

Single-cell RNA sequencing reveals immunosuppressive pathways associated with metastatic breast cancer

Furkan Ozmen^{1†*}, Tugba Y. Ozmen^{1†}, Aysegul Ors^{2†}, Mahnaz Janghorban¹, Matthew J. Rames¹, Xi Li¹, Fariba Behbod³, Gordon B. Mills¹, Hisham Mohammed²

¹Division of Oncological Sciences Knight Cancer Institute, Oregon Health & Science University, Portland, Oregon 97201, USA

²Cancer Early Detection Advanced Research Center, Knight Cancer Institute, Oregon Health & Science University, Portland, Oregon 97201, USA,

³University of Kansas Medical Center, Kansas City, KS

†These authors contributed equally to this work.

* Correspondence: ozmen@ohsu.edu

Abstract

Metastatic breast cancer remains largely incurable, and the mechanisms underlying the transition from primary to metastatic breast cancer remain elusive. We analyzed the complex landscape of primary and metastatic breast cancer using scRNA-seq data from twenty-three female patients with either primary or metastatic disease to elucidate the genetic and molecular mechanisms underlying changes in the metastatic tumor ecosystem. We identify specific subtypes of stromal and immune cells critical to forming a pro-tumor microenvironment in metastatic lesions, including CCL2⁺ macrophages, cytotoxic T cells with an exhausted gene signature, and FOXP3⁺ regulatory T cells. Analysis of cell-cell communication highlights a marked decrease in tumor-immune cell interactions in metastatic tissues, likely strengthening the immunosuppressive microenvironment. In contrast, primary breast cancer samples displayed increased activation of the TNF- α signaling pathway via NF- κ B, indicating a potential therapeutic target. Our study comprehensively characterizes the transcriptional landscape encompassing primary and metastatic breast cancer.

Introduction

Breast cancer, which remains the most prevalent cancer in women, is a diverse and complex disease with a wide range of clinical manifestations and outcomes. The shift from an early localized primary tumor to metastatic lesions in distant organs represents a pivotal moment in the clinical course and prognosis of the disease. Despite advances in early detection and treatment, progression to metastatic disease continues to pose a significant clinical challenge with an unfavorable prognosis¹. Individuals diagnosed with

localized breast cancer typically exhibit an overall survival rate exceeding 90%². Conversely, the prognosis drastically declines when cancer progresses to distant metastasis, with survival rates plummeting to around 25%³. Therefore, understanding the complex mechanisms and differences between primary and metastatic breast cancer is essential for informing treatment approaches and improving patient outcomes.

Tumor metastasis requires a complex, orchestrated cascade involving the inherent characteristics of tumor cells, such as genetic mutations, and the intricate interplay between cancer cells and different cellular elements within the tumor microenvironment (TME). This dynamic interaction encompasses a range of participants, including immune cells, tumor-associated macrophages (TAMs) and lymphoid cells, cancer-associated fibroblasts (CAFs), and components of the extracellular matrix (ECM)⁴⁻⁶. Understanding the complex interactions between different cellular components in the TME is crucial for comprehending the mechanisms of tumor initiation, progression to metastasis, and prognosis. Studies have uncovered mutational and transcriptional signatures that are more frequent in breast cancer metastases using bulk genomic sequencing methods^{7,8}. However, genetic signatures arising from bulk sequencing cannot decipher the sources of observed differences or the dynamic interplay between the cell types shaping the metastatic microenvironment in advanced breast cancer.

Single-cell RNA sequencing (scRNA-seq) can reveal the distinct transcriptional profiles of individual malignant and non-malignant cells in the tumor ecosystem. This has enabled analysis of complex intra-tumoral heterogeneity among TME interactions in BCs, such as triple-negative breast cancer⁹. In particular, previous studies have suggested that FOXP3⁺ regulatory T cells (Tregs) in breast cancer may lead to immune tolerance and poorer overall survival¹⁰, while cytotoxic T cells with exhausted gene expression patterns might characterize an immunosuppressive TME¹¹. While immune deregulations are undeniably a core component of the transition to metastatic disease, comprehensive transcriptomic profiling comparing the primary and metastatic breast cancer TME at single-cell resolution has only been applied to a limited number of cases and metastatic sites¹².

Here, we conduct scRNA-seq analysis to deconvolve the TME landscape in primary and metastatic ER⁺ breast cancer. We elucidate distinct gene expression profiles of tumor cells in primary and metastatic breast cancer while identifying specific subtypes of stromal and immune cells that may collectively contribute to developing an immunosuppressive microenvironment within metastatic tumors. Our study provides an overview of the underlying functional landscape of primary and metastatic breast cancer cells while shedding light on the heterogeneity and transcriptomic TME patterns underpinning disease progression.

Results

Landscape of primary and metastatic breast cancer via scRNA-seq

Despite the limited survival of patients with metastatic breast cancer, little is known about the evolution that occurs between normal and malignant cells in the tumor ecosystem of primary and metastatic breast cancer. To investigate this issue, scRNA-seq was performed on an all-female patient cohort comprising individuals diagnosed with either primary (n=12) or metastatic ER+ breast cancer (n=11). Multiple metastatic sites, including the liver, bone, lymph nodes, mediastinum, adrenal gland, and skin, were sampled (Fig. 1a). All patients were classified as estrogen receptor-positive (ER+) based on IHC analysis (Supp. Table 1).

After quality control, removal of batch effects, and principal component analysis, 56,384 single cells from primary breast cancer tissues and 42,813 single cells from metastatic breast cancer tissues, a total of 99,197 cells, were visualized using UMAP for downstream analysis. The cells were partitioned into fifty-four clusters, consisting of seven main cell types: malignant cells, myeloid cells, T cells, natural killer (NK) cells, B cells, endothelial cells, and fibroblasts (Fig. 1b, 1c). Each cell type was characterized using established gene expression markers¹³⁻¹⁵ (Fig. 1d). Additional details and corresponding references for the methodologies employed can be found in 'Methods'.

Copy number variation (CNV) profiles were determined using gene expression data and used to identify normal and malignant cells (see Methods). While both primary and metastatic samples exhibited the same main cell types, the proportions of each cell type varied widely between patients (Supp. Fig. 1a, 1b). However, when we explored the differences in minor cell types (subtypes) within each group, we found a clear distinction in the proportion of cellular subtypes associated with primary and metastatic disease (Fig. 1e, 1f, Supp. Fig. 1c). Malignant epithelial cells were present in similar proportions in both primary and metastatic samples. In primary samples, FOLR2 and CXCR3 positive macrophages, which have been associated with a pro-inflammatory phenotype¹⁶⁻¹⁹, were predominantly observed. In contrast, macrophages positive for CCL2 and SPP1, which have been associated with a pro-tumorigenic subtype^{20,21}, were more abundant in metastatic samples. Our observations highlight changes in the TME that are associated with primary and metastatic tumors, potentially underlying the transition toward the metastatic state.

Genomic and phenotypic alterations within malignant cells

To identify cell types that exhibit high levels of gene expression heterogeneity, we performed differential gene expression analysis between patients for each major lineage. We found that malignant cells exhibited the most remarkable diversity of differentially expressed genes (DEGs), indicating pronounced transcriptional dynamics within these cellular populations across varying environmental conditions (Fig. 2a). To validate our malignant cell assignments and further define the tumor phenotype and clonal substructure, we estimated CNV using InferCNV²² and CaSpER²³. T cells were

used as a reference for each condition (primary/metastasis) (Fig. 2b, Supp. Fig. 2a, Supp. Document 1).

CNV analysis revealed substantial copy number alterations in both primary and metastatic disease. Our analysis identified substantial inter-patient differences within the primary or metastatic group. We then compared the overall CNV structure in both groups and identified substantial inter-site differences. Notably, we observed significant variations in CNVs on chromosomes 1, 6, 11, 12, 16, and 17 when comparing primary and metastatic samples. These unique genomic alterations could potentially play a role in facilitating the transition to metastasis²⁴.

To comprehend the distinct clonal structures between primary and metastatic breast cancer patients, we integrated CNVs from all tumor subpopulations per patient. We compared the overall pattern of copy number alterations across chromosomal arms. Our analysis led us to identify the top 25 CNVs within chromosomal arms that were specific to metastatic or primary subclones (Supp. Fig. 2b). In particular, CNVs in specific chromosomal regions, namely, chr7q34-q36, chr2p11-q11, chr16q13-q24, chr11q21-q25, chr12q13, chr7p22, and chr1q21-q44, were more frequent in the metastatic samples. Intriguingly, these regions encompass genes that have previously been associated with progression and aggressiveness of different cancer types, including ARNT, BIRC3, EIF2AK1, EIF2AK2, FANCA, HOXC11, KIAA1549, MSH2, MSH6, and MYCN²⁵⁻³². These genes are associated with various aspects of cancer development and progression, including cell growth, proliferation, metabolism, and survival²⁵⁻³². Elevated expression of several of the genes affected by these chromosomal alterations has also been reported to be associated with poorer survival rates in ER+ breast tumors³³ (Supp. Fig. 3a).

We then calculated CNV scores for each cell (see Methods) using InferCNV, which represents the extent of copy number variations within a cell and reflects genomic instability. We found higher CNV scores in tumor cells from metastatic patients compared to primary breast samples (Fig. 2c). This finding is consistent with previous studies that have linked high CNV scores to poor prognosis in various types of cancer^{34,35}. Besides intertumoral heterogeneity, intratumoral heterogeneity presents another significant challenge in accurately depicting the genomic landscape³⁶. To further investigate intratumoral heterogeneity of copy number alterations, we used the SCEVAN³⁷ algorithm to identify tumor sub-populations with different copy number alterations for each sample. Our observations indicate that metastatic tumors have a higher intratumoral heterogeneity gene expression score (ITHGEX)³⁸ compared to primary tumors, consistent with metastatic breast cancer samples exhibiting higher levels of intratumoral heterogeneity (Fig. 2d). This is further evidenced by the existence of more tumor subclones within metastatic samples (Supp. Fig. 2b, Supp. Table 2).

Gene regulatory signatures across primary and metastatic breast cancer

Next, we aimed to uncover the cellular processes potentially involved in primary and metastatic breast cancer by constructing lineage-specific gene regulatory networks (GRNs) based on transcription factor (TF) activity and their associated targets using SCENIC³⁹. Through GRN analysis, we identified specific regulatory mechanisms in each significant cell type for primary and metastatic breast cancer. Our study revealed distinct expression patterns of the transcription factors associated with breast cancer pathophysiology (Fig. 2e).

Primary breast cancer malignant cells demonstrated higher regulon activities of transcription factors such as ETS2, EPAS1, BATF, NFIL3, TCF7L1, KLF6, MAFF, KLF10, CEBPD, and ATF3. These transcription factors are involved in various cellular functions implicated in tumor development, apoptosis, immune cell differentiation, energy metabolism, and regulation of essential signaling pathways^{40–46}. For instance, ETS2 is known to regulate both tumor initiation and apoptosis⁴⁰, while EPAS1 plays a crucial role in cellular responses to hypoxia⁴¹. BATF controls the differentiation of immune cells⁴², while NFIL3 is involved in energy metabolism and immune cell differentiation^{43,44}. TCF7L1 mediates the Wnt signaling pathway⁴⁵, and KLF6 modulates metabolism, immunity, and oncogenesis⁴⁶.

In metastatic breast cancer, higher regulon activities of transcription factors, including HOXC13, GATA2, IRF9, MLX, CREB3L4, NFATC4, STAT1, HTATIP2, USF1, and CREB3, indicate their potential roles in facilitating cancer metastasis by influencing various biological processes. For instance, HOXC13 and GATA2 have been linked to a poor prognosis and aggressive phenotypes in a number of cancer lineages^{47,48}. IRF9, an interferon response marker, is primarily known for its role in anti-viral immunity and has been linked to tumor growth and metastasis⁴⁹. MLX coordinates lipid storage with metabolic gene expression regulation and is linked to poor prognosis⁵⁰, whereas CREB3L4 is involved in unfolded protein response and has been associated with breast carcinoma progression⁵¹.

The differential activities of specific regulons in malignant cells reveal distinct mechanisms that could potentially be targeted in primary and metastatic tumor states.

Metastatic breast cancer displays an enrichment for more aggressive Integrative Clusters.

We scrutinized gene expression data and classified individual cells using Integrative Clusters (IntClust) derived from the METABRIC study⁵². The classification was based on raw expression data derived from the METABRIC study, using the top two hundred DEGs for each IntClust. To compute scores for each IntClust and classify each malignant cell, we employed the Cluster Independent Annotation (CIA) tool⁵³. We performed a statistical proportion analysis for each cluster to determine whether any cell type was preferentially enriched or depleted for each method.

Malignant cells in primary tumors showed an increased presence of signatures associated with IntClust3, 4ER+, 4ER-, 5, and 10. IntClust3 is known for distinct patterns of chemosensitivity, and IntClust10 is associated with a stable probability of relapse-free cases among ER patients after five years⁵⁴. In contrast, in metastatic malignant cells, we observed significant enrichment of signatures associated with IntClust1, 2, and 9. These three IntClust types are associated with more aggressive tumor behavior⁵²⁻⁵⁴. IntClust1 is characterized by late-recurring ER-positive genomic subgroups⁵⁴, IntClust2 is associated with genomic alterations that contribute to its aggressive behavior⁵⁴, and IntClust9 is distinguished by amplification of the MYC oncogene at 8q24 in 89% of tumors⁵⁴ (Fig. 2f). Furthermore, we observed significant intratumor heterogeneity in the expression of IntClust types suggesting that tumors may contain cells with differing IntClust subtypes. The dominant IntClust types that represent the most frequent cellular subtypes varied for each sample, highlighting the complexity of the underlying cellular landscape (Supp. Fig. 3b).

Primary breast cancer samples displayed increased activation of the TNF- α signaling pathway via NF κ B

TNF- α and NF κ B signaling pathway activities were elevated in malignant cells of primary breast cancer samples compared to those of metastatic samples. In contrast, we observed higher JAK-STAT pathway activity in metastatic samples compared to primary samples using Progeny pathway activity⁵⁵ (Fig. 2g, 2h). We further investigated highly expressed genes from the NF- κ B and JAK-STAT pathways in malignant cells from primary and metastatic samples. We found that many of the JAK-STAT-related genes enriched in metastatic malignant cells were also associated with interferon activity, while malignant cells from primary breast samples were enriched for genes associated positively with chemotaxis or immune regulation such as CCL20 and CXCL2^{56,57} (Supp. Fig. 3c, 3d, 3e). A similar trend was observed at the single-cell level, where malignant cells from primary samples were mostly in a cell state defined by higher TNF- α signaling via NF- κ B, whereas malignant cells from metastatic samples were enriched in either hypoxia or interferon-alpha response cell states.

Given the importance of these pathways in regulating the immune response and altering the tumor microenvironment, we further investigated changes within non-malignant cells to assess potential deregulations in the immune microenvironment.

Macrophage subtypes are polarized toward immunosuppression across metastatic breast cancer

Myeloid cells play pivotal roles in the TME. We identified twenty-five distinct myeloid subclusters in both primary and metastatic cancer types. These include fifteen subtypes of macrophages, six subtypes of dendritic cells (DCs), three subtypes of monocytes,

and a single subtype of mast cells (Fig. 3a). Specific macrophage subtypes, including FOLR2-expressing macrophages, CX3CR-expressing macrophages, and FCN1-expressing monocytes, were notably enriched in primary tumors (Fig. 3b, 3c). FOLR2-positive macrophages in the tumor microenvironment have been observed to have increased communication with CD8+ T cells. This interaction is facilitated by the high expression of CXCL9 and CXCR3, which are closely associated with FOLR2^{18,58}. Importantly, elevated expression of CXCL9 and CX3CR and also elevated FOLR2+ macrophages in tumors has been correlated with a better prognosis in breast cancer patients^{16,18}. Importantly, CX3CR-expressing macrophages play a pivotal role in immune activation and have emerged as a promising target for cancer therapies due to their regulatory role in immune responses^{58,59}. FCN1-positive monocytes have been associated with increased inflammatory function in several other cancer types, indicating a potential role for FCN1-expressing monocytes in modulating immune responses in cancer contexts⁶⁰.

In the metastatic tumors, our study predominantly found an increase in macrophages expressing CCL2, MGP, SPP1, and MMP9. CCL2-expressing macrophages can potentially contribute to facilitating tumor cell invasion and metastasis^{20,61–63}. This is supported by the role of the CCL2/CCR2 axis in TAM development, the direct impact of CCR2 signaling on tumor cell survival/growth and invasion/metastasis, and the association between high CCL2 expression and poor prognosis in cancer patients²⁰. Macrophages expressing MGP have been found to contribute to upregulating pro-tumorigenic factors associated with promoting immunoresistance⁶⁴. SPP1+ tumor-associated macrophages were associated with poor prognosis in various cancers, potentially by contributing to tumor invasion by degrading the basement membrane through MMP expression⁶⁰. Additionally, MMP9 expressing TAMs have been found to significantly contribute to creating a favorable environment for cancer metastasis⁶⁵. These macrophages have been linked to promoting aggressiveness and poor prognosis in various types of cancer^{60,65}.

Using pseudobulk differential gene expression analysis, we observed a distinct partitioning of marker enrichment across myeloid cell populations exemplified by a volcano plot representation of differentially expressed genes. Intriguingly, the top-ranking genes associated with metastasis were predominantly linked to the interferon (IFN) response, underscoring the potential role of this pathway in the metastatic cascade. Indeed, chronic IFN signaling has been associated with an immunosuppressive immune contexture⁶⁶. Conversely, the genes most prominently expressed in primary tumors were correlated with the TNF- α /NF- κ B signaling pathway, suggesting a differential regulatory landscape between the primary and metastatic cellular contexts (Supp. Fig. 4a).

