Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Single-cell RNA sequencing reveals route-specific remodeling of the tumor microenvironment and identifies predictive factors of breast cancer metastasis

Congmei Hu
Harbin Medical University
Kaiyue Yang
Harbin Medical University
Rui Fu
Harbin Medical University
Yining Ma
Harbin Medical University
Yongjin Long
Harbin Medical University
Maohao Li
Harbin Medical University
Mingyao Sun
Harbin Medical University
Zhenzhen Wang
Hainan Medical University
Xia Li
Harbin Medical University
Yunpeng Zhang
zhangyp@hrbmu.edu.cn
Harbin Medical University
Xingda Zhang
Harbin Medical University

Article

Keywords: Single cell RNA sequencing, breast cancer, metastasis, cell reprogramming, cell interactions

Posted Date: May 28th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9551458/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Breast cancer metastasis is the primary cause of patient mortality, driven by dynamic remodeling of the tumor microenvironment and the synergistic regulation of key factors. This study integrates single-cell transcriptomic data from 71,805 breast cancer cells, systematically mapping the dynamic evolution of tumor ecosystems across non-metastatic, lymph node-metastatic, and bone-metastatic stages. Analysis reveals that while cell lineages remain stable in the transcriptomic space, the proportions of malignant epithelial cells, endothelial cells, and macrophages significantly increase along the metastatic trajectory. This is accompanied by enhanced oxidative phosphorylation, hypoxic adaptation, and activation of epithelial-mesenchymal transition. Concurrently, CD4 + T cells undergo functional reprogramming toward regulatory and exhausted states, with marked upregulation of immunosuppressive molecules in bone metastasis samples. Macrophages exhibit continuous polarization from proinflammatory to anti-inflammatory phenotypes. The stroma exhibits marked heterogeneity during metastasis; notably, bone metastatic tumors are enriched with vascular-associated fibroblasts and mature endothelial cells that synergistically enhance angiogenesis and invasive potential. Intercellular communication analysis further reveals that malignant cells actively remodel the microenvironment via signaling pathways involving HBEGF, TGF-β, and PDGFB/C, thereby coordinating immune suppression, metabolic reprogramming, and metastatic niche stability. In summary, this study constructed a high-resolution single-cell ecosystem map of breast cancer metastasis, systematically revealing the multi-level mechanisms driven by cellular state remodeling and intercellular collaboration during metastasis. It also identified potential predictive factors for metastasis, providing crucial theoretical support for microenvironment-targeted interventions in metastatic breast cancer.

Introduction

Breast cancer is the most common malignant tumor among women worldwide and the leading cause of cancer-related deaths[1]. Despite advances in early detection and treatment, approximately 20% to 30% of patients develop metastatic disease following diagnosis and therapy[2]. Lymph nodes—particularly axillary lymph nodes—are common sites for early metastasis[3,4], from which cancer cells may further spread to distant organs including bones, lungs, liver, and brain. Among these, bone metastases are the most common type and are largely incurable[5–7]. Therefore, the 5-year overall survival rate for patients with breast cancer bone metastases is only 22.8%[8], significantly lower than the 5-year overall survival rate exceeding 80% observed in patients without metastatic disease in regions such as Australia and the United States[9]. Single-cell RNA sequencing (scRNA-seq) enables high-precision analysis of the cellular composition and functional states within the tumor microenvironment at single-cell resolution, providing a powerful tool for systematically elucidating the key cell populations and their molecular regulatory mechanisms involved in breast cancer metastasis[10].

The alteration of cellular states is considered a core adaptive mechanism in tumor progression and metastasis. Traditional views suggest that malignant evolution primarily relies on lineage replacement

driven by genetic mutations. However, increasing evidence indicates that tumor cells and their microenvironmental counterparts can rapidly respond to diverse ecological stresses through reversible functional state remodeling while maintaining relatively stable lineage identities[11–13]. This phenotypic conversion, characterized by transcriptional, metabolic, and signaling pathway reprogramming, endows existing cell types with novel biological functions—such as enhanced invasiveness, immune evasion capabilities, or therapeutic tolerance—thus playing a pivotal role in tumor heterogeneity and metastatic adaptation.

Within the tumor microenvironment, immune cells, fibroblasts, and vascular endothelial cells also exhibit significant phenotypic plasticity. Their state transitions are co-regulated by hypoxia, inflammatory signals, and tumor-derived secreted factors. For example, tumor-associated macrophages (TAMs) promote angiogenesis, induce immune suppression, and drive tumor progression by secreting signaling molecules such as VEGF[14,15]; under hypoxic conditions, TAMs-mediated T cell suppression is further enhanced, with its molecular mechanism dependent on HIF-1α activation[16]. Beyond immune regulation, tumor cells can also promote immune evasion by upregulating PD-L1 to suppress T cell-mediated cytotoxic responses[17]. Concurrently, they activate endothelial cells (ECs) by secreting VEGF-A, inducing the formation of abnormal vascular networks that provide structural and metabolic support for tumor growth and metastasis[18]. During tumor angiogenesis, expression of anti-tumor factors (e.g., TSP-1) gradually diminishes, while pro-tumor factors such as TGF-β1 and POSTN significantly accumulate around neovasculature. Together, these factors shape a pro-metastatic microenvironment niche, accelerating malignant tumor progression[19,20].

This study systematically dissected the differences in tumor ecosystems across distinct metastatic pathways in breast cancer, revealing significant molecular and functional heterogeneity among metastatic malignant cells. Compared to lymph node metastasis, malignant cells in bone metastasis exhibited more pronounced hypoxic adaptation and invasive phenotypes, accompanied by transcriptional reprogramming of immune and stromal cells toward immunosuppressive and pro-invasive states. Further analysis of cell-cell communication revealed that the bone metastasis-associated tumor microenvironment possesses a more complex and highly remodeled intercellular crosstalk network. Malignant cells systematically drive immune suppression, metabolic reprogramming, and stabilization of the metastatic niche through synergistic interactions with multiple microenvironment cell subpopulations, thereby promoting the sustained progression of bone metastasis.

Results

Result 1 Single-cell analysis reveals the tumor ecosystem during breast cancer metastasis

To systematically characterize the compositional features and dynamic changes of the tumor ecosystem during breast cancer metastasis, this study integrated single-cell transcriptomic sequencing

data from nine breast cancer samples, including two cases of non-metastatic primary breast cancer, four cases of lymph node metastatic breast cancer, and three cases of bone metastatic breast cancer (Supplementary Table S1). After removing low-quality cells through rigorous quality control procedures, a final cohort of 71,805 high-quality cells was obtained for subsequent analysis. Using principal component analysis (PCA) for dimensionality reduction combined with unsupervised clustering via uniform manifold approximation and projection (UMAP), we systematically classified all cells and annotated them based on expression profiles of classical marker genes. Ten major cell lineages were identified, including malignant epithelial cells, normal epithelial cells, T/NK cells, B cells, fibroblasts, endothelial cells, plasma cells, mast cells, parietal cells, and macrophages (Fig. 1A, 1B, and 1C; Supplementary Fig. 1A and 1B). This reveals that the breast cancer tumor microenvironment, across different stages, is composed of diverse immune, stromal, and vascular-associated cells, exhibiting a highly complex and multi-layered ecosystem.

To assess the potential impact of different metastasis directions on cell population composition, after mapping cell origin information onto the UMAP space, we observed that cells from different metastasis stages did not show distinct separation based on metastasis direction. Instead, they clustered primarily according to their cellular lineage (Fig. 1D). This finding indicates that regardless of whether metastasis occurs, various cell types maintain relatively stable and conserved lineage-specific molecular characteristics at the transcriptomic level. This phenomenon suggests that metastatic progression in breast cancer is not achieved through fundamental reshaping of cellular identity, but rather relies on systematic reconfiguration of the proportions and functional states of existing cell populations, thereby driving the overall evolution of the tumor ecosystem. Based on this, we further compared the composition ratios of different cell types across samples with varying metastasis directions (Fig. 1E and 1F; Supplementary Fig. 1C). Results indicate that compared to non-metastatic breast cancer, lymph node and bone metastasis samples exhibit significantly elevated proportions of malignant tumor cells, endothelial cells, and macrophages. Furthermore, overall cellular abundance demonstrates a progressive increase along the disease progression trajectory. This pattern of continuous change indicates that during metastasis, tumor cells themselves, along with their closely interacting vascular systems and immune cells, undergo synergistic amplification. This suggests that these cellular populations play a crucial role in supporting tumor spread, growth, and the maintenance of metastasis.

Further functional analysis of the molecular characteristics of different cell populations from various transfer sources (Fig. 1G, Supplementary Table S2) revealed that the biological pathways enriched in each major cell lineage highly align with their known biological functions. Immune cells are significantly enriched in pathways related to immune response and inflammation regulation, highlighting their central role in tumor immune surveillance and antitumor responses[21]. Endothelial cells, on the other hand, predominantly accumulate in functions associated with angiogenesis and cell migration, suggesting they play a crucial role in tumor progression by promoting neovascularization and providing sustained oxygen and nutrient support for tumor growth and metastasis[22]. Notably, the cell population originating from lymph node metastasis exhibits significantly enhanced activity in energy metabolism-related

pathways such as oxidative phosphorylation and aerobic respiration. This phenomenon suggests that the tumor microenvironment associated with metastasis may undergo significant metabolic reprogramming, shifting from an energy acquisition pattern relatively dependent on glycolysis to a state more reliant on mitochondrial oxidative metabolism[23]. This metabolic remodeling may reflect the heightened demands for stable energy supply and metabolic efficiency placed on tumor cells and their microenvironment during metastasis. It also suggests that enhanced oxidative metabolism could serve as a crucial supportive mechanism enabling tumor cells to successfully adapt to the microenvironments of secondary organs and sustain continuous growth. Combined with systemic alterations in cellular composition, these findings collectively demonstrate that breast cancer metastasis is not an isolated tumor cell event, but rather an ecosystem-level process involving the coordinated remodeling of malignant cells, the vascular system, and the immune microenvironment.

Overall, this study employs a single-cell resolution systematic analysis to reveal that breast cancer exhibits a tumor ecosystem characterized by relatively stable cellular lineage architecture during metastasis, yet with significantly remodeled cellular composition and functional states. The synergistic expansion of malignant cells, endothelial cells, and macrophages, coupled with enhanced oxidative metabolism, collectively form the critical biological foundation supporting metastatic progression. This provides a novel theoretical framework for understanding and intervening in metastatic breast cancer through tumor microenvironment regulation and metabolic vulnerability.

Result 2 Enhanced hypoxic adaptation and invasive programs in bone metastatic malignant cells

To accurately distinguish malignant from non-malignant epithelial cells of different origins, we first employed the inferCNV method for systematic inference of copy number variations (CNVs) in epithelial cells (Supplementary Fig. 2A, 2B, 2C, and 2D). This approach reliably identified malignant cell populations exhibiting significant genomic abnormalities at single-cell resolution. Building upon this, we extracted all malignant epithelial cells and classified them into three tumor cell clusters based on their source samples: non-metastatic, lymph node metastatic, and bone metastatic (Fig. 2A). This enabled a systematic comparison of molecular characteristics and functional states among malignant cells from different metastatic pathways. Quantitative analysis revealed significant differences in overall composition proportions and Ro/e indices among malignant cells from different origins (Fig. 2B), indicating a distinct origin-dependent distribution pattern across metastatic directions. This suggests malignant cells have evolved distinct cellular states with divergent biological properties. Concurrently, the expression patterns of classic tumor hallmark genes in malignant cells further revealed molecular differences among the malignant cell clusters (Fig. 2C), supporting the notion that malignant cells have undergone relatively stable state differentiation at the molecular level.

A systematic molecular analysis of malignant cell clusters from different metastatic sources revealed (Fig. 2D, Supplementary Table S2) that lymph node metastasis-associated malignant cells exhibited

significantly enhanced activation of the ERK1/2 cascade and PI3K/Akt signaling pathways compared to non-metastatic tumor cells. Given the central role of these pathways in regulating key biological processes such as cell proliferation, growth, migration, and survival, their sustained activation may confer functional advantages for tumor cell expansion and persistence within the lymph node microenvironment[24,25]. In contrast, malignant cells originating from bone metastases exhibit distinct molecular characteristics, showing significant enrichment in zinc ion transport-related pathways. Among these, the key zinc transporter LIV-1 has been demonstrated to enhance tumor cell invasiveness and distant metastatic potential by regulating intracellular zinc homeostasis and inducing epithelial-mesenchymal transition (EMT)[26].

