and BoM.

Results and Conclusions Our investigation identifies a metastatic niche in BoM marked by an increased abundance of cancer-associated fibroblasts (CAFs) and reduced immune cell presence. A distinct subtype (State 1) of BRCA BoM cells associated with adverse prognosis is identified. State 1, displaying heightened stemness traits, may represent an initiation phase for BoM in BRCA. Complex cell communications involving tumor, stromal, and immune cells are revealed. Interactions of FN1, SPP1, and MDK correlate with elevated immune cells in BoM. CD46, MDK, and PTN interactions drive myofibroblast activation and proliferation, contributing to tissue remodeling. Additionally, MDK, PTN, and FN1 interactions influence FAP+ CAF activation, impacting cell adhesion and migration in BoM. These insights deepen our understanding of the metastatic niche in breast cancer BoM.

Keywords Breast cancer · Bone metastasis · Single-cell RNA sequencing · Immune microenvironment · Myofibroblast-immune interactions

Introduction

Breast cancer (BRCA) represents a formidable global public health challenge, taking the forefront in 2020 as the preeminent global cancer. Approximately 5% of BRCA patients manifest bone metastases (BoM) at the initial diagnosis,

with an elevated 75% risk of developing BoM over the subsequent decade [1]. Advanced BRCA exhibits a strikingly high incidence of BRCA BoM, ranging from 65 to 75%. Notably, bone tissue emerges as the primary site for distant metastasis in BRCA, affecting 60% to 75% of all metastatic BRCA cases, particularly in hormone receptor-positive BRCA patients [2]. However, due to pathophysiological impairment and lack of specificity, therapeutic agents are

Xiangyu Li and Ziyu Gao have made equal contributions to the paper.

✉ Xiangyu Li
lixiangyu@czmc.edu.cn

✉ Jie Zhou
sunnysmile@gzhmu.edu.cn

¹ The Stem Cell and Tissue Engineering Research Center, College of Pharmacy, Changzhi Medical College, Changzhi 046000, Shanxi, People's Republic of China

² Guangdong Cardiovascular Institute, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou 510080, China

³ Medical Research Institute, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou 510080, China

⁴ Medical Oncology, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou 510080, China

⁵ Department of Breast Oncology Surgery, Affiliated Cancer Hospital & Institute of Guangzhou Medical University, Guangzhou 510095, Guangdong, China

difficult to accumulate in metastatic bone [3]. Consequently, the analysis of pathological features and related biological parameters of bone metastases proves invaluable for predicting patient survival rates and recurrence risks. This profound understanding not only underscores the critical need for effective therapeutic strategies but also sheds light on the intricate interplay between BRCA and its metastatic cascade.

Main treatment strategy for BoM is to inhibit the growth of tumor cells, while ignoring the influence of the tumor stromal microenvironment (TSM) on the progression of BoM [4]. The intricate landscape of the tumor microenvironment (TME) is composed of diverse non-cellular factors and a myriad of cell types, including cancer-associated fibroblasts (CAFs), immune cells, endothelial cells, pericytes, and adipocytes. The multifaceted crosstalk among tumor, stromal cells, and immune cells not only underlies treatment resistance but also propels tumor progression and progression to overt BoM [3]. Hence, a nuanced comprehension of these extensive interactions assumes paramount importance in advancing the efficacy of tumor treatments. Previous investigations have unveiled that CAFs, predominantly activated fibroblasts influenced by the tumor, play instrumental roles in propelling BRCA progression. Their involvement spans a spectrum of functions, including fostering tumor cell proliferation, facilitating cancer cell invasion and metastasis, orchestrating extracellular matrix remodeling and deposition, promoting angiogenesis, instigating drug resistance, generating circulating CAFs (cCAFs), and secreting pro-tumor factors. Notably, CAFs contribute to the establishment of an immunosuppressive microenvironment, thus evading immune surveillance [5–8]. These insights, drawn from prior studies, underscore the pivotal role of CAFs in shaping the complex intercellular network within the TME, illuminating potential avenues for therapeutic interventions.

As predominant stromal constituents within the TME, CAFs intricately engage in dynamic dialogues with diverse immune cells. Employing a variety of paracrine mechanisms, CAFs meticulously secrete soluble factors that efficaciously impede anti-tumor immune responses. Playing a pivotal role, CAFs are central to the recruitment of Tumor-Associated Macrophages (TAMs), fostering a pro-tumor phenotype. In addition, they contribute significantly to the recruitment and differentiation of Tumor-Associated Neutrophils (TANs). In advanced tumor stages, TANs facilitate metastasis through extracellular trap release, immune response suppression, and production of cytokines and proteases. Furthermore, CAFs actively promote the migration and generation of Myeloid-Derived Suppressor Cells (MDSCs) via the secretion of cytokines and chemokines, exerting immunosuppressive effects on acquired and innate immunity. Integral to immune suppression, CAFs play a key role in converting CD4⁺ T cells to Regulatory T cells (Tregs) and T Helper lymphocytes

(Th) cells to Th2 cells. By regulating the differentiation and maturation of Dendritic Cells (DCs), CAFs inhibit antigen presentation, thus limiting T-cell activation. Moreover, CAFs hinder the infiltration of Cytotoxic T Lymphocytes (CTLs) into tumors, attenuating their tumoricidal potential. The intricate orchestration of these immunosuppressive mechanisms by CAFs, encompassing upregulation of immune checkpoint molecules, extracellular matrix remodeling via collagen, fibronectin, MMPs, and activation of the FAK signaling pathway, underscores their central role in mediating tumor immune escape through metabolic reprogramming and the production of immunosuppressive metabolites [9–16]. However, the precise involvement of CAFs in BRCA BoM remains elusive.

In this study, we utilized single-cell RNA sequencing (scRNA-seq) and conducted an extensive analysis of The Cancer Genome Atlas (TCGA) data. Through a comparative evaluation of primary tumors (PT), lymph node metastasis (LN), and bone metastasis (BoM), our investigation reveals a distinctive metastatic niche characterized by an increase in CAFs and a reduction in immune cell populations in BoM. Notably, we identified a unique subtype of BRCA BoM cells strongly associated with an adverse prognosis. Our analysis spans the exploration of genes, signaling pathways, and variations in the immune microenvironment across PT, LN, and BoM. By uncovering intricate cellular dialogues among tumor, stromal, and immune cells, we pinpoint pivotal interactions involving FN1, SPP1, and MDK that correlate with an augmented presence of immune cells in BoM. These findings provide insights into the complexities of the immune microenvironment in BRCA BoM and offer perspectives for therapeutic interventions targeting this specific metastatic manifestation of BRCA.

Materials and methods

Data acquisition

This study obtained approval from the Medical Ethics Committee of the Affiliated Cancer Hospital & Institute of Guangzhou Medical University, and all subjects provided informed consent preoperatively. Two separate data sets were employed for the scRNA-seq analysis.

Immunohistochemistry (IHC) staining assays were performed on formalin-fixed, paraffin-embedded tissue blocks retrieved from one BRCA BoM case in the eleventh thoracic vertebra after a thorough review of archived materials. This BoM data set included the expression profiles of 32,738 genes across 9181 individual cells.

The data set, GSE225600, was sourced from the Gene Expression Omnibus (GEO, <https://www.ncbi.nlm.nih.gov/>)

geo) on October 17, 2023. It encompassed gene expression data from a total of 81,683 cells across four PT and their corresponding four paired LN, providing insights into the expression patterns of 36,601 genes.

Bulk data for BRCA patients were acquired from The Cancer Genome Atlas Genomic Data Commons (TCGA GDC) via UCSC Xena (<https://xenabrowser.net/>, accessed on 2023/12/10, cohort: GDC TCGA Breast Cancer). This gene expression data set included count data for 60,488 genes across 1217 samples. In addition, survival data for 1260 BRCA samples and clinical data for 1248 BRCA samples were obtained from the same source.

Single-cell RNA-seq data preprocessing

High-quality sequencing reads were processed using “Cell Ranger” (version: 3.0.2) for tasks including sequence alignment, filtering, barcoding, and unique molecular index counting. The analysis utilized the hg19 reference genome. Subsequent scRNA-seq data examination was performed with the “Seurat” package (version: 5.0.1) in R software (version: 4.3.1) (<https://satijalab.org/seurat/>).

