Unravelling the Metastatic Niche in Breast Cancer Patients with Bone Metastasis through Single-cell RNA Sequencing

Xiangyu Li (lixiangyu@czmc.edu.cn)
The Stem Cell and Tissue Engineering Research Center, College of Pharmacy, Changzhi Medical College, Changzhi, Shanxi 046000, P. R. China

Ziyu Gao
The Stem Cell and Tissue Engineering Research Center, College of Pharmacy, Changzhi Medical College, Changzhi, Shanxi 046000, P. R. China

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

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

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

Wenhui Cui
The Stem Cell and Tissue Engineering Research Center, College of Pharmacy, Changzhi Medical College, Changzhi, Shanxi 046000, P. R. China

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

Jie Zhou (sunnysmile@gzhmu.edu.cn)
Department of Breast Oncology Surgery, Affiliated Cancer Hospital & Institute of Guangzhou Medical University, Guangzhou, Guangdong, 510095, China

Research Article

Keywords: breast cancer, bone metastasis, single-cell RNA sequencing, immune microenvironment, myofibroblast-immune interactions
Abstract

Breast cancer (BRCA) is characterized by a unique metastatic pattern and often presents with bone metastasis (BoM), which poses significant clinical challenges. This study employed single-cell RNA sequencing (scRNA-seq) and TCGA data analysis to compare primary tumor (PT), lymph node metastasis (LN), and BoM data. Our investigation identified a metastatic niche in BoMs marked by an increased abundance of cancer-associated fibroblasts (CAFs) and a reduced immune cell count. A distinct subtype (State 1) of BRCA BoM cells associated with adverse prognosis was identified. State 1, characterized by heightened stemness traits, may represent an initiation phase for BoM in BRCA. Complex cell communications involving tumor, stromal, and immune cells were revealed. Interactions between FN1, SPP1, and MDK correlate with elevated immune cells in the BoM. CD46, MDK, and PTN interactions drive myofibroblast activation and proliferation, contributing to tissue remodelling. Additionally, MDK, PTN, and FN1 interactions influence FAP⁺ CAF activation, impacting cell adhesion and migration in BoMs. These insights deepen our understanding of the metastatic niche in breast cancer BoMs.

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 exhibit bone metastases (BoMs) at initial diagnosis, with an elevated 75% risk of developing BoM over the subsequent decade. Advanced BRCA patients exhibit a strikingly high incidence of BRCA BoM, ranging from 65–75%. Notably, bone tissue is the primary site for distant metastasis in BRCA patients, affecting 60–75% of all metastatic BRCA patients, particularly hormone receptor-positive BRCA patients. However, due to pathophysiological impairment and lack of specificity, therapeutic agents cannot easily accumulate in metastatic bone. Consequently, the analysis of pathological features and related biological parameters of bone metastases proves invaluable for predicting patient survival and recurrence risk. 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.

The main treatment strategy for BoM is to inhibit the growth of tumor cells while ignoring the influence of the tumor stromal microenvironment on the progression of BoM. The intricate landscape of the tumor microenvironment (TME) is composed of diverse noncellular factors and a myriad of cell types, including cancer-associated fibroblasts (CAFs), immune cells, endothelial cells, pericytes, and adipocytes. The multifaceted crosstalk among tumors, stromal cells, and immune cells not only underlies treatment resistance but also propels tumor progression and progression to overt BoM. Hence, a nuanced comprehension of these extensive interactions is of paramount importance in advancing the efficacy of tumor treatments. Previous investigations have revealed that CAFs, which are predominantly activated fibroblasts influenced by tumors, play instrumental roles in promoting BRCA progression. Its involvement spans a spectrum of functions, including fostering tumor cell proliferation, facilitating cancer cell invasion and metastasis, orchestrating extracellular matrix remodelling and deposition, promoting
angiogenesis, instigating drug resistance, generating circulating CAFs (cCAFs), and secreting protumor factors. Notably, CAFs contribute to the establishment of an immunosuppressive microenvironment, thus allowing the evasion of 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 antitumour immune responses. Playing a pivotal role, CAFs are central to the recruitment of tumor-associated macrophages (TAMs), fostering a protumor phenotype. Additionally, 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 cytokine and protease production. 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. In addition 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, CAFs inhibit antigen presentation, thus limiting T-cell activation. Moreover, CAFs hinder the infiltration of cytotoxic T lymphocytes into tumors, attenuating their tumoricidal potential. The intricate orchestration of these immunosuppressive mechanisms by CAFs, encompassing the upregulation of immune checkpoint molecules; extracellular matrix remodelling via collagen, fibronectin, and MMPs; and activation of the FAK signalling pathway, underscores the central role of these cells in mediating tumor immune escape through metabolic reprogramming and the production of immunosuppressive metabolites$^{9-16}$. However, the specific 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 (PTs), lymph node metastases (LNs), and bone metastases (BoMs), our investigation revealed a distinctive metastatic niche characterized by an increase in CAFs and a reduction in immune cell populations in the BoM. Notably, we identified a unique subtype of BRCA BoM cells strongly associated with an adverse prognosis. Our analysis included the exploration of genes, signalling pathways, and variations in the immune microenvironment across PTs, LNs, and BoMs. By revealing the intricate cellular dialogues among tumor, stromal, and immune cells, we pinpointed pivotal interactions involving FN1, SPP1, and MDK that correlate with an increased presence of immune cells in the BoM. These findings provide insights into the complexities of the immune microenvironment in BRCA BoMs and offer perspectives for therapeutic interventions targeting this specific metastatic manifestation of BRCA.