To gain a better understanding of the biological characteristics of myeloid cells, we investigated hallmark pathway activities within these cells. Our analysis revealed a significant increase in TNF- α signaling through the NF κ B pathway in myeloid cells from

primary tumors (Fig. 3d). In contrast, myeloid cells from metastatic tumors displayed pathways associated with oxidative phosphorylation, as well as responses to IFN-alpha and gamma. This suggests a diverse functional landscape across different myeloid cell states, which was also confirmed at the single-cell level (Fig. 3d, Supp. Fig. 4b).

We also utilized the CancerSEA⁶⁷ database to examine cancer-related pathway activity. Our findings revealed a notable increase in invasion and metastasis traits in metastatic breast cancer macrophages, which implies a more aggressive phenotype in these cells, potentially contributing to the dissemination of the disease. In contrast, primary breast cancer myeloid cells showed a general increase in inflammation characteristics, particularly in the macrophage subset. This might suggest an active immune response within the primary tumor microenvironment, which could potentially influence tumor progression and treatment outcomes (Fig. 3e, Supp. Fig. 4c).

To delineate functional mechanisms in the immune microenvironment, we utilized CellChat⁶⁸ to identify alterations in signaling interactions and cell-cell communications within the TME. Using CellChat, we focused on the top 15% predominant cell-cell interactions in both the primary and metastatic groups (Fig. 3f, 3g).

When comparing primary and metastatic datasets, we noticed unique patterns of cell-cell communication. In the primary dataset, there was a greater frequency of signaling from malignant cells, Treg cells, CD4+ T cells, and CD8+ T cells to macrophages, suggesting a strong interaction between these cell types in the primary tumor context. However, in the metastatic dataset, we observed a change in this communication pattern with an increase in signaling from macrophages to Treg cells and CD8+ T cells. This reversal in signaling direction suggests a potential change in the role of macrophages within the metastatic TME, possibly affecting the behavior of Treg and CD8+ T cells. These observations emphasize the dynamic nature of cell-cell interactions within the TME and the importance of understanding these changes in relation to tumor progression and metastasis. Collectively, these observed differences in myeloid/macrophage interactions point to a gain in the immunosuppressive TME within metastatic sites (Fig 3h).

We observed an enrichment of ligands that could potentially mediate cellular interactions in both primary and metastatic cancers (Supp. Fig. 4d, 4e). We noted elevated pro-inflammatory activity interactions, specifically with markers such as CXCL1, CXCL2, and CXCL8 in primary breast cancer. In the case of metastatic breast cancer, our study revealed aberrant interactions involving markers such as FN1, COL1A1, COL1A2, COL6A1, and THY1. These markers are integral to the formation and regulation of the ECM, a critical component of the tumor microenvironment. Aberrant interactions involving these proteins could lead to alterations in the ECM, potentially promoting cancer progression and metastasis^{69,70} (Supp. Fig. 4d, 4e).

Aberrant lymphoid response within metastatic breast cancer TME

To expand upon our prior findings, which highlighted the interactions between immune cells and the TME, we aimed to identify possible alterations in lymphoid subtypes during breast cancer metastasis. We conducted an unsupervised clustering of lymphoid cells, resulting in the identification of twelve T cell, three B cell, and two NK cell subclusters (Fig. 4a) that clearly distinguished the subtypes in both primary and metastatic tumors.

Although all T-cell subtypes were present in both primary and metastatic samples (Supp. Fig. 5a), the relative proportions of certain T-cell subtypes changed significantly from primary to metastatic tumors (Fig. 4b). Primary tumors generally displayed a higher proportion of pro-inflammatory T-cell subtypes, including CD4⁺ central memory T-cells that contribute to the rapid expansion of antigen-specific CD4⁺ T-cells and promote inflammation⁷¹. In contrast, metastatic samples exhibited mixed deregulation of T-cell subtypes, characterized by increases in CD8⁺ exhausted T-cells, CD4⁺ IFI6 T-cells, CD4⁺ Native T-cells, CD8⁺ MKI67⁺ T-cells, and FOXP3⁺ Tregs (Fig. 4b). The IFI6-expressing activated T cell (Tact) subtype is associated with disease progression post-chemotherapy in TNBC patients and correlates with a lower metastasis-free survival rate in breast cancer patients²¹. The upregulation of the IFI6 gene is linked to the control of mitochondrial ROS production, potentially contributing to unfavorable clinical outcomes in breast cancer patients²¹. These associations might also extend to the ER⁺ breast cancers in this study.

In our study, proliferative MKI67⁺ T cells were present exclusively in metastatic ER-positive patients. Studies in patients with triple-negative breast cancer (TNBC) with disease progression after chemotherapy have also shown proliferative MKI67⁺ T-cells (Tprf-MKI67)²¹. Additionally, MKI67 gene expression was significantly correlated with lymph node metastases, tumor invasion, and adverse survival outcomes in TNBC²¹.

In primary tumors, there is a higher proportion of CD16-positive NK cells, whereas in metastatic breast cancer, there is a higher proportion of CD56-positive NK cells (Fig. 4c). CD16-positive NK cells contribute to antibody-dependent cellular cytotoxicity (ADCC), a process where specialized immune cells, such as NK cells, recognize and kill cells coated with antibodies through interaction with CD16 receptors on their surface^{72,73}. Additionally, our data revealed that primary tumors show significant enrichment in ADCC-related gene signatures compared to metastatic samples published previously⁷⁴ (Supp. Fig. 5b).

Regarding B-cell subtypes, we found a decrease in memory B cells and an increase in naive B-cells within the metastatic samples (Fig. 4d). This observation is consistent with the active recruitment of B cells to the TME without prior priming against malignant cells.

Lymphoid metabolic reprogramming results in an immunosuppressive metastatic TME

To further probe lymphoid subtypes, we performed pathway enrichment and also interaction analysis to decipher phenotypic changes. When considering all types of lymphoid cells, T cells in primary tumors show a preference for glycolytic signaling, while T cells in metastatic tumors are more enriched for fatty acid metabolism, reactive oxygen species production, and oxidative phosphorylation (Fig. 4e).

While our data shows higher hypoxic signaling in metastatic malignant cells, in primary samples, we rather observe this enrichment in the non-epithelial cell types. Indeed, we see a collective increase in hypoxic signaling, along with a decrease in oxidative phosphorylation in T cells from primary tumors compared to T cells from metastatic tumors (Fig. 4f, 4g). While low oxygen tension in T cells has been associated with inhibiting their function, recent studies also show that this characteristic can enhance aspects of the adaptive immune response⁷⁵. Additionally, hypoxia during antigen recognition may differentially prime T cells, leading to improved antitumor activity^{76,77}.

Network analysis revealed a dramatic shift in lymphoid cell-related signaling in metastatic tumors (Fig. 4h, 4i). Malignant and lymphoid interactions in primary tumors were dominated by an immunostimulatory regime, which largely signaled CD8+ T effector cells, while metastatic interactions were dominated by substantive signaling into CD8+ T effector memory (TEM) cells. The loss of the CD8+ cytotoxic population with a concomitant increase in CD8+ TEM may indicate a chronic immunosuppressive TME where CD8+ TEMs recognize malignant cells yet cannot effectively mediate tumor clearance^{78,79}.

Our study observed varying levels of immune checkpoint inhibition across different cell types in both primary and metastatic breast cancers. Specifically, Tregs and Th1-like CXCL13+CD4+ T cells showed elevated expression of immune checkpoint transcripts. Also, metastatic lesions displayed a stronger immune checkpoint inhibition signal than primary tumors (Supp. Fig. 5c).

Th1-like CXCL13+ CD4+ T cells are known to interact with Tregs and are involved in downregulating genes associated with TCR signaling, similar to Tregs⁸⁰. We also found that both cell types were more prevalent in metastatic samples (Supp. Fig. 5a). Our findings suggest that ER+ metastatic breast cancer patients may benefit from immune checkpoint inhibitor therapy, which could potentially enhance antitumor immune responses.

Stromal cell remodeling occurs in the TME in metastatic sites

As key modulators of the stromal landscape of TMEs, we then focused on the characterization of fibroblasts. Unsupervised clustering of stromal cells revealed changes between primary and metastatic samples across one endothelial and 10 CAF subtypes (Fig. 5a) with a loss of both antigen-presenting (apCAF) and inflammatory

CAFs (iCAF) coupled with a gain of both matrix CAFs(mCAF) and pericytes (Fig. 5a, 5b, Supp. Fig. 6a) in metastatic tumors.

While CAFs are usually considered to have an immunosuppressive role⁸¹, antigen-presenting cancer-associated fibroblasts can enhance antigen presentation in the tumor microenvironment through MHC-II⁸². In primary breast cancer samples, there were higher levels of inflammatory CAFs, which were identified by high expression of cytokines such as CXCL12, CXCL14, and IL6⁸³. These cytokines are indicators of inflammation and suggest an increased pathway involving PDGF, STAT3, KRAS signaling, and complement activation⁸³.

In metastatic tumors, we observed an increase in matrix cancer-associated fibroblasts, which exhibited gene expression associated with invasion, such as matrix metalloproteinases (MMP), consistent with the alteration of the metastatic breast cancer TME⁸³. Vascular pericytes are also found in a higher proportion in metastatic breast tumors and have previously been found to play a part in cancer invasion and metastasis through several pathways⁸⁴.

As seen with myeloid and lymphoid lineages, stromal cells similarly exhibited higher levels of hypoxic signaling within the primary TME (Fig. 5c), supporting the notion that hypoxic stress shifts from non-tumor cells to tumor cells in the metastatic sites. Our findings indicate that hypoxic signaling varies between these sites, reflecting differing responses to extracellular oxygen levels rather than uniformity across cell types⁸⁵.

Metastatic stromal cells displayed a notable increase in genes associated with angiogenesis and metastasis, indicating their potential involvement in promoting the growth and spread of tumors (Fig. 5c, 5d, 5e). In contrast, fibroblasts from primary TME showed an increase in inflammation and quiescence-related gene signatures.

Additionally, primary breast cancer fibroblasts displayed distinct cellular states at a cellular level, with subtypes including iCAFs within the primary TME predominantly exhibiting a quiescent state (Fig. 5d, 5e).

Next, we investigated the impact of transcription factors on fibroblasts. We observed high regulon activity of DLX5, a member of the MMP family⁸⁶, especially in vascular cancer-associated fibroblasts (Pericytes and vascular CAFs) (Fig. 5f). Notably, we observed significantly higher DLX5 regulon activity in metastatic samples compared to primary tumors (Fig. 5g).

We evaluated communication between stromal cells and other cell types using the CellChat framework (Fig. 5h). Specifically, we found that myeloid cells exhibited heightened intercellular communication with fibroblasts in primary tumors, which aligns with our analysis of communication between malignant and tumor cells. Additionally, we discovered that fibroblasts established communication channels with other cell types and within themselves, particularly in the metastatic TME.

To gain a deeper understanding of the signaling pathways influencing cell-cell interactions in the TME, we conducted network analysis utilizing the MultiNicheNet⁸⁷ approach. Our analysis revealed significant integrin and collagen-mediated interactions among stromal cells, as depicted in (Supp. Fig. 6b, 6c). Integrins are known to play a critical role in cell-cell and cell-extracellular matrix interactions, while collagen is a major component of the extracellular matrix⁸⁸. The presence of these interactions underscores the importance of integrin and collagen-mediated signaling in the communication and behavior of stromal cells within the TME, particularly in metastatic tumors⁸⁹.

Discussion

As primary breast cancers progress and transition to metastatic breast cancer, an imbalance of TME signaling occurs that ultimately leads to poorer outcomes⁹⁰. Unraveling the interactions within the TME is an evolving need, where tools, including single-cell sequencing, can reveal complex interactions between cell types. Several recent studies have used sc-RNAseq to reveal changes among cellular players in TME, including in breast and pancreatic cancers^{12,91} with the breast cancer study being restricted to lymph node metastasis. In this study, we analyzed the cellular composition and signaling pathways in primary breast cancers in comparison to tumors occurring in the most prevalent metastatic sites, including lymph nodes, liver, bone, adrenal gland, and subcutaneous (Fig. 1a). Our differential gene expression analysis did not identify substantial differences between tumor metastatic sites. This could be due to the small sample size analyzed, as well as the possibility that metastasis-specific gene expression patterns have a high degree of conservation across the different sites. With our analysis, we shed light on the core cellular and functional differences between primary and metastatic breast cancer niches.

While the main cell types were consistent between primary and metastatic samples, there were significant differences in the proportions of cellular subtypes within each group. The substantial cellular heterogeneity highlights the need for precise treatment approaches, with the differences between the subtypes potentially indicating therapeutic opportunities.

We observed pronounced transcriptional dynamics within malignant cells, indicating significant diversity in gene expression between primary and metastatic tumors. Our analysis of CNVs revealed substantial inter-patient differences in CNV profiles within both primary and metastatic tumor cell populations, further supporting the presence of intertumoral heterogeneity. We identified specific chromosomal regions that were pronounced in metastatic samples, encompassing genes associated with the progression and aggressiveness of breast cancer²⁵⁻³² (Supp. Fig. 2b). Consistent with previous studies⁹², metastatic tumor cells exhibited a higher average number of clones compared to primary tumor sites.

Additionally, our study examined the classification of malignant cells using Integrative Clusters (IntClust) derived from the METABRIC study⁵². We observed an enrichment of more aggressive IntClust types in metastatic breast cancer, including IntClust1, 2, and 9. These IntClust types have been associated with late-recurring ER-positive genomic subgroups, genomic alterations contributing to aggressive behavior, and amplification of the MYC oncogene, respectively. This enrichment further supports the notion that metastatic breast cancer displays a more aggressive tumor behavior compared to primary breast cancer. Interestingly, single-cell analysis indicated that the primary and metastatic breast cancers contain cells from multiple different IntClust types, with the designated IntClust likely representing the dominant cell type in the tumors. This also suggests that some aspects of the IntClust may be due to cell state differences rather than determined by the mutation status of the cell.

While studying differences in activity of signaling pathways between the two disease states, we found elevated TNF- α via NF κ B signaling pathway activities in primary breast cancer samples compared to metastatic samples. In contrast, metastatic samples showed higher JAK-STAT and IFN pathway activities. The activation of the TNF- α pathway in primary tumors may contribute to the recruitment of immune cells and the production of pro-inflammatory cytokines and chemokines⁹³. Notably, the prominence of inflammatory signaling in primary tumors implies a potential role in initiating key processes conducive to tumor progression, including tumor cell invasion and migration⁹⁴. Specifically, this aberrant signaling may serve as an early driver in the pathogenesis of primary tumors, setting the stage for subsequent metastatic dissemination⁹⁵. The observed loss of NF κ B signaling in metastatic malignant cells may contribute to the establishment of an immunosuppressive microenvironment permissive for growth in the metastatic niche⁹⁶.

Our findings highlight the heterogeneity of myeloid cells, particularly macrophages, in the TME and point to their role in tumor progression and metastasis. In primary tumors, we observed that the enrichments of specific macrophage and monocyte subtypes are linked to better prognosis in cancer. While FOLR2+ macrophages were found to communicate more effectively with CD8+ T cells¹⁶, CXCR3+ macrophages, on the other hand, play a pivotal role in immune activation and have emerged as a promising target in cancer therapies. The diminished expression of CXCR3 in macrophages in metastatic breast cancer may be linked to changes in the immune microenvironment that support tumor survival in the metastatic niche⁵⁹. FCN1+ monocytes have also been associated with increased inflammatory function in other cancer types, suggesting their potential role in modulating immune responses in cancer contexts⁶⁰.

Our study further revealed the presence of macrophages expressing CCL2, MGP, SPP1, and MMP9 in metastatic tumors. As described above, these TAMs have been associated with aspects of tumor metastasis and may prepare metastatic sites for malignant cell growth.

Our examination of marker enrichment and pathway activities within myeloid cells revealed clear differences between primary and metastatic tumors. Primary tumors showed heightened TNF- α signaling through the NF κ B pathway, indicating an active immune response within the primary tumor microenvironment. On the other hand, metastatic myeloid cells exhibited pathways related to oxidative phosphorylation and responses to interferon-alpha and gamma. Acute Type I IFN signaling response is considered a key driver for inflammation⁹⁷. Studies suggest that Type I IFN has a dual role in cancer and chronic inflammation. The Type I IFN response in the Tumor Microenvironment may promote pro-tumorigenic TAM infiltration in metastatic lesions⁹⁸. However, chronic interferon stimulation has been proposed to be immunosuppressive⁶⁶ and may contribute to the immunosuppressive microenvironment observed in metastatic ER+ breast cancers in our study. This diverse functional landscape across different myeloid cell states highlights the dynamic nature of the TME and its potential influence on tumor progression and treatment outcomes.

Furthermore, our investigation of cell-cell interactions within the TME has uncovered significant changes in cellular interactions in metastatic tumors. In the case of metastatic breast cancer, cell-cell interactions predominantly revolved around the formation and regulation of the ECM, a pivotal component of the TME. This suggests a significant shift in the role of macrophages within the metastatic TME, potentially influencing the behavior of other immune cells. These findings underscore the vital significance of understanding the dynamic nature of cell-cell interactions and their impact on tumor progression and metastasis.

Within the lymphoid context, primary tumors exhibited a higher proportion of pro-inflammatory T-cell subtypes, while metastatic samples displayed mixed deregulation of T-cell subtypes, including increases in exhausted T cells, IFI6 T cells, naive T cells, MKI67+ T cells, and Tregs suggesting a shift towards a chronic immunosuppressive tumor microenvironment.