At the functional level, malignant cells from different metastatic origins exhibit significant differences. Cancer cell clusters originating from bone metastases (Malignantcells_BM1, Malignantcells_BM3) exhibited markedly enhanced hypoxic responses and epithelial-mesenchymal transition (EMT) states (Fig. 2E). Their expression levels of hypoxia-related genes (e.g., BHLHE40, PFKFB3) and EMT-related genes (e.g., EPB41L5) were significantly higher than those of malignant cells originating from lymph node metastases (Fig. 2F). Previous studies have demonstrated that BHLHE40 can be activated by HIF-1α under hypoxic conditions and promote tumor progression[27]. This finding suggests that hypoxic adaptation may serve as a key functional basis for bone metastatic malignant cells to acquire enhanced invasiveness and distant seeding capabilities. In contrast, malignant cells originating from lymph node metastases tend to enrich functional pathways such as oxidative phosphorylation and may exhibit significant activation of inflammation-related pathways due to oxygen availability[28], reflecting distinct adaptive strategies in energy metabolism and immune-related functions. Additionally, malignant cell clusters originating from lymph node and bone metastases activate angiogenesis-related programs, supporting sustained tumor cell growth within the metastatic microenvironment (Fig. 2E, Supplementary Fig. 2H). Notably, regardless of origin, all malignant epithelial cells exhibit a significantly enhanced proliferative state compared to normal epithelial cells, with widespread upregulation of cell proliferation-associated genes (Supplementary Fig. 2F and 2G). This indicates that enhanced autonomous proliferative capacity is a common feature of malignant cells and a fundamental basis for tumor progression[29,30].

To systematically integrate functional differences among malignant cells with distinct metastatic patterns, we further performed differential expression analysis on cancer cells from various sources and constructed functional interaction networks (Fig. 2G and 2H, Supplementary Table S3, Table S4 and Table S5). Network analysis revealed that genes differentially expressed between lymph node metastases and non-metastatic tumors were significantly enriched in the “Fibroblast growth factor (FGF)-stimulated cellular response” pathway, which may promote metastasis from primary tumors to lymph nodes by inducing EMT[31,32]; In contrast, the comparison between bone-metastatic and non-metastatic tumors showed significant enrichment in pathways related to “osteoblast differentiation,” suggesting that malignant cells originating from bone metastasis may actively regulate bone-related

biological processes to remodel the bone microenvironment, thereby promoting tumor colonization and proliferation within bone tissue[33].

Result 3 Exhaustion signatures emerge in CD4 + T cells along metastatic progression.

T lymphocytes serve as the core effector cells mediating antitumor immunity and are also key targets for immune checkpoint therapy. Alterations in their functional state are considered a crucial prerequisite for tumor immune escape and metastasis[34]. To systematically decipher the dynamic remodeling of T cell functional states during breast cancer progression, we first isolated T cells from tissues with different metastatic patterns and identified multiple immune cell subsets, including CD4 + T cells at distinct differentiation stages (naive, memory, regulatory, and exhausted), CD8 + T cells (effector memory and cytotoxic), proliferating T cells, and NK cells (Fig. 3A and 3B, Supplementary Fig. 3A). Global cell composition analysis revealed distinct T cell distribution patterns across different metastatic sites (Fig. 3C). Proliferative T cells and memory CD4 + T cells were relatively enriched in lymph node metastasis samples, whereas bone metastasis samples exhibited significant increases in nearly all CD4 + and CD8 + T cell subsets, suggesting selective shaping of T cell subsets by distinct metastatic microenvironments. Compared to non-metastatic samples, T cells in bone metastases exhibited higher enrichment in antigen presentation-related pathways, indicating these cells retain antigen recognition and immune activation functions within the bone metastasis microenvironment (Fig. 3D). Due to limited numbers of non-proliferative T cells in lymph node metastasis samples, subsequent analyses focused primarily on bone metastasis and non-metastatic samples (Supplementary Fig. 3B).

Further comparison revealed that naive CD4 + T cells predominated in non-metastatic samples, whereas in bone metastasis the proportions of regulatory T cells and exhausted T cells were relatively increased (Fig. 3E). This finding suggests that tumor metastasis is not simply accompanied by changes in immune cell abundance, but rather drives a functional transition of T cells from an immunocompetent state toward immunosuppressive and dysfunctional states. Given that the functional status of T cells is closely associated with patient prognosis, we further compared the differences in effector and exhausted states between cells from non-metastatic and bone-metastatic conditions. The results showed that, compared with the non-metastatic state, both effector and exhausted scores were significantly increased in bone metastasis (Fig. 3F). This pattern was highly consistent across different T-cell subtypes (Fig. 3G), indicating that T cells in the metastatic tumor microenvironment exhibit a generally stronger activation status than those in the primary site, highlighting the importance of considering organ-specific immune contexts when applying immunotherapy.

Within the CD4 + T-cell population, the major potential tumor-reactive subsets include CD4_FOXP3_Treg and CD4_PDCD1_Tex. Pseudotime analysis revealed that CD4 + T cells can differentiate from a naive state toward either regulatory T cells or exhausted T cells (Fig. 3H, 3I, and 3J; Supplementary Fig. 3C and 3D). This differentiation trajectory was highly consistent between non-metastatic and bone-metastatic samples, suggesting that CD4 + T-cell dysfunction does not arise solely during metastasis but rather represents an intrinsic immune evolutionary program that gradually develops during tumor

progression[35]. This evolution may be driven by persistent tumor-derived antigens and immunomodulatory signals, which expose T cells to chronic stimulation and progressively shift them from an effector state toward immunosuppressive or dysfunctional phenotypes. However, in bone-metastatic samples, the proportion of cells at the later stages of the differentiation trajectory, as well as their exhaustion scores, were markedly increased (Fig. 3K and 3L), indicating that CD4 + T cells in the bone metastatic microenvironment are more prone to develop toward terminal immunosuppressive phenotypes. Consistently, immunosuppressive molecules including PDCD1, CXCL13, FOXP3, and CTLA4 were progressively upregulated along the differentiation trajectory (Supplementary Fig. 3E). Among them, PDCD1 and CTLA4 are classical immune checkpoint molecules, and their elevated expression suggests suppressed T-cell activity. The upregulation of FOXP3 reflects an increased proportion of regulatory T cells, which may promote immune evasion by secreting immunosuppressive factors or inhibiting effector T-cell function. Meanwhile, CXCL13 expression is commonly associated with chronic antigen stimulation and the development of exhausted T cells[36,37]. The sustained upregulation of these molecules indicates that the bone metastatic microenvironment imposes stronger immunosuppressive pressure, driving CD4 + T cells into a state of deep exhaustion and thereby weakening antitumor immune responses. Similar exhaustion-related transformation was observed in CD8 + T cells, suggesting that functional exhaustion is not confined to CD4 + T cells but represents a widespread immune evasion mechanism in breast cancer progression (Supplementary Fig. 3F-L).

To further elucidate the transcriptional characteristics underlying functional differences in CD4 + T cells across distinct metastatic pathways, we performed functional enrichment analysis on significantly differentially expressed gene sets across different differentiation states (Fig. 3M). CD4 + T cells in bone metastases exhibited enrichment in pathways related to exogenous antigen presentation, T cell-mediated immune responses, and cell adhesion across different differentiation states. This suggests that immune activation signals coexist with immune suppression and functional exhaustion programs[38]. This aligns with previous single-cell studies in breast cancer, which revealed T cells in the tumor microenvironment simultaneously exhibiting both activated and suppressed features. It further demonstrates that during tumor progression, immune activation does not directly translate into effective antitumor responses but is profoundly disrupted by complex immunosuppressive mechanisms[39]. Overall, CD4 + T cells in breast cancer bone metastases exhibit distinct functional duality: while retaining antigen recognition and immune response potential, the early establishment of regulatory and exhaustion programs provides a potential immunological basis for tumor immune escape and the establishment of a stable metastatic niche in bone. In summary, we have elucidated the immunoregulatory cellular reprogramming process of CD4 + T cells, transitioning from initial to exhausted characteristics during the malignant progression of breast cancer.

Result 4 Pro- to Anti-Inflammatory macrophage conversion drives immunosuppression in bone metastasis

Myeloid cells in the tumor microenvironment are recognized as key regulators driving tumor metastasis and directed dissemination[40]. They synergistically promote metastasis formation through multiple mechanisms, including mediating immune suppression[41], remodeling the extracellular matrix (ECM)[42], and inducing angiogenesis[43]. The myeloid cell population is primarily composed of macrophages, followed by dendritic cells, with macrophages exhibiting significant functional remodeling and state differences across distinct migratory pathways. Based on co-expression patterns of genes associated with inflammation and lipid metabolism, we identified a macrophage subtype marked by Phosphorylated Protein 1 (SPP1) (Fig. 4A). This subset co-expresses typical pro-inflammatory molecules alongside anti-inflammatory and immunosuppressive-associated apolipoprotein E (APOE) (Fig. 4B and 4C), suggesting it does not reside in a classical pro-inflammatory or anti-inflammatory binary state but rather exhibits a highly plastic phenotype. Building on previous studies highlighting the critical role of SPP1 + macrophages in immunosuppression and breast cancer metastasis[44], we further subdivided SPP1 + macrophages into SPP1-Pro, SPP1-Pro-Anti, and SPP1-Anti subtypes, thereby providing a more refined characterization of the continuum of macrophage states.

Comparisons across samples from different metastatic directions revealed markedly distinct distribution patterns among the three SPP1 + macrophage subtypes (Fig. 4D). The SPP1-Pro subtype was significantly enriched in lymph node metastasis, while the proportions of SPP1-Pro-Anti and SPP1-Anti subtypes were markedly elevated in bone metastasis. This suggests that as tumors progress from no metastasis to proximal metastasis and then to distant metastasis, the functional state of macrophages does not remain static but undergoes a systematic transformation, gradually shifting from pro-inflammatory to anti-inflammatory. To validate this hypothesis, we constructed a developmental trajectory of SPP1 + macrophages (Fig. 4E). The results revealed that macrophages exhibiting the most pronounced pro-inflammatory characteristics were primarily located in the early differentiation stage. Subsequently, they gradually transitioned to SPP1-Pro-Anti and SPP1-Anti states, ultimately forming three terminal differentiation states (State3, State4, and State5). These terminal states are predominantly composed of pro-inflammatory/anti-inflammatory mixed and anti-inflammatory SPP1 + macrophages, exhibiting significant source preferences in non-metastatic and bone metastatic settings (Fig. 4F and 4G, Supplementary Fig. 4A). Notably, this differentiation pathway exhibits high consistency across different metastatic directions (Supplementary Fig. 4B), indicating that the transition from pro-inflammatory to anti-inflammatory properties is not an incidental event specific to certain metastatic directions, but rather an essential pathway in the breast cancer metastasis process. Concurrently, this process generates pro-inflammatory/anti-inflammatory mixed cells with distinct transcriptional characteristics. Further functional annotation of dynamically changing gene modules during pro-inflammatory-to-anti-inflammatory differentiation (Fig. 4H and 4I, Supplementary Fig. 4C) revealed systematic functional remodeling of macrophages: early stages primarily featured cytokine responses and antigen presentation[45], while progression toward the terminal state progressively shifted toward cell migration, promoting the formation of the metastatic microenvironment[46]. Although the overall translational trajectories are highly similar, there are distinct differences in the key transcription factors driving this process (Fig. 4J). The pro-inflammatory/anti-inflammatory mixed macrophage-specific

transcription factor BHLHE40 promotes extensive pro-inflammatory and glycolysis-related gene expression by upregulating HIF-1α levels in macrophages, enhancing pro-inflammatory functions[47]. It also provides the metabolic foundation for subsequent transition to an anti-inflammatory state[48], exhibiting specific activation in bone metastasis samples (Supplementary Fig. 4D). Another significantly upregulated transcription factor, CEBPB, drives pro-tumorigenic inflammatory-immune suppression programs. By inducing the secretion of immunoregulatory cytokines and growth factors such as its target gene TGFβ1, it activates myeloid differentiation-related pathways to maintain a highly plastic functional state. Concurrently, it promotes the proliferation of epithelioid tumor cells, enhancing tumor adaptation and expansion within the bone microenvironment (Fig. 4J, Supplementary Fig. 4E)[49]. The enrichment of osteogenic differentiation pathways further reveals the critical role of bone metastasis-associated macrophages in bone remodeling. Coupled with negative regulation of inflammatory responses, they collectively shape an immunotolerant yet highly tumor-promoting and bone metastasis-facilitating microenvironment.

Result 5 Fibroblasts display enhanced invasive properties in bone metastases

To systematically characterize the remodeling features of stromal cell states in breast cancer across different metastatic pathways, we performed subtype segmentation and functional annotation of fibroblasts and endothelial cells within the tumor microenvironment. Based on the integrated expression of classical cell marker genes and functionally relevant core genes, we identified seven fibroblast subtypes and three vascular endothelial cell subtypes (Fig. 5A). Among fibroblasts, matrix-remodeling fibroblasts (mCAFs) are characterized by extracellular matrix and fibrosis-related genes such as COL11A1, FN1, and POSTN, representing typical matrix-remodeling fibroblasts; antiviral fibroblasts (avCAFs) highly express interferon-stimulated genes including RSAD2, MX1, and STAT1, suggesting antiviral and immune response features; Vascular fibroblasts (vCAFs) are defined by TIMP1, APOE, and the pericellular-associated gene RGS5, reflecting their potential role in vascular stability and angiogenesis regulation; Immune fibroblasts (iCAFs) are marked by immune and inflammation-related genes such as CXCL14, C3, and IGF1, suggesting their function in immune regulation; Metabolic fibroblasts (meCAFs) highly express FABP5, ATP5PO, and COX8A, exhibiting active glycolytic and oxidative phosphorylation metabolic characteristics; Myofibroblasts (myCAFs) are defined by cytoskeletal and contractile-related genes such as ACTA2, TPM1, and TAGLN, demonstrating enhanced contractile capacity; proliferative fibroblasts (pCAFs) are identified through proliferation marker genes including MKI67, TOP2A, and BIRC5, representing a highly proliferative state. Concurrently, we identified three endothelial cell subtypes: arterial endothelial cells (ArtECs), venous endothelial cells (VenECs), and capillary endothelial cells (CapECs) (Fig. 5B). The consistency of these subtypes in functional annotation further validated the reliability of the results and revealed significant heterogeneity between fibroblasts and endothelial cells in different metastasis directions.