Materials and methods
Data acquisition

This study received approval from the Medical Ethics Committee of the Affiliated Cancer Hospital & Institute of Guangzhou Medical University, and all participating subjects provided informed consent before undergoing surgery. For the scRNA-seq analysis, two distinct datasets were utilized.

Immunohistochemistry (IHC) was performed on formalin-fixed, paraffin-embedded tissue blocks retrieved from one patient with BRCA BoM in the eleventh thoracic vertebra after a thorough review of the archived materials. The BoM dataset included the expression profiles of 32,738 genes across 9,181 individual cells.

The GSE225600 dataset was obtained from the Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo) on October 17, 2023. The dataset included gene expression data from a total of 81,683 cells across four PTs and their corresponding four paired LNs, 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 dataset included count data for 60,488 genes across 1,217 samples. Additionally, survival data for 1,260 BRCA samples and clinical data for 1,248 BRCA samples were obtained from the same source.

Single-cell RNA-seq Data Preprocessing

The high-quality reads obtained from sequencing experiments were subjected to meticulous processing using "Cell Ranger" (version 3.0.2). This approach encompasses essential tasks such as sequence alignment, filtering, barcoding, and unique molecular index counting. The reference genome employed for this analysis was hg19. Subsequently, a thorough examination of the scRNA-seq data was conducted utilizing the "Seurat" package (version 5.0.1; https://satijalab.org/seurat/) in R software (version 4.3.1).

The comprehensive analysis included multiple stages, including data quality control, normalization, and differential gene expression analysis. Initially, each scRNA-seq dataset underwent a stringent filtering process, excluding cells with fewer than 200 genes and those with more than 10% of the total expressed genes being mitochondrial genes. Additionally, genes detected in fewer than 10 cells were excluded. The normalization process employed the "NormalizeData" function with default parameters. Subsequently, dimensionality reduction was performed using principal component analysis (PCA), generating a 13-dimensional output for the two datasets.

Clustering analysis was performed using the “FindClusters” function with a resolution of 10 for the BoM data and 7 for the GEO data. The identification of doublets was performed using the “DoubletFinder” R package (version 2.0.3; https://github.com/chris-mcginnis-ucsf/DoubletFinder). Finally, the "IntegrateData" function was employed to correct batch effects by integrating the data from the two datasets for subsequent analyses. Two samples from the GEO database with insufficient cells were
excluded, resulting in an integrated dataset comprising 40,333 genes in 34,375 cells derived from seven samples (3 PT, 3 LN, and 1 BoM).

Cell clustering and annotation

The values of the integrated dataset underwent z score conversion using the "ScaleData" command, and highly variable genes were meticulously selected through the "FindVariableGenes" function with default parameters. Principle components were calculated based on these selected genes and subsequently projected onto all other genes using the "RunPCA" function. Subsequently, the "FindNeighbors" and "FindClusters" commands were used to detect clusters of similar cells, constructing a shared nearest neighbor map with an empirically set resolution.