Our study highlights the potential prognostic use of specific T cell subtypes in breast cancer. The presence of IFI6-expressing activated T cells was associated with disease progression post-chemotherapy and lower metastasis-free survival rates²¹. Additionally, the presence of proliferative MKI67+ T cells was exclusive to TNBC patients experiencing disease progression after chemotherapy, and their gene expression correlated with lymph node metastases, tumor invasion, and adverse survival outcomes. Our study suggests that similar processes may be involved in ER+ tumors.

In terms of NK cells, CD16-positive NK cells, which play a role in ADCC, were more prevalent in primary breast cancer. However, in metastatic breast cancer, there was a higher proportion of CD56-positive NK cells. This shift may suggest a loss of NK cell ADCC ability during the metastatic transition, consistent with previous studies^{72,73}.

We observed a decrease in memory B cells and an increase in naive B cells within the metastatic samples. This shift suggests active recruitment of naive B cells to the

metastatic tumor microenvironment without prior priming against malignant cells. These findings provide further evidence of the complex interactions between lymphoid cells and the tumor microenvironment during breast cancer metastasis.

The identification of specific T cell subtypes associated with disease progression and prognosis, as well as the dynamic changes in NK cell and B cell subtypes, provide valuable insights for the development of targeted therapies and immunotherapies in metastatic breast cancer. Further studies are warranted to elucidate the underlying mechanisms driving these alterations in lymphoid subtypes and their functional implications in the tumor microenvironment. We also investigated the metabolic reprogramming of lymphoid cells. We found that T cells in primary tumors showed a preference for glycolytic signaling, while T cells in metastatic tumors were more enriched for fatty acid metabolism, reactive oxygen species production, and oxidative phosphorylation. This indicates a shift in metabolic pathways during breast cancer metastasis. While low oxygen tensions in the T cell microenvironment have been associated with inhibiting their function, recent studies suggest that this characteristic can enhance aspects of the adaptive immune response and improve antitumor activity. Network analysis demonstrated a dramatic shift in lymphoid cell-related signaling upon metastasis. Primary breast cancer lymphoid interactions were dominated by an immunostimulatory regime, primarily signaling CD8+ T effector cells. In contrast, metastatic interactions were dominated by substantive signaling into CD8+ T effector memory cells. This suggests a chronic immunosuppressive TME in metastatic breast cancer, where CD8+ TEMs recognize malignant cells but cannot effectively mediate tumor clearance.

Finally, our study revealed varying levels of immune checkpoint inhibition across different cell types in both primary and metastatic breast cancer. Tregs and Th1-like CXCL13+CD4+ T cells showed elevated expression of immune checkpoint mediators. Metastatic lesions displayed a stronger immune checkpoint inhibition signal than primary tumors. These findings suggest that breast cancer patients with metastatic ER+ tumors may benefit from immune checkpoint inhibitor therapy, which could enhance antitumor immune responses⁹⁹.

It has been established that endothelial and fibroblast cells play crucial roles in the tumor microenvironment. In the context of the TME, fibroblasts, specifically cancer-associated fibroblasts, have been recognized as key modulators of the stromal landscape. They interact with tumor cells and other stromal components, influencing tumor progression, immune response, and therapy resistance. Building upon this knowledge, our study delved deeper into the characterization of fibroblasts within the TME, particularly during metastasis. We observed a loss of apCAF and iCAF, coupled with a gain of mCAF and pericytes in metastatic breast cancer. While CAFs are generally considered to have an immunosuppressive role, apCAFs can enhance antigen presentation in the TME⁸². Notably, we observed higher levels of inflammatory CAFs, identified by high expression of cytokines associated with inflammation in primary

breast cancer samples, indicating their potential involvement in promoting an inflammatory environment conducive to tumor growth. A subset of iCAFs in the primary samples showed an increase in quiescence-related gene signatures. The identification of these quiescent iCAFs suggests that these fibroblasts may have specific functions related to cellular dormancy and inflammation regulation within the primary TME¹⁰⁰. In metastatic breast cancer samples, we observed an increase in matrix CAFs, which exhibited gene expression patterns associated with invasion, such as MMPs. This observation aligns with previous studies highlighting the role of fibroblasts in facilitating tumor invasion and metastasis. Moreover, the increased presence of vascular pericytes in metastatic breast cancer further implicates their contribution to cancer progression through various pathways.

We observed elevated hypoxic signaling in stromal cells within the primary TME, consistent with the increased hypoxia seen in myeloid and lymphoid lineages. Notably, hypoxic signaling levels varied across different cell types, underscoring the critical influence of oxygen availability in shaping the dynamics of the TME. This variation highlights the potential role of cell type-specific responses to hypoxia in driving tumor progression and adaptation.

One of the key findings in our study was the high activity of the DLX5 regulon, particularly in vascular CAFs and pericytes. DLX5 is known to regulate the expression of MMPs, which are enzymes involved in degrading the extracellular matrix and facilitating tumor cell invasion. The upregulation of DLX5 in metastatic stromal cells indicates a potential mechanism by which these cells contribute to the growth and spread of tumors. Overall, the high DLX5 regulon activity observed in vascular CAFs and pericytes in metastatic samples suggests its potential as a therapeutic target⁸⁶. The prominent presence of integrin and collagen-mediated interactions in the metastatic stroma highlights their importance in establishing metastasis (Supp. Fig. 6b, 6c).

In conclusion, our study provides valuable insights into the characteristics and interactions of stromal cells within the TME. The differential hypoxic signaling, gene expression profiles, cellular states, and intercellular communication observed among stromal cells contribute to our understanding of tumor progression and metastasis. These findings may have implications for the development of targeted therapies aimed at disrupting the tumor-stroma interactions in cancer treatment.

Methods

Sample collection and preparation

This study was approved by the Oregon Health & Science University (OHSU) Institutional Review Board (IRB). All biospecimens were collected and analyzed under the OHSU IRB-approved CEDAR (IRB #20750) or MMTERT observational study (Mitri 2018) (IRB #16113).

Participant eligibility was determined by the enrolling physician, and informed written consent was obtained from all subjects. All biopsy biospecimen used in this study were prospectively collected freshly for single-cell RNA sequencing library construction.

Tissue dissociation

Tumor tissues were disaggregated by gentleMACS kits (Miltenyi Biotec, #130-095-929) following the manufacturer's protocol. The lysate was resuspended and filtered through a 70- μ m cell strainer (130-098-462; Miltenyi Biotec Germany). Cells were collected by centrifuging (300 \times g for 7 min at 4 °C) and resuspended at 700–1200 cells/ μ l. Live cells were isolated by EasySep Dead Cell Removal (Annexin V) Kit (STEMCELL Technologies, #17899).

Single-cell RNA sequencing library construction

Single-cell suspensions were processed according to the 10xGenomics scRNAseq sample preparation protocol (Chromium Next GEM Single Cell 3' Kit v3.1, 10xGenomics). The entire mixed cell population was further analyzed without sorting or enrichment for specific cell subtypes. Cell suspensions were uploaded into the Chromium controller, capturing GEMs that encapsulated an estimated 5,000-10,000 single cells per channel. Libraries were constructed from the amplified cDNA, and sequencing was performed on the Illumina NovaSeq 6000 platform. All steps were performed according to the manufacturer's standard protocol.

Processing and quality control of scRNA-seq data

To ensure high-quality data, we implemented three quality control measures on the raw gene-cell-barcode matrix for each cell: the proportion of mitochondrial genes ($\leq 20\%$), unique molecular identifiers (UMIs), and gene count (ranging from 400 to 100,000 and 200 to 10,000, respectively) using Scanpy¹⁰¹. Doublets were identified and removed using the Scrublet¹⁰² package for each sample. Normalization of total counts per cell was performed using the *normalize_total* function in the Scanpy¹⁰¹ package in Python, followed by log-normalization with the *log1p* function. Clustering was conducted using the Leiden algorithm at a resolution of 1, as provided by Scanpy¹⁰¹.

Alignment and raw expression matrix construction

Raw sequencing data were aligned to GRCh38 genome reference using 10X software Cell Ranger (Version 6.1.2) with default parameters.

RNA velocity analysis

For RNA velocity analysis, the spliced and unspliced reads were counted using the velocity.py¹⁰³ package (v0.17.17) from aligned bam files generated by Cell Ranger. A separate loom file was generated and used to process each sample further.

Annotation

Marker genes for each cluster were identified using a t-test implemented in Scanpy¹⁰¹. These marker genes were then used to annotate each cluster using publicly available databases such as CellMarker¹³ and PanglaoDB¹⁴ To determine the cell identities. These annotations were later refined using the CellTypist¹⁵ package and its annotate function. The model 'Immune_All_High.pkl' was specified, and majority voting was enabled. Through this process, seven major cell type clusters were annotated: epithelial cells, natural killer cells, myeloid cells, T cells, B cells, fibroblasts, and endothelial cells. To distinguish malignant cells, we calculated and identified large-scale chromosomal copy number variation (CNV) by inferCNV²² and CaSpER²³ tools for each sample based on transcriptomes. T cells and myeloid cells were considered reference cells; epithelial cells that had differing CNV patterns and exhibited higher CNV scores were annotated as malignant cells.

Datasets were merged across different samples using SCVI¹⁰⁴ based integration, the top 4000 highly variable genes were used to train the VAE models, with each biopsy as a covariate key. After training the initial VAE model, the annotated cell types were used to build an extended model with scANVI¹⁰⁵ for better integration.

After integration, NK cells, myeloid cells, T cells, B cells, and fibroblasts were further classified using the following publicly available single-cell RNA seq datasets: NK cells (GSE212890)¹⁰⁶, T, myeloid, and B cells (GSE169426)¹⁰⁷, and fibroblasts (GSE103322, GSE132465, GSE154778, and GSE212966)⁸³. This label transfer was performed using the scArches¹⁰⁸ algorithm, following the best practices described previously¹⁰⁹.

Proportion Analysis

A proportion analysis was conducted using the scProportionTest¹¹⁰ package to compare the fractions of cells within different cell populations. This involved performing a permutation test to calculate a p-value for each cluster and obtaining a confidence interval through bootstrapping.

Constructing gene regulatory networks

We used pySCENIC³⁹ to construct gene regulatory networks. This involved employing the GRNboost2 method for network inference. We used the *cisTarget* function with the Human motif database v10 to enrich gene signatures and pruned based on cis-regulatory cues using default settings. AUC scores were used to assess regulon enrichment across single cells, and the regulon specificity score was used to compute differential regulon activity.

Copy number profile and subclone inference

We used the SCEVAN³⁷ to determine clonal structures from inferred copy-number alteration profiles. The multiSampleComparisonClonalCN pipeline was employed for intratumoral comparison among multiple samples, with T cells and myeloid cells considered as the reference cells.

Pathway analysis

To assess the pathway activities, we employed *decoupler-py*¹¹¹ and retrieved the gene sets from the Molecular Signatures Database (MSigDB)⁵⁶, PROGENy⁵⁵ database, and CancerSEA⁶⁷ databases.

For each single-cell, pathway activity was inferred using the *decouple* function with default parameters applied to gene sets that included weight information for each gene. In cases where gene sets did not have weight information, we employed the *aucell* function with default parameters, also within the *decoupler-py*¹¹¹ package.

Data visualization

We used essential Scanpy¹⁰¹ functions to generate UMAPs, box plots, heatmaps, dot plots, and violin plots. Proportional change was analyzed and visualized by using *scProportionTest*¹¹⁰ and *Pertpy*¹¹².

For each condition under consideration, we generated pseudobulk profiles in accordance with the official vignette for pseudo-bulk functional analysis using the *decoupler-py* algorithm. The *decoupler-py*¹¹¹ *plot_barplot* function was employed to illustrate the top absolute value activities for pathway activities in each condition. Additionally, we utilized the *plot_targets* function to compare primary and metastatic gene expression profiles, focusing on the target genes associated with each PROGENy⁵⁵ pathway.

Visualization of the dimension reduction of cellular states for pathway activities at the cellular level was performed using the SCpubr¹¹³ *do_CellularStatesPlot*. Gene signatures sourced from the Molecular Signatures Database (MSigDB)⁵⁶ were utilized to create these visual representations.

Differential abundance testing

Graph-based differential abundance analysis was performed using the Milo¹¹⁴ framework in *Pertpy*¹¹² to compare the cellular composition between primary and metastatic samples. The analysis followed the tutorial provided in the *Pertpy* package and utilized integrated datasets for each major cell type.

Cell-cell communication analysis

We investigated cell-cell communication using the CellChat v2⁶⁸ tool to infer, visualize, and analyze cell-cell communication networks from scRNA-seq data. The analysis was conducted following the recommended pipeline provided by the CellChat authors⁶⁸.

Furthermore, we utilized the MultiNicheNet⁸⁷ package to conduct cell interaction analysis, predicting ligand-receptor interactions between different cell types in both primary and metastatic tumor microenvironments. The analysis followed the guidelines outlined in the MultiNicheNet vignette. We visualized each group's top twenty-five ligand-receptor interactions using the *make_circos_group_comparison* function.

Survival Analysis

Kaplan-Meier survival analysis was performed using publicly available RNA-Seq datasets in the KM plotter³³, focusing specifically on ER+ tumors among 2,575 patients. We assessed the collective high or low expression of the following genes: ARNT, BIRC3, EIF2AK1, EIF2AK2, FANCA, HOXC11, KIAA1549, MSH2, MSH6, and MYCN.

Methodology for IntClust Classification

We utilized the Cluster Independent Annotation (CIA)⁵³ tool to compute scores for each IntClust and to classify each cell accordingly. The top 200 differentially expressed genes (DEGs) for each IntClust, as identified in the METABRIC⁵² study, were utilized as input for the CIA tool. Specifically, we utilized the *CIA_classify* function to calculate scores based on the expression levels of the selected DEGs and assign each cell to the most probable IntClust.

Data availability

Raw single-cell RNA sequencing data for this study are available in the NCBI BioProject database under accession number PRJNA1140267. Processed single-cell RNA sequencing data and single-cell RNA seq objects can be found as follows:
<https://doi.org/10.5281/zenodo.13743373>

Code availability

The authors declare that all scripts supporting the findings of this study are available from the corresponding author upon reasonable request.

Author Information

F.O., T.Y.O., and A.O. contributed equally to this work. G.B.M. and H.M. conceived and supervised the project. F.O., T.Y.O., A.O., G.B.M., and H.M. designed the study. F.O. performed the computational bioinformatics analysis. F.O. and T.Y.O. wrote the original draft, with A.O., M.J.R., G.B.M., and H.M. providing editing contributions. A.O., T.Y.O., M.J., F.B., and X.L. were responsible for collecting clinical samples and conducting experiments. G.B.M and H.M provided the final review of the manuscript. G.B.M. and H.M. acquired the funding. All authors contributed to data generation and interpretation. All authors read and approved the final manuscript.

Author List: Furkan Ozmen, Tugba Y. Ozmen, Aysegul Ors, Mahnaz Janghorban, Matthew J. Rames, Xi Li, Fariba Behbod, Gordon B. Mills, Hisham Mohammed.

Competing Interests

The authors declare no competing interests.

Ethics Declarations

This study was approved by the institutional ethics review board of OHSU. Written informed consent was signed for all experiments involving patients.

Supplementary Information

All supplementary materials are available upon publication. The supplementary materials include three spreadsheets, and one document as follows:

Supplementary Table 1: Clinical information related to samples in the study (Supplementary Table 1.xlsx)

Supplementary Table 2: Subclone information of malignant cells for individual samples (Supplementary Table 2.xlsx)

Supplementary Table 3: Proportions, counts, and pathway activity comparisons for various cell types across primary and metastatic breast cancer samples. This includes the proportion and count of major and minor cell types, differentially expressed markers, CNV and ITHGEX score distributions, IntClust subtype proportions, regulon activity scores, pathway activity comparisons (PROGENy, Hallmark, CancerSEA), and gene signatures used in the analysis (Supplementary Table 3.xlsx)

Supplementary Document 1: CNV alteration heatmaps for individual samples (Supplementary Document 1.pdf)

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Furkan Ozmen(ozmen@ohsu.edu).

References

1. SEER*Explorer: An interactive website for SEER cancer statistics. SEER Incidence Data, November 2023 Submission (1975-2021).
2. Allemani, C. *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. *The Lancet* **391**, 1023–1075 (2018).

3. Valastyan, S. & Weinberg, R. A. Tumor Metastasis: Molecular Insights and Evolving Paradigms. *Cell* **147**, 275–292 (2011).
4. Mu, X. *et al.* Tumor-derived lactate induces M2 macrophage polarization via the activation of the ERK/STAT3 signaling pathway in breast cancer. *Cell Cycle* **17**, 428–438 (2018).
5. Bahrami, A. *et al.* The Therapeutic Potential of Targeting Tumor Microenvironment in Breast Cancer: Rational Strategies and Recent Progress. *J. Cell. Biochem.* **119**, 111–122 (2018).
6. Yang, J., Bahcecioglu, G. & Zorlutuna, P. The Extracellular Matrix and Vesicles Modulate the Breast Tumor Microenvironment. *Bioengineering* **7**, 124 (2020).
7. Jiang, G. *et al.* Comprehensive comparison of molecular portraits between cell lines and tumors in breast cancer. *BMC Genomics* **17**, 525 (2016).
8. Sørli, T. The Impact of Gene Expression Patterns in Breast Cancer. *Clin. Chem.* **62**, 1150–1151 (2016).
9. Shome, R., Sen, P., Sarkar, S. & Ghosh, S. S. Single-cell transcriptomics reveals the intra-tumoral heterogeneity and SQSTM1/P62 and Wnt/ β -catenin mediated epithelial to mesenchymal transition and stemness of triple-negative breast cancer. *Exp. Cell Res.* **438**, 114032 (2024).
10. Qiu, Y. *et al.* FOXP3+ regulatory T cells and the immune escape in solid tumours. *Front. Immunol.* **13**, 982986 (2022).
11. Jiang, W. *et al.* Exhausted CD8+ T Cells in the Tumor Immune Microenvironment: New Pathways to Therapy. *Front. Immunol.* **11**, 622509 (2020).
12. Liu, T. *et al.* Single cell profiling of primary and paired metastatic lymph node tumors in breast cancer patients. *Nat. Commun.* **13**, 6823 (2022).
13. Hu, C. *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.* **51**, D870–D876 (2023).
14. Franzén, O., Gan, L.-M. & Björkegren, J. L. M. PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data. *Database* **2019**, baz046 (2019).
15. Xu, C. *et al.* Automatic cell-type harmonization and integration across Human Cell Atlas datasets. *Cell* **186**, 5876-5891.e20 (2023).
16. Nalio Ramos, R. *et al.* Tissue-resident FOLR2+ macrophages associate with CD8+ T cell infiltration in human breast cancer. *Cell* **185**, 1189-1207.e25 (2022).
17. Xiang, C. *et al.* Single-cell profiling reveals the trajectory of FOLR2-expressing tumor-associated macrophages to regulatory T cells in the progression of lung adenocarcinoma. *Cell Death Dis.* **14**, 1–12 (2023).
18. Jiang, B. *et al.* FOLR2-Positive Macrophages Suppress Breast Cancer Invasion Through T Cell Activation. SSRN Scholarly Paper at <https://doi.org/10.2139/ssrn.4667619> (2023).