After comparing sample compositions across different metastatic sites, we observed significant enrichment of meCAFs and myCAFs in lymph node metastases (Fig. 5C and 5D). Differential gene and pathway analysis revealed that these fibroblasts highly activate metabolic processes such as oxidative phosphorylation, suggesting they provide sustained metabolic support to tumor cells by enhancing energy metabolism[50]. Concurrently, myCAF-mediated cytoskeletal remodeling and contractile capabilities promote tumor cell migration and invasion (Fig. 5E and 5F)[51,52]. In contrast, the most prominent feature in bone metastasis samples was the marked expansion of vCAFs, a phenomenon highly consistent with the concurrent enrichment of vascular-associated cells within the endothelial cell subtype. Functional scoring further indicates that bone metastasis samples exhibit significantly higher activity in pathways such as angiogenesis and vascular homeostasis regulation compared to non-metastatic and lymph node metastasis samples. This suggests that tumors in the bone microenvironment are highly dependent on vascular-associated stromal remodeling to sustain growth and invasion. This finding reveals that bone-metastatic breast cancer exhibits a pro-invasive microenvironment characterized by vascular-associated fibroblasts (Fig. 5E and 5F)[43,53,54].

To further elucidate the regulatory mechanisms underlying the differential states of fibroblasts across different metastatic trajectories, we performed transcription factor regulatory network analysis on fibroblasts from distinct samples. Fibroblasts exhibited a continuous and orderly transcriptional reprogramming, with their core functions progressively evolving alongside tumor progression. In non-metastatic fibroblasts, TWIST2 and TGIF1 exert dominant regulatory control, with their target genes significantly enriched in cellular metabolism, growth, and development-related processes. This suggests these cells respond to microenvironmental stresses such as hypoxia and nutrient fluctuations by enhancing energy metabolism and environmental adaptability. Concurrently, structural and functional remodeling occurs, providing sustained support for tumor long-term survival and stable growth (Fig. 5G and 5H)[55–57]. In lymph node metastasis, the regulatory focus of fibroblasts shifts to TCF12 and FOXP1, with downstream gene functions enriched in processes such as cytoskeletal remodeling and epithelial tissue development. Notably, the significant activation of actin nucleation and Arp2/3 complex-mediated cytoskeletal reorganization pathways confirms the dominant role of myCAFs in lymph node metastasis at the regulatory level. This mechanism reveals that myCAFs promote tumor cell colonization and proliferation within lymph nodes by enhancing contractility and hormone responsiveness (Fig. 5G and 5H)[58–60]. In contrast, during the bone metastasis stage, fibroblasts are primarily driven by NR2F2 and JUNB. Their target genes are concentrated in pathways related to cell proliferation and responses to hormonal and fibroblast growth factor (FGF) stimulation, reflecting their heightened sensitivity to growth factors and endocrine signals within the bone microenvironment. This reveals a tightly coupled positive feedback regulatory circuit formed between tumor cells and fibroblasts via growth factor signaling, continuously amplifying pro-proliferative and pro-invasive signals to drive the stable expansion of bone metastases (Fig. 5G and 5H)[61,62]. Overall, fibroblasts dynamically evolve during breast cancer progression, highlighting their irreplaceable tumor-promoting functions across different metastatic stages.

Result 6 Angiogenesis driven by venous endothelial cells

Tumor angiogenesis refers to the process of forming new blood vessels on the basis of an existing vascular bed, serving as a key driver of tumor growth and progression[63,64]. To elucidate the molecular basis of tumor angiogenesis in breast cancer, we reconstructed the angiogenesis process through transcriptomic analysis of 2,243 endothelial cells (including venous endothelial cells (VenECs), capillary endothelial cells (CapECs), and arterial endothelial cells (ArtECs)) (Fig. 5A). In non-metastatic tissues, VenECs were significantly enriched, whereas the proportions of ArtECs and CapECs progressively increased in lymph node and bone metastasis samples as metastasis progressed (Fig. 5C and 5D). This suggests that tumor angiogenesis becomes more active and vascular structures become less stable during metastasis, thereby accelerating tumor progression[65].

Differentiation potential analysis revealed that VenECs exhibit the highest cellular plasticity, suggesting they may serve as the initial population for endothelial cell state transitions (Fig. 5I). Pseudotemporal trajectory analysis further demonstrated that endothelial cells follow a continuous differentiation pathway from VenECs to CapECs, ultimately differentiating into ArtECs (Fig. 5J, 5K, 5L, and 5M). Within this trajectory, non-metastatic samples predominantly reside in early stages, whereas metastatic samples—particularly bone metastasis samples—gradually enrich in mid-to-late stages as the trajectory progresses (Fig. 5L and 5N). This suggests metastasis-associated endothelial cells exhibit a more mature and functionally enhanced vascular state. Functional enrichment analysis indicates that the early stage of angiogenesis is characterized by an inflammatory vein-like state in endothelial cells, significantly enriched in processes such as wound healing, leukocyte adhesion, and MHC-II complex assembly, reflecting its potential role in inflammatory responses and immune cell recruitment. The intermediate phase exhibits transitional transcriptional reprogramming marked by protein folding and translational regulation; by the trajectory's end, pathways related to endothelial cell development, differentiation, and arterial morphogenesis are significantly enhanced, accompanied by high expression of arterial and mature endothelial marker genes such as EFNB2, ESM1, and CXCL12, signifying the cells' transition into a terminal arterial-like endothelial state (Fig. 5O and 5P). Among these, the endothelial cell-specific molecule ESM1 (endocan), acting as a proangiogenic secreted protein, has been demonstrated to exert a key carcinogenic role in multiple cancers by regulating tumor angiogenesis[66]. This further supports the biological significance of endothelial cell transformation toward a pro-tumor vascular phenotype in bone metastasis samples.

Result 7 Complex immunosuppressive and pro-invasive patterns in bone metastatic TME

To systematically analyze the tumor microenvironments of primary tumors with different metastatic directions in breast cancer, we first quantitatively characterized the communication relationships between malignant and non-malignant cells in non-metastatic, lymph node metastatic, and bone metastatic samples based on the number of potential intracellular ligand-receptor interactions

(Supplementary Fig. 5A). Results indicate that regardless of metastasis status, malignant cells form stable interactive networks with tumor-associated macrophages (TAMs), cancer-associated fibroblasts (CAFs), and endothelial cells, suggesting these cell types serve as core hubs mediating tumor progression within the breast cancer tumor microenvironment (TME). However, under identical selection conditions, the tumor microenvironment of bone metastases exhibits significantly more complex and dense intercellular communication patterns. Malignant cells with tissue-specific origins serve as primary ligand signal emitters, while non-malignant cells act as signal receptors that respond synergistically, thereby constructing a specific metastatic microenvironment with invasive potential. Building upon this foundation, we employed “NicheNet” to infer hypothetical crosstalk between cells, predicting the expression of ligands with the highest activity across different metastatic directions and their corresponding receptors, along with their downstream target gene networks. This approach confirmed the key mediators and ultimate outcomes of interactions between malignant cells and other cell types (Fig. 6A–6I).

In non-metastatic tumors, annexin A2 (ANXA2) exhibits high activity in malignant cells. This molecule has been demonstrated to play a critical role in cell membrane repair, cell migration, and invasion processes. Suppressing its expression significantly blocks tumor invasion and distant metastasis. Concurrently, PDGFB secreted by cancer cells synergistically regulates cell proliferation and angiogenesis by acting on corresponding receptors on CAFs. This process, coupled with the significant activation of pathways related to extracellular matrix (ECM) composition, collectively promotes tumor progression. This phenomenon is highly consistent with previous findings on primary breast cancer matrix remodeling (Fig. 6B and 6C, Supplementary Fig. 5B and 5C)[67]. In contrast, the tumor microenvironment of lymph node metastases exhibits more pronounced immunoregulatory characteristics. We observed that serum amyloid A1 (SAA1) demonstrates specific high activity in malignant cells of this origin (Fig. 6E). which binds to Toll-like receptor 2 (TLR2) on macrophage surfaces (Fig. 6J). This interaction activates TLR2/MYD88-mediated PI3K/NF-κB and p38 MAPK signaling pathways, inducing neutrophils to produce immunosuppressive cytokines that synergistically promote tumor progression[68]. Additionally, the immune checkpoint molecule VTCN1 and the chemokine CXCL10 were also highly active in lymph node metastasis samples. These molecules have been extensively reported to be closely associated with tumor cell migration, EMT, and poor prognosis [69–71]. Functional enrichment analysis of their predicted target genes further indicates that pathways related to immune cell differentiation, activation, and immune response are significantly enhanced in this tumor microenvironment (Fig. 6L, Supplementary Table S6). This suggests that lymph node metastasis exist in a state of highly remodeled immunity but have not yet reached complete immune exhaustion.

Notably, bone metastasis samples exhibited the most complex and synergistically enhanced pro-invasive signaling networks. We found that heparin-binding epidermal growth factor (HBEGF) was significantly upregulated in these malignant cells (Fig. 6H). Under hypoxic conditions, this molecule undergoes enhanced expression through BHLHE40 mediated transcriptional activation and is secreted via the exosomal pathway, significantly boosting tumor cell survival and migration capacity[72]. This

finding strongly aligns with our previous observations of enhanced hypoxic adaptation in malignant cells and the synchronous upregulation of BHLHE40 during macrophage reprogramming in bone metastasis samples, suggesting that BHLHE40 holds potential value as a predictive factor for metastatic status. Furthermore, transforming growth factor β (TGF-β) has been identified as a key mediator of immune suppression and invasive phenotype remodeling in bone metastasis. By activating the TGF-β signaling pathway, it promotes immune evasion, angiogenesis, EMT, and distant metastasis during advanced tumor progression[73,74]. Additionally, PDGFB and PDGFC synergistically drive matrix remodeling and angiogenesis by acting on PDGFRB on the surface of CAFs, thereby further amplifying tumor-matrix interactions (Fig. 6H). Pathway enrichment analysis of predicted target genes for the aforementioned signaling axes revealed significant clustering in pathways related to metabolic reprogramming and extracellular matrix composition (Fig. 6L, Supplementary Table S6). Among these, abnormal phosphorylation states and nitric oxide (NO) production are recognized as key drivers inducing cell migration, invasion, and angiogenesis[75,76]. Collectively, these findings reveal that bone-metastatic breast cancer, through cross-cell-type signaling synergism, shapes a highly pro-invasive, immunosuppressive tumor microenvironment characterized by significant metabolic remodeling.

Discussion

Previous studies have demonstrated that breast cancer metastasis occurs alongside the activation of key oncogenic pathways and profound remodeling of the tumor microenvironment[77]. However, constrained by bulk-level analysis, direct evidence remains lacking as to whether tumor progression relies on cellular lineage replacement or functional reprogramming of existing cellular states. In recent years, single-cell studies have revealed tumor heterogeneity[78], but most have focused on a single metastatic site or terminal metastases[79–81]. Systematic comparisons of tumor ecosystems in primary tumors with different metastatic directions remain insufficient. Based on single-cell transcriptomic analysis of primary breast cancer tumors with different metastatic directions, this study reveals that breast cancer does not promote tumor progression by altering cell lineage composition across different metastatic pathways. Instead, it drives the overall evolution of the tumor ecosystem through fine-tuned regulation of the quantity, functional state, and interactive relationships within existing cell populations[82]. This discovery provides a crucial conceptual framework for understanding breast cancer metastasis: metastasis is not an isolated malignant cellular event, but rather a systemic process driven by the coordinated remodeling of multiple cell types and pathways[83].

Our findings indicate that despite maintaining relatively stable lineage identities in their transcriptomes across different metastatic stages, tumor cells undergo significant functional reprogramming. This supports the notion that “state remodeling rather than identity replacement” is the primary driver of breast cancer progression[84–86]. Specifically, between lymph node metastasis and bone metastasis, the tumor ecosystem undergoes a functional shift from prioritizing enhanced metabolic efficiency and proliferation support toward hypoxic adaptation and increased invasiveness. This suggests fundamentally different microenvironmental selection pressures confronting tumor cells at distinct

metastatic stages: lymph nodes, as immunologically active and relatively oxygen-rich transit organs, favor the expansion of proliferative and metabolically enhanced cellular states[23]; whereas the bone microenvironment exerts intense selective pressure through hypoxic conditions and unique cellular composition, driving tumor cells to acquire heightened invasive potential and immune evasion capabilities[87,88].