The analysis proceeded through several stages, including data quality control, normalization, and differential gene expression analysis. Each scRNA-seq data set underwent rigorous filtering to exclude cells with fewer than 200 genes or more than 10% mitochondrial gene expression. In addition, genes detected in fewer than 10 cells were removed. Normalization was performed using the “NormalizeData” function with default parameters, followed by dimensionality reduction through principal component analysis (PCA), resulting in a 13-dimensional output for the two data sets.

Clustering analysis utilized the “FindClusters” function with resolutions of 10 and 7 for BoM and GEO data sets, respectively. The identification of doublets employed the “DoubletFinder” R package (version: 2.0.3; <https://github.com/chris-mcginnisucsf/DoubletFinder>). Batch effects were corrected using the “IntegrateData” function to integrate the two data sets. Two GEO data samples with insufficient cells were excluded, resulting in a final integrated data set of 40,333 genes in 34,375 cells from seven samples (3 PT, 3 LN, and 1 BoM).

Cell clustering and annotation

The integrated data set values were normalized using z-score conversion via the “ScaleData” command. Highly variable genes were identified using the “FindVariableGenes” function with default parameters. Principal components were computed based on these selected genes and projected onto all other genes using the “RunPCA” function. Subsequently, the “FindNeighbors” and “FindClusters” commands were

utilized to identify clusters of similar cells, creating a shared nearest neighbor map with a resolution empirically set.

In analyzing the integrated data set, principal components were utilized to capture heterogeneity, primarily reflecting distinctions in tissue compartments. Clustering based on the expression of cell type markers resulted in the categorization of four clusters: epithelial cells, endothelial cells, immune cells, and fibroblasts. The same analytical approach was employed for immune cells, identifying cell types by matching cluster-specific gene sets with known signature genes from relevant literature [17–24].

Clusters lacking significant marker gene expression were classified according to their most differentially expressed genes. This comprehensive method facilitated a detailed exploration of cellular diversity within the integrated data set, yielding insights into tissue-specific compartments and a variety of cell types within the examined samples.

Quantification of epithelial cell copy number variation

We employed the “infercnv” R package (v1.16.0) to assess copy number variation (CNV) in individual epithelial cells, using fibroblasts as reference normal cells. Setting a cutoff of 0.1 and enabling denoising, we computed CNV scores systematically for a comprehensive evaluation of copy number alterations in the epithelial cell population.

Identification and functional enrichment of differentially expressed genes

Differentially expressed genes (DEGs) were identified using the “FindMarkers” function within the “Seurat” R package, employing the Wilcoxon Rank Sum test with a log₂ fold change threshold of 0.1. Stringent filtering criteria were applied, requiring an absolute average log₂ fold change > 1 and a *p* value < 0.05 to refine the results.

For Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enrichment analysis, the “clusterProfiler” R package (version: 4.8.3) was utilized. Significantly enriched GO terms and KEGG pathways were determined based on a *q* value < 0.05.

Cancer hallmark enrichment analysis employed a reference set of 50 gene sets from MSigDB (accessed on 2023/11/01). Enrichment scores were computed using the “UCell” (v2.7.1; <https://github.com/carmonalab/UCell>) and “irGSEA” (v2.1.5; <https://github.com/chuiqin/irGSEA>) R packages, providing a detailed exploration of the functional significance of differentially expressed genes in relation to cancer hallmarks.

Pseudotime analysis and stemness inference

Pseudotime analysis using the “monocle” R package (v2.28.0) identified variable genes to delineate pseudotime stages, pinpoint trajectory differentiation genes, and scrutinize branch point gene alterations. This analysis revealed functional disparities among pseudotime stages, defining distinct BoM states. Stemness scores for each state, computed with established markers (OCT4, SOX2, NANOG), offered insights into associated stem cell-like characteristics.

Identification of markers for early stage BoM

To characterize state identities, we initiated the identification of DEGs specific to BoM state 1 using the “FindAllMarkers” function within the “Seurat” package. Significant genes were filtered based on a stringent threshold, requiring an absolute average log₂ fold change exceeding 1.5 and an adjusted *p* value below 0.05. To further refine marker selection, we employed the R package “rpart” (version: 4.1.23) to construct a tree model, utilizing the method parameter and pruning the tree with “cp=0.01”. This meticulous process resulted in the identification of markers specifically associated with the early state of BoM.

Prognostic values of BoM early state markers

We assessed the prognostic significance of identified markers in TCGA BRCA patients by calculating enrichment scores. These scores, obtained by multiplying gene expression values with weights assigned by a constructed tree model, were used to stratify patients into two groups based on median values. Prognostic effectiveness in predicting 10-year survival was evaluated using R packages “survival” (version: 3.5–7) and “survminer” (version: 0.4.9). In addition, we investigated dynamic changes in these markers across various clinical features of TCGA BRCA samples.

Cell–cell interaction analysis

In examining BoM intercellular interactions, we employed the “CellChat” R package (v1.6.1; <https://github.com/sqjin/CellChat>), tailored for carcinoma cells, immune cells, and fibroblasts. This tool includes a molecular signaling database with 60% paracrine/autocrine, 21% ECM-receptor, and 19% cell–cell contact interactions. Through comprehensive cellular communication analysis, we gained valuable insights into intercellular communication networks, contributing to a nuanced understanding of diverse cell-type interactions during BRCA BoM development.

Clinical validation

Using the KM plotter (<https://www.kmplot.com/analysis/>), we conducted a survival analysis of differentially expressed genes in the population of patients with osteosarcoma. Through the Kaplan–Meier analysis method, we obtained survival curves for the relevant genes.

Statistical analysis

All statistical analyses were conducted using R software (v4.3.1). Differentially expressed genes (DEGs) were identified using the Wilcoxon Rank Sum test with a log₂ fold change threshold of 0.1 and an adjusted *p* value of <0.05. Batch effect correction was performed using the “IntegrateData” function in Seurat. Functional enrichment analysis was conducted with significance thresholds set at *q* <0.05. Cancer hallmark enrichment analysis was performed with enrichment scores calculated using the U-statistic, and significance was determined by permutation tests. Pseudotime trajectory analysis was carried out, with genes identified as variable using a dispersion-based method, and the significance of trajectory changes assessed using the likelihood ratio test. Survival analysis was conducted using log-rank tests. For cell–cell interaction analysis, signaling pathway enrichment was determined using permutation tests, with significance thresholds set at *p* <0.01. Statistical significance was set at *p* <0.05 unless otherwise stated. The flowchart shows the concise workflow of the research (Fig. 1).

Results

Single-cell RNA-seq profiling of BRCA primary tumor, lymph node, and bone metastasis

In elucidating BRCA BoM, we integrated two single-cell sequencing data sets and conducted a thorough analysis (Fig. 2A). Following rigorous quality control, we obtained transcriptome data from 34,375 cells, comprising 12,427 PT cells, 15,740 LN cells, and 6208 BoM cells. Subsequent to dimension reduction, clustering, and cell annotation (Fig. 2B–D), a predominant ratio of immune cells was observed in LN samples. Notably, a decrease in the ratio of immune cells, coupled with a noticeable increase in epithelial cells and fibroblasts, was noted in comparison to PT and LN samples. This observation leads us to propose that immune cells and fibroblasts may have crucial roles in the BoM process.

Next, we employed the “infercnv” R package to infer copy number variations in epithelial cells. Comparative analysis with fibroblasts unveiled extensive mutations in epithelial