Upon clustering all cells within the integrated dataset, the principal components delineating heterogeneity were found to predominantly represent differences in tissue compartments. Consequently, the clusters were grouped based on the expression of distinct cell type markers, leading to the classification of cells into four clusters (epithelial cells, endothelial cells, immune cells, and fibroblasts). A similar analytical pipeline was applied to immune cells, where the identification of immune cell types was achieved by matching each cluster-specific gene set with known signature genes of cell populations reported in previous literature\textsuperscript{17–24}.

Clusters that did not significantly express marker genes were categorized based on their most differentially expressed genes. This comprehensive approach ensured a nuanced understanding of cellular heterogeneity within the integrated dataset, providing valuable insights into tissue-specific compartments and diverse cell types present in the studied samples.

Quantification of epithelial cell copy number variation

To assess copy number variation (CNV) in individual epithelial cells, we utilized the "infercnv" R package (version 1.16.0; https://github.com/broadinstitute/infercnv). For this analysis, fibroblasts were designated as the reference normal cells, providing a baseline copy number for comparison. Employing a cut-off of 0.1 and setting denoising to TRUE, we systematically computed CNV scores, enabling 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 employing the "FindMarkers" function within the "Seurat" R package and utilizing the Wilcoxon rank sum test with a log2-fold change threshold of 0.1. To refine the results, stringent filtering criteria were applied, necessitating an absolute average log2-fold change > 1 and a p value < 0.05.

For the analysis of Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment, we utilized the "clusterProfiler" R package (version 4.8.3). Significantly enriched GO terms and KEGG pathways were determined based on a q value less than 0.05.
To investigate cancer hallmark enrichment, we obtained a reference set of 50 cancer hallmark gene sets from the Molecular Signature Database (MSigDB, accessed on 2023/11/01). Enrichment scores were calculated using the "UCell" (version 2.7.1; https://github.com/carmonalab/UCell) and "irGSEA" (version 2.1.5; https://github.com/chuiqin/irGSEA) R packages, providing a comprehensive exploration of the functional significance of differentially expressed genes in the context of cancer hallmarks.

**Pseudotime analysis and stemness inference**

Pseudotime analysis was conducted with the R package "monocle" (version: 2.28.0), after which variable genes were harnessed to delineate pseudotime stages, pinpoint trajectory differentiation-related genes, and scrutinize alterations in branch point genes. This analysis enabled the exploration of functional disparities among pseudotime stages, defining distinct BoM states in this study. Furthermore, stemness scores for each state were estimated by leveraging established markers such as OCT4, SOX2, and NANOG, providing valuable insights into the stem cell-like characteristics associated with each pseudotime stage.

**Identification of Markers for Early-stage BoM**

To characterize the state identities, we identified DEGs specific to BoM state 1 using the "FindAllMarkers" function of 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 less than 0.05. To further refine the marker selection, we employed the R package "rpart" (version: 4.1.23) to construct a tree model utilizing the method parameters and pruning the tree with "cp = 0.01". This meticulous process resulted in the identification of markers specifically associated with the early stage of BoM.

**Prognostic value of BoM early-stage markers**

To evaluate the prognostic significance of the identified markers, we computed enrichment scores for these marker genes in individual TCGA BRCA patients. These scores were derived by multiplying gene expression values with weights assigned by the constructed tree model. Patients were subsequently stratified into two groups based on median values. We then assessed the prognostic effectiveness of these markers in predicting the 10-year survival of TCGA BRCA patients using the R packages "survival" (version 3.5-7) and "survminer" (version 0.4.9). Additionally, we explored the dynamic changes in the expression of these markers across patients with different clinical features in the TCGA cohort.

**Cell–cell interaction analysis**

To investigate intercellular interactions within the BRCA BoM, we utilized the R package “CellChat” (version 1.6.1; https://github.com/sqjin/CellChat/) specifically tailored for carcinoma cells, immune cells, and fibroblasts. This package encompasses a database of molecular signalling interactions comprising 60% of paracrine/autocrine signalling interactions, 21% of extracellular matrix receptor interactions, and 19% of cell–cell contact interactions. Through comprehensive cellular communication analysis, we obtained valuable insights into cell–cell interactions and elucidated intercellular communication.
networks. This approach contributes to a holistic understanding of the intricate interactions among various cell types during the development of BRCA BoM.