At the immune level, this study further reveals that T cell exhaustion and the shift of macrophages from pro-inflammatory to anti-inflammatory states are not isolated events, but rather highly conserved immune reprogramming processes during breast cancer progression. Notably, immune activation signals coexist with immunosuppressive programs in bone metastases, creating a unique “functional paradox.” This state may provide crucial insights into the limited response of bone metastasis patients to immunotherapy: despite sustained immune stimulation, the early-established exhaustion state and immunosuppressive networks effectively constrain antitumor immune responses. Furthermore, our research highlights the pivotal role of fibroblasts and the vascular system within the metastasis-associated tumor microenvironment. Fibroblasts undergo differential regulation by distinct transcription factor networks across metastatic stages, progressively shifting from a state dominated by metabolic support and contractile functions toward a vascular-associated and growth factor-responsive state. Endothelial cells, meanwhile, evolve along a continuous differentiation trajectory toward a mature, highly proangiogenic state. This synergistic remodeling of the stroma and vascular system not only provides structural and nutritional support for tumor growth and distant implantation[89], but also amplifies the signaling output dominated by malignant cells, enhancing the overall invasiveness of the metastatic ecosystem.

Through a comprehensive analysis of the intercellular ligand-receptor communication network, we further discovered that in bone metastasis, malignant cells serve as central nodes for signal transmission. Collaborating with immune cells, fibroblasts, and endothelial cells, they construct a highly synergistic tumor microenvironment characterized by pronounced immunosuppression and dense pro-invasive signaling. This ecosystem systematically supports distant bone metastasis in breast cancer through the synergistic action of multi-level mechanisms including metabolic reprogramming, immune evasion, matrix remodeling, and angiogenesis. This discovery not only expands our understanding of breast cancer metastasis biology but also provides a theoretical basis for intervention strategies targeting the tumor ecosystem rather than individual cell types.

Materials and methods

Participants

Primary tumor tissues, lymph node metastases, and bone metastasis samples were collected from four patients at the Second Affiliated Hospital of Harbin Medical University. Among them, two samples were taken from the primary lesion of breast cancer, two samples were taken from the primary lesion of breast cancer with bone metastasis, and one sample was taken from the paired axillary lymph nodes. All

participants provided written informed consent, and the study was conducted in compliance with Institutional Review Board-approved protocols of the Second Affiliated Hospital of Harbin Medical University (IRB: KY2025-83). The research adhered to the principles outlined in the Declaration of Helsinki. The samples of these four patients were all used for RNA sequencing.

Single-cell suspension preparation

Fresh tissue was washed three times with Hanks balanced salt solution and cut into small pieces, with 3 ml tissue dissociation digestion under 37°C for 15 min. The cell suspension was collected and filtered through a 40-micron sterile filter. Then, red blood cell lysis buffer was added, and the mixture [cell: RCLB = 1:2 (volume ratio)] was incubated at room temperature for 5–8 min to remove the red blood cells. We then centrifuged the mixture at 300 × g at 4°C for 5 min, removed the supernatant, and used PBS for soft suspension.

Library preparation and scRNA-seq

A single single-cell suspension (1×10^5 cells /ml) was barcoded, and an scRNA-seq library was generated using the Chromium single-cell 3'Library, Gel Bead & Multiplex Kit (10x Genomics) according to the manufacturer's instructions. In the ChromiumTM Controller instrument, individual cells were segmented into gel beads in the emulsion for cell lysis and barcode RNA reverse transcription. All libraries were next generation sequenced (NSG) on Illumina Novaseq 6000 with a paired end read length of 150 bp.

Integrated analysis of scRNA-seq data

Single-cell RNA-seq (scRNA-seq) datasets from primary lesion of breast cancer with lymph node metastasis were retrieved from four publicly available cohorts (GSE180286). Together with our in-house generated breast cancer scRNA-seq data, a total of 9 samples were included in the main analysis. Scanpy v1.9.3[90] was used for quality control, dimensionality reduction and clustering under Python 3.10. For each sample dataset, we filtered expression matrix by the following criteria: 1) cells with gene count less than 200 or with top 2% gene count were excluded; 2) cells with top 2% UMI count were excluded; 3) cells with mitochondrial content 50% were excluded; 4) genes expressed in less than 5 cells were excluded. After filtering, 71,805 cells were retained for the downstream analyses. The raw count matrix was normalized by total counts per cell and logarithmically transformed into normalized data matrix. Top 2000 variable genes were selected by setting flavor = ‘seurat_v3’. Principle Component Analysis (PCA) was performed on the scaled variable gene matrix, and top 2000 principle components were used for clustering and dimensional reduction. DecontX v1.6.1[91] was used to estimate and remove contamination. Cells were separated into 34 clusters by using Louvain algorithm and setting resolution parameter at 1.2. Cell clusters were visualized by using Uniform Manifold Approximation and Projection (UMAP) .

Cell annotation

The cell type of each cluster was determined based on the expression of the typical markers in the reference database SynEcoSys™ (Singleron Biotechnology, Germany). SynEcoSys™ contains a collection of standard cell type markers for single-cell sequence data from CellMaker 2.0[92], PanglaoDB[93], and recently published literature. Cell doublets were estimated based on the expression pattern of canonical cell markers. Any clusters rich in markers specific to multiple cell types were excluded for downstream analysis.

Comparison dendrograms

For an unsupervised comparison of the sub-populations identified from primary lesions with different metastatic directions, the following steps were performed. Here, we used the batch-corrected expression value from Harmony.

1. Identify a set of highly variable genes across different celltype.
2. Calculate the mean expression of genes in each cluster.
3. For hierarchical clustering, the distance defined as (1-Pearson correlation coefficient)/2 was used.

For major cell lineages comparison across primary lesions with different metastatic directions, we used the top 2,000 highly variable genes.

InferCNV identifies malignant cells

We used intercnv (https://github.com/broadinstitute/infercnv) [94] to infer the copy number variation scores of epithelial cells in the breast cancer dataset, setting the parameters to the default values, with cutoff = 0.1: filtering out genes with an average count less than 0.1 in the cells, and denoise = T: performing denoising. After obtaining the scores for each cell, the matrices corresponding to the reference and observed panels were extracted, and K-means clustering was performed to group them into seven categories. Then, the heat map of CNV scores was observed, and it was found that the seventh category clustered with the reference cells. After removing the cells in the seventh category, the malignant epithelial cells were obtained.

Heatmap of functional differences in genes among subtypes of cancer cells

We used the R package "ClusterGVis" (https://github.com/junjunlab/ClusterGVis) to analyze the top 5 functional enrichment terms of the characteristic gene expressions of cell subgroups as shown in the tutorial (parameters: type = "BP"). Among them, we used the "FindAllMarkers" function in Seurat to identify the genes specifically expressed by subgroups, setting the parameters: min.pct = 0.25, logfc.threshold = 0.25.

Describe the functional state of cancer cells

We downloaded the gene sets describing including stemness, invasion, metastasis, proliferation, EMT, angiogenesis, apoptosis, cell cycle, differentiation, DNA damage, DNA repair, hypoxia, inflammation and

quiescence from the CancerSEA website (http://biocc.hrbmu.edu.cn/CancerSEA)[95]. Then, we used the AUcell algorithm to score the cancer cells in different sample groups and different subgroups successively. At the same time, irGSEA (https://github.com/chuiqin/irGSEA) was used for visualization.

Different transfer types of subtypes

To quantify the enrichment of each subtypes across transfer types, we compared the observed and expected cell numbers in each subtypes by computing the $R_{o/e}$ value using the epitools (v.0.5–10.1) R package according to the following formula15:

$$R_{o/e} = \frac{Observed}{Expected'}$$

where the expected cell numbers of each subtypes and transfer types were obtained from the χ^2 test. We assumed that one subtype was enriched in a specific grade if $R_{o/e} > 1$.

Constructing gene function network

The R packages ClusterProfiler[96], org.Hs.eg.db, and aPEAR (https://github.com/kerseviciute/aPEAR)[97] are used to construct the functional network of different cancer cell sources' differential genes. Here, the differential genes are identified by the "FindMarkers" function of the R 'Seurat' package, with specific parameter settings as follows: ident.1 = "LyM", ident.2 = "PT", min.pct = 0.2, logfc.threshold = 0; ident.1 = "BM", ident.2 = "PT", min.pct = 0.2, logfc.threshold = 0. On this basis, genes that meet the condition of p_val_adj ≤ 0.05 and |avg_log2FC| ≥ 2 are further selected as significantly differentially expressed genes. We use enrichGO to obtain the gene ontology (GO) functions of the differential genes, and use the enrichmentNetwork function to create a network diagram for the top 50 BP that are significantly enriched, setting the parameters simMethod = 'jaccard', clustMethod = 'markov' to visualize the enrichment degree of genes and the connection between genes and functions.

Monocle2 trajectory analysis

We used Monocle2[98] to infer the evolutionary trajectories of each cell subpopulation. The newCellDataSet in the Monocle package was used to construct CellDataSet objects for specific cell types, followed by the application of the dispersionTable and setOrderingFilter functions to identify the variable genes. The reduceDimension function (DDRTree) was used to reduce the dimensionality, and finally, the orderCells function was applied to arrange the cells and obtain the cell trajectories. We used plot_cell_trajectory to visualize the colored pseudo-time trajectories of each sample group.

Enrichment analysis

We used the GSEA function to obtain the common functional pathways of the differential gene expression profiles of cancer cells in primary and metastatic breast cancer. We used the R "clusterProfiler" package for KEGG pathway enrichment analysis and GO enrichment analysis, applying

the enrichKEGG and enrichGO functions, specifying the p-value correction method as pAdjustMethod = 'fdr'. Single-cell level enrichment analysis applied GSVA scoring. The GO functional classification was obtained from the website (https://www.geneontology.org/). The gene signature of CD8 + cytotoxic-related genes includes "IL2", "GZMA", "GNLY", "PRF1", "GZMB", "GZMK", "IFNG", "NKG7". Macrophage pro-inflammatory-related genes include "IL1B", "IL6", "S100A8", "S100A9" anti-inflammatory-related genes are "CD163", "SEPP1", "APOE", "MAF".

Transcription factor regulatory network analysis

The transcriptional regulatory networks and the activity of regulators in different sample groups and different cell subpopulations were analyzed using pySCENIC (https://github.com/aertslab/pySCENIC)[99]. Regulators with high activity were determined based on the default threshold of AUCell. After obtaining the loom file, the characteristic transcription factors and corresponding target genes in different dimensional sample groups were obtained. The network composed of transcription factors and the genes they regulate was constructed by Cytoscape (https://cytoscape.org).

Cell-cell interaction network analysis

We used CellPhoneDB (www.cellphonedb.org)[100] to analyze the ligand-receptor pairs of intercellular interactions within different sample combinations. We obtained the expression matrices of the sample combinations and the metadata information of the cell subpopulations they belonged to. Based on the gene expression levels, we inferred the potential interaction strength between the two cell subpopulations and obtained significance through permutation tests.

NicheNet[101] is used to infer the interactions between tumor-specific cancer cells and the TME within different sample groups. Set "ligand cells" as malignant cells and "receptor cells" as TME cells, and evaluate the differential ligand activity among different sources of malignant cells. The make_heatmap_ggplot function is used to visualize the AUPR values of the top 10 active ligands, the regulatory potential between the ligands and the predicted target genes, and the visualization of the interaction potential between the ligands and the receptors.

Declarations

Author contributions

Zhang Yunpeng and Li Xia: Funding acquisition, Resources. Hu Congmei: Project administration, Supervision. Yang Kaiyue: Conceptualization, Methodology, Validation, Writing - Original Draft. Yang Kaiyue and Fu Rui: Software. Yang Kaiyue and Ma Yining: Investigation. Yang Kaiyue and Long Yongjin: Data Curation. Yang Kaiyue and Li Maohao: Formal analysis, Visualization. Yang Kaiyue, Wang Zhenzhen and Sun Mingyao: Writing - Review & Editing. All the authors approved the final manuscript.

Availability of data and materials

Some data used in this study are publicly available from GEO (https://ncbi.nlm.nih.gov/geo/). The datasets analyzed are existing publicly accessible datasets.

Acknowledgments

This work was supported by grants from the National Science and Technology Major Program [2024ZD0530500]; and National Natural Science Foundation of China [62472131, U23A20166, 32570792, 82503871], and the Key Research and Development Program of Heilongjiang Province [2024ZX12C27], and Natural Science Foundation of Heilongjiang Province [QC2025H013].

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interest statement

The authors declare that they have no competing interests.

References

1. Filho AM, Laversanne M, Ferlay J, et al. The GLOBOCAN 2022 cancer estimates: Data sources, methods, and a snapshot of the cancer burden worldwide. *Int J Cancer.* Apr 1 2025;156(7):1336-1346.
2. Cancer Genome Atlas N. Comprehensive molecular portraits of human breast tumours. *Nature.* Oct 4 2012;490(7418):61-70.
3. Xu K, Wang R, Xie H, et al. Single-cell RNA sequencing reveals cell heterogeneity and transcriptome profile of breast cancer lymph node metastasis. *Oncogenesis.* Oct 5 2021;10(10):66.
4. Davis RT, Blake K, Ma D, et al. Transcriptional diversity and bioenergetic shift in human breast cancer metastasis revealed by single-cell RNA sequencing. *Nat Cell Biol.* Mar 2020;22(3):310-320.
5. Weigelt B, Peterse JL, van 't Veer LJ. Breast cancer metastasis: markers and models. *Nat Rev Cancer.* Aug 2005;5(8):591-602.
6. Tulotta C, Ottewell P. The role of IL-1B in breast cancer bone metastasis. *Endocr Relat Cancer.* Jul 2018;25(7):R421-R434.
7. Lei PJ, Pereira ER, Andersson P, et al. Cancer cell plasticity and MHC-II-mediated immune tolerance promote breast cancer metastasis to lymph nodes. *J Exp Med.* Sep 4 2023;220(9).
8. Chaffer CL, Weinberg RA. A perspective on cancer cell metastasis. *Science.* Mar 25 2011;331(6024):1559-1564.