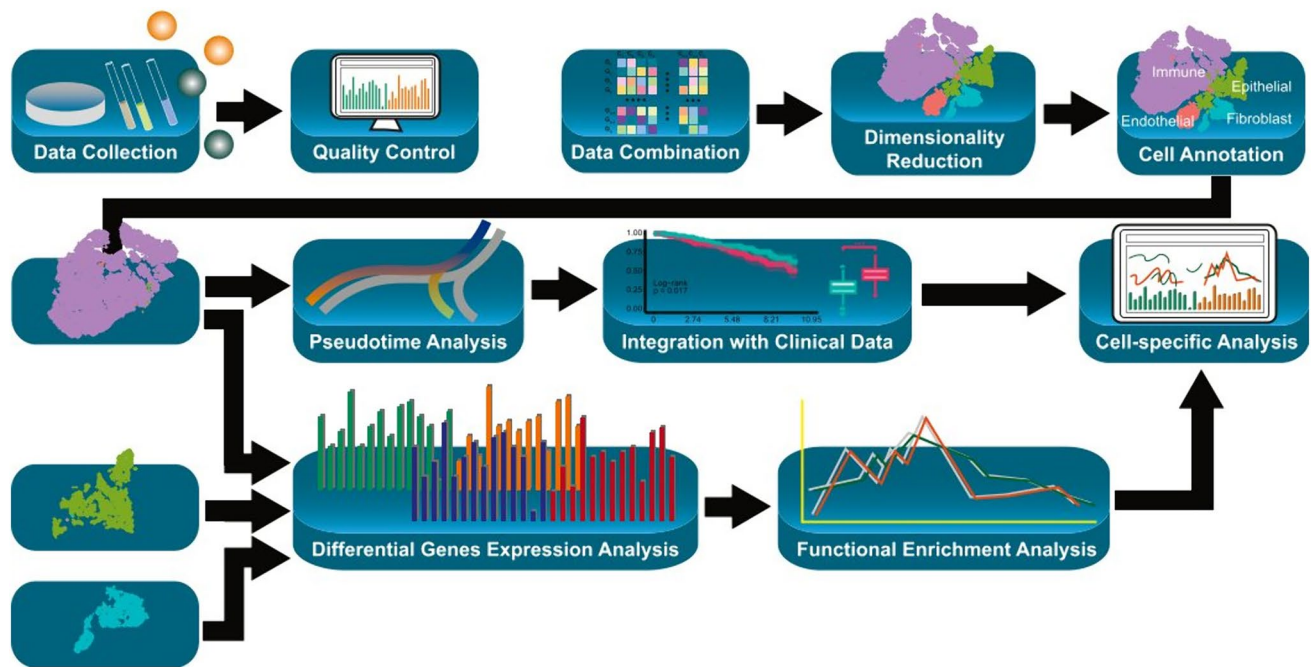


Fig. 1 Detailed Flowchart of the Research Design, Implementation, and Analysis Process. Data collection: This stage involves the acquisition and preparation of data. Quality control: Ensuring the integrity and reliability of the collected data. Data combination: Integrating different data sets to create a comprehensive data set. Dimensionality reduction: Simplifying the data to highlight the most critical features. Cell annotation: Categorizing cells based on specific characteristics. Pseudotime analysis: Analyzing cellular developmental trajectories.

cells, indicating consistent characteristics with carcinoma across all epithelial cells (Fig. 2E).

Metabolic and immune signaling pathways up-regulated in BRCA BoM

Utilizing the “FindMarkers” function in the “Seurat” R package, we identified DEGs in tumor cells originating from distinct BRCA types. Simultaneously, GSEA was performed, revealing significant enrichment of upregulated genes in the Toll-like receptor signaling pathway, PI3K/Akt/mTOR pathway, and specific metabolic programs in BRCA cells from BoM, highlighting the pivotal roles of immune response and cell proliferation in BRCA BoM progression (Fig. 3A, B).

Employing pseudotime trajectory analysis on PT and BoM cells, excluding LN samples with insufficient tumor cells, revealed three distinct states in BRCA BoM progression (Fig. 3C, D). State 1, prominent at the trajectory's onset, exhibited the highest stemness score (Fig. 3E), while state 2, characterized by the lowest stemness score, predominated at the trajectory's conclusion. Functional enrichment analysis of up-regulated KEGG pathways in these states indicated that state 1 primarily engages in biological processes related to cell growth, development, proliferation, differentiation,

Integration with clinical data: Correlating findings with clinical data for validation. Differential gene expression analysis: Identifying differentially expressed genes. Functional enrichment analysis: Understanding the biological processes and pathways involved. Cell-specific analysis: Investigating interactions and communications among various cell types, including immune, epithelial, endothelial, and fibroblast cells

and cell adhesion (Fig. 3F). As state 1 manifests in both PT and BoM samples early in the trajectory and displays heightened stemness traits, it is suggested to represent an initiation phase for BoM in BRCA.

Identification of marker genes for early stage BRCA BoM

The “rpart” R package was utilized to construct a classification and regression tree for identifying pivotal genes in the progression of BRCA BoM. The recursive partitioning and regression tree revealed three marker genes (ZNF831, CTLA4, and GIMAP7) and their positions within the decision tree model (Fig. 4A). TCGA BRCA samples were scored based on gene expression in RNA-seq data and tree model-derived weights. BRCA samples were then categorized into State 1 and non-State 1 based on the median score (Fig. 4B). Kaplan–Meier analysis demonstrated a significantly lower overall survival for BRCA patients in State 1 compared to non-State 1 (Fig. 4C).

BRCA status distribution across clinical features exhibits its variable State 1 percentages (Fig. 4D), indicative of dynamic BRCA progression. This implies potential for

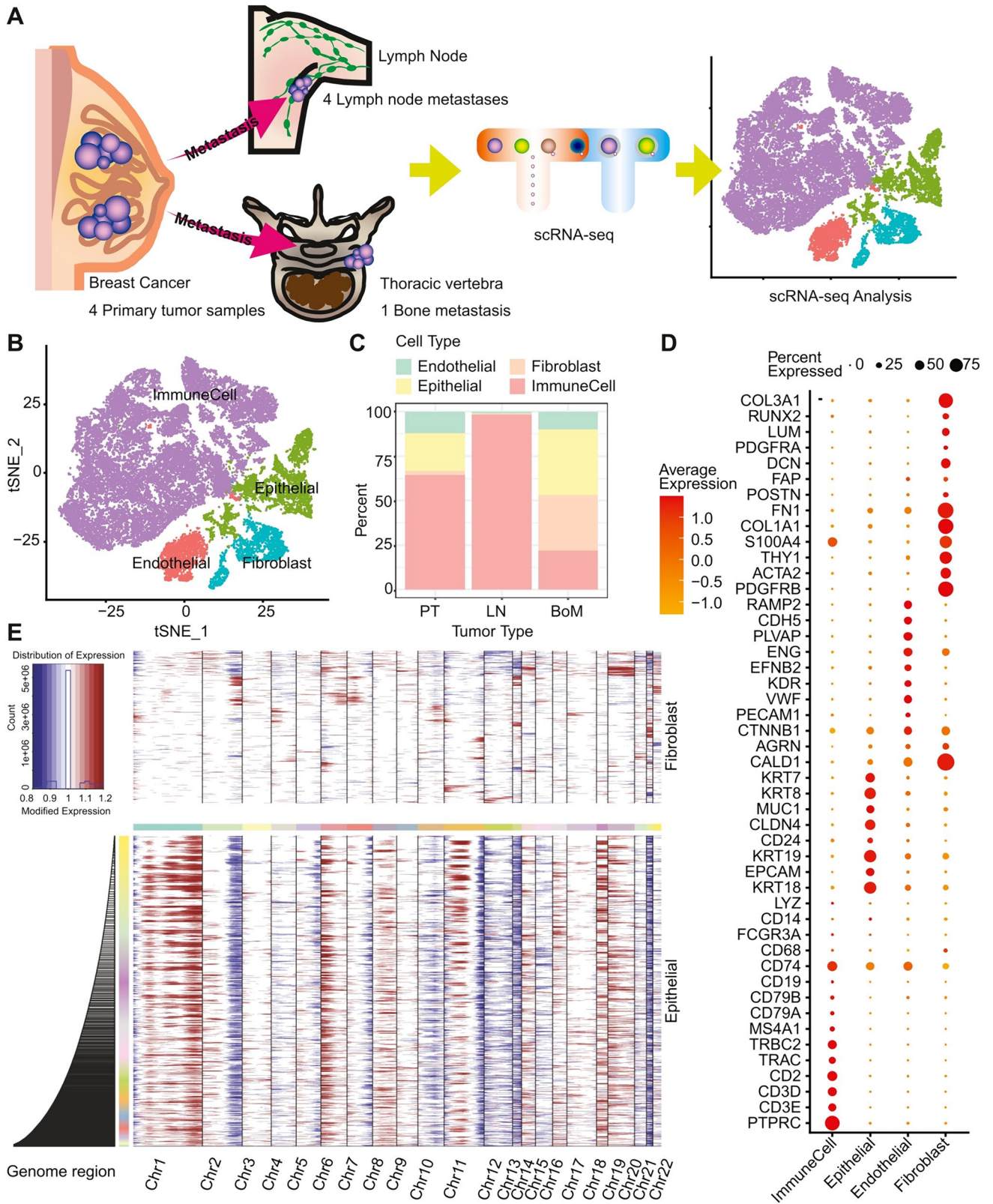


Fig. 2 Single-cell RNA-seq Atlas of BRCA Primary Tumor and Metastasis. **A** Workflow overview illustrating the decoding of BRCA progression using single-cell RNA-seq (scRNA-seq). Single-cell suspensions from PT, LN, and BoM were subjected to scRNA-seq using the 10×Genomics platform. **B** A t-distributed Stochastic Neighbor Embedding (t-SNE) plot, derived from integrated BRCA data ($n=34,375$ cells), visually delineates principal cell types. **C** Proportional representation of each cell type across different tumor types in BRCA is shown in a bar chart. **D** Dot plot presenting marker gene expression levels in indicated cell types. Dot size indicates the proportion of cells expressing the marker within the group, while color represents marker expression levels. **E** Heatmap depicting results from “infercnv” provides insights into copy number variations across samples

distant metastasis in primary tumor cells as BRCA advances, leading to subsequent transformation into BoM tumor cells.