19. Oghumu, S. *et al.* CXCR3 deficiency enhances tumor progression by promoting macrophage M2 polarization in a murine breast cancer model. *Immunology* **143**, 109–119 (2014).
20. Fei, L., Ren, X., Yu, H. & Zhan, Y. Targeting the CCL2/CCR2 Axis in Cancer Immunotherapy: One Stone, Three Birds? *Front. Immunol.* **12**, (2021).
21. Kholod, O. *et al.* Immune-Related Gene Signatures to Predict the Effectiveness of Chemoimmunotherapy in Triple-Negative Breast Cancer Using Exploratory Subgroup Discovery. *Cancers* **14**, 5806 (2022).
22. Timothy Tickle and Itay Tirosh and Christophe Georgescu and Maxwell Brown and Brian Haas. inferCNV of the Trinity CTAT Project. *Klarman Cell Obs. Broad Inst. MIT Harv. Cambridge, MA, USA*, (2019).
23. Serin Harmanci, A., Harmanci, A. O. & Zhou, X. CaSpER identifies and visualizes CNV events by integrative analysis of single-cell or bulk RNA-sequencing data. *Nat. Commun.* **11**, 89 (2020).
24. Gao, Y. *et al.* Single-cell sequencing deciphers a convergent evolution of copy number alterations from primary to circulating tumor cells. *Genome Res.* **27**, 1312–1322 (2017).
25. Schafer, J. M. *et al.* Targeting MYCN-expressing triple-negative breast cancer with BET and MEK inhibitors. *Sci. Transl. Med.* **12**, eaaw8275 (2020).
26. Zhao, Y., Han, F., Zhang, X., Zhou, C. & Huang, D. Aryl hydrocarbon receptor nuclear translocator promotes the proliferation and invasion of clear cell renal cell carcinoma cells potentially by affecting the glycolytic pathway. *Oncol. Lett.* **20**, 56 (2020).
27. Koboldt, D. C. *et al.* Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).
28. Frazzi, R. BIRC3 and BIRC5: multi-faceted inhibitors in cancer. *Cell Biosci.* **11**, 8 (2021).
29. Du, H.-X. *et al.* Eukaryotic translation initiation factor 2 α kinase 2 in pancreatic cancer: An approach towards managing clinical prognosis and molecular immunological characterization. *Oncol. Lett.* **26**, 1–21 (2023).
30. Bravo-Navas, S., Yáñez, L., Romón, Í., Pipaón, & Carlos. Elevated FANCA expression determines a worse prognosis in chronic lymphocytic leukemia and interferes with p53 function. *FASEB J.* **33**, 10477–10489 (2019).
31. Gong, N. *et al.* HOXC11 positively regulates the long non-coding RNA HOTAIR and is associated with poor prognosis in colon adenocarcinoma. *Exp. Ther. Med.* **22**, 1–6 (2021).
32. Zhou, L.-Z., Xiao, H.-Q. & Chen, J. Mismatch repair gene MSH6 correlates with the prognosis, immune status and immune checkpoint inhibitors response of endometrial cancer. *Front. Immunol.* **15**, (2024).
33. Gyórfy, B. Integrated analysis of public datasets for the discovery and validation of survival-associated genes in solid tumors. *The Innovation* **5**, 100625 (2024).

34. Li, F. *et al.* Relationship Between CNVs and Immune Cells Infiltration in Gastric Tumor Microenvironment. *Front. Genet.* **13**, 869967 (2022).
35. Chen, S. *et al.* Quantified CIN Score From Cell-free DNA as a Novel Noninvasive Predictor of Survival in Patients With Spinal Metastasis. *Front. Cell Dev. Biol.* **9**, (2021).
36. Ramón y Cajal, S. *et al.* Clinical implications of intratumor heterogeneity: challenges and opportunities. *J. Mol. Med.* **98**, 161–177 (2020).
37. De Falco, A., Caruso, F., Su, X.-D., Iavarone, A. & Ceccarelli, M. A variational algorithm to detect the clonal copy number substructure of tumors from scRNA-seq data. *Nat. Commun.* **14**, 1074 (2023).
38. Wu, F. *et al.* Single-cell profiling of tumor heterogeneity and the microenvironment in advanced non-small cell lung cancer. *Nat. Commun.* **12**, 2540 (2021).
39. Van de Sande, B. *et al.* A scalable SCENIC workflow for single-cell gene regulatory network analysis. *Nat. Protoc.* **15**, 2247–2276 (2020).
40. Segditsas, S. *et al.* Putative direct and indirect Wnt targets identified through consistent gene expression changes in APC-mutant intestinal adenomas from humans and mice. *Hum. Mol. Genet.* **17**, 3864–3875 (2008).
41. Kristan, A., Debeljak, N. & Kunej, T. Integration and Visualization of Regulatory Elements and Variations of the EPAS1 Gene in Human. *Genes* **12**, 1793 (2021).
42. Jia, C. *et al.* Evidence of Omics, Immune Infiltration, and Pharmacogenomics for BATF in a Pan-Cancer Cohort. *Front. Mol. Biosci.* **9**, (2022).
43. Yu, X. *et al.* The basic leucine zipper transcription factor NFIL3 directs the development of a common innate lymphoid cell precursor. *eLife* **3**, e04406 (2014).
44. Kubo, M. Diurnal Rhythmicity Programs of Microbiota and Transcriptional Oscillation of Circadian Regulator, NFIL3. *Front. Immunol.* **11**, (2020).
45. Yuan, Y.-H. *et al.* Identification of key genes and pathways downstream of the β -catenin-TCF7L1 complex in pancreatic cancer cells using bioinformatics analysis. *Oncol. Lett.* **18**, 1117–1132 (2019).
46. Lin, J. *et al.* Comprehensive analysis of KLF family reveals KLF6 as a promising prognostic and immune biomarker in pancreatic ductal adenocarcinoma. *Cancer Cell Int.* **24**, 177 (2024).
47. Li, C., Cui, J., Zou, L., Zhu, L. & Wei, W. Bioinformatics analysis of the expression of HOXC13 and its role in the prognosis of breast cancer. *Oncol. Lett.* **19**, 899–907 (2020).
48. Böhm, M., Locke, W. J., Sutherland, R. L., Kench, J. G. & Henshall, S. M. A role for GATA-2 in transition to an aggressive phenotype in prostate cancer through modulation of key androgen-regulated genes. *Oncogene* **28**, 3847–3856 (2009).
49. Sharma, B. R. *et al.* The Transcription Factor IRF9 Promotes Colorectal Cancer via Modulating the IL-6/STAT3 Signaling Axis. *Cancers* **14**, 919 (2022).

50. Guo, W. *et al.* Super-enhancer-driven MLX mediates redox balance maintenance via SLC7A11 in osteosarcoma. *Cell Death Dis.* **14**, 1–13 (2023).
51. Pu, Q. *et al.* The Novel Transcription Factor CREB3L4 Contributes to the Progression of Human Breast Carcinoma. *J. Mammary Gland Biol. Neoplasia* **25**, 37–50 (2020).
52. Ali, H. R. *et al.* Genome-driven integrated classification of breast cancer validated in over 7,500 samples. *Genome Biol.* **15**, 431 (2014).
53. Ferrari, I. *et al.* CIA: a Cluster Independent Annotation method to investigate cell identities in scRNA-seq data. 2023.11.30.569382 Preprint at <https://doi.org/10.1101/2023.11.30.569382> (2023).
54. Rueda, O. M. *et al.* Dynamics of breast-cancer relapse reveal late-recurring ER-positive genomic subgroups. *Nature* **567**, 399–404 (2019).
55. Schubert, M. *et al.* Perturbation-response genes reveal signaling footprints in cancer gene expression. *Nat. Commun.* **9**, 20 (2018).
56. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci.* **102**, 15545–15550 (2005).
57. Mootha, V. K. *et al.* PGC-1 α -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* **34**, 267–273 (2003).
58. Ozga, A. J. *et al.* CXCL10 chemokine regulates heterogeneity of the CD8+ T cell response and viral set point during chronic infection. *Immunity* **55**, 82–97.e8 (2022).
59. Wang, X. *et al.* The role of CXCR3 and its ligands in cancer. *Front. Oncol.* **12**, 1022688 (2022).
60. Wang, J., Zhu, N., Su, X., Gao, Y. & Yang, R. Novel tumor-associated macrophage populations and subpopulations by single cell RNA sequencing. *Front. Immunol.* **14**, (2024).
61. Jin, J. *et al.* CCL2: An Important Mediator Between Tumor Cells and Host Cells in Tumor Microenvironment. *Front. Oncol.* **11**, (2021).
62. Hao, Q., Vadgama, J. V. & Wang, P. CCL2/CCR2 signaling in cancer pathogenesis. *Cell Commun. Signal.* **18**, 82 (2020).
63. Qian, B.-Z. *et al.* CCL2 recruits inflammatory monocytes to facilitate breast-tumour metastasis. *Nature* **475**, 222–225 (2011).
64. Chang, K. *et al.* MGP+ and IDO1+ tumor-associated macrophages facilitate immunoresistance in breast cancer revealed by single-cell RNA sequencing. *Int. Immunopharmacol.* **131**, 111818 (2024).
65. Tsukamoto, S. *et al.* Matrix Metalloproteinase 9 Induced in Esophageal Squamous Cell Carcinoma Cells via Close Contact with Tumor-Associated Macrophages Contributes to Cancer Progression and Poor Prognosis. *Cancers* **15**, 2987 (2023).

66. Snell, L. M., McGaha, T. L. & Brooks, D. G. Type I Interferon in Chronic Virus Infection and Cancer. *Trends Immunol.* **38**, 542–557 (2017).
67. Yuan, H. *et al.* CancerSEA: a cancer single-cell state atlas. *Nucleic Acids Res.* **47**, D900–D908 (2019).
68. Jin, S., Plikus, M. V. & Nie, Q. CellChat for systematic analysis of cell-cell communication from single-cell and spatially resolved transcriptomics. 2023.11.05.565674 Preprint at <https://doi.org/10.1101/2023.11.05.565674> (2023).
69. Cai, X. *et al.* Down-regulation of FN1 inhibits colorectal carcinogenesis by suppressing proliferation, migration, and invasion. *J. Cell. Biochem.* **119**, 4717–4728 (2018).
70. Ma, H.-P. *et al.* Collagen 1A1 (COL1A1) Is a Reliable Biomarker and Putative Therapeutic Target for Hepatocellular Carcinogenesis and Metastasis. *Cancers* **11**, 786 (2019).
71. Gray, J. I., Westerhof, L. M. & MacLeod, M. K. L. The roles of resident, central and effector memory CD4 T-cells in protective immunity following infection or vaccination. *Immunology* **154**, 574–581 (2018).
72. Capuano, C. *et al.* Harnessing CD16-Mediated NK Cell Functions to Enhance Therapeutic Efficacy of Tumor-Targeting mAbs. *Cancers* **13**, 2500 (2021).
73. Yeap, W. H. *et al.* CD16 is indispensable for antibody-dependent cellular cytotoxicity by human monocytes. *Sci. Rep.* **6**, 34310 (2016).
74. Zhang, L. *et al.* Identification of a novel ADCC-related gene signature for predicting the prognosis and therapy response in lung adenocarcinoma. *Inflamm. Res. Off. J. Eur. Histamine Res. Soc. Al* **73**, 841–866 (2024).
75. Xu, Y. *et al.* Glycolysis determines dichotomous regulation of T cell subsets in hypoxia. *J. Clin. Invest.* **126**, 2678–2688 (2016).
76. Facciabene, A. *et al.* Tumour hypoxia promotes tolerance and angiogenesis via CCL28 and Treg cells. *Nature* **475**, 226–230 (2011).
77. McNamee, E. N., Korn Johnson, D., Homann, D. & Clambey, E. T. Hypoxia and hypoxia-inducible factors as regulators of T cell development, differentiation, and function. *Immunol. Res.* **55**, 58–70 (2013).
78. Muroyama, Y. & Wherry, E. J. Memory T-Cell Heterogeneity and Terminology. *Cold Spring Harb. Perspect. Biol.* **13**, a037929 (2021).
79. Wolint, P., Betts, M. R., Koup, R. A. & Oxenius, A. Immediate Cytotoxicity But Not Degranulation Distinguishes Effector and Memory Subsets of CD8+ T Cells. *J. Exp. Med.* **199**, 925–936 (2004).
80. Han, D. *et al.* Microenvironmental network of clonal CXCL13⁺CD4⁺ T cells and Tregs in pemphigus chronic blisters. *J. Clin. Invest.* **133**, (2023).

81. Jenkins, L. *et al.* Cancer-Associated Fibroblasts Suppress CD8+ T-cell Infiltration and Confer Resistance to Immune-Checkpoint Blockade. *Cancer Res.* **82**, 2904–2917 (2022).
82. Kerdidani, D. *et al.* Lung tumor MHCII immunity depends on in situ antigen presentation by fibroblasts. *J. Exp. Med.* **219**, e20210815 (2022).
83. Cords, L. *et al.* Cancer-associated fibroblast classification in single-cell and spatial proteomics data. *Nat. Commun.* **14**, 4294 (2023).
84. Molnár, K. *et al.* Pericyte-secreted IGF2 promotes breast cancer brain metastasis formation. *Mol. Oncol.* **14**, 2040–2057 (2020).
85. Hompland, T., Fjeldbo, C. S. & Lyng, H. Tumor Hypoxia as a Barrier in Cancer Therapy: Why Levels Matter. *Cancers* **13**, 499 (2021).
86. Endres, M. *et al.* Regulation of matrix metalloproteinases (MMPs) expression and secretion in MDA-MB-231 breast cancer cells by LIM and SH3 protein 1 (LASP1). *Oncotarget* **7**, 64244–64259 (2016).
87. Browaeys, R. *et al.* MultiNicheNet: a flexible framework for differential cell-cell communication analysis from multi-sample multi-condition single-cell transcriptomics data. 2023.06.13.544751 Preprint at <https://doi.org/10.1101/2023.06.13.544751> (2023).
88. Sun, L., Guo, S., Xie, Y. & Yao, Y. The characteristics and the multiple functions of integrin β 1 in human cancers. *J. Transl. Med.* **21**, 787 (2023).
89. Huang, R. & Rofstad, E. K. Integrins as therapeutic targets in the organ-specific metastasis of human malignant melanoma. *J. Exp. Clin. Cancer Res.* **37**, 92 (2018).
90. Takebe, N., Warren, R. Q. & Ivy, S. P. Breast cancer growth and metastasis: interplay between cancer stem cells, embryonic signaling pathways and epithelial-to-mesenchymal transition. *Breast Cancer Res.* **13**, 211 (2011).
91. Zhang, S. *et al.* Single cell transcriptomic analyses implicate an immunosuppressive tumor microenvironment in pancreatic cancer liver metastasis. *Nat. Commun.* **14**, 5123 (2023).
92. Kumar, S. *et al.* PathFinder: Bayesian inference of clone migration histories in cancer. *Bioinformatics* **36**, i675–i683 (2020).
93. Zelová, H. & Hošek, J. TNF- α signalling and inflammation: interactions between old acquaintances. *Inflamm. Res.* **62**, 641–651 (2013).
94. Roca, H. & McCauley, L. K. Inflammation and skeletal metastasis. *BoneKEy Rep.* **4**, 706 (2015).
95. Oskarsson, T., Battle, E. & Massagué, J. Metastatic Stem Cells: Sources, Niches, and Vital Pathways. *Cell Stem Cell* **14**, 306–321 (2014).
96. Amato, C. M. *et al.* Pre-Treatment Mutational and Transcriptomic Landscape of Responding Metastatic Melanoma Patients to Anti-PD1 Immunotherapy. *Cancers* **12**, 1943 (2020).