9. Allemani C, Matsuda T, Di Carlo V, et al. Global surveillance of trends in cancer survival 2000-14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. *Lancet.* Mar 17 2018;391(10125):1023-1075.
10. Ding S, Chen X, Shen K. Single-cell RNA sequencing in breast cancer: Understanding tumor heterogeneity and paving roads to individualized therapy. *Cancer Commun (Lond).* Aug 2020;40(8):329-344.
11. Bahcecioglu G, Basara G, Ellis BW, Ren X, Zorlutuna P. Breast cancer models: Engineering the tumor microenvironment. *Acta Biomater.* Apr 1 2020;106:1-21.
12. Mittal S, Brown NJ, Holen I. The breast tumor microenvironment: role in cancer development, progression and response to therapy. *Expert Rev Mol Diagn.* Mar 2018;18(3):227-243.
13. Zarrilli G, Businello G, Dieci MV, et al. The Tumor Microenvironment of Primitive and Metastatic Breast Cancer: Implications for Novel Therapeutic Strategies. *Int J Mol Sci.* Oct 30 2020;21(21).
14. Wyckoff J, Wang W, Lin EY, et al. A paracrine loop between tumor cells and macrophages is required for tumor cell migration in mammary tumors. *Cancer Res.* Oct 1 2004;64(19):7022-7029.
15. Su S, Liu Q, Chen J, et al. A positive feedback loop between mesenchymal-like cancer cells and macrophages is essential to breast cancer metastasis. *Cancer Cell.* May 12 2014;25(5):605-620.
16. Doedens AL, Stockmann C, Rubinstein MP, et al. Macrophage expression of hypoxia-inducible factor-1 alpha suppresses T-cell function and promotes tumor progression. *Cancer Res.* Oct 1 2010;70(19):7465-7475.
17. Kythreotou A, Siddique A, Mauri FA, Bower M, Pinato DJ. Pd-L1. *J Clin Pathol.* Mar 2018;71(3):189-194.
18. Sa-Nguanraksa D, P OC. The role of vascular endothelial growth factor a polymorphisms in breast cancer. *Int J Mol Sci.* Nov 13 2012;13(11):14845-14864.
19. Ghiabi P, Jiang J, Pasquier J, et al. Endothelial cells provide a notch-dependent pro-tumoral niche for enhancing breast cancer survival, stemness and pro-metastatic properties. *PLoS One.* 2014;9(11):e112424.
20. Ghajar CM, Peinado H, Mori H, et al. The perivascular niche regulates breast tumour dormancy. *Nat Cell Biol.* Jul 2013;15(7):807-817.
21. Michelet X, Dyck L, Hogan A, et al. Metabolic reprogramming of natural killer cells in obesity limits antitumor responses. *Nat Immunol.* Dec 2018;19(12):1330-1340.
22. Roskoski R, Jr. Vascular endothelial cells and angiogenesis. *Pharmacol Res.* Nov 2025;221:107983.
23. Liu YM, Ge JY, Chen YF, et al. Combined Single-Cell and Spatial Transcriptomics Reveal the Metabolic Evolvement of Breast Cancer during Early Dissemination. *Adv Sci (Weinh).* Feb 2023;10(6):e2205395.
24. Ge XY, Shao LL, Gao XL, He RX. [Extracellular Signal-regulated Kinase 1/2 Signaling Regulates Cell Invasion:a Review]. *Zhongguo Yi Xue Ke Xue Yuan Xue Bao.* Feb 2023;45(1):155-160.

25. Hers I, Vincent EE, Tavare JM. Akt signalling in health and disease. *Cell Signal.* Oct 2011;23(10):1515-1527.
26. Unno J, Satoh K, Hirota M, et al. LIV-1 enhances the aggressive phenotype through the induction of epithelial to mesenchymal transition in human pancreatic carcinoma cells. *Int J Oncol.* Oct 2009;35(4):813-821.
27. Kiss Z, Mudryj M, Ghosh PM. Non-circadian aspects of BHLHE40 cellular function in cancer. *Genes Cancer.* 2020;11(1-2):1-19.
28. Wang X, Liu JZ, Hu JX, et al. ROS-activated p38 MAPK/ERK-Akt cascade plays a central role in palmitic acid-stimulated hepatocyte proliferation. *Free Radic Biol Med.* Jul 15 2011;51(2):539-551.
29. Li T, Forbes ME, Fuller GN, Li J, Yang X, Zhang W. IGFBP2: integrative hub of developmental and oncogenic signaling network. *Oncogene.* Mar 2020;39(11):2243-2257.
30. Liu G, Lin W, Zhang K, et al. Elucidating the prognostic and therapeutic significance of TOP2A in various malignancies. *Cancer Genet.* Nov 2024;288-289:68-81.
31. Xie Y, Su N, Yang J, et al. FGF/FGFR signaling in health and disease. *Signal Transduct Target Ther.* Sep 2 2020;5(1):181.
32. Yu Z, Lou L, Zhao Y. Fibroblast growth factor 18 promotes the growth, migration and invasion of MDA-MB-231 cells. *Oncol Rep.* Aug 2018;40(2):704-714.
33. Zarrer J, Taipaleenmaki H. The osteoblast in regulation of tumor cell dormancy and bone metastasis. *J Bone Oncol.* Apr 2024;45:100597.
34. Kim N, Kim HK, Lee K, et al. Single-cell RNA sequencing demonstrates the molecular and cellular reprogramming of metastatic lung adenocarcinoma. *Nat Commun.* May 8 2020;11(1):2285.
35. Zheng L, Qin S, Si W, et al. Pan-cancer single-cell landscape of tumor-infiltrating T cells. *Science.* Dec 17 2021;374(6574):abe6474.
36. Tay C, Tanaka A, Sakaguchi S. Tumor-infiltrating regulatory T cells as targets of cancer immunotherapy. *Cancer Cell.* Mar 13 2023;41(3):450-465.
37. Vignali PDA, DePeaux K, Watson MJ, et al. Hypoxia drives CD39-dependent suppressor function in exhausted T cells to limit antitumor immunity. *Nat Immunol.* Feb 2023;24(2):267-279.
38. Guo M, Liu MYR, Brooks DG. Regulation and impact of tumor-specific CD4(+) T cells in cancer and immunotherapy. *Trends Immunol.* Apr 2024;45(4):303-313.
39. Miggelbrink AM, Jackson JD, Lorrey SJ, et al. CD4 T-Cell Exhaustion: Does It Exist and What Are Its Roles in Cancer? *Clin Cancer Res.* Nov 1 2021;27(21):5742-5752.
40. Lee HO, Hong Y, Etlioglu HE, et al. Lineage-dependent gene expression programs influence the immune landscape of colorectal cancer. *Nat Genet.* Jun 2020;52(6):594-603.
41. Li K, Shi H, Zhang B, et al. Myeloid-derived suppressor cells as immunosuppressive regulators and therapeutic targets in cancer. *Signal Transduct Target Ther.* Oct 7 2021;6(1):362.
42. Sangaletti S, Chiodoni C, Tripodo C, Colombo MP. Common extracellular matrix regulation of myeloid cell activity in the bone marrow and tumor microenvironments. *Cancer Immunol*

Immunother. Aug 2017;66(8):1059-1067.

43. De Palma M, Biziato D, Petrova TV. Microenvironmental regulation of tumour angiogenesis. *Nat Rev Cancer.* Aug 2017;17(8):457-474.
44. Castello LM, Raineri D, Salmi L, et al. Osteopontin at the Crossroads of Inflammation and Tumor Progression. *Mediators Inflamm.* 2017;2017:4049098.
45. Shapouri-Moghaddam A, Mohammadian S, Vazini H, et al. Macrophage plasticity, polarization, and function in health and disease. *J Cell Physiol.* Sep 2018;233(9):6425-6440.
46. Cheng S, Li Z, Gao R, et al. A pan-cancer single-cell transcriptional atlas of tumor infiltrating myeloid cells. *Cell.* Feb 4 2021;184(3):792-809 e723.
47. Zafar A, Ng HP, Kim GD, Chan ER, Mahabeleshwar GH. BHLHE40 promotes macrophage pro-inflammatory gene expression and functions. *FASEB J.* Oct 2021;35(10):e21940.
48. Zheng Y, Chen Z, Han Y, et al. Immune suppressive landscape in the human esophageal squamous cell carcinoma microenvironment. *Nat Commun.* Dec 8 2020;11(1):6268.
49. Yang Y, Jin X, Xie Y, et al. The CEBPB(+) glioblastoma subcluster specifically drives the formation of M2 tumor-associated macrophages to promote malignancy growth. *Theranostics.* 2024;14(10):4107-4126.
50. Lewinska M, Zhuravleva E, Satriano L, et al. Fibroblast-Derived Lysyl Oxidase Increases Oxidative Phosphorylation and Stemness in Cholangiocarcinoma. *Gastroenterology.* May 2024;166(5):886-901 e887.
51. Sandbo N, Dulin N. Actin cytoskeleton in myofibroblast differentiation: ultrastructure defining form and driving function. *Transl Res.* Oct 2011;158(4):181-196.
52. Yi B, Chen L, Zeng J, et al. Ezrin Regulating the Cytoskeleton Remodeling is Required for Hypoxia-Induced Myofibroblast Proliferation and Migration. *Front Cardiovasc Med.* 2015;2:10.
53. Zeng H, Hou Y, Zhou X, et al. Cancer-associated fibroblasts facilitate premetastatic niche formation through lncRNA SNHG5-mediated angiogenesis and vascular permeability in breast cancer. *Theranostics.* 2022;12(17):7351-7370.
54. Li P, Zhang H, Chen T, Zhou Y, Yang J, Zhou J. Cancer-associated fibroblasts promote proliferation, angiogenesis, metastasis and immunosuppression in gastric cancer. *Matrix Biol.* Sep 2024;132:59-71.
55. Li S, Chen K, Zhang Y, et al. Twist2 amplification in rhabdomyosarcoma represses myogenesis and promotes oncogenesis by redirecting MyoD DNA binding. *Genes Dev.* Jun 1 2019;33(11-12):626-640.
56. Pan J, Zhang L, Huang J. Prostaglandin D2 synthase/prostaglandin D2/TWIST2 signaling inhibits breast cancer proliferation. *Anticancer Drugs.* Nov 1 2021;32(10):1029-1037.
57. Wang B, Ma Q, Wang X, Guo K, Liu Z, Li G. TGIF1 overexpression promotes glioma progression and worsens patient prognosis. *Cancer Med.* Dec 2022;11(24):5113-5128.
58. Yang J, Zhang L, Jiang Z, et al. TCF12 promotes the tumorigenesis and metastasis of hepatocellular carcinoma via upregulation of CXCR4 expression. *Theranostics.* 2019;9(20):5810-5827.

59. Wang G, Wu Y, Su Y, et al. TCF12-regulated GRB7 facilitates the HER2+ breast cancer progression by activating Notch1 signaling pathway. *J Transl Med.* Aug 7 2024;22(1):745.
60. Chen X, Hao Y, Liu Y, et al. NAT10/ac4C/FOXP1 Promotes Malignant Progression and Facilitates Immunosuppression by Reprogramming Glycolytic Metabolism in Cervical Cancer. *Adv Sci (Weinh).* Nov 2023;10(32):e2302705.
61. Tao M, Ruan X, Yu J, et al. Enhancer-mediated NR2F2 recruitment activates BGN to promote tumor growth and shape tumor microenvironment in papillary thyroid cancer. *Theranostics.* 2026;16(1):298-324.
62. Sundqvist A, Morikawa M, Ren J, et al. JUNB governs a feed-forward network of TGFbeta signaling that aggravates breast cancer invasion. *Nucleic Acids Res.* Feb 16 2018;46(3):1180-1195.
63. Potente M, Makinen T. Vascular heterogeneity and specialization in development and disease. *Nat Rev Mol Cell Biol.* Aug 2017;18(8):477-494.
64. Jain RK, di Tomaso E, Duda DG, Loeffler JS, Sorensen AG, Batchelor TT. Angiogenesis in brain tumours. *Nat Rev Neurosci.* Aug 2007;8(8):610-622.
65. Pan X, Li X, Dong L, et al. Tumour vasculature at single-cell resolution. *Nature.* Aug 2024;632(8024):429-436.
66. Zhang J, Ouyang F, Gao A, et al. ESM1 enhances fatty acid synthesis and vascular mimicry in ovarian cancer by utilizing the PKM2-dependent warburg effect within the hypoxic tumor microenvironment. *Mol Cancer.* May 8 2024;23(1):94.
67. Zhang Y, Manouchehri Doulabi E, Herre M, et al. Platelet-Derived PDGFB Promotes Recruitment of Cancer-Associated Fibroblasts, Deposition of Extracellular Matrix and Tgfbeta Signaling in the Tumor Microenvironment. *Cancers (Basel).* Apr 12 2022;14(8).
68. Niu X, Yin L, Yang X, et al. Serum amyloid A 1 induces suppressive neutrophils through the Toll-like receptor 2-mediated signaling pathway to promote progression of breast cancer. *Cancer Sci.* Apr 2022;113(4):1140-1153.
69. Gao A, Zhang L, Chen X, et al. Effect of VTCN1 on progression and metastasis of ovarian carcinoma in vitro and vivo. *Biomed Pharmacother.* Jul 2015;73:129-134.
70. Podojil JR, Miller SD. Potential targeting of B7-H4 for the treatment of cancer. *Immunol Rev.* Mar 2017;276(1):40-51.
71. Kim M, Choi HY, Woo JW, Chung YR, Park SY. Role of CXCL10 in the progression of in situ to invasive carcinoma of the breast. *Sci Rep.* Sep 9 2021;11(1):18007.
72. Sethuraman A, Brown M, Krutilina R, et al. BHLHE40 confers a pro-survival and pro-metastatic phenotype to breast cancer cells by modulating HBEGF secretion. *Breast Cancer Res.* Oct 1 2018;20(1):117.
73. Ali S, Rehman MU, Yatoo AM, et al. TGF-beta signaling pathway: Therapeutic targeting and potential for anti-cancer immunity. *Eur J Pharmacol.* May 15 2023;947:175678.