Critical involvement of immune cells in BRCA-driven BoM

To unravel the dynamics of immune cells in BRCA BoM, we employed re-clustering techniques and marker gene annotations (Fig. 5A, B). Seven distinct immune cell clusters emerged, with two identified clusters devoid of significant immune cell type marker gene expression, highlighted by their highest DEGs. Analysis of immune cell percentages in PT, LN, and BoM revealed a notable increase in myeloid cells and a relative rise in cytotoxic NK-T cells in BoM (Fig. 5C). To gain comprehensive insights into the functions of myeloid cells in BoM, we further re-clustered them into six distinct subsets termed TAM 1–6 (Fig. 5D). The bar plot illustrates that, except for TAM3, all clusters were predominantly present in BoM (Fig. 5E), actively engaging in biological processes such as cell adhesion, immune response, and immune regulation (Fig. 5F).

We analyzed immune cell dynamics in BRCA BoM using re-clustering techniques and marker gene annotations (Fig. 5A, B), identifying seven immune cell clusters. Two clusters lacked significant immune cell marker gene expression, highlighted by their highest DEGs. Examination of immune cell percentages in PT, LN, and BoM indicated increased myeloid cells and a relative rise in cytotoxic NK-T cells in BoM (Fig. 5C). Further re-clustering of myeloid cells revealed six subsets (TAM 1–6) with all but TAM3 predominantly present in BoM, actively participating in cell adhesion, immune response, and immune regulation (Fig. 5D–F).

Deciphering immune cell interactions in BoM TME

In BoM, we utilized “CellChat” to assess immune cell impact on the TME. Numerous interactions were identified (Fig. 6A), highlighting predominant pathway presence (Fig. 6B). Our investigation unveiled that cytotoxic NK-T cells in BoM release CD8A, engaging in communication

with B cells, naïve T cells, myeloid cells, regulatory T cells, and carcinoma cells through interaction with HLA-A, HLA-B, and HLA-C (Fig. 6C), potentially augmenting immune responses in BoM.

BoM state 1 carcinoma cells primarily interact with cytotoxic NK-T cells, myeloid cells, and B cells (Fig. 6D). The engagement is facilitated by FN1, HLA genes, and MDK, interacting with their respective target genes. Notably, the MDK-NCL interaction (Fig. 6E) has a substantial impact, potentially suppressing the immune response. Concurrently, myeloid cells and B cells predominantly target BoM state 1 tumor cells through interactions with FN1, SPP1, GRN, and MK, along with their target genes (Fig. 6F). The activation of signaling pathways by FN1 may contribute to cell survival, metastasis, and the progression of BoM (Fig. 6G).

Decoding fibroblast dynamics and interactions in BRCA BoM

In investigating fibroblast dynamics in BRCA BoM, we utilized re-clustering techniques and marker gene annotations, revealing three distinct fibroblast cell clusters (Fig. 7A, B). Analysis of fibroblast proportions in PT, LN, and BoM indicated a notable increase in myofibroblast cells and FAP⁺ inflammatory cells in BoM (Fig. 7C). Functional exploration demonstrated that myofibroblast cells and FAP⁺ inflammatory cells in BoM are actively involved in processes such as cell proliferation, adhesion, and complement and coagulation cascades, as indicated by KEGG functional enrichment (Fig. 7D).

Fibroblast cell influence in BoM was evaluated using “CellChat”, revealing numerous interactions in BoM (Fig. 7E). Our investigation identified FAP⁺ inflammatory cells releasing FN1, engaging in communication with carcinoma cells, myofibroblast cells, FAP⁻ inflammatory cells, and self-interactions (Fig. 7F).

In BoM, state 1 carcinoma cells engage primarily with myofibroblast cells and FAP⁺ inflammatory cells (Fig. 7G), facilitated by MDK and CD46, activating their target genes (Fig. 7H). MDK-SDC1/SDC4 interactions play a significant role, potentially enhancing cell proliferation, angiogenesis, and epithelial-mesenchymal transition (EMT), promoting metastasis from the primary site (Fig. 7H). The overall survival of the high-expression group of MDK was significantly higher than that of the low-expression group (Fig. 7K). Simultaneously, myofibroblast cells and FAP⁻ inflammatory cells target BoM state 1 tumor cells through THBS, PTN, and NOTCH pathways (Fig. 7I). PTN activation may contribute to stimulating new blood vessel formation and tumor angiogenesis (Fig. 7J). The overall survival of the high-expression group of PTN was significantly lower than that of the low-expression group (Fig. 7L).

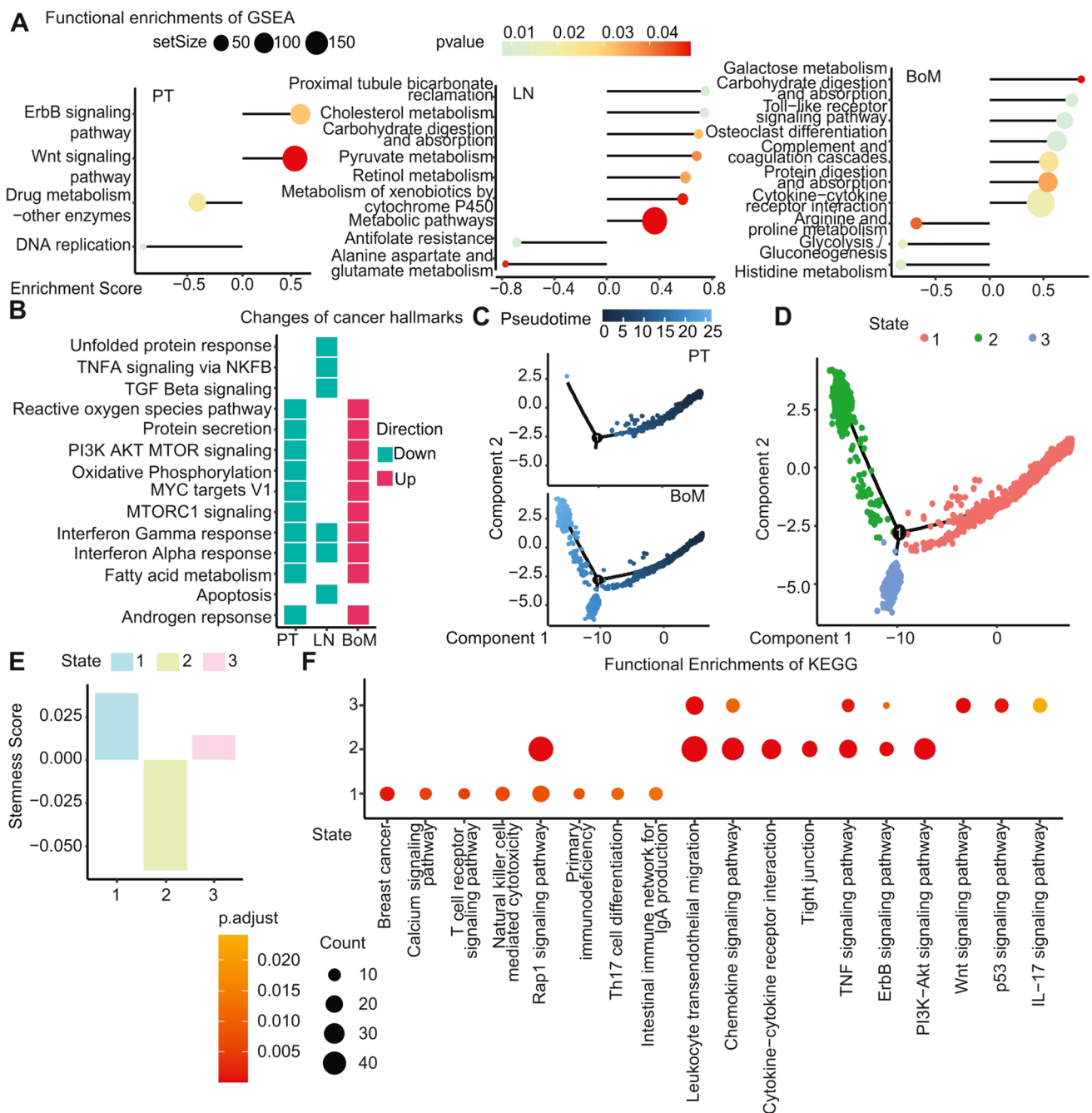


Fig. 3 Functional Enrichment and Pseudotime Analysis. **A** Dot plot illustrating Gene Set Enrichment Analysis (GSEA) functional enrichment results across BRCA types. Dot sizes correspond to gene set sizes, and colors indicate enrichment *p* values. **B** Heatmap presenting changes in cancer hallmarks within each BRCA tumor type. **C** Pseudotime-ordered analysis of tumor cells from PT and BoM, with the spectrum of blue indicating the temporal order. **D** Pseudotime states

are color-labeled, with each dot representing a single cell. **E** Bar plot displaying inferred stemness scores of pseudotime states. **F** Dot plot showcasing Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment results for each state, where dot size reflects the gene set size, and color indicates the *p* value of the result