97. Goritzka, M. *et al.* Alpha/beta interferon receptor signaling amplifies early proinflammatory cytokine production in the lung during respiratory syncytial virus infection. *J. Virol.* **88**, 6128–6136 (2014).
98. Ma, W. *et al.* Type I interferon response in astrocytes promotes brain metastasis by enhancing monocytic myeloid cell recruitment. *Nat. Commun.* **14**, 2632 (2023).
99. Geurts, V. & Kok, M. Immunotherapy for Metastatic Triple Negative Breast Cancer: Current Paradigm and Future Approaches. *Curr. Treat. Options Oncol.* **24**, 628–643 (2023).
100. Lendahl, U., Muhl, L. & Betsholtz, C. Identification, discrimination and heterogeneity of fibroblasts. *Nat. Commun.* **13**, 3409 (2022).
101. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).
102. Wolock, S. L., Lopez, R. & Klein, A. M. Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. *Cell Syst.* **8**, 281-291.e9 (2019).
103. La Manno, G. *et al.* RNA velocity of single cells. *Nature* **560**, 494–498 (2018).
104. Gayoso, A. *et al.* A Python library for probabilistic analysis of single-cell omics data. *Nat. Biotechnol.* **40**, 163–166 (2022).
105. Xu, C. *et al.* Probabilistic harmonization and annotation of single-cell transcriptomics data with deep generative models. *Mol. Syst. Biol.* **17**, e9620 (2021).
106. Tang, F. *et al.* A pan-cancer single-cell panorama of human natural killer cells. *Cell* **186**, 4235-4251.e20 (2023).
107. Dussiau, C. *et al.* Hematopoietic differentiation is characterized by a transient peak of entropy at a single-cell level. *BMC Biol.* **20**, 60 (2022).
108. Lotfollahi, M. *et al.* Mapping single-cell data to reference atlases by transfer learning. *Nat. Biotechnol.* **40**, 121–130 (2022).
109. Heumos, L. *et al.* Best practices for single-cell analysis across modalities. *Nat. Rev. Genet.* **24**, 550–572 (2023).
110. Miller, S. A. *et al.* LSD1 and Aberrant DNA Methylation Mediate Persistence of Enteroendocrine Progenitors That Support BRAF-Mutant Colorectal Cancer. *Cancer Res.* **81**, 3791–3805 (2021).
111. Badia-i-Mompel, P. *et al.* decoupleR: ensemble of computational methods to infer biological activities from omics data. *Bioinforma. Adv.* **2**, vbac016 (2022).
112. Heumos, L. *et al.* Pertpy: an end-to-end framework for perturbation analysis. 2024.08.04.606516 Preprint at <https://doi.org/10.1101/2024.08.04.606516> (2024).

113. Blanco-Carmona, E. Generating publication ready visualizations for Single Cell transcriptomics using SCpubr. 2022.02.28.482303 Preprint at <https://doi.org/10.1101/2022.02.28.482303> (2022).

114. Dann, E., Henderson, N. C., Teichmann, S. A., Morgan, M. D. & Marioni, J. C. Differential abundance testing on single-cell data using k-nearest neighbor graphs. *Nat. Biotechnol.* **40**, 245–253 (2022).

Figures

It is made available under a [CC-BY-NC 4.0 International license](https://creativecommons.org/licenses/by-nc/4.0/).

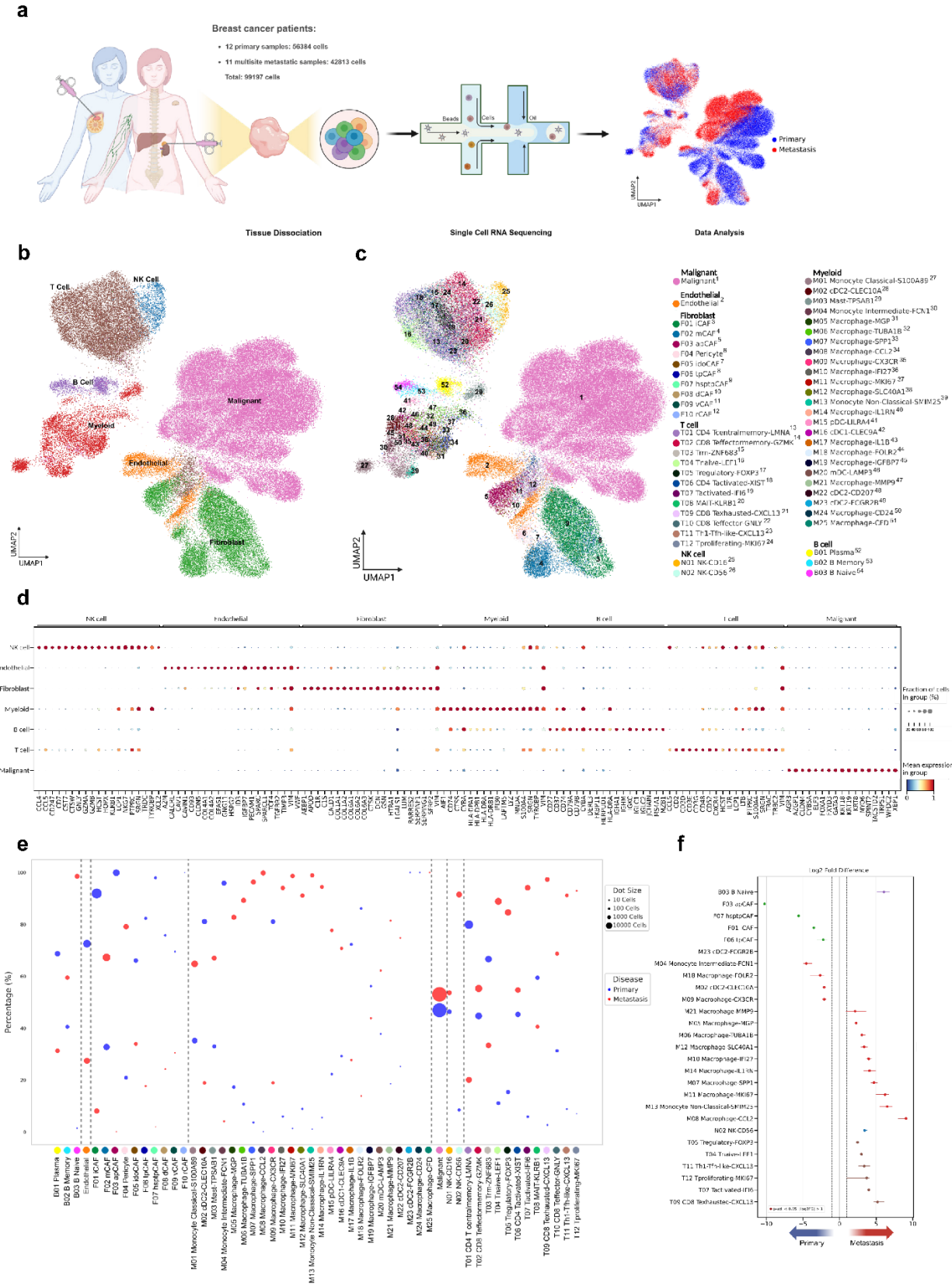


Fig.1 Single-Cell Landscape of Primary and Metastatic ER+ Breast Cancer: Characterization and Cellular Subtype Differences

a. Sample collection workflow and data analysis overview. (Created in BioRender. BioRender.com/y20n356) **b.** UMAP visualization of the unified cell map, consisting of 7 major cell types colored by clusters. **c.** UMAP visualization of the unified cell map, sub-segmenting by 54 minor cell types uniquely colored within the 7 major cell types. **d.** Dot plot representation of differentially expressed genes across the 7 major cell type clusters. **e.** Relative percentage of 54 minor cell types by metastatic status and corresponding cell number indicated by dot size. **f.** Significant changes in minor cell type proportions between primary and metastatic breast cancer

Fig.1 Single-Cell Landscape of Primary and Metastatic ER+ Breast Cancer: Characterization and Cellular Subtype Differences

- a.** Sample collection workflow and data analysis overview. (Created in BioRender. BioRender.com/y20n356)
- b.** UMAP visualization of the unified cell map, consisting of 7 major cell types colored by clusters.
- c.** UMAP visualization of the unified cell map, sub-segmenting by 54 minor cell types uniquely colored within the 7 major cell types.
- d.** Dot plot representation of differentially expressed genes across the 7 major cell type clusters.
- e.** Relative percentage of 54 minor cell types by metastatic status and corresponding cell number indicated by dot size.
- f.** Significant changes in minor cell type proportions between primary and metastatic breast cancer

It is made available under a [CC-BY-NC 4.0 International license](https://creativecommons.org/licenses/by-nc/4.0/).

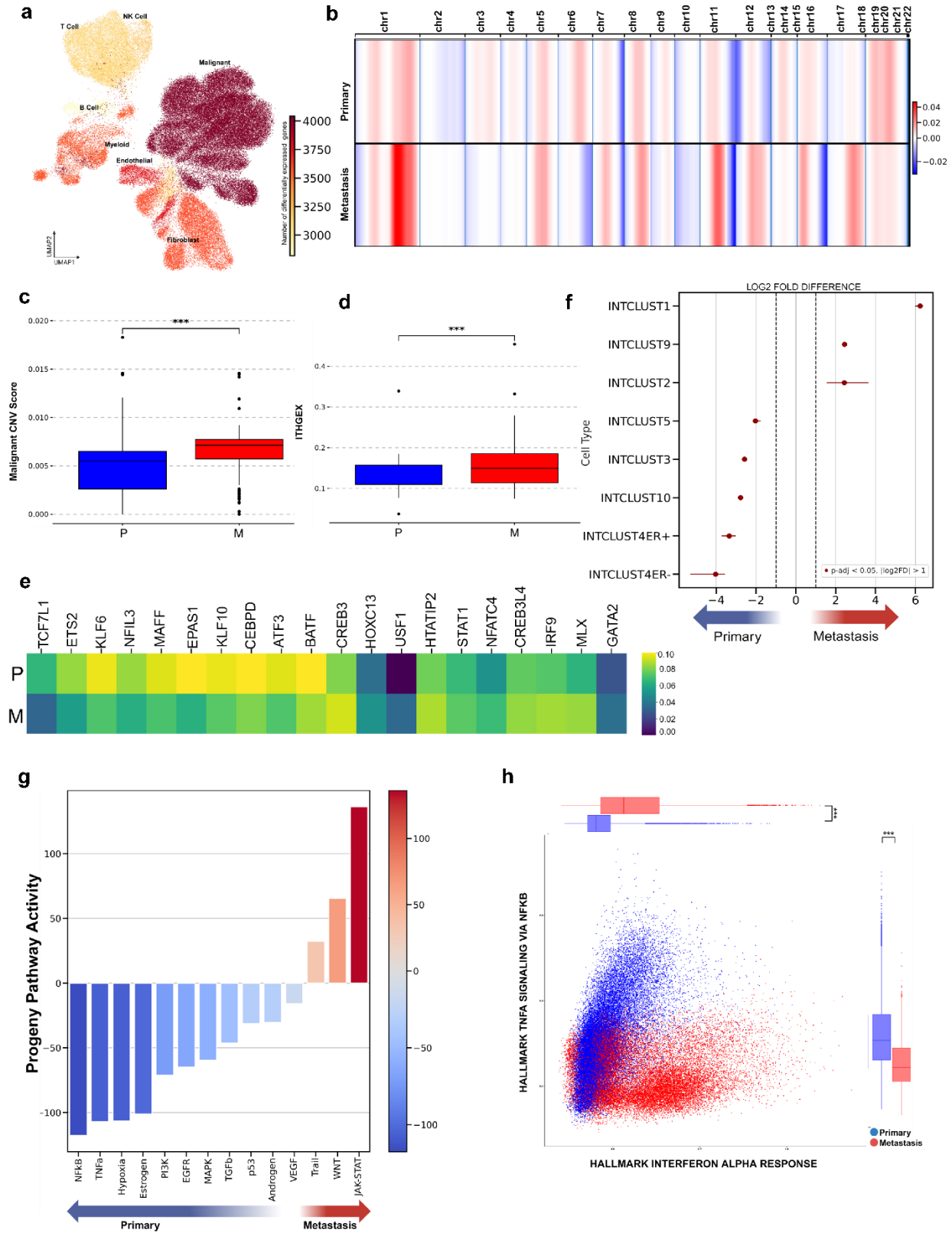


Fig.2 Genomic and Phenotypic Alterations in Malignant Cells of Primary vs. Metastatic Breast Cancer: Insights from CNV Analysis, Gene Regulatory Networks, and Integrative Clusters

a, UMAP visualization of the number of differentially expressed genes within all major cell type clusters. **b**, Chromosomal CNV alteration score by metastatic status, with copy number gain and copy number loss indicated by red and blue, respectively. **c**, CNV score comparison between primary and metastatic breast cancer malignant cells. Significance indicated as (***, $p < 0.001$). **d**, ITGEX score comparison between primary and metastatic breast cancer malignant cells. Significance indicated as (***, $p < 0.001$). **e**, Regulation activity scores by metastatic status within malignant cells. **f**, Significant ($\text{Log}_2 \text{FC}$) integrative cluster differences between primary and metastatic malignant cells ($p\text{-adj} < 0.05$). **g**, Differential progeny pathway activity scores within malignant cells across metastatic status. **h**, Cellular states plot showing hallmark gene signature activity for Interferon Alpha response versus TNFa signaling via NFKB in malignant cell populations.

Fig.2 Genomic and Phenotypic Alterations in Malignant Cells of Primary vs. Metastatic Breast Cancer: Insights from CNV Analysis, Gene Regulatory Networks, and Integrative Clusters

- a.** UMAP visualization of the number of differentially expressed genes within all major cell type clusters.
- b.** Chromosomal CNV alternation score by metastatic status, with copy number gain and copy number loss indicated by red and blue, respectively.
- c.** CNV score comparison between primary and metastatic breast cancer malignant cells. Significance indicated as (***, $p < 0.001$).
- d.** ITGEX score comparison between primary and metastatic breast cancer malignant cells. Significance indicated as (***, $p < 0.001$).
- e.** Regulon activity scores by metastatic status within malignant cells.
- f.** Significant ($\text{Log}_2 \text{FC}$) integrative cluster differences between primary and metastatic malignant cells ($p\text{-adj} < 0.05$).
- g.** Differential progeny pathway activity scores within malignant cells across metastatic status.
- h.** Cellular states plot showing hallmark gene signature activity for Interferon Alpha response versus TNF- α signaling via NFkB in malignant cell populations.

It is made available under a [CC-BY-NC 4.0 International license](https://creativecommons.org/licenses/by-nc/4.0/).

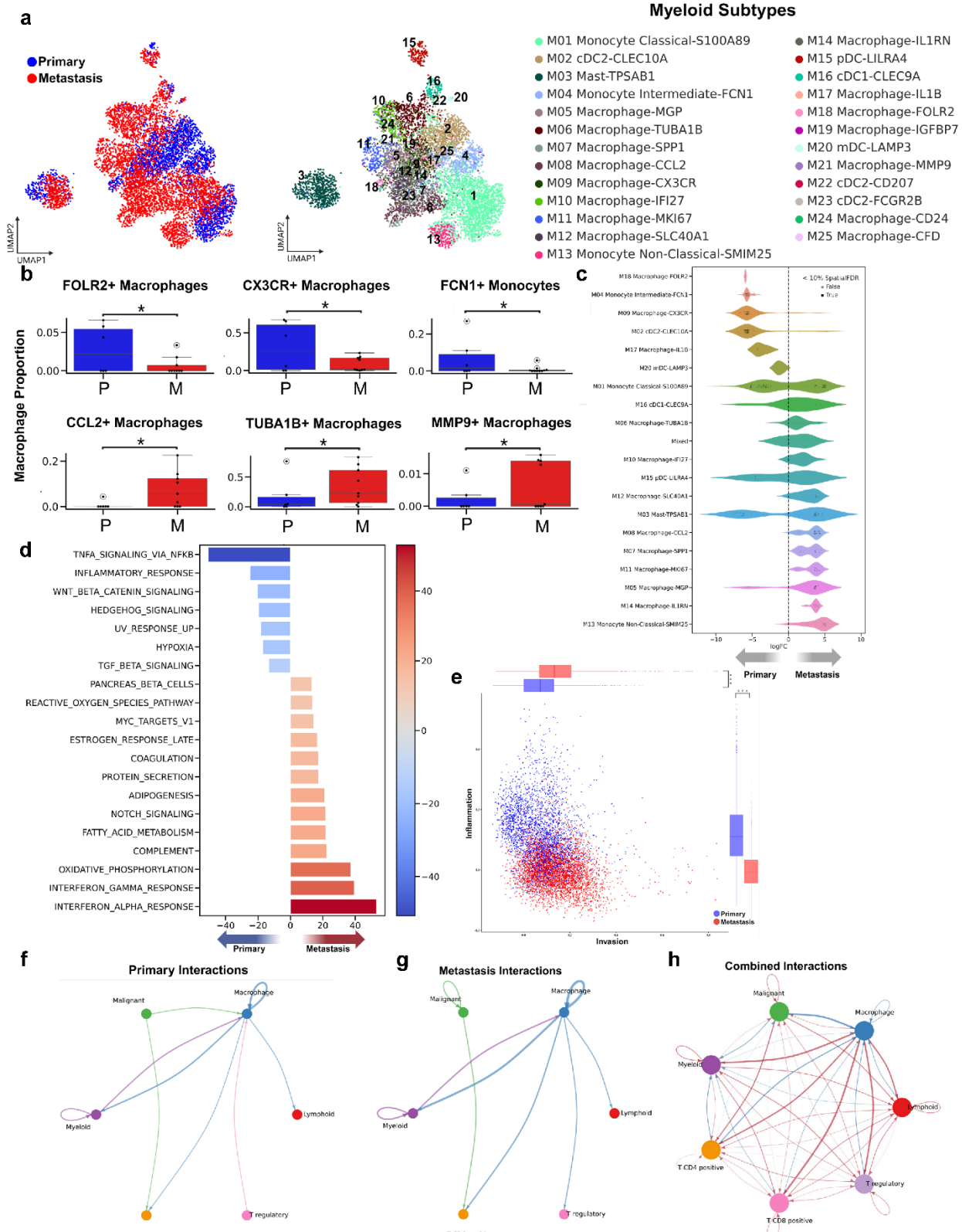


Fig.3 Distinct Myeloid Cell Populations and Their Roles in Primary and Metastatic Breast Cancer: Insights into Macrophage Polarization and Tumor Microenvironment Interactions

a. UMAP visualizations of 7651 myeloid cells. Cells are colored by tissue class of origin (left) and grouped by one of 25 identified sub-types (right), including 6 Dendritic, 3 Monocytic, 1 Mast, and 15 Macrophage lineages. **b.** Bar plots showing the proportions of cells from each tissue class of origin (P=Primary, M=Metastatic) for macrophage subtypes with a significant difference in proportion. (permutation test, $p < 0.05$, $\log_2FC > 2$) **c.** The beeswarm plot illustrates the differential abundance analysis of myeloid cell subtypes, revealing distinct variations in cell abundance across neighborhoods, stratified by tissue classes of origin in primary and metastatic breast cancer. **d.** Differential Hallmark gene signature pathway activity scores within malignant cells across metastatic status. **e.** Cellular states plot showing CANCERSEA gene signature activity for Inflammation versus Invasion in myeloid cell populations. **f.** Circle plot displaying the top 15% of primary tumor microenvironment cell-cell interactions. Arrow colors indicate sender and thickness represents interaction strength. **g.** Circle plot displaying the top 15% of metastatic tumor microenvironment cell-cell interactions. Arrow colors indicate sender and thickness represents interaction strength. **h.** Comparative circle plot visualizing differential cell-cell communication between primary and metastatic tumor microenvironments, where red edges indicate increased interactions in metastatic tumors and blue edges denote increased interactions in primary tumors