Results

Single-cell RNA-seq Profiling of BRCA Primary Tumors, Lymph Nodes, and Bone Metastasis Data

To enhance our understanding of BRCA BoM, we integrated two single-cell sequencing datasets and performed a comprehensive analysis (Fig. 1A). Following rigorous quality control, we obtained transcriptome data from 34,375 cells, including 12,427 cells from four PT samples, 15,740 cells from four LN samples, and 6,208 cells from one BoM sample. After dimension reduction, clustering, and cell annotation (Fig. 1B-D), we observed a predominant ratio of immune cells in the LN samples. Notably, compared to those in PT and LN tissues, the ratio of immune cells in these tissues was lower, while the ratio of epithelial cells and fibroblasts was noticeably greater. This observation led us to hypothesize that immune cells and fibroblasts may play pivotal roles in the BoM process.

Next, we employed the "infercnv" R package to infer copy number variations in epithelial cells. Comparative analysis of fibroblasts revealed extensive mutations in epithelial cells, leading us to conclude that all the epithelial cells exhibited characteristics consistent with those of carcinoma (Fig. 1E).

Upregulated Metabolic and Immune Signalling Pathways in BRCA BoMs

Using the "FindMarkers" function within the "Seurat" R package, we identified DEGs in tumor cells originating from distinct BRCA types. In parallel, we conducted GSEA and examined cancer hallmark enrichment to gain deeper insights into these DEGs. Our analysis revealed significant enrichment of upregulated genes in the Toll-like receptor signalling pathway, the PI3K/Akt/mTOR pathway, and specific metabolic programs in BRCA cells derived from the BoM cohort (Fig. 2A, B). These findings underscore the crucial roles of the immune response and cell proliferation in the progression of BRCA BoM.

Subsequently, by performing pseudotime trajectory analysis on PT and BoM cells, excluding LN samples with insufficient tumor cells, we elucidated dynamic cell transitions. Our analysis identified three distinct states involved in BRCA BoM progression (Fig. 2C, D). State 1, prevalent at the trajectory’s outset, exhibited the highest stemness score (Fig. 2E), while state 2, characterized by the lowest stemness score, was predominantly located at the trajectory’s conclusion. Functional enrichment analysis of upregulated KEGG pathways in these states revealed that state 1 primarily involved biological processes encompassing cell growth, development, proliferation, differentiation, and cell adhesion (Fig. 2F). Given the manifestation of state 1 in both PT and BoM samples early in the trajectory and the display of heightened stemness traits, we propose that state 1 may represent an initiation phase for BoM in BRCA.

Identification of Marker Genes for Early-stage BRCA BoM
The "rpart" R package provides a robust framework for constructing classification and regression trees. To pinpoint pivotal genes involved in the progression of BRCA BoM, we employed the "rpart" package to generate a recursive partitioning and regression tree. The resulting model highlighted three marker genes (ZNF831, CTLA4, and GIMAP7) and their respective positions within the decision tree model (see Fig. 3A). The TCGA BRCA cohort was scored based on the expression of each gene in the RNA-seq data and the corresponding weight derived from the tree model. Subsequently, BRCA samples were categorized into State 1 and non-State 1 status based on the median score (Fig. 3B). The Kaplan–Meier plot indicated that overall survival was significantly lower for BRCA patients in State 1 than for those in non-State 1 (Fig. 3C).

The distribution of BRCA status across various clinical features revealed a fluctuating trend in State 1 percentages, indicating the dynamic progression of BRCA. This finding suggested that primary tumor cells may acquire the potential for distant metastasis as BRCA advances, subsequently leading to 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 reclustering techniques and marker gene annotations (Fig. 4A, 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 PTs, LNs, and BoMs revealed a notable increase in myeloid cells and a relative increase in cytotoxic NK-T cells in BoMs (Fig. 4C). To gain comprehensive insights into the functions of myeloid cells in BoMs, we further reclustered them into six distinct subsets termed TAMs 1–6 (Fig. 4D). The bar plot illustrates that, except for TAM3, all the clusters were predominantly present in BoMs (Fig. 4E) and actively engaged in biological processes such as cell adhesion, immune response, and immune regulation (Fig. 4F).

**Deciphering Immune Cell Interactions in the BoM TME**

To assess the impact of immune cells on the TME in BoMs, we used "CellChat" to analyse cell communication networks. Notably, a substantial number of interactions were observed in BoM (Fig. 5A), with predominant pathway presence (Fig. 5B). Our investigation revealed that cytotoxic NK-T cells in BoM release CD8A, which engages in communication with B cells, naive T cells, myeloid cells, regulatory T cells, and carcinoma cells through interactions with HLA-A, HLA-B, and HLA-C, potentially enhancing immune responses in BoM (Fig. 5C).