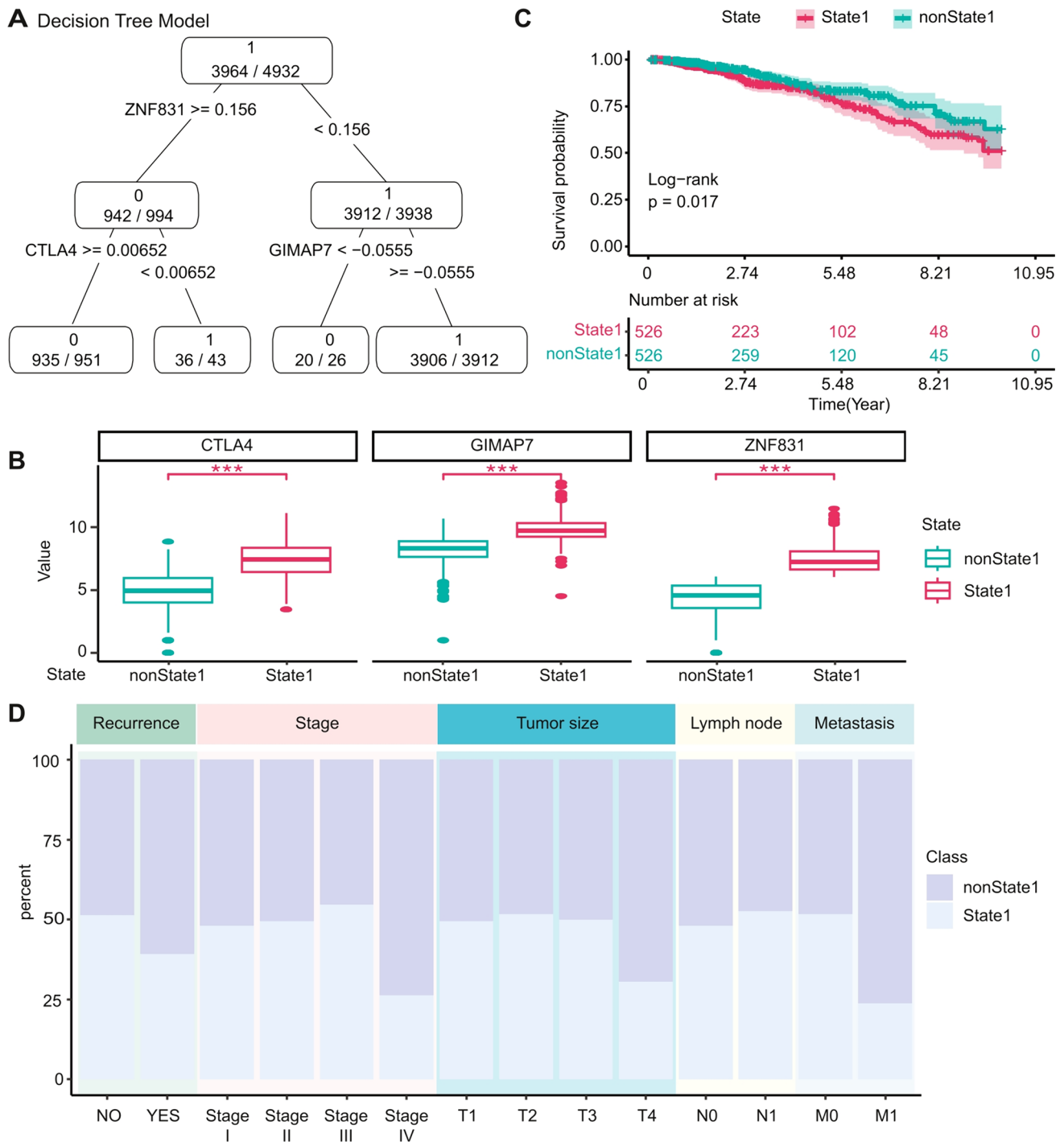


Fig. 4 BoM State 1 Markers Identification and Its Association with BRCA Clinical Features. **A** Decision tree model illustrating the identification of BoM State 1 markers. **B** Box plot displaying the expression levels of individual markers in The Cancer Genome Atlas

(TCGA) BRCA samples. **C** Kaplan–Meier plot revealing significant differences between two BRCA states. **D** Bar plot illustrating the distribution of the two BRCA states across distinct clinical features