Fig.3 Distinct Myeloid Cell Populations and Their Roles in Primary and Metastatic Breast Cancer: Insights into Macrophage Polarization and Tumor Microenvironment Interactions

- a.** UMAP visualizations of 7651 myeloid cells. Cells are colored by tissue class of origin (left) and grouped by one of 25 identified sub-types (right), including 6 Dendritic, 3 Monocytic, 1 Mast, and 15 Macrophage lineages.
- b.** Bar plots showing the proportions of cells from each tissue class of origin (P=Primary, M=Metastatic) for macrophage subtypes with a significant difference in proportion. (permutation test, $p < 0.05$, $\log_2FC > 2$)
- c.** The beeswarm plot illustrates the differential abundance analysis of myeloid cell subtypes, revealing distinct variations in cell abundance across neighborhoods, stratified by tissue classes of origin in primary and metastatic breast cancer.
- d.** Differential Hallmark gene signature pathway activity scores within malignant cells across metastatic status.
- e.** Cellular states plot showing CANCERSEA gene signature activity for Inflammation versus Invasion in myeloid cell populations.
- f.** Circle plot displaying the top 15% of primary tumor microenvironment cell-cell interactions. Arrow colors indicate sender and thickness represents interaction strength.
- g.** Circle plot displaying the top 15% of metastatic tumor microenvironment cell-cell interactions. Arrow colors indicate sender and thickness represents interaction strength.
- h.** Comparative circle plot visualizing differential cell-cell communication between primary and metastatic tumor microenvironments, where red edges indicate increased interactions in metastatic tumors and blue edges denote increased interactions in primary tumors

It is made available under a [CC-BY-NC 4.0 International license](https://creativecommons.org/licenses/by-nc/4.0/).

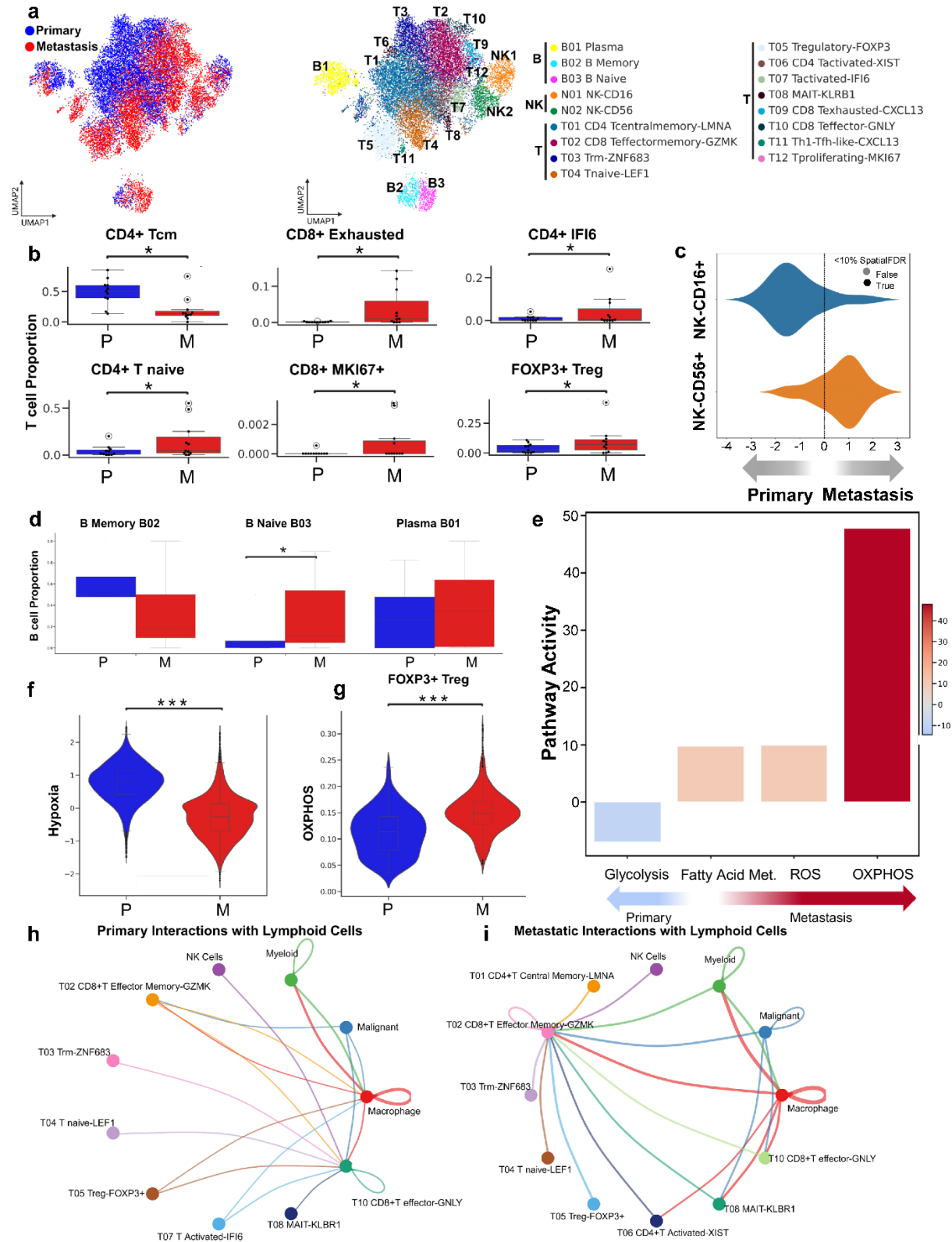


Fig.4 Lymphoid Cell Landscape and Functional Shifts in Primary vs. Metastatic Breast Cancer Tumors

a. UMAP visualizations of 19104 lymphoid cells. Cells are colored by tissue class of origin on the left and grouped by one of 17 identified subtypes on the right, including 3 B cell subtypes, 2 NK cell subtypes, and 12 T cell subtypes. **b.** Bar plots showing the proportions of cells from each tissue class of origin (P=Primary, M=Metastatic) for T cell subtypes with a significant difference in proportion. (permutation test, $p < 0.05$, $\log_2FC > 2$) **c.** The beeswarm plot illustrating the differential abundance analysis of NK cells, uncovering variations in cell abundance across neighborhoods, with a focus on specific NK cell subtypes, categorized by tissue classes of origin in primary and metastatic breast cancer. **d.** Bar plots showing the proportions of cells from each tissue class of origin (P=Primary, M=Metastatic) for B cell subtypes with a significant difference in proportion. (permutation test, $p < 0.05$, $\log_2FC > 2$) **e.** Differential Hallmark gene signature pathway activity scores within T cells across metastatic status. **f.** Hypoxia score comparison between primary and metastatic breast cancer T cells. Significance indicated as (***, $p < 0.001$). **g.** OXPHOS score comparison between primary and metastatic breast cancer T cells. Significance indicated as (***, $p < 0.001$). **h.** Circle plot showing the top 10% of predominant cell-cell interactions in the primary tumor microenvironment, with arrow colors indicating the source (sender) of the interaction, and arrow thickness representing the strength of interaction. **i.** Circle plot showing the top 10% of predominant cell-cell interactions in the metastatic tumor microenvironment, with arrow colors indicating the source (sender) of the interaction, and arrow thickness representing the strength of interaction.

Fig.4 Lymphoid Cell Landscape and Functional Shifts in Primary vs. Metastatic Breast Cancer Tumors"

- a.** UMAP visualizations of 19104 lymphoid cells. Cells are colored by tissue class of origin on the left and grouped by one of 17 identified subtypes on the right, including 3 B cell subtypes, 2 NK cell subtypes, and 12 T cell subtypes.
- b.** Bar plots showing the proportions of cells from each tissue class of origin (P=Primary, M=Metastatic) for T cell subtypes with a significant difference in proportion. (permutation test, $p < 0.05$, $\log_2FC > 2$)
- c.** The beeswarm plot illustrating the differential abundance analysis of NK cells, uncovering variations in cell abundance across neighborhoods, with a focus on specific NK cell subtypes, categorized by tissue classes of origin in primary and metastatic breast cancer.
- d.** Bar plots showing the proportions of cells from each tissue class of origin (P=Primary, M=Metastatic) for B cell subtypes with a significant difference in proportion. (permutation test, $p < 0.05$, $\log_2FC > 2$)
- e.** Differential Hallmark gene signature pathway activity scores within T cells across metastatic status.
- f.** Hypoxia score comparison between primary and metastatic breast cancer T cells. Significance indicated as (***, $p < 0.001$).
- g.** OXPHOS score comparison between primary and metastatic breast cancer T cells. Significance indicated as (***, $p < 0.001$).
- h.** Circle plot showing the top 10% of predominant cell-cell interactions in the primary tumor microenvironment, with arrow colors indicating the source (sender) of the interaction, and arrow thickness representing the strength of interaction.
- i.** Circle plot showing the top 10% of predominant cell-cell interactions in the metastatic tumor microenvironment, with arrow colors indicating the source (sender) of the interaction, and arrow thickness representing the strength of interaction.

It is made available under a [CC-BY-NC 4.0 International license](https://creativecommons.org/licenses/by-nc/4.0/).

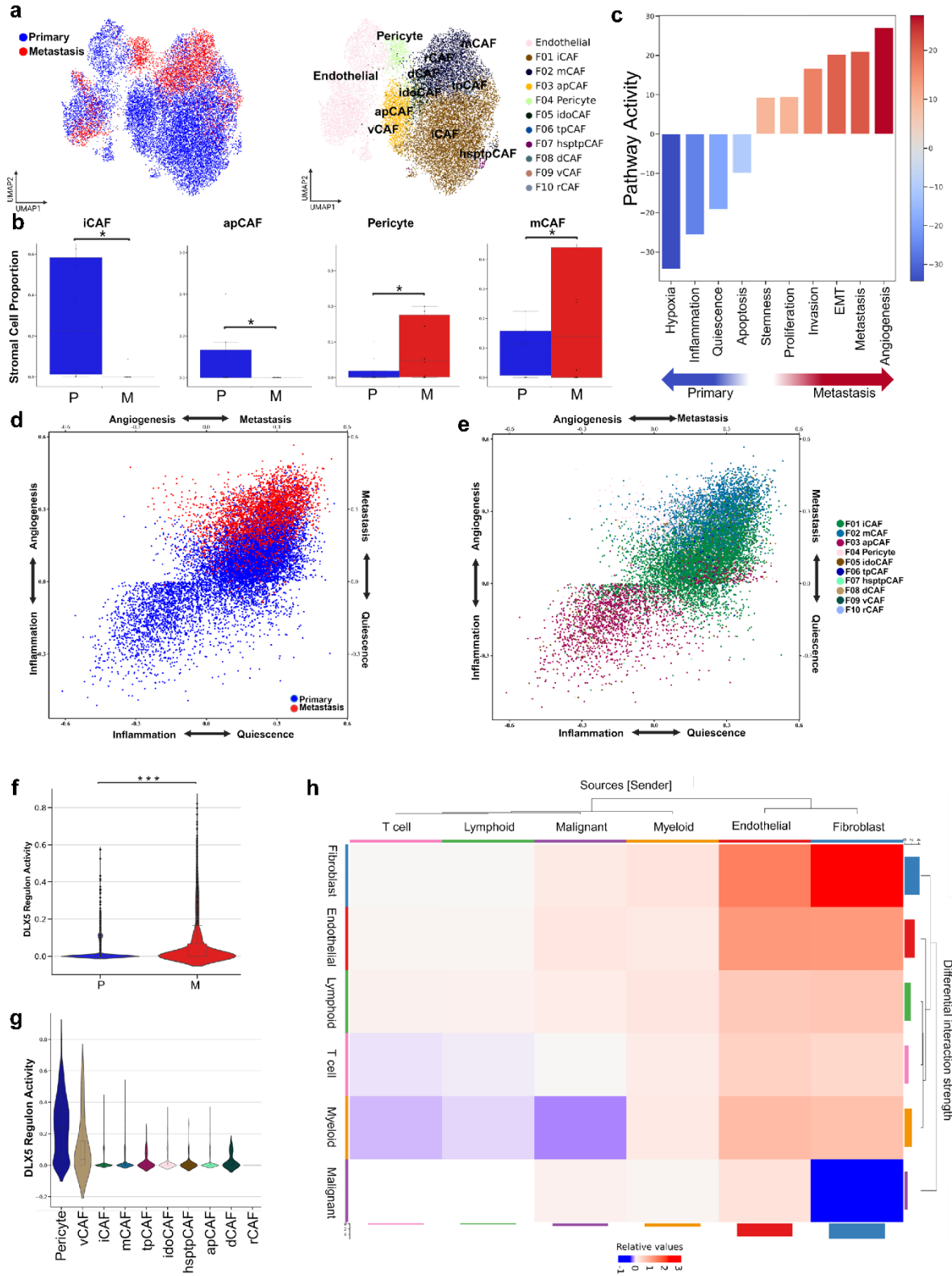


Fig.5 Stromal Cell Remodeling and Fibroblast Subtype Dynamics in Metastatic vs. Primary Breast Cancer TME

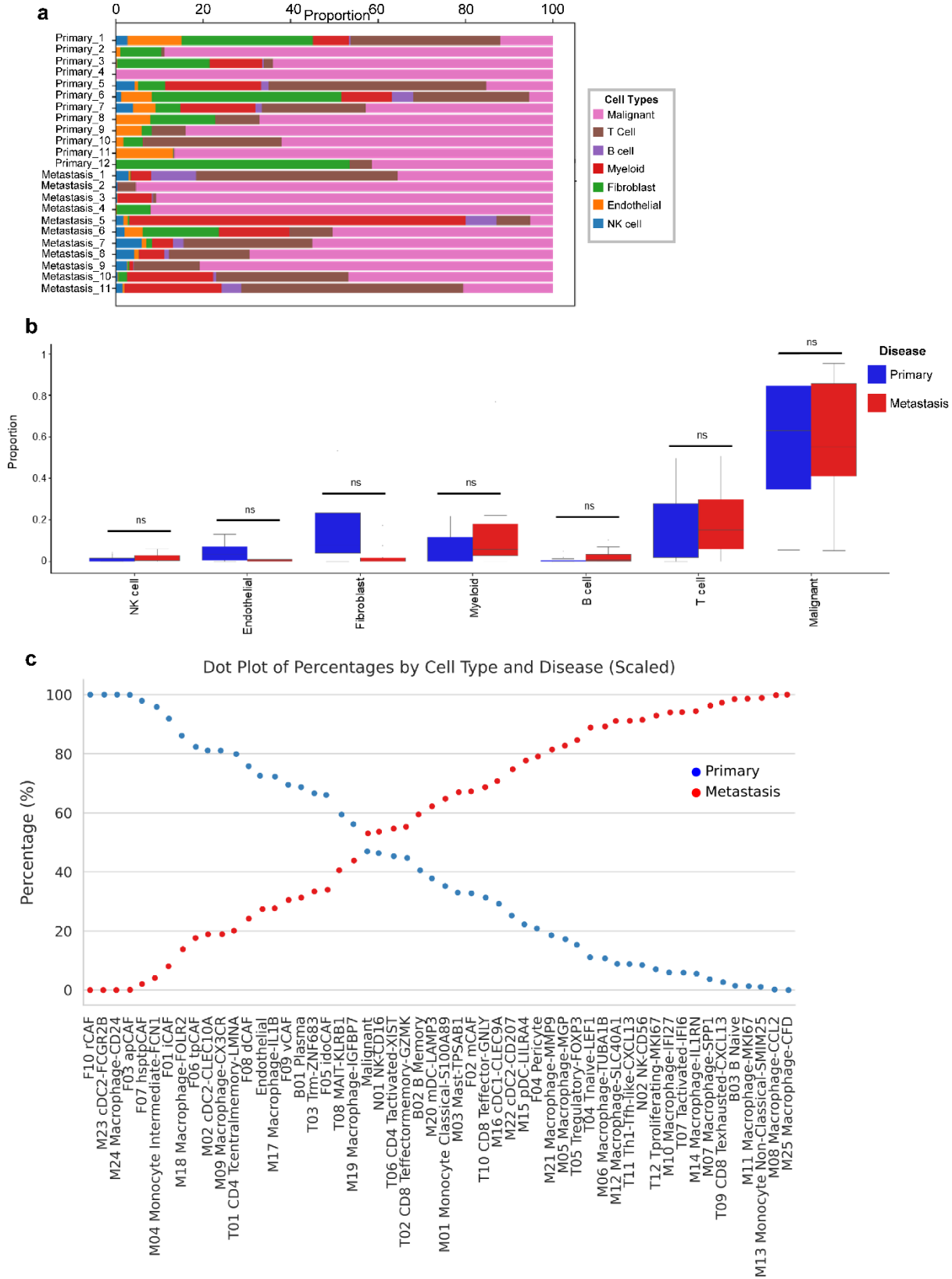
a. UMAP visualizations of 17339 stromal cells. Cells are colored by tissue class of origin on the left and grouped by one of 11 identified subtypes on the right, including 10 Fibroblast and an endothelial cell subtype. **b.** Bar plots showing the proportions of cells from each tissue class of origin (P=Primary, M=Metastatic) for Fibroblast cell subtypes with a significant difference in proportion (permutation test, $p < 0.05$, $\log_2FC > 2$). **c.** Differential CANCERSEA gene signature pathway activity scores within fibroblast cells across metastatic status. **d.** Cellular states plot to show CANCERSEA gene signature activity for Angiogenesis, Metastasis, Inflammation, and Quiescence across the multiple fibroblast subtypes. **e.** Cellular states plot to show CANCERSEA gene signature activity for Angiogenesis, Metastasis, Inflammation, and Quiescence across the multiple fibroblast subtypes. **f.** DLX5 Regulon Activity score comparison between primary and metastatic breast cancer stromal cells. Significance indicated as (***, $p < 0.001$). **g.** DLX5 Regulon Activity score comparison between primary and metastatic breast cancer stromal cells. Significance indicated as (***, $p < 0.001$). **h.** Heatmap comparing the differential number of interactions and interaction strength between primary and metastatic tumor microenvironments. The top-colored bar plot indicates incoming signaling, while the right-colored bar plot represents outgoing signaling for each cell type. Red indicates increased signaling in metastatic tumors, whereas blue represents increased signaling in primary tumors. Bar height reflects the magnitude of change in interaction number or strength between the two conditions.