74. Liu S, Ren J, Ten Dijke P. Targeting TGFbeta signal transduction for cancer therapy. *Signal Transduct Target Ther.* Jan 8 2021;6(1):8.
75. Fukumura D, Kashiwagi S, Jain RK. The role of nitric oxide in tumour progression. *Nat Rev Cancer.* Jul 2006;6(7):521-534.
76. Lala PK, Chakraborty C. Role of nitric oxide in carcinogenesis and tumour progression. *Lancet Oncol.* Mar 2001;2(3):149-156.
77. Liu K, Han H, Xiong K, et al. Single-cell landscape of intratumoral heterogeneity and tumor microenvironment remolding in pre-nodal metastases of breast cancer. *J Transl Med.* Aug 29 2024;22(1):804.
78. Wagner J, Rapsomaniki MA, Chevrier S, et al. A Single-Cell Atlas of the Tumor and Immune Ecosystem of Human Breast Cancer. *Cell.* May 16 2019;177(5):1330-1345 e1318.
79. Kaur J, Adhikari M, Sabol HM, et al. Single-Cell Transcriptomic Analysis Identifies Senescent Osteocytes That Trigger Bone Destruction in Breast Cancer Metastasis. *Cancer Res.* Dec 2 2024;84(23):3936-3952.
80. Winkler J, Tan W, Diadhiou CM, et al. Single-cell analysis of breast cancer metastasis reveals epithelial-mesenchymal plasticity signatures associated with poor outcomes. *J Clin Invest.* Sep 3 2024;134(17).
81. Xu K, Wang R, Chen Q, et al. Microenvironment components and spatially resolved single-cell transcriptome atlas of breast cancer metastatic axillary lymph nodes. *Acta Biochim Biophys Sin (Shanghai).* Sep 25 2022;54(9):1336-1348.
82. Camargo S, Moskowitz O, Giladi A, et al. Neutrophils physically interact with tumor cells to form a signaling niche promoting breast cancer aggressiveness. *Nat Cancer.* Mar 2025;6(3):540-558.
83. Ding S, Qiao N, Zhu Q, et al. Single-cell atlas reveals a distinct immune profile fostered by T cell-B cell crosstalk in triple negative breast cancer. *Cancer Commun (Lond).* Jun 2023;43(6):661-684.
84. Hu Q, Hong Y, Qi P, et al. Atlas of breast cancer infiltrated B-lymphocytes revealed by paired single-cell RNA-sequencing and antigen receptor profiling. *Nat Commun.* Apr 12 2021;12(1):2186.
85. Azizi E, Carr AJ, Plitas G, et al. Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment. *Cell.* Aug 23 2018;174(5):1293-1308 e1236.
86. Liu SQ, Gao ZJ, Wu J, et al. Single-cell and spatially resolved analysis uncovers cell heterogeneity of breast cancer. *J Hematol Oncol.* Mar 3 2022;15(1):19.
87. Deng R, Zhang HL, Huang JH, et al. MAPK1/3 kinase-dependent ULK1 degradation attenuates mitophagy and promotes breast cancer bone metastasis. *Autophagy.* Oct 2021;17(10):3011-3029.
88. Gilkes DM. Implications of Hypoxia in Breast Cancer Metastasis to Bone. *Int J Mol Sci.* Sep 30 2016;17(10).
89. Xu W, Liang T, Fang H, et al. Single-Cell RNA Sequencing Identifies MMP11(+) Cancer-Associated Fibroblasts as Drivers of Angiogenesis and Bladder Cancer Progression. *Adv Sci (Weinh).* Aug 2025;12(32):e02774.

90. Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* Feb 6 2018;19(1):15.
91. Yang S, Corbett SE, Koga Y, et al. Decontamination of ambient RNA in single-cell RNA-seq with DecontX. *Genome Biol.* Mar 5 2020;21(1):57.
92. Hu C, Li T, Xu Y, et al. CellMarker 2.0: an updated database of manually curated cell markers in human/mouse and web tools based on scRNA-seq data. *Nucleic Acids Res.* Jan 6 2023;51(D1):D870-D876.
93. Franzen O, Gan LM, Bjorkegren JLM. PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data. *Database (Oxford).* Jan 1 2019;2019.
94. Tirosh I, Izar B, Prakadan SM, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science.* Apr 8 2016;352(6282):189-196.
95. Yuan H, Yan M, Zhang G, et al. CancerSEA: a cancer single-cell state atlas. *Nucleic Acids Res.* Jan 8 2019;47(D1):D900-D908.
96. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS.* May 2012;16(5):284-287.
97. Kerseviciute I, Gordevicius J. aPEAR: an R package for autonomous visualization of pathway enrichment networks. *Bioinformatics.* Nov 1 2023;39(11).
98. Qiu X, Mao Q, Tang Y, et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods.* Oct 2017;14(10):979-982.
99. Aibar S, Gonzalez-Blas CB, Moerman T, et al. SCENIC: single-cell regulatory network inference and clustering. *Nat Methods.* Nov 2017;14(11):1083-1086.
100. Efremova M, Vento-Tormo M, Teichmann SA, Vento-Tormo R. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat Protoc.* Apr 2020;15(4):1484-1506.
101. Browaeys R, Saelens W, Saeys Y. NicheNet: modeling intercellular communication by linking ligands to target genes. *Nat Methods.* Feb 2020;17(2):159-162.

Figures

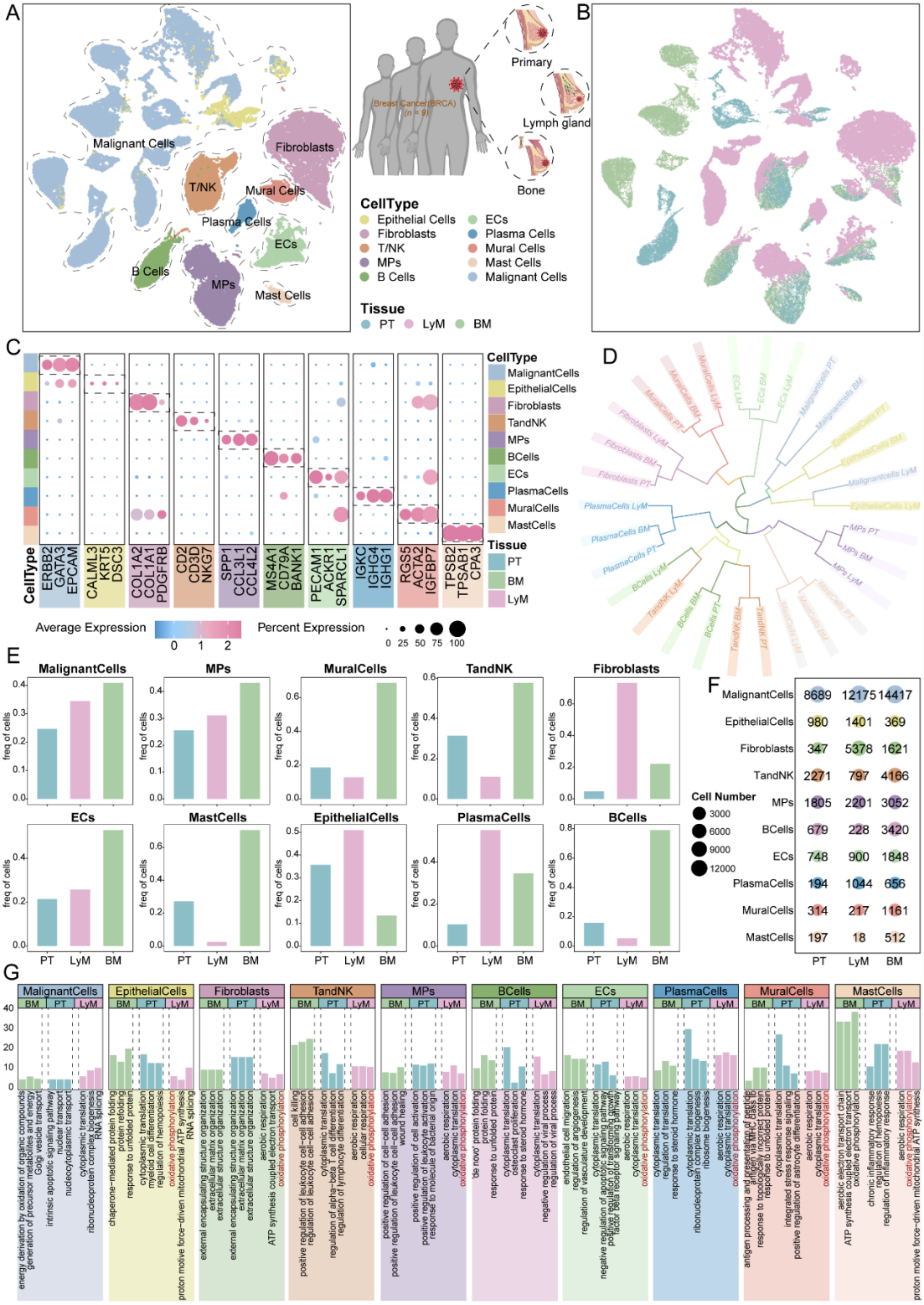


Figure 1

Single-cell map of different metastatic directions of breast cancer. A. UMAP plot of 71,805 cells from breast cancer tumor tissue shows 10 cell lineages. Each cell lineage is represented by a different color. **B.** UMAP plot shows cells from different tumor sources. **C.** Dotplot shows the average expression of known markers in the specified cell clusters. Size of the points represents the percentage of cells expressing each gene in the cluster. Expression intensity of the markers is as shown in the figure. **D.**

Hierarchical clustering analysis of various cell types from different tumor sources. **E.** Bar plot shows the relative frequencies of each cell type in different tumor sources. **F.** Number of cells of each cell type in different tumor sources. **G.** Differential biological process (BP) enrichment analysis of cell types from different tumor sources.