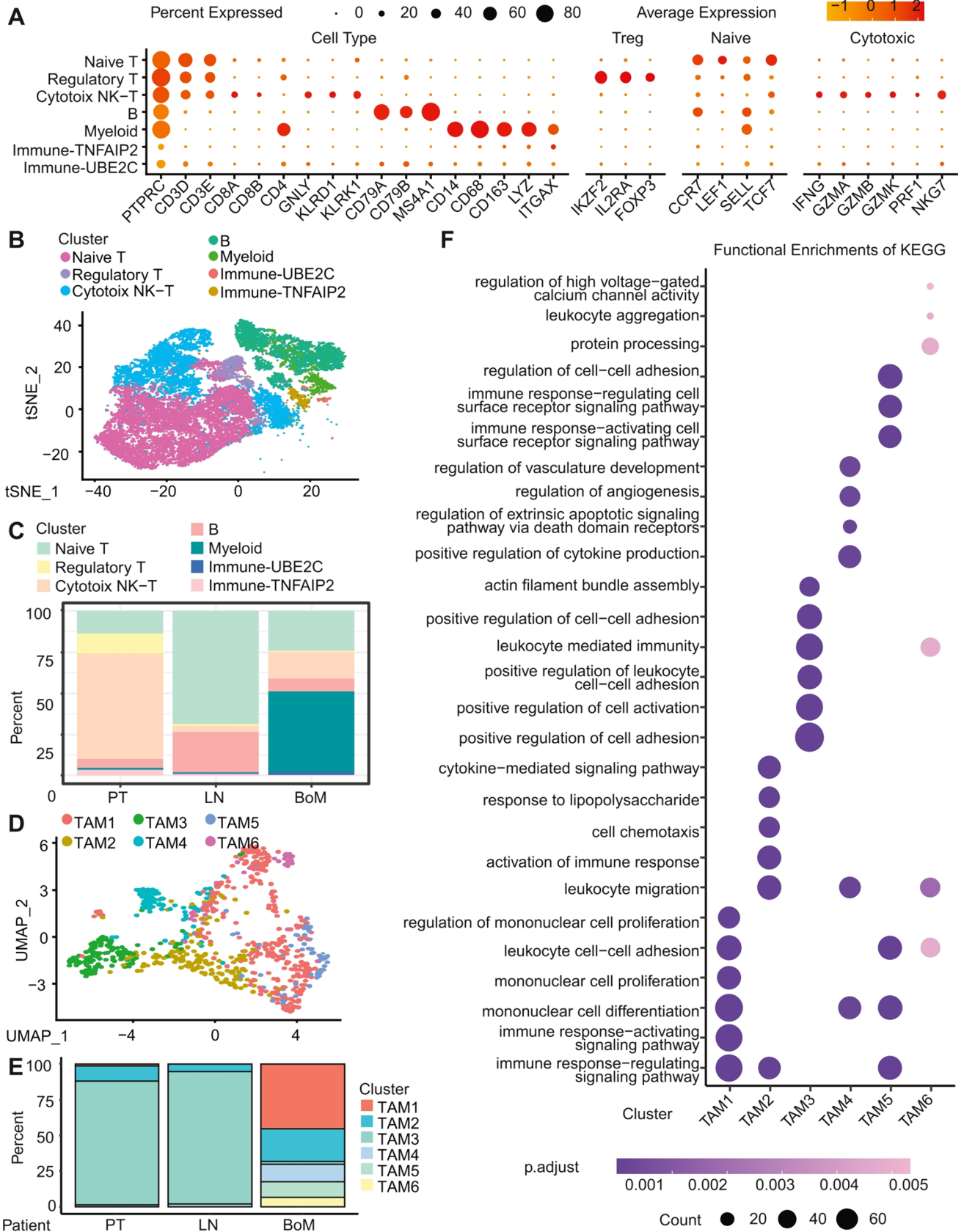


Fig. 5 Immune Cell Annotation and Myeloid Cell Analysis. **A** Dot plot displaying marker gene expression in specified cell types, with dot size indicating the proportion of expressing cells and color representing marker expression levels. **B** Utilizing t-SNE, a plot generated from integrated immune data visually delineates principal cell types. **C** Bar chart illustrating the proportional representation of each immune cell type across various BRCA tumor types. **D** Uniform Manifold Approximation and Projection (UMAP) plot depicts the primary cell types of TAM cells extracted from immune cell data. **E** Bar chart presenting the proportional representation of each TAM cell type across diverse BRCA tumor types. **F** Dot plot revealing Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment results for each TAM cluster, with dot size indicating gene set size and color denoting the *p* value of the result

Myofibroblast-immune interactions in BRCA BoM progression

In examining cell communications between fibroblast cells and immune cells, numerous interactions were identified in the extracellular matrix (Fig. 8A). Quantitative analysis revealed a heightened frequency of interactions between myofibroblast cells and immune cells compared to other cell types (Fig. 8B). Notably, cytotoxic NK-T cells and myeloid cells were the predominant interactors with myofibroblast cells (Fig. 8C), facilitated by PTN, MDK, and LAMININ signaling pathways (Fig. 8D). Simultaneously, myofibroblast cells and FAP⁺ inflammatory cells targeted myofibroblast cells through interactions involving PTN, MDK, and SPP1 signaling pathways (Fig. 8E). The PTN-NCL interaction significantly influenced both incoming and outgoing interactions with myofibroblast cells, potentially influencing cell proliferation, angiogenesis, metastasis, and resistance to apoptosis in cancer cells (Fig. 8F).

Discussion

The emergence of BoM is of significant prognostic importance in BRCA, underscoring the necessity to delve into the intricate pathogenesis and molecular regulatory networks governing this phenomenon. In this study, we undertook a comprehensive investigation of BRCA, with a particular focus on both LN and BoM. Our meticulous analysis provides valuable insights into the nuanced intricacies of BRCA progression, with special attention to the metastatic niche of BoM. Specifically, we observed a distinct elevation in CAFs alongside a reduction in immune cells within the bone metastatic microenvironment. These findings enhance our understanding of the disease and present potential avenues for therapeutic interventions.

Several recent studies have extensively explored the intricate microenvironments within BoM, shedding light on the niche that supports tumor colonization [25–27]. The dynamic interplay of tumor-stromal interactions orchestrates

the progression from initial seeding to the development of overt macrometastasis. Consistent with these findings, our investigation into the early-stage colonization of breast cancer bone metastasis (BRCA BoM) aligns with the observed overexpression of heterotypic adherens junctions and an up-regulation of calcium influx. A pivotal outcome of our research is the identification of a distinct subtype of BRCA BoM cells. This specific subtype demonstrates a close correlation with the occurrence of BRCA BoM and serves as an indicator of an unfavorable prognosis. Through a comparative analysis of cancer hallmarks between BoM and PT, as well as LN, we unveiled a predominant upregulation of metabolic and Toll-like receptor signaling pathways in BoM. This highlights significant molecular distinctions in the metastatic microenvironment.

To gain deeper insights, we further stratified BoM into three distinct states using stemness scores. Intriguingly, State 1, characterized by the highest stemness, was found to coexist in both primary and metastatic sites, acting as the initiating point for BoM. KEGG functional enrichment analysis of State 1 underscored its involvement in critical biological processes, including cell growth, development, proliferation, differentiation, and cell adhesion. By comparing our findings with the conclusions drawn in recent publications [28], we contribute to the ongoing discourse on bone metastatic microenvironments. Our identification of a specific BRCA BoM cell subtype and the delineation of distinct functional pathways provide novel perspectives for understanding and potentially targeting the unique aspects of metastasis within the bone microenvironment.

Ma et al. recently identified a specific subset of protumorigenic macrophages which derived from CCL2-recruited inflammatory monocytes, promoting BRCA BoM in an IL-4R-dependent manner [29]. Our exploration has illuminated the intricate dynamics of communication between BRCA cells and immune cells, providing a nuanced understanding of the immune landscape. Employing distinct biomarkers for immune cell identification, we observed a significant upregulation of myeloid cells in BoM as opposed to PT and LN. Further elucidating the myeloid cell landscape through dimensionality reduction and clustering revealed that specific clusters of TAM were notably elevated in BoM. These clusters were found to predominantly engage in processes associated with cell adhesion and immune response, as substantiated by KEGG functional enrichment analysis.

Remarkably, our examination of cell–cell interactions has revealed a substantial augmentation in interactions specific to the formation of the metastatic niche within BoM. Noteworthy is the identification of Major Histocompatibility Complex class I (MHC-I) as a central mediator in facilitating communication between tumor cells and immune cells, as well as orchestrating intercellular interactions among