BoM state 1 carcinoma cells predominantly include cytotoxic NK-T cells, myeloid cells, and B cells (Fig. 5D). This interaction is facilitated through the involvement of FN1, HLA genes, and MDK, which interact with their respective target genes. Notably, the MDK-NCL interaction (Fig. 5E) appeared to exert a more significant impact, potentially suppressing the immune response. Concurrently, myeloid cells and B cells predominantly target BoM state 1 tumor cells via interactions with FN1, SPP1, GRN, and MK and
their target genes (Fig. 5F). The activation of signalling pathways by FN1 may contribute to cell survival, metastasis, and the progression of BoM (Fig. 5G).

**Decoding Fibroblast Dynamics and Interactions in BRCA BoMs**

To unravel the intricate dynamics of fibroblasts in BRCA BoMs, we employed re-clustering techniques and marker gene annotations, revealing three distinct fibroblast clusters (Fig. 6A, B). Analysis of fibroblast proportions in PTs, LNs, and BoMs highlighted the significant increase in the number of myofibroblasts and FAP$^+$ inflammatory cells in the BoM cohort (Fig. 6C). By exploring the functional roles of fibroblasts in BoM, KEGG functional enrichment revealed the active involvement of myofibroblasts and FAP$^+$ inflammatory cells in processes such as cell proliferation, adhesion, and complement and coagulation cascades (Fig. 6D).

To assess the influence of fibroblasts on BoMs, we utilized "CellChat" to analyse cell communication networks, revealing a significant number of interactions in BoMs (Fig. 6E). Our study revealed that FAP$^+$ inflammatory cells in the BoM release FN1, which participates in communication with carcinoma cells, myofibroblasts, FAP$^-$ inflammatory cells, and self-interaction (Fig. 6F).

In the BoM, state 1 carcinoma cells exhibited predominant interactions with myofibroblasts and FAP$^+$ inflammatory cells (Fig. 6G), facilitated by MDK and CD46, which engaged their respective target genes (Fig. 6H). The MDK-SDC1/SDC4 interaction exerts a significant influence on tumor progression, potentially enhancing cell proliferation, angiogenesis, and epithelial–mesenchymal transition, thereby promoting tumor metastasis from the primary site to distant locations. Simultaneously, myofibroblasts and FAP$^-$ inflammatory cells predominantly target BoM state 1 tumor cells through interactions involving the THBS, PTN, and NOTCH pathways (Fig. 6I). PTN activation of signalling pathways may contribute to stimulating new blood vessel formation and tumor angiogenesis (Fig. 6J).

**Myofibroblast-Immune Interactions in BRCA BoM Progression**

By conducting an in-depth analysis of cell communications between fibroblasts and immune cells, we revealed a substantial number of inferred interactions in the BoM (Fig. 7A). Quantitative analysis demonstrated that myofibroblasts exhibit a greater frequency of interactions with immune cells than with other cell types (Fig. 7B). The predominant interactions involved cytotoxic NK-T cells and myeloid cells with myofibroblasts (Fig. 7C) and were mediated through the PTN, MDK, and LAMININ signalling pathways (Fig. 7D). Concurrently, myofibroblasts and FAP$^+$ inflammatory cells primarily target myofibroblasts through interactions involving the PTN, MDK, and SPP1 signalling pathways (Fig. 7E). The PTN-NCL interaction significantly influences interactions from and toward myofibroblasts, potentially promoting cell proliferation, angiogenesis, metastasis, and heightened resistance to apoptosis in cancer cells.
Discussion

The emergence of BoM is of significant prognostic importance in BRCA, underscoring the necessity of elucidating 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 given to the metastatic niche of the BoM. Specifically, we observed a distinct increase in CAFs and a decrease 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 the BoM, shedding light on the niche that supports tumor colonization. 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 of early-stage colonization of BRCA BoM aligns with the observed increase in heterotypic adherens junctions and increase in calcium influx. A pivotal outcome of our research was 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 unfavourable prognosis. Through a comparative analysis of cancer hallmarks between BoMs and PTs, as well as between BoMs and PTs, we revealed a predominant upregulation of metabolic and Toll-like receptor signalling pathways in BoMs. This highlights significant molecular distinctions in the metastatic microenvironment.