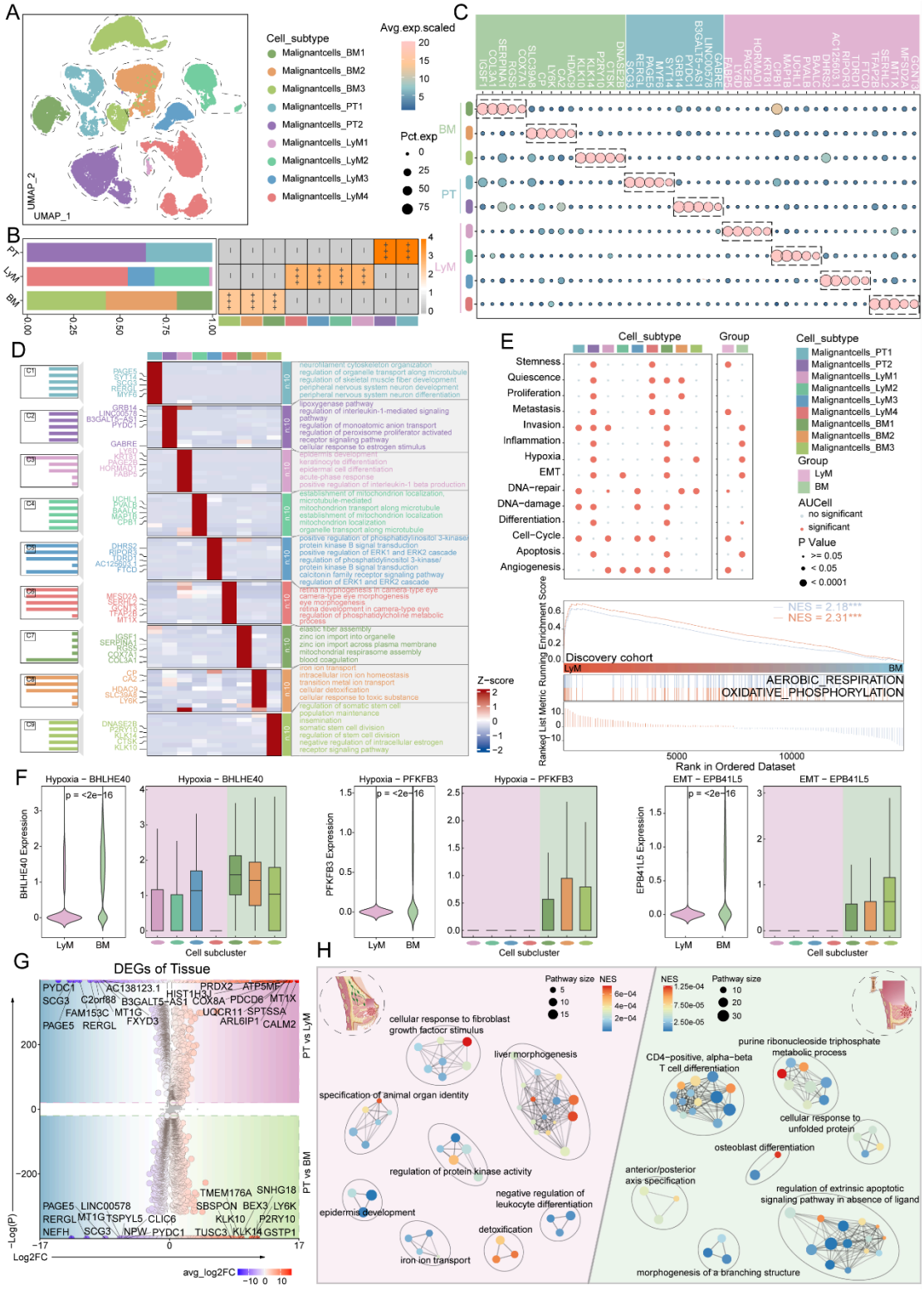


Figure 2

Molecular and functional heterogeneity of tumor-specific cancer cells. A. UMAP plot of breast cancer cells, colored according to different sample origins. **B.** Proportion of cancer cell clusters in different metastatic directions and the Ro/e index. **C.** Dotplot shows the top 5 characteristic genes for each cell cluster. The size of each dot represents the percentage of cells expressing the gene within that cluster. The average expression intensity of each marker is shown in the figure. **D.** Composite heatmap displays the top 10 characteristic genes and top 5 GO functional enrichment terms for each cancer cell subtype. Bar charts represent functional enrichment counts, with subtype types distinguished by different colors. **E.** Dotplot displays AUcell scores for 14 functional states of CancerSEA cancer cells across all tumor cell subtypes and sources (Lymph node metastasis, bone metastasis). Point size indicates P-value, while point color denotes statistical significance (top panel). GSEA identified two major enriched pathways for differentially expressed genes between lymph node metastasis-derived and bone metastasis-derived cancer cells (bottom panel). **F.** Differential expression of hypoxia-related and EMT-related genes in lymph nodes and bone metastasis directions (left) and in cell subclusters (right); Wilcoxon signed-rank test. **G.** Volcano plot of differential expression in ccancer cells with different metastatic directions. **H.** The nodes in the figure represent the top 50 biological processes (BP) significantly enriched by differentially expressed genes from cancer cells of different tumor origins. These nodes are clustered based on similarities among pathway gene sets.

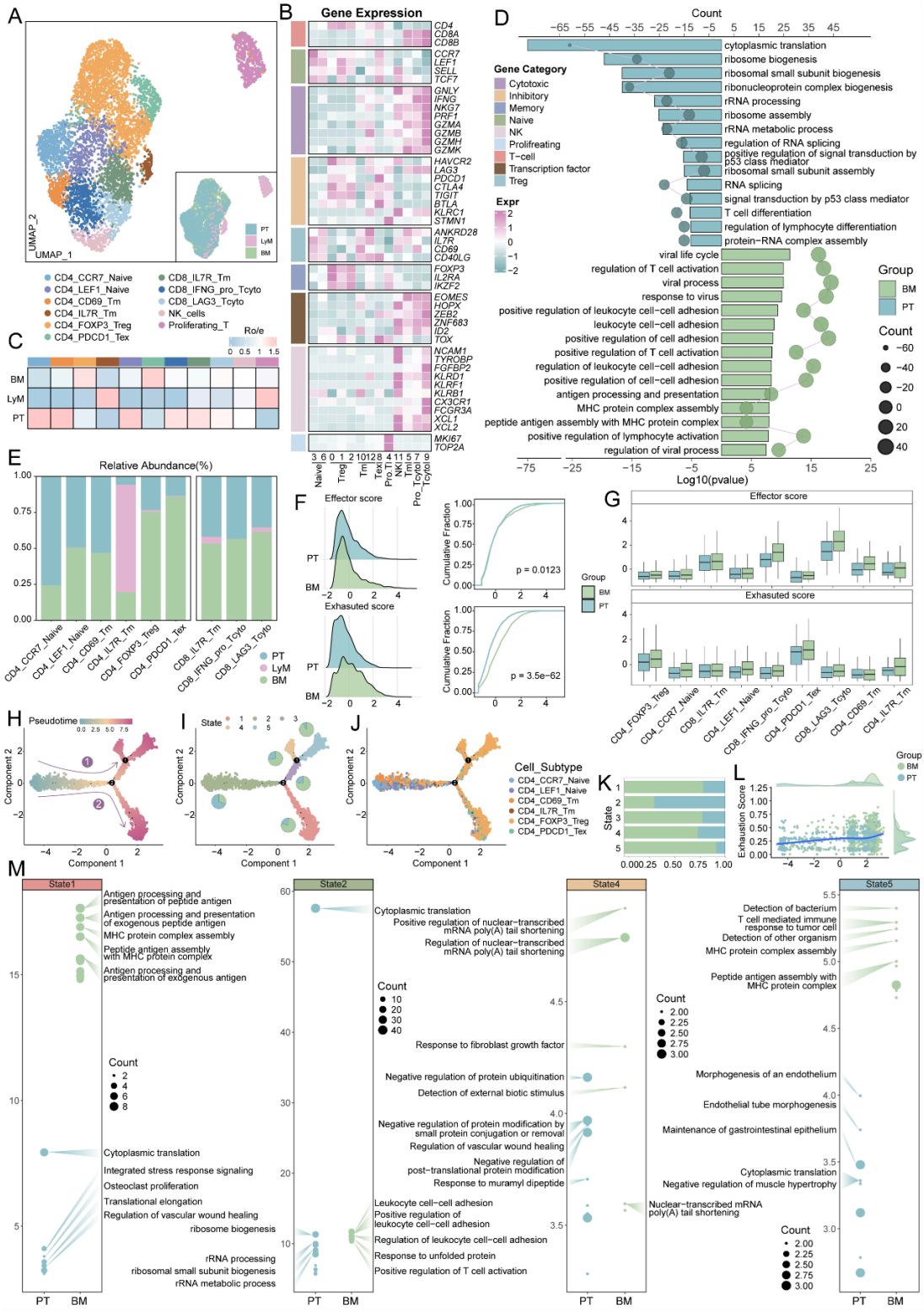


Figure 3

Molecular characteristics and cellular proportions of T cell subtypes. A. UMAP plots of T cells and NK cells, colored according to different cell subtypes and metastatic directions. **B.** Heatmap of typical lineage marker genes for T cells and NK cells. Displays the average gene expression heatmap of selected typical markers for cell subtypes within each cluster. Centered on the average expression across all single cells, with a range from 2 to −2. **C.** Ro/e Index of T cells and NK cells in different

metastatic directions. **D.** The grouped bar plot displays the top 15 GO pathways enriched in differentially expressed genes between the non-metastatic and bone metastasis groups. The size of each dot indicates the number of genes. **E.** Proportion of CD4+ and CD8+ T cells in different metastatic directions. **F.** Ridgeline plots (left) and cumulative distribution functions (right) showing the distributions of effector (top) and exhausted (bottom) state scores between non-metastatic and bone-metastatic samples. **G.** Distributions of effector (top) and exhausted (bottom) state scores between non-metastatic and bone-metastatic samples. **H.** Investigating CD4+ T cell subtypes using semi-supervised pseudo-time trajectories with Monocle2. Trajectories are colored by pseudotime. **I.** Investigating CD4+ T cell subtypes using semi-supervised pseudo-time trajectories with Monocle2. Trajectories are colored by cell state. **J.** Investigating CD4+ T cell subtypes using semi-supervised pseudo-time trajectories with Monocle2. Trajectories are colored by cell subtype. **K.** Bar plot displays the proportion of metastatic directions in different states. **L.** Correlation between monocle components and T cell functional exhaustion. Each point represents an individual cell colored by its metastatic direction (left). The blue solid line and density plot represent the LOESS fit and density of each cell subtype, respectively (top and right). **M.** GO pathways for four cell states in non-metastatic and bone metastatic conditions, with dot size indicating gene count.

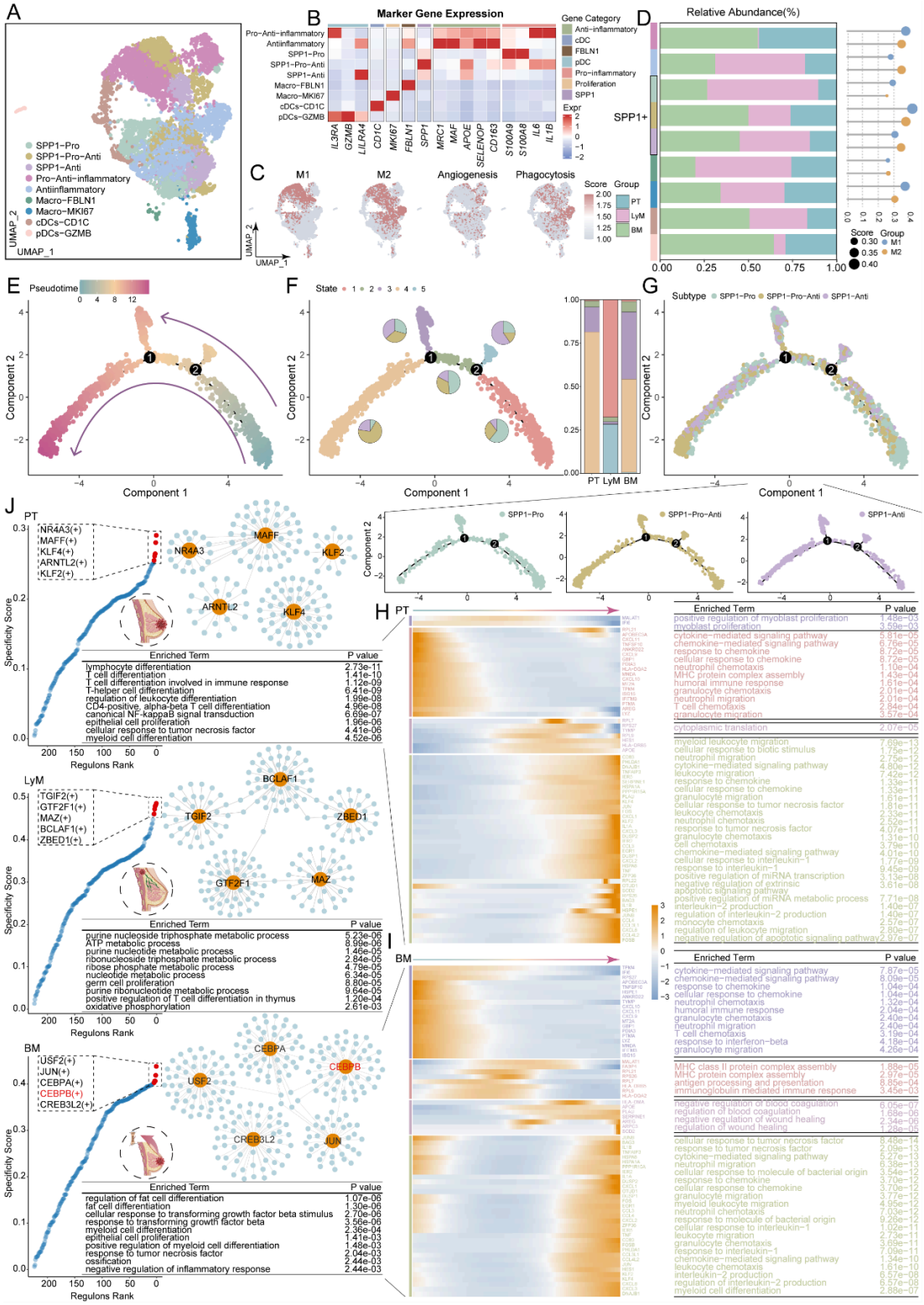


Figure 4

Molecular characteristics and evolutionary trajectories of myeloid cells. A. UMAP plot of myeloid cells, colored according to different cell subtypes and metastatic directions. **B.** Heatmap of typical lineage marker genes for macrophages and dendritic cells. Displays the average gene expression heatmap for selected typical markers across cell subtypes within each cluster. Centered on the average expression across all single cells, with a range from 2 to −2. **C.** Functional characteristic score of myeloid cells. **D.**

Proportions of myeloid cell subtypes across different metastatic directions. Comparison of inflammatory and non-inflammatory feature levels across each cluster indicates pro-inflammatory or anti-inflammatory macrophage subtypes. **E.** Investigating SPP1+ macrophage subtypes using Monocle2 with semi-supervised pseudo-time trajectories. Trajectories are colored by pseudotime. **F.** Investigating SPP1+ macrophage subtypes using semi-supervised pseudo-time trajectories with Monocle2. Trajectories are colored by cell state. Bar charts show the proportion of cells in different states across various metastatic directions. **G.** Investigating SPP1+ macrophage subtypes using semi-supervised pseudo-time trajectories with Monocle2. Trajectories are colored by cell subtype. **H.** Heatmap showing genes associated with the evolutionary trajectory from SPP1+ pro-inflammatory to anti-inflammatory macrophages in non-metastasis, along with corresponding GO terms. **I.** Heatmap showing genes associated with the evolutionary trajectory from SPP1+ pro-inflammatory to anti-inflammatory macrophages in bone metastasis, along with corresponding GO terms. **J.** Top 5 regulators specific to SPP1+ macrophages across different metastatic directions (left), target genes regulated by top 5 transcription factors (upper right), and GO terms enriched in target genes (lower right).