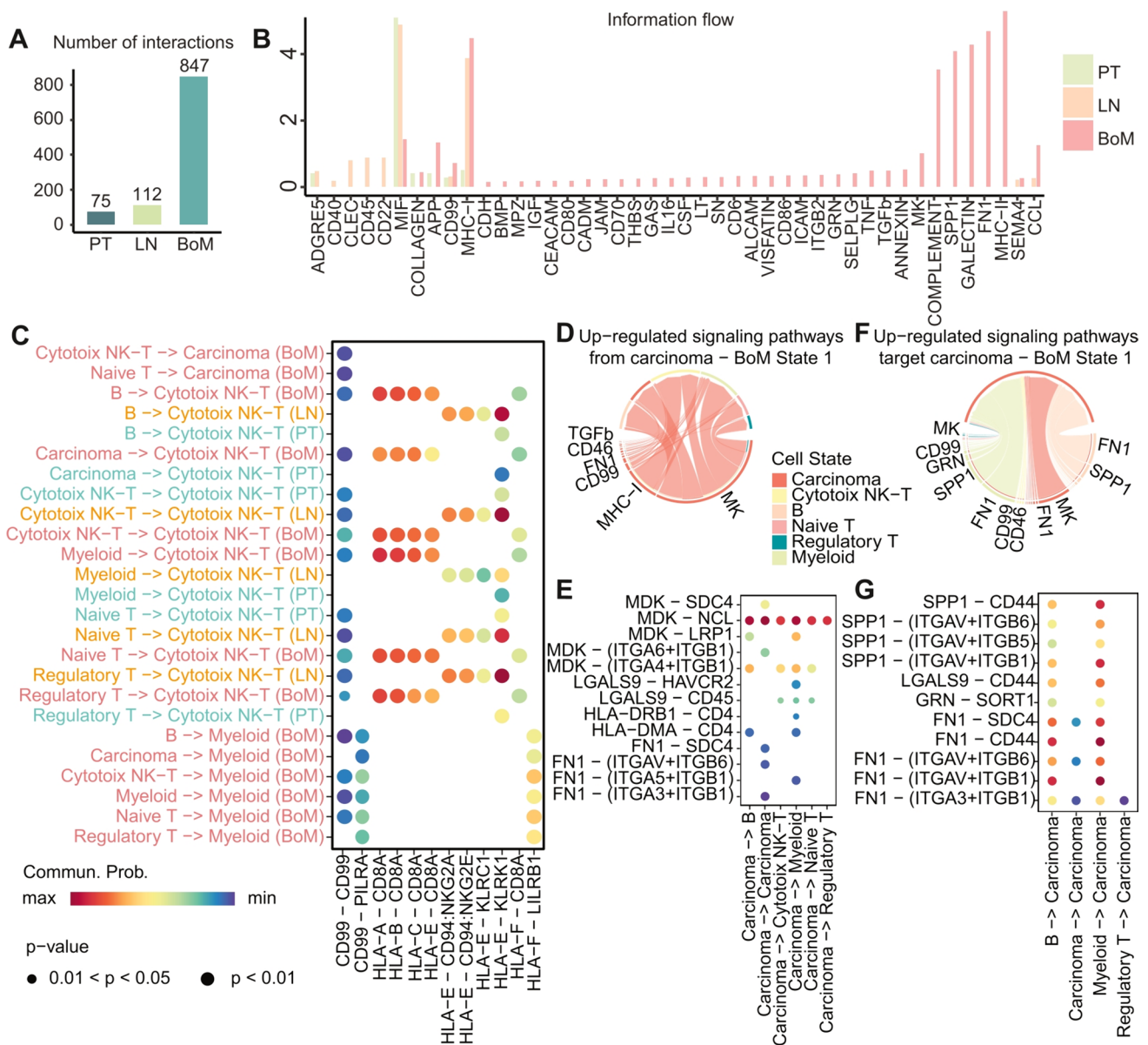
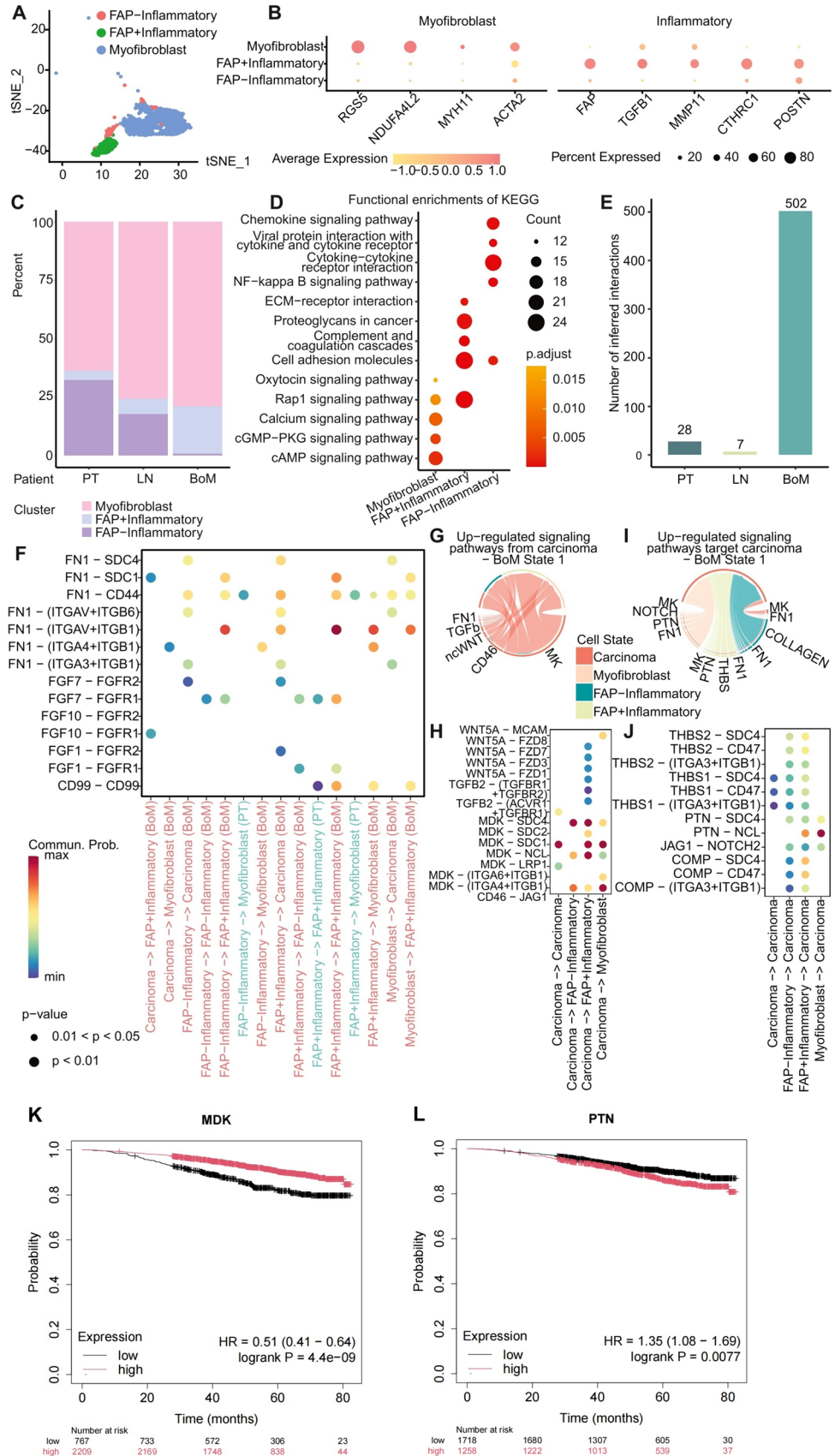


Fig. 6 Interactions between BRCA Tumor Cells and Immune Cells. **A** Quantification of carcinoma cell and immune cell interactions in PT, LN, and BoM depicted in a bar plot. **B** Bar plot presenting the cell-cell interaction count for each signaling pathway in PT, LN, and BoM. **C** Dot plot visualizing cell-cell communication probabilities within the MHC-I and CD99 pathways. **D** Chord plot showcasing up-regulated signaling pathways originating from BoM state 1 carcinoma cells and connecting to various immune cell types. **E** Dot plot revealing communication probabilities of BoM-specific pathways (MK, MHC-II, and FN1) from carcinoma cells. **F** Chord plot presenting up-regulated signaling pathways targeting BoM state 1 carcinoma cells from diverse immune cell types. **G** Dot plot demonstrating communication probabilities of BoM-specific pathways (SPP1 and FN1) targeting carcinoma cells

immune cells specifically within the BoM microenvironment, a phenomenon not as prominently observed in the PT or LN. Intriguingly, our focused analysis of interactions involving FN1, SPP1, and MDK with their target genes has yielded additional insights. These interactions were found to significantly contribute to an augmentation in myeloid cells, B cells, Naive T cells, and Cytotoxic T cells within the dynamic milieu of the BoM microenvironment. This

intricately orchestrated interplay emphasizes the influential role of specific signaling pathways in shaping the immune landscape of BRCA BoM. Our study brings forth a nuanced understanding of the roles played by cancer-associated myofibroblasts and inflammatory CAFs within the metastatic niche. Particularly in BoM, a significant augmentation of myofibroblasts and FAP⁺ inflammatory CAFs was observed in comparison

Fig. 7 Fibroblast Insights in BRCA BoM. **A** t-SNE-based dimensionality reduction clustering plot of fibroblasts. **B** Dot plot showcasing marker gene expression in specific cell types, with dot size indicating proportion and color denoting expression levels. **C** Bar chart illustrating the proportional representation of fibroblast cell types across diverse BRCA tumor types. **D** Dot plot revealing Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment in fibroblast clusters, with dot size indicating gene set size and color representing the *p* value. **E** Bar plot quantifying interactions between carcinoma cells and fibroblasts in PT, LN, and BoM. **F** Dot plot visualizing cell-cell communication probabilities within FN1, CD99, and LAMININ pathways. **G** Chord plot highlighting up-regulated signaling pathways originating from BoM state 1 carcinoma cells and connecting to various fibroblast cell types. **H** Dot plot displaying communication probabilities of BoM-specific pathways (MK, CD46, ncWNT, and TGFb) from carcinoma cells. **I** Chord plot presenting up-regulated signaling pathways targeting BoM state 1 carcinoma cells from diverse fibroblast cell types. **J** Dot plot demonstrating communication probabilities of BoM-specific pathways (THBS, PTN, and NOTCH) targeting carcinoma cells. **K** The survival curves of MDK using the KM plotter. **L** The survival curves of PTN using the KM plotter



group of MDK was significantly higher than that of the low-expression group.

In conclusion, our investigation has meticulously constructed a comprehensive single-cell map, providing a detailed portrayal of the metastatic niche throughout the spectrum of BRCA progression, encompassing in situ conditions, LN, and BoM. The systematic delineation of the metastatic niche in BoM has uncovered distinctive features, unraveling the intricate mechanisms that govern the immunosuppression induced by cancer cells upon metastasizing to the bone.

Supplementary materials

BoMrawData: The eleventh thoracic vertebra of the BRCA BoM case in the BoM data set contains expression profiles of 32,738 genes across 9181 individual cells.