To gain further insight, we further stratified BoMs into three distinct states using stemness scores. Intriguingly, State 1, characterized by the highest stemness, was found to coexist at both the primary and metastatic sites, acting as the initiating point for BoM. KEGG functional enrichment analysis of State 1 revealed 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, 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 derived from CCL2-recruited inflammatory monocytes that promote BRCA BoM in an IL-4R-dependent manner. 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 significant upregulation of myeloid cells in BoMs compared with PTs and LNs. Further elucidating the myeloid cell landscape through dimensionality reduction and clustering revealed that specific clusters of TAMs were notably elevated in BoMs. These clusters were found to
predominantly engage in processes associated with cell adhesion and the immune response, as substantiated by KEGG functional enrichment analysis.

Remarkably, our examination of cell–cell interactions revealed a substantial augmentation in interactions specific to the formation of the metastatic niche within the BoM. Notably, major histocompatibility complex class I (MHC-I) was identified as a central mediator facilitating communication between tumor cells and immune cells, as well as orchestrating intercellular interactions among 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 yielded additional insights. These interactions were found to significantly contribute to an increase 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 signalling pathways in shaping the immune landscape of BRCA BoMs.

Our study provides a nuanced understanding of the roles played by cancer-associated myofibroblasts and inflammatory CAFs within the metastatic niche. Particularly in the BoM cohort, a significant increase in myofibroblasts and FAP− inflammatory CAFs was observed in comparison to those in the PT and LN cohorts, while the number of FAP+ inflammatory CAFs was reduced. These identified myofibroblasts and FAP+ inflammatory CAFs emerged as pivotal contributors that are primarily involved in crucial cellular functions such as proliferation, adhesion, and extracellular matrix organization. The intricate interplay orchestrated by CD46, MDK, PTN, and their target genes emerged as a driving force behind the activation and proliferation of myofibroblasts, significantly contributing to tissue remodelling within the BoM. Furthermore, the interactions facilitated by MDK, PTN, FN1, and their respective target genes were found to stimulate the activation and proliferation of FAP+ CAFs, concurrently promoting cell adhesion and migration within the BoM microenvironment. Our in-depth exploration of immune-stromal cell communication revealed critical genes, including PTN, MK, SPP1, and FN1. Through interactions with their target genes, these genes were implicated in fostering the activation and proliferation of myofibroblasts while concurrently playing a pivotal role in orchestrating inflammatory responses within the dynamic context of BoM.

In conclusion, in our study, a comprehensive single-cell map was constructed, providing a detailed portrayal of the metastatic niche throughout the spectrum of BRCA progression, encompassing in situ conditions, LNs, and BoMs. The systematic delineation of the metastatic niche in BoMs has revealed distinctive features, revealing the intricate mechanisms that govern the immunosuppression induced by cancer cells upon metastasizing to bone.

**Declarations**

**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 the National Natural Science Foundation of China (No. 82003805, No. 82002776), the Guangzhou Science and Technology Project (No. 202124A03J0649), the Shanxi Province Science Foundation for Youths (No. 201901D211471), and the Young Academic and Technical Leaders Project of Changzhi Medical College (No. XSQ202101). We are grateful to all the 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 the literature. D.W. and J.Z. provided samples from the BRCA BoM patient. X.L. conceptualized and designed the study. X.L. wrote the introduction and discussion sections. All the authors contributed to the paper and approved the work submitted.

Competing interests

The authors declare no conflicts of interest.

Data availability

Publicly available datasets were analysed in this study. The data can be found at https://www.ncbi.nlm.nih.gov/geo/(accessed on 17 October 2023) with access number GSE225600. The bulk data can be found at https://xenabrowser.net/(accessed on 10 December 2023).

Ethics declarations

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

Supplementary Materials

BoMrawData: The eleventh thoracic vertebra of the BRCA BoM patient in the BoM dataset contains expression profiles of 32,738 genes across 9,181 individual cells.

References


Figures
Figure 1

Single-cell RNA-seq Atlas of BRCA Primary Tumor and Metastasis Data

(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 10x Genomics platform.
(B) A t-distributed stochastic neighbor embedding (t-SNE) plot, derived from integrated BRCA data (n=34,375 cells), visually delineates the principal cell types.