Fig.5 Stromal Cell Remodeling and Fibroblast Subtype Dynamics in Metastatic vs. Primary Breast Cancer TME

- a.** UMAP visualizations of 17339 stromal cells. Cells are colored by tissue class of origin on the left and grouped by one of 11 identified subtypes on the right, including 10 Fibroblast and an endothelial cell subtype.
- b.** Bar plots showing the proportions of cells from each tissue class of origin (P=Primary, M=Metastatic) for Fibroblast cell subtypes with a significant difference in proportion. (permutation test, $p < 0.05$, $\log_2FC > 2$)
- c.** Differential CANCERSEA gene signature pathway activity scores within fibroblast cells across metastatic status.
- d.** Cellular states plot to show CANCERSEA gene signature activity for Angiogenesis, Metastasis, Inflammation, and Quiescence across the metastatic status within fibroblast cells.
- e.** Cellular states plot to show CANCERSEA gene signature activity for Angiogenesis, Metastasis, Inflammation, and Quiescence across the multiple fibroblast subtypes.
- f.** DLX5 Regulon Activity score comparison between primary and metastatic breast cancer stromal cells. Significance indicated as (***, $p < 0.001$).
- g.** DLX5 Regulon Activity score comparison between primary and metastatic breast cancer stromal cells. Significance indicated as (***, $p < 0.001$).
- h.** Heatmap comparing the differential number of interactions and interaction strength between primary and metastatic tumor microenvironments. The top-colored bar plot indicates incoming signaling, while the right-colored bar plot represents outgoing signaling for each cell type. Red indicates increased signaling in metastatic tumors, whereas blue represents increased signaling in primary tumors. Bar height reflects the magnitude of change in interaction number or strength between the two conditions.

Supplementary Figures

It is made available under a [CC-BY-NC 4.0 International license](https://creativecommons.org/licenses/by-nc/4.0/).



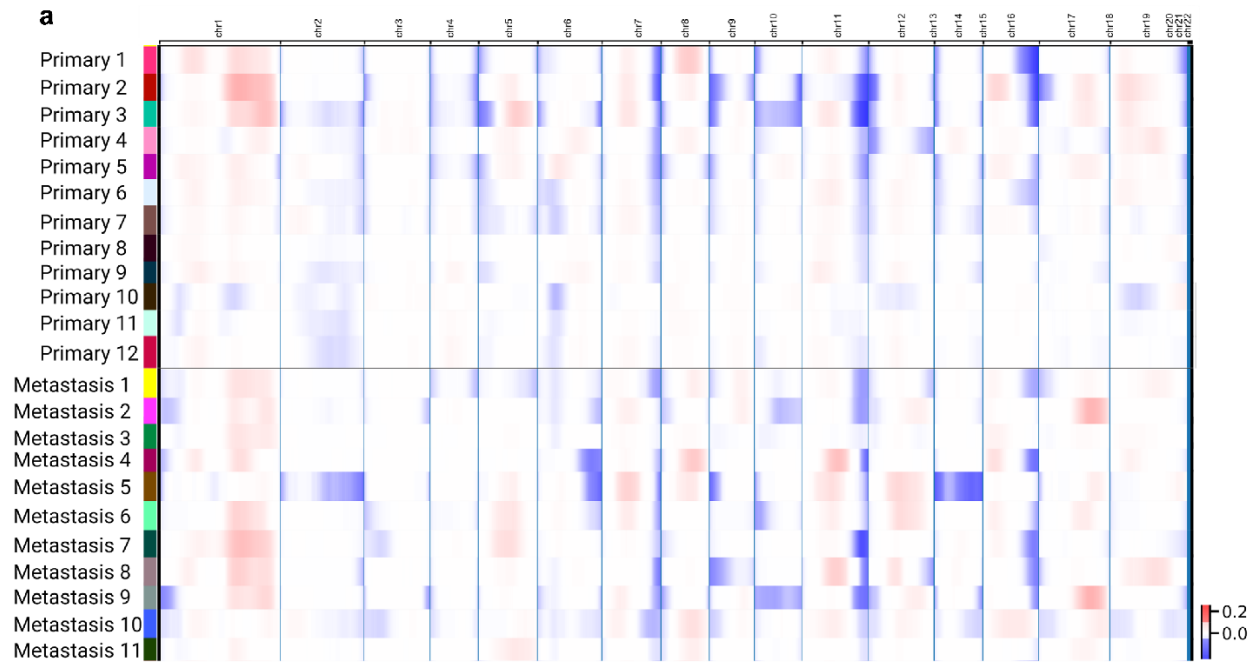
Supp Fig.1

a. Stacked barplot showing the proportion of seven major cell types across individual samples. b. Boxplots illustrating changes in the proportions of seven major cell types between primary and metastatic breast cancer samples. (permutation test, $p < 0.05$, $\log_2FC > 2$) c. Dot plot illustrating the percentage of each cell subtype from primary versus metastatic tumors, highlighting each cell subtype's distribution and comparative abundance across the two conditions.

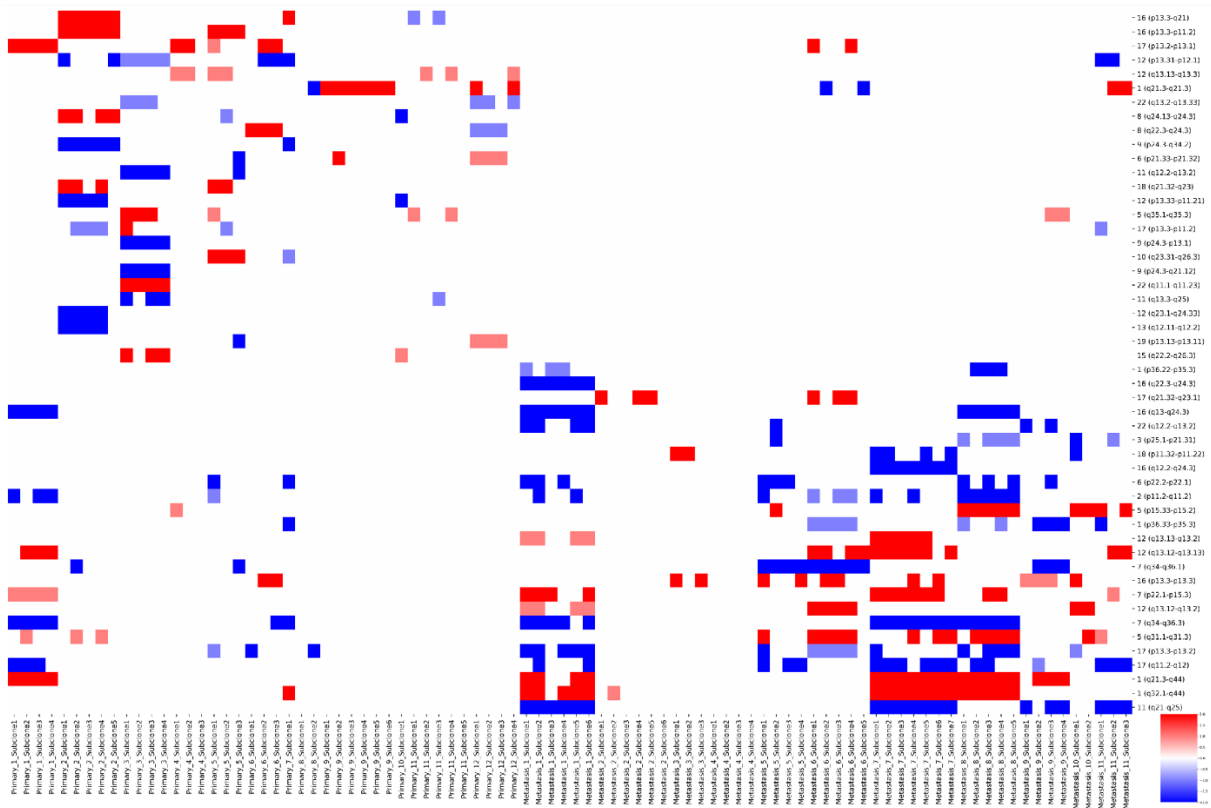
Supp Fig.1

- a.** Stacked barplot showing the proportion of seven major cell types across individual samples.
- b.** Boxplots illustrating changes in the proportions of seven major cell types between primary and metastatic breast cancer samples. (permutation test, $p < 0.05$, $\log_2FC > 2$)
- c.** Dot plot illustrating the percentage of each cell subtype from primary versus metastatic tumors, highlighting each cell subtype's distribution and comparative abundance across the two conditions.

It is made available under a [CC-BY-NC 4.0 International license](https://creativecommons.org/licenses/by-nc/4.0/).



b



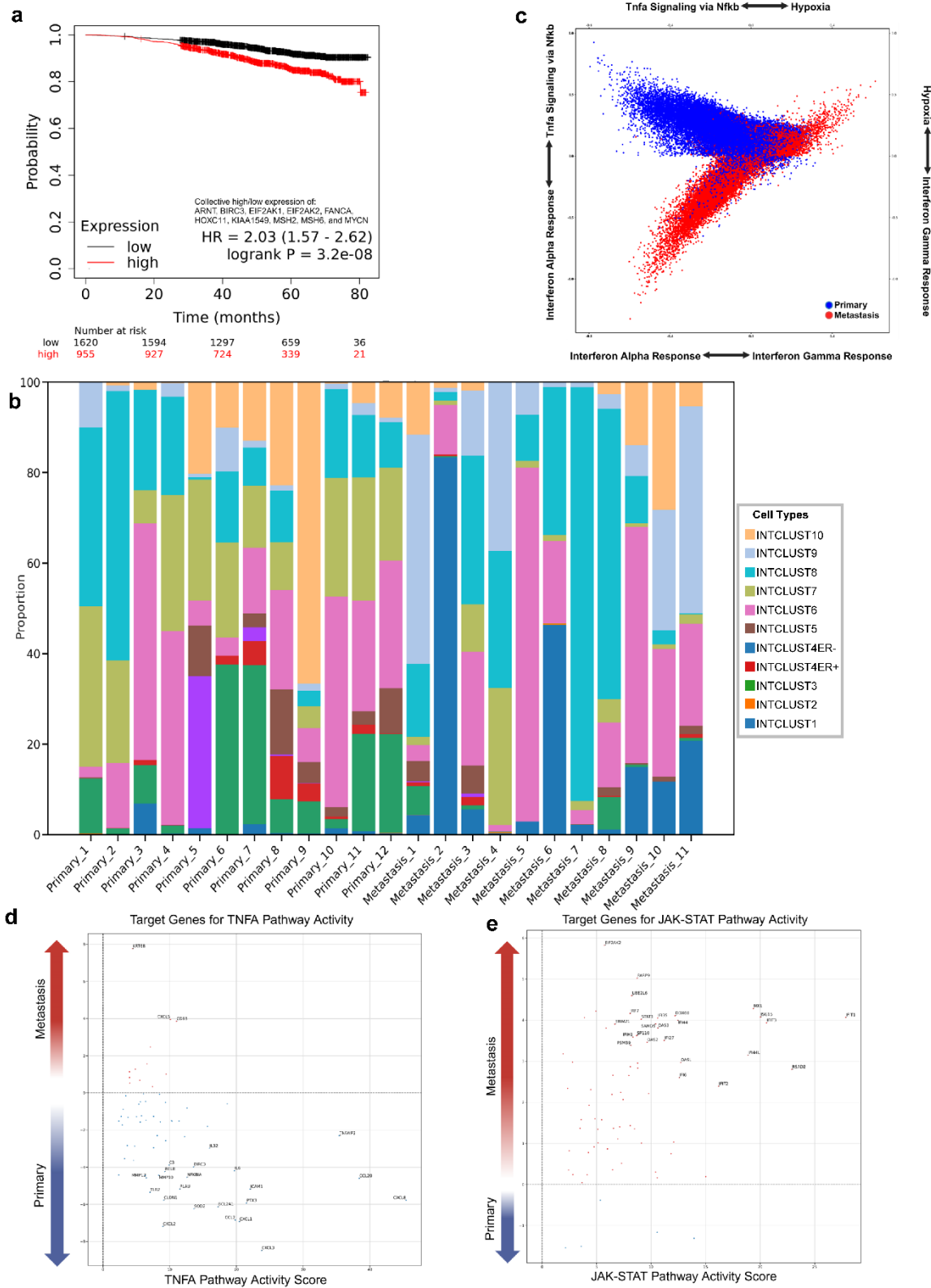
Supp Fig.2

a. Heatmap showing CNV profiles for malignant cells in individual primary and metastatic tumor samples, with T cells as a reference. **b.** Heatmap depicting CNV profiles from all tumor subpopulations per patient, comparing the top 25 copy number alterations across chromosomal arms between primary and metastatic breast cancer subclones.

Supp Fig.2

- a.** Heatmap showing CNV profiles for malignant cells in individual primary and metastatic tumor samples, with T cells as a reference.
- b.** Heatmap depicting CNV profiles from all tumor subpopulations per patient, comparing the top 25 copy number alterations across chromosomal arms between primary and metastatic breast cancer subclones.

It is made available under a [CC-BY-NC 4.0 International license](https://creativecommons.org/licenses/by-nc/4.0/).



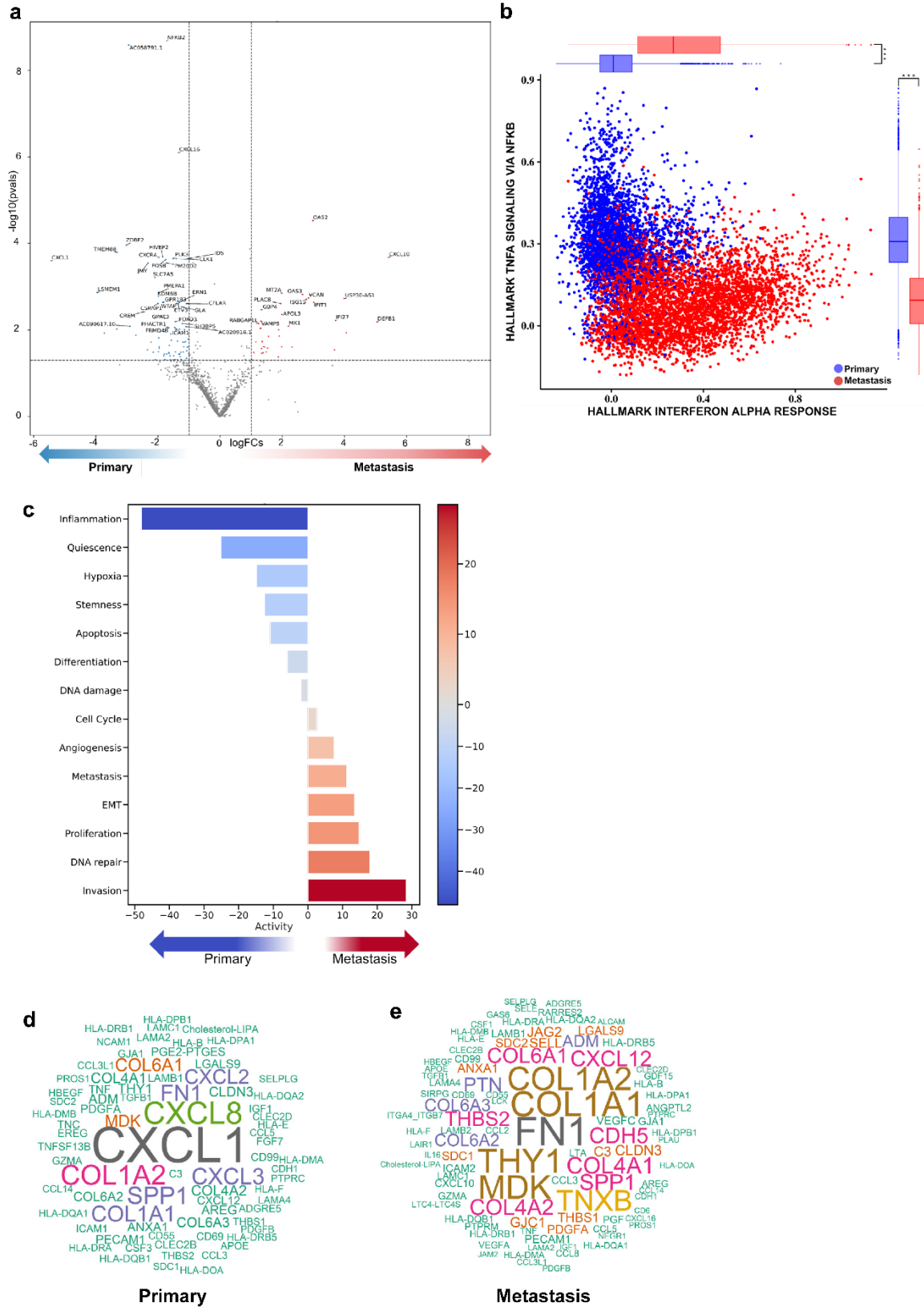
Supp Fig.3

a. Kaplan-Meier survival curves for estrogen receptor-positive tumors among 2,575 patients, comparing survival outcomes associated with high versus low expression of the genes ARNT, BIRC3, EIF2AK1, EIF2AK2, FANCA, HOXC11, KIAA1549, MSH2, MSH6, and MYCN. **b.** Stacked bar plot showing the proportion of Integrative Clusters (IntClust) across individual samples, based on classification using the top 200 differentially expressed genes from the METABRIC study. **c.** Cellular states plot to show Hallmark gene signature activity for Hypoxia, Interferon Gamma Response, Interferon Alpha Response, and TNFA Signaling via NFKB across the metastatic status within malignant cells. **d.** Plot of the top 25 target genes affected by TNFA pathway activity, highlighting their weight and statistical significance in primary versus metastatic malignant cells based on PROGENY pathway activity. **e.** Plot of the top 25 target genes affected by JAK-STAT pathway activity, highlighting their weight and statistical significance in primary versus metastatic malignant cells based on PROGENY pathway activity.