Acknowledgements We would like to express our gratitude to Professor Wei Zhao from the Zhongshan School of Medicine at Sun Yat-sen University for his guidance and assistance. This work was supported by National Natural Science Foundation of China (No.82003805, No.82002776), Guangzhou Science and Technology Project (No.2024A03J0649), and Shanxi Province Science Foundation for Youths (No.201901D211471), Young Academic and Technical Leaders Project of Changzhi Medical College (No.XSQ202101). Acknowledgement to all funding sources.

Author contributions Z.G. performed the scRNA-seq analyses and wrote the materials and methods section. C.Y. and D.Y. collected and validated the data. M.Y. provided professional assistance. W.C. retrieved literature. D.W. and J.Z. provided samples of BRCA BoM case. X.L. conceptualized and designed the study. X.L. wrote the introduction and discussion section. All authors contributed to the paper and approved the work submitted.

Funding National Natural Science Foundation of China, No.82003805, Xiangyu Li, No.82002776, Jie Zhou, Guangzhou Science and Technology Project, No.2024A03J0649, Jie Zhou, Shanxi Province Science Foundation for Youths, No.201901D211471, Xiangyu Li, Nanjing Military Region Medical Scientific and Technical Innovation Foundation Projects of People's Liberation Army of China, No.XSQ202101, Xiangyu Li

Data availability Publicly available data sets were analyzed in this study. The data can be found here: <https://www.ncbi.nlm.nih.gov/geo/> (accessed on 17 October 2023) with access number GSE225600. The bulk data can be found here: <https://xenabrowser.net/> (accessed on 10 December 2023). The bone metastasis data of breast cancer is available in the National Genomics Data Center (<https://ngdc.cncb.ac.cn/>), the project number is HRA007576. The code data in this research is available in the github website (<https://github.com/goododo/BCBM/>).

Declarations

Competing interests The authors declare no conflict of interest.

Ethical approval This study was conducted in accordance with the Medical Ethics Committee of the Affiliated Cancer Hospital & Institute of Guangzhou Medical University.

Informed consent All participants provided informed consent prior to their participation.

References

1. Anwar SL, Avanti WS, Dwianingsih EK, Cahyono R, Suwardjo S. Risk factors, Patterns, and distribution of bone metastases and skeletal-related events in high-risk breast cancer patients. *Asian Pac J Cancer Prev.* 2022;23(12):4109–17.
2. Pantel K, Hayes DF. Disseminated breast tumour cells: biological and clinical meaning. *Nat Rev Clin Oncol.* 2018;15(3):129–31.
3. Hofbauer LC, Bozec A, Rauner M, Jakob F, Perner S, Pantel K. Novel approaches to target the microenvironment of bone metastasis. *Nat Rev Clin Oncol.* 2021;18(8):488–505.
4. Shen Y, Zou Y, Bie B, Lv Y. Hierarchically released liquid metal nanoparticles for mild photothermal therapy/chemotherapy of breast cancer bone metastases via remodeling tumor stromal microenvironment. *Adv Healthc Mater.* 2023;12(27): e2301080.
5. Monteran L, Erez N. The dark side of fibroblasts: cancer-associated fibroblasts as mediators of immunosuppression in the tumor microenvironment. *Front Immunol.* 2019;10:1835.
6. Neophytou CM, Panagi M, Stylianopoulos T, Papageorgis P. The role of tumor microenvironment in cancer metastasis: molecular mechanisms and therapeutic opportunities. *Cancers (Basel).* 2021;13(9):2053.
7. Wang J, Akter R, Shahriar MF, Uddin MN. Cancer-associated stromal fibroblast-derived transcriptomes predict poor clinical outcomes and immunosuppression in colon cancer. *Pathol Oncol Res.* 2022;28:1610350.
8. Zhang H, Yue X, Chen Z, Liu C, Wu W, Zhang N, et al. Define cancer-associated fibroblasts (CAFs) in the tumor microenvironment: new opportunities in cancer immunotherapy and advances in clinical trials. *Mol Cancer.* 2023;22(1):159.
9. Inoue C, Miki Y, Saito R, Hata S, Abe J, Sato I, et al. PD-L1 induction by cancer-associated fibroblast-derived factors in lung adenocarcinoma cells. *Cancers.* 2019;11(9):1257.
10. Xiang H, Ramil CP, Hai J, Zhang C, Wang H, Watkins AA, et al. Cancer-associated fibroblasts promote immunosuppression by inducing RO S-generating monocytic MDSCs in lung squamous cell carcinoma. *Cancer Immunol Res.* 2020;8(4):436–50.
11. Song M, He J, Pan Q-Z, Yang J, Zhao J, Zhang Y-J, et al. Cancer-associated fibroblast-mediated cellular crosstalk supports hepa tocellular carcinoma progression. *Hepatology.* 2021;73(5):1717–35.
12. Mhaidly R, Mechta-Grigoriou F. Role of cancer-associated fibroblast subpopulations in immune infiltration, as a new means of treatment in cancer. *Immunol Rev.* 2021;302(1):259–72.
13. Gunaydin G. CAFs interacting with TAMs in tumor microenvironment to enhance tumorigenesis and immune evasion. *Front Oncol.* 2021;11:668349.
14. Rømer AMA, Thorseth M-L, Madsen DH. Immune modulatory properties of collagen in cancer. *Front Immunol.* 2021;12:791453.
15. Liang L, Li W, Li X, Jin X, Liao Q, Li Y, et al. 'Reverse Warburg effect' of cancer-associated fibroblasts (review). *Int J Oncol.* 2022;60(6):67.
16. Huang H, Wang Z, Zhang Y, Pradhan RN, Ganguly D, Chandra R, et al. Mesothelial cell-derived antigen-presenting cancer-associated fibroblasts induce expansion of regulatory T cells in pancreatic cancer. *Cancer Cell.* 2022;40(6):656–673.e657.
17. Alen BO, Estévez-Pérez LS, Otero Alén M, Domínguez Hormaetxe S, Simón L, Concha A. Expression of epithelial and mesenchymal markers in plasmatic extracellular vesicles as a diagnostic tool for neoplastic processes. *Int J Mol Sci.* 2023;24(4):3578.

18. Charbonneau H, Tonks NK, Walsh KA, Fischer EH. The leukocyte common antigen (CD45): a putative receptor-linked protein tyrosine phosphatase. *Proc Natl Acad Sci U S A*. 1988;85(19):7182–6.
19. Chistiakov DA, Killingsworth MC, Myasoedova VA, Orekhov AN, Bobryshev YV. CD68/macrosialin: not just a histochemical marker. *Lab Invest*. 2017;97(1):4–13.
20. Li Y, Li C, Liu Q, Wang L, Bao AX, Jung JP, et al. Loss of Acta 2 in cardiac fibroblasts does not prevent the myofibroblast differentiation or affect the cardiac repair after myocardial infarction. *J Mol Cell Cardiol*. 2022;171:117–32.
21. Maetzel D, Denzel S, Mack B, Canis M, Went P, Benk M, et al. Nuclear signalling by tumour-associated antigen EpCAM. *Nat Cell Biol*. 2009;11(2):162–71.
22. Nurmik M, Ullmann P, Rodriguez F, Haan S, Letellier E. In search of definitions: cancer-associated fibroblasts and their markers. *Int J Cancer*. 2020;146(4):895–905.
23. Raskov H, Orhan A, Christensen JP, Gögenur I. Cytotoxic CD8⁺ T cells in cancer and cancer immunotherapy. *Br J Cancer*. 2021;124(2):359–67.
24. Wang K, Wei G, Liu D. CD19: a biomarker for B cell development, lymphoma diagnosis and therapy. *Exp Hematol Oncol*. 2012;1(1):36.
25. Esposito M, Guise T, Kang Y. The biology of bone metastasis. *Cold Spring Harb Perspect Med*. 2018;8(6): a031252.
26. Chen F, Han Y, Kang Y. Bone marrow niches in the regulation of bone metastasis. *Br J Cancer*. 2021;124(12):1912–20.
27. Satcher RL, Zhang XHF. Evolving cancer-niche interactions and therapeutic targets during bone metastasis. *Nat Rev Cancer*. 2022;22(2):85–101.
28. Nolan E, Kang Y, Malanchi I. Mechanisms of organ-specific metastasis of breast cancer. *Cold Spring Harb Perspect Med*. 2023;13(11): a041326.
29. Ma RY, Zhang H, Li XF, Zhang CB, Selli C, Tagliavini G, et al. Monocyte-derived macrophages promote breast cancer bone metastasis outgrowth. *J Exp Med*. 2020;217(11): e20191820.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.