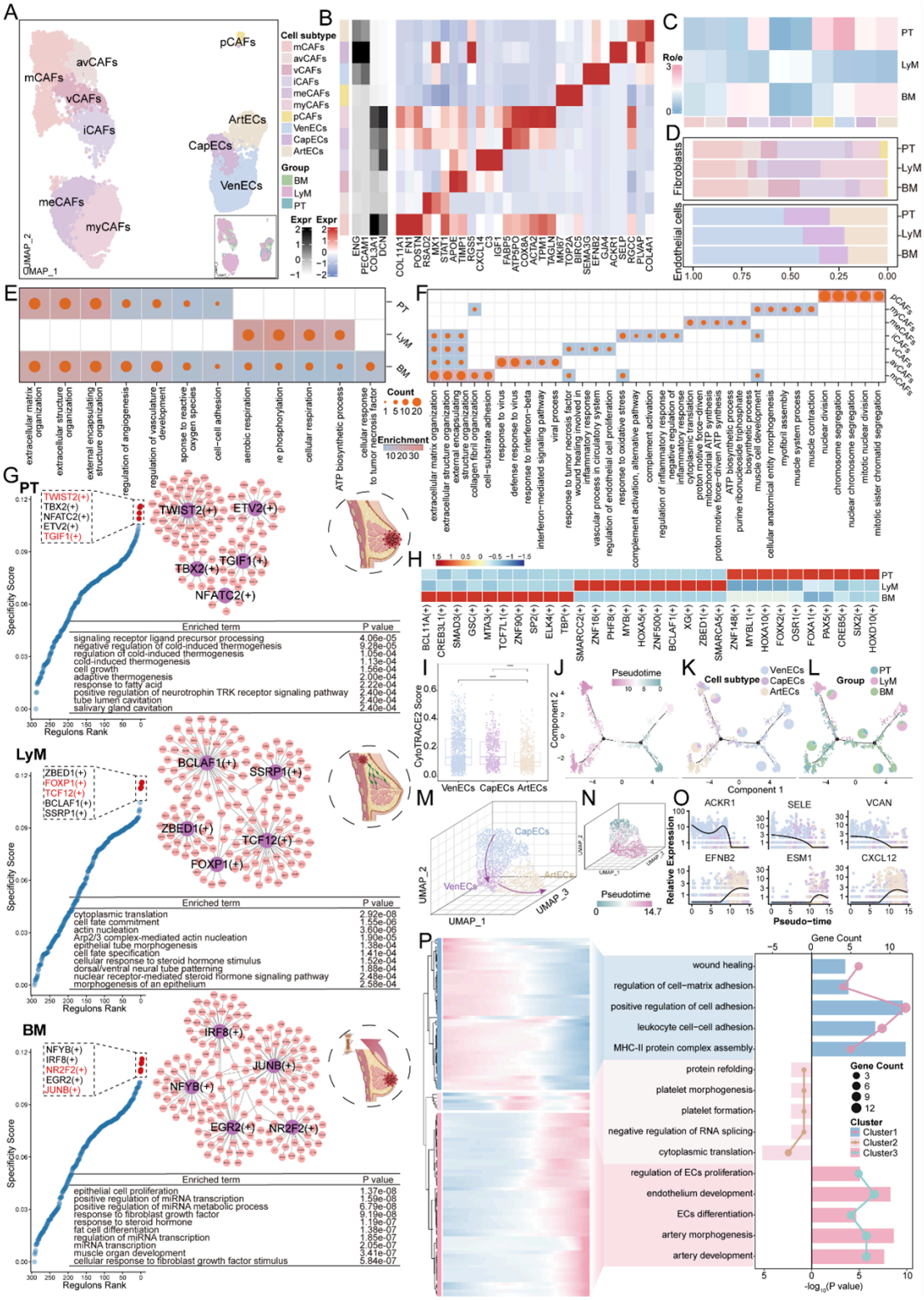


Figure 5

Molecular characteristics and evolutionary trajectories of stromal cells. A. UMAP plot of fibroblasts and endothelial cells, colored according to different cell subtypes and metastatic directions. **B.** Heatmap of typical lineage marker genes for fibroblasts and endothelial cells. Displays the average gene expression heatmap for selected typical markers across cell subtypes within each cluster. Centered on the average expression across all single cells, with a range from 2 to −2. **C.**Ro/e Index in fibroblasts and endothelial

cells across different metastatic directions. **D.** Proportions of fibroblasts to endothelial cells in different metastatic directions. **E.** GO terms enriched for characteristic genes of fibroblasts with different metastatic directions. **F.** GO terms enriched with characteristic genes of different fibroblast subtypes. **G.** Top 5 fibroblast-specific regulators across different metastatic directions (left), target genes regulated by top5 transcription factors (upper right), and GO terms enriched in target genes (lower right). **H.** Transcription factor activity in fibroblasts of different metastatic directions. **I.** Analysis of the differentiation potential of endothelial cell subtypes. **J.**Investigating endothelial cell subtypes using semi-supervised pseudo-time trajectories with Monocle2. Trajectories are colored by pseudotime. **K.** Investigating endothelial cell subtypes using semi-supervised pseudo-time trajectories with Monocle2. Trajectories are colored by cell subtypes. **L.** Investigating endothelial cell subtypes using semi-supervised pseudo-time trajectories with Monocle2. Trajectories are colored by cell state. **M.** 3D UMAP visualization of endothelial cell subtypes, colored according to distinct cell subtypes. **N.** 3D UMAP visualization of endothelial cell subtypes, colored according to different metastatic directions. **O.** Gene expression varies with pseudo-time values, Pseudo-time (x-axis), Gene expression (y-axis). **P.** Heatmap displays genes associated with the evolutionary trajectory from venous endothelial cells to arterial endothelial cells and their corresponding GO terms.

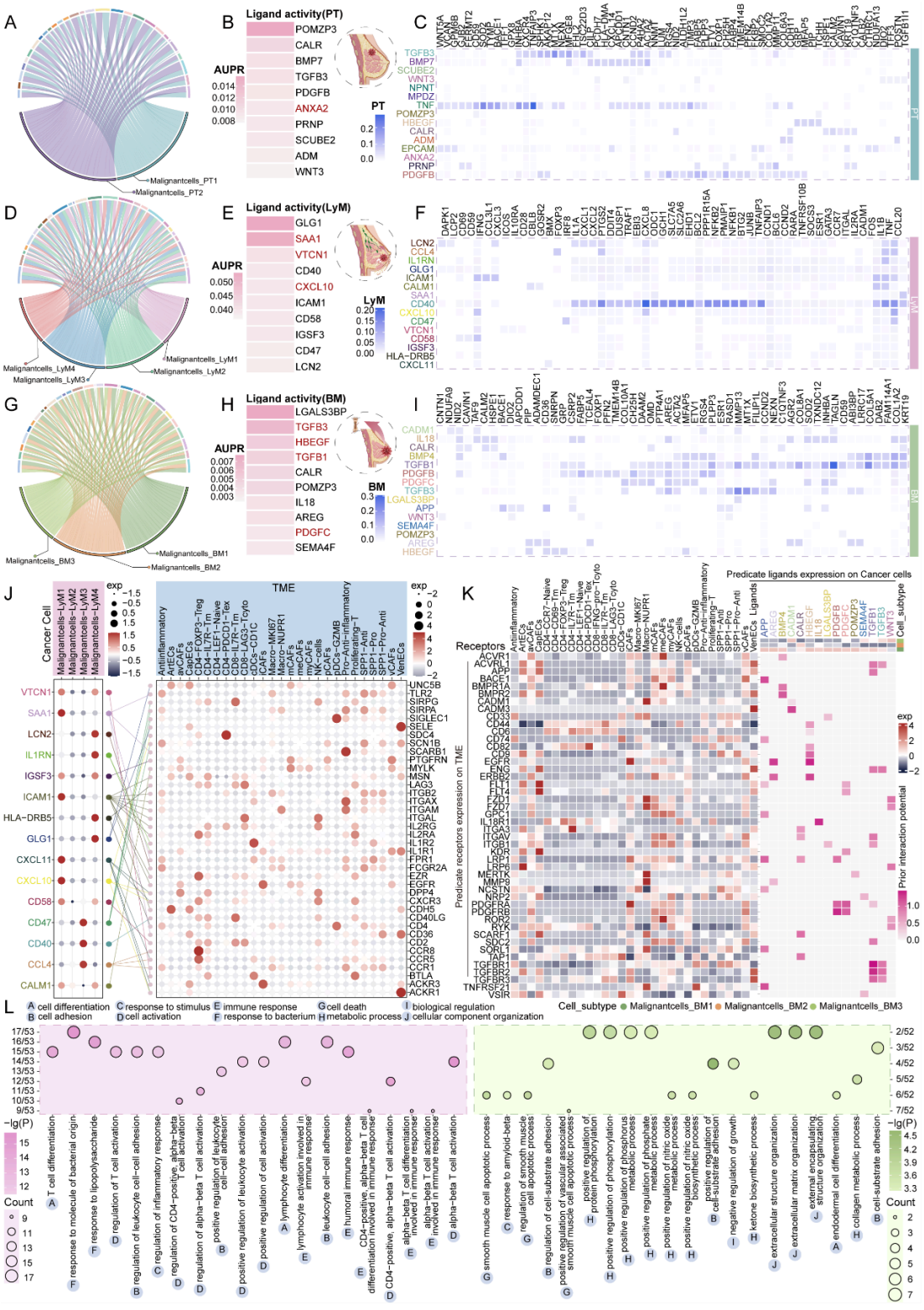


Figure 6

Tumor-specific microenvironments characterized by cell interactions in different metastatic directions. **A.** String diagrams reveal intercellular interactions between clusters of non-metastasis cancer cells and subtypes of stromal cells and immune cells. **B.** Top 10 ligands inferred by regulating non-cancer cells through cancer cells in non-transferred samples, according to NicheNet. **C.** Thermogram reveals the regulatory potential of downstream target genes in clusters of non-metastatic cancer cells. **D.** String

diagrams reveal intercellular interactions between lymph node metastasis cancer cell clusters and subtypes of stromal and immune cells. **E.** Top 10 ligands inferred from regulating non-cancer cells by cancer cells in lymph node metastasis samples, according to NicheNet. **F.** Thermogram reveals the regulatory potential of downstream target genes within clusters of metastatic cancer cells in lymph nodes. **G.** String diagrams reveal intercellular interactions between bone metastasis cancer cell clusters and subtypes of stromal cells and immune cells. **H.** Top 10 ligands inferred by regulating non-cancer cells through cancer cells in bone metastasis samples, according to NicheNet. **I.** Thermogram reveals regulatory potential of downstream target genes within bone metastasis cancer cell clusters. **J.**NicheNet predicts ligands regulating non-cancerous cell expression by cancer cells in lymph node metastasis samples (left). Heatmap of receptor expression associated with different TME subtypes (right). Dot size indicates expression percentage, and dot color intensity reflects expression strength. Key ligand-receptor pairs between tumor cells and TME cells (center). **K.** NicheNet predicts ligands regulating non-cancerous cell expression by cancer cells in bone metastasis samples (top). Heatmap of receptor expression associated with different TME subtypes (left). Heatmap of key ligand-receptor pairs between tumor cells and TME cells (center). **L.** Bubble plot displaying GO functional enrichment results (top 20) for target genes interacting between lymph node metastasis and bone metastasis cells. Left panel: Lymph node metastasis; Right panel: Bone metastasis. Letters denote major functional description categories.

Supplementary Files

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

- GraphicalAbstractImage.tif
- SupplementaryFigures.docx
- SupplementaryTableS1.xlsx
- SupplementaryTableS2.xlsx
- SupplementaryTableS3.xlsx
- SupplementaryTableS4.xlsx
- SupplementaryTableS5.xlsx
- SupplementaryTableS6.xlsx