Supp Fig.3

- a.** Kaplan-Meier survival curves for estrogen receptor-positive tumors among 2,575 patients, comparing survival outcomes associated with high versus low expression of the genes ARNT, BIRC3, EIF2AK1, EIF2AK2, FANCA, HOXC11, KIAA1549, MSH2, MSH6, and MYCN.
- b.** Stacked bar plot showing the proportion of Integrative Clusters (IntClust) across individual samples, based on classification using the top 200 differentially expressed genes from the METABRIC study.
- c.** Cellular states plot to show Hallmark gene signature activity for Hypoxia, Interferon Gamma Response, Interferon Alpha Response, and TNF- α Signaling via NF κ B across the metastatic status within malignant cells.
- d.** Plot of the top 25 target genes affected by TNF- α pathway activity, highlighting their weight and statistical significance in primary versus metastatic malignant cells based on PROGENY pathway activity.
- e.** Plot of the top 25 target genes affected by JAK-STAT pathway activity, highlighting their weight and statistical significance in primary versus metastatic malignant cells based on PROGENY pathway activity.

It is made available under a [CC-BY-NC 4.0 International license](https://creativecommons.org/licenses/by-nc/4.0/).



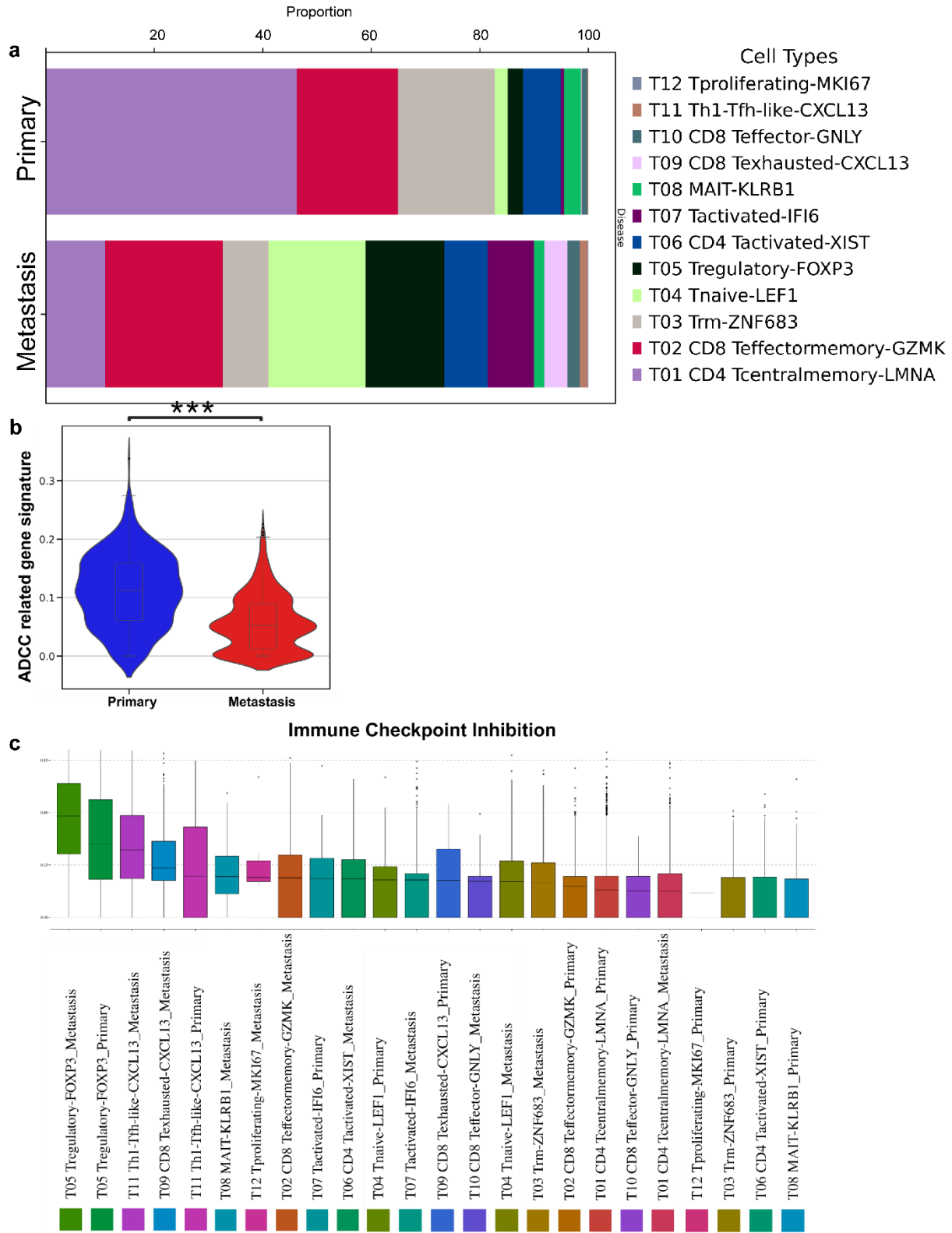
Supp Fig.4

a. Volcano plot illustrating the differential expression of genes in myeloid cells when comparing primary versus metastatic tumors. **b.** Cellular states plot showing Hallmark gene signature activity for Interferon Alpha Response versus TNFA Signaling via NFkB in myeloid cell populations. **c.** Differential CANCERSEA gene signature pathway activity scores within myeloid cells across metastatic status. **d.** Word cloud plot illustrating the enriched ligands in primary breast tumor microenvironment compared to metastatic cancer. **e.** Word cloud plot illustrating the enriched ligands in metastatic breast tumor microenvironment compared to primary breast cancer.

Supp Fig.4

- a.** Volcano plot illustrating the differential expression of genes in myeloid cells when comparing primary versus metastatic tumors.
- b.** Cellular states plot showing Hallmark gene signature activity for Interferon Alpha Response versus TNF- α Signaling via NF κ B in myeloid cell populations.
- c.** Differential CANCERSEA gene signature pathway activity scores within myeloid cells across metastatic status.
- d.** Word cloud plot illustrating the enriched ligands in primary breast tumor microenvironment compared to metastatic cancer
- e.** Word cloud plot illustrating the enriched ligands in metastatic breast tumor microenvironment compared to primary breast cancer

It is made available under a [CC-BY-NC 4.0 International license](https://creativecommons.org/licenses/by-nc/4.0/).



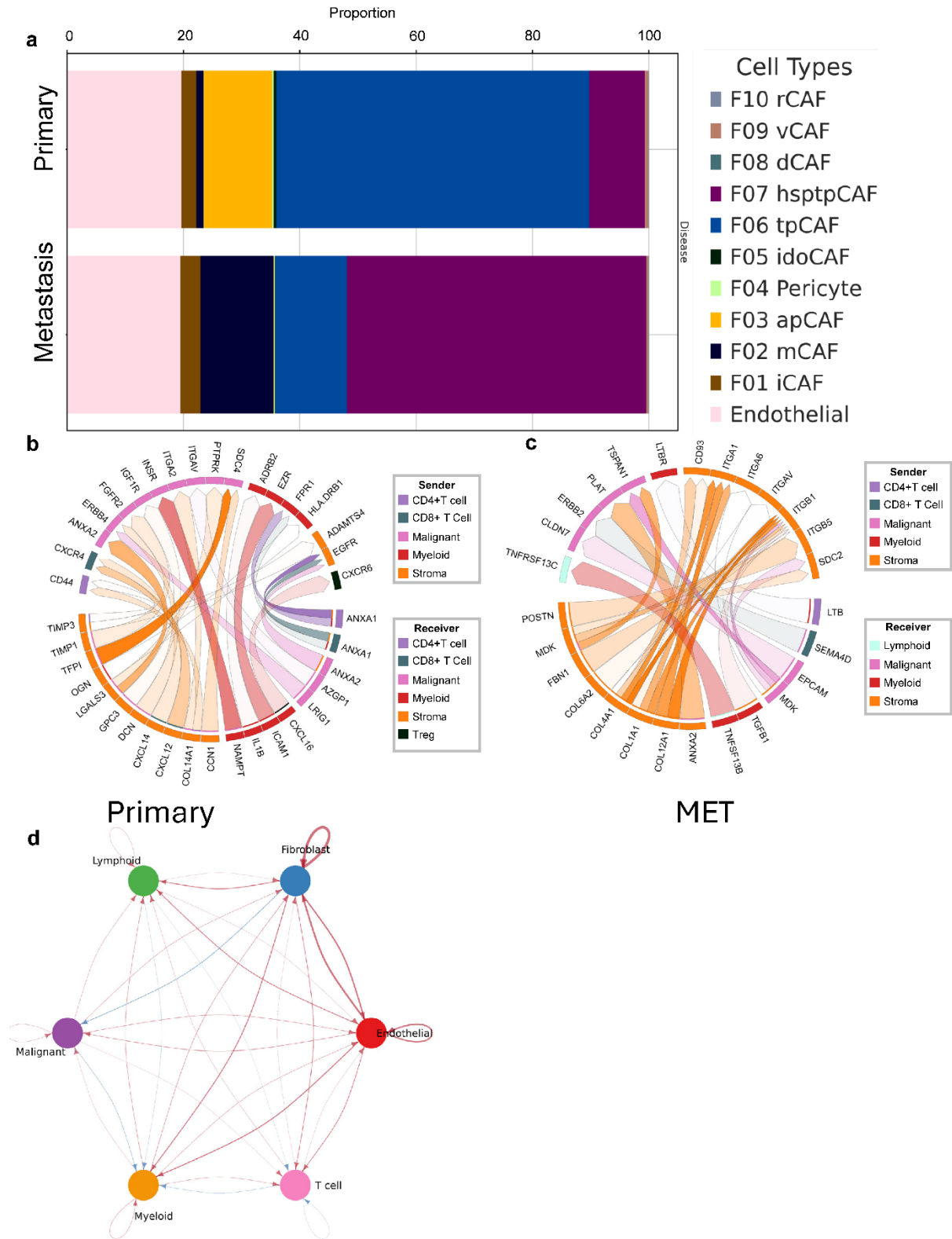
Supp Fig.5

a. Stacked bar plot showing the proportion of T cell subtypes across the metastatic status. b. Antibody-dependent cellular cytotoxicity gene signature comparison between primary and metastatic breast cancer NK cells. Significance indicated as (***, $p < 0.001$). c. Immune checkpoint inhibition levels compared across different cell types in primary and metastatic breast cancers

Supp Fig.5

- a. Stacked bar plot showing the proportion of T cell subtypes across the metastatic status.
- b. Antibody-dependent cellular cytotoxicity gene signature comparison between primary and metastatic breast cancer NK cells. Significance indicated as (***, $p < 0.001$).
- c. Immune checkpoint inhibition levels compared across different cell types in primary and metastatic breast cancers

It is made available under a [CC-BY-NC 4.0 International license](https://creativecommons.org/licenses/by-nc/4.0/).



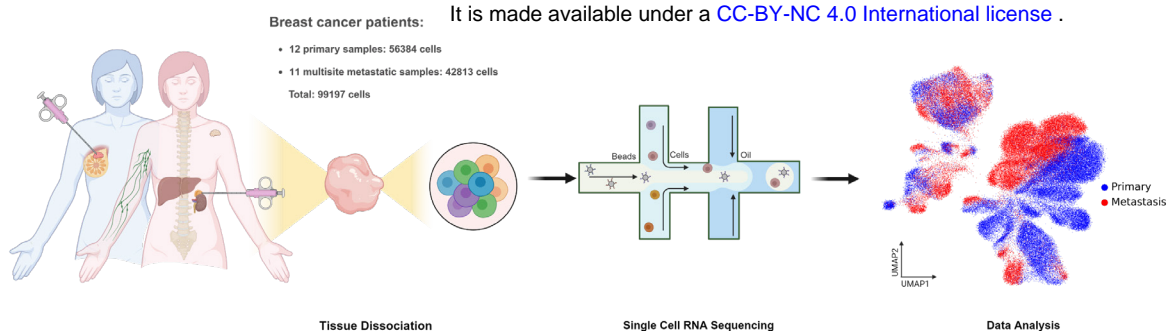
Supp Fig.6

a. Stacked bar plot showing the proportion of stromal cell subtypes across the metastatic status. b. Visualization of the top 25 ligand-receptor interactions between different cell types in the primary tumor microenvironment, as analyzed using the MultiNicheNet. c. Visualization of the top 25 ligand-receptor interactions between different cell types in the metastatic tumor microenvironment, as analyzed using the MultiNicheNet. d. Comparative circle plot visualizing differential cell-cell communication between primary and metastatic tumor microenvironments, where red edges indicate increased interactions in metastatic tumors and blue edges denote increased interactions in primary tumors.

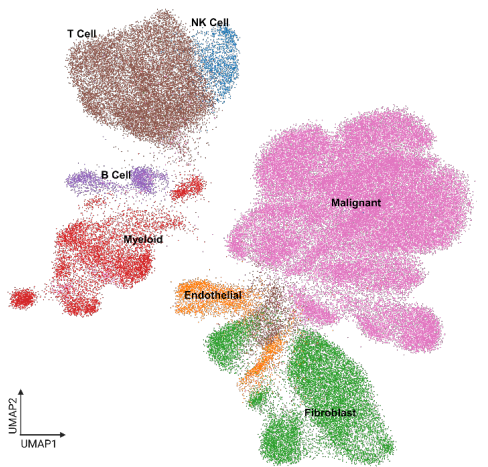
Supp Fig.6

- a.** Stacked bar plot showing the proportion of stromal cell subtypes across the metastatic status.
- b.** Visualization of the top 25 ligand-receptor interactions between different cell types in the primary tumor microenvironment, as analyzed using the MultiNicheNet.
- c.** Visualization of the top 25 ligand-receptor interactions between different cell types in the metastatic tumor microenvironment, as analyzed using the MultiNicheNet.
- d.** Comparative circle plot visualizing differential cell-cell communication between primary and metastatic tumor microenvironments, where red edges indicate increased interactions in metastatic tumors and blue edges denote increased interactions in primary tumors

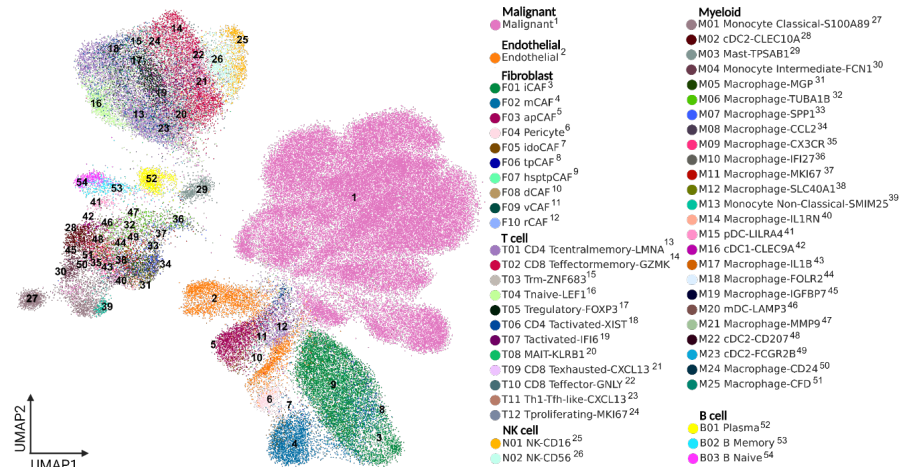
a