(C) Proportional representation of each cell type across different tumor types in BRCA patients is shown in the bar chart.

(D) Dot plot presenting marker gene expression levels in the indicated cell types. The dot size indicates the proportion of cells expressing the marker within the group, while the color represents the marker expression level.

(E) Heatmap depicting results from "infercnv" providing insights into copy number variations across samples.
Figure 2

Functional enrichment and pseudotime analysis.

(A) Dot plot illustrating the 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 PTs and BoMs, with the blue spectrum indicating the temporal order.

(D) Pseudotime states are color-labelled, with each dot representing a single cell.

(E) Bar plot displaying inferred stemness scores of pseudotime states.

(F) Dot plot showing Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment results for each state, where the dot size reflects the gene set size and the color indicates the p value of the result.
Figure 3

BoM State 1 Marker 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 TCGA BRCA patient samples.
(C) Kaplan–Meier plot revealing significant differences between the two BRCA subtypes.

(D) Bar plot illustrating the distribution of the two BRCA states across distinct clinical features.

Figure 4

Immune cell annotation and myeloid cell analysis.
(A) Dot plot displaying marker gene expression in specified cell types, with the dot size indicating the proportion of expressing cells and the color representing marker expression levels.

(B) Utilizing t-SNE, a plot generated from integrated immune data was generated to visually delineate the 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 depicting the primary cell types of TAMs 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 the dot size indicating the gene set size and the color denoting the p value.
Figure 5

Interactions between BRCA Tumor Cells and Immune Cells.

(A) Quantification of carcinoma cell and immune cell interactions in PTs, LNs, and BoMs depicted in a bar plot.

(B) Bar plot presenting the cell–cell interaction count for each signalling pathway in PTs, LNs, and BoMs.

(C) Dot plot visualizing cell–cell communication probabilities within the MHC-I and CD99 pathways.
(D) Chord plot showing upregulated signalling pathways originating from BoM state 1 carcinoma cells and connecting to various immune cell types.

(E) Dot plot revealing the probabilities of communicating with carcinoma cells via BoM-specific pathways (MK, MHC-II, and FN1).

(F) Chord plot presenting upregulated signalling pathways targeting BoM state 1 carcinoma cells from diverse immune cell types.

(G) Dot plot demonstrating the probabilities of communicating between BoM-specific pathways (SPP1 and FN1) and carcinoma cells.
Figure 6

Fibroblast Insights in BRCA BoMs.

(A) t-SNE-based dimensionality reduction clustering plot of fibroblasts.
(B) Dot plot showing marker gene expression in specific cell types, with the dot size indicating the proportion and color denoting the expression level.

(C) Bar chart illustrating the proportional representation of fibroblast types across diverse BRCA tumor types.

(D) Dot plot revealing Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment in fibroblast clusters, with the dot size indicating the gene set size and the color representing the p value.

(E) Bar plot quantifying interactions between carcinoma cells and fibroblasts in PTs, LNs, and BoMs.

(F) Dot plot visualizing cell–cell communication probabilities within the FN1, CD99, and LAMININ pathways.

(G) Chord plot highlighting upregulated signalling pathways originating from BoM state 1 carcinoma cells and connecting to various fibroblast types.

(H) Dot plot displaying the probabilities of communicating with carcinoma cells via BoM-specific pathways (MK, CD46, ncWNT, and TGFb).

(I) Chord plot presenting upregulated signalling pathways targeting BoM state 1 carcinoma cells from diverse fibroblast types.

(J) Dot plot demonstrating the probabilities of communicating with carcinoma cells via BoM-specific pathways (THBS, PTN, and NOTCH).
Figure 7

Intricate interactions between fibroblasts and immune cells.

(A) Bar plot quantifying interactions between fibroblasts and immune cells across PTs, LNs, and BoMs.

(B) Visualization of upregulated signalling pathways in BoMs using a chord plot.

(C) Chord plot highlighting upregulated pathways originating from BoM myofibroblasts connected with other cell types.

(D) Communication probabilities of BoM-specific pathways (PTN, MK, and LAMININ) from myofibroblasts are shown in a dot plot.

(E) Chord plot showing upregulated signalling pathways targeting BoM myofibroblasts from other cell types.
(F) Dot plot illustrating the probability of communicating BoM-specific pathways (SPP1, PTN, and MK) targeting myofibroblasts.

**Supplementary Files**

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

- BoMrawData.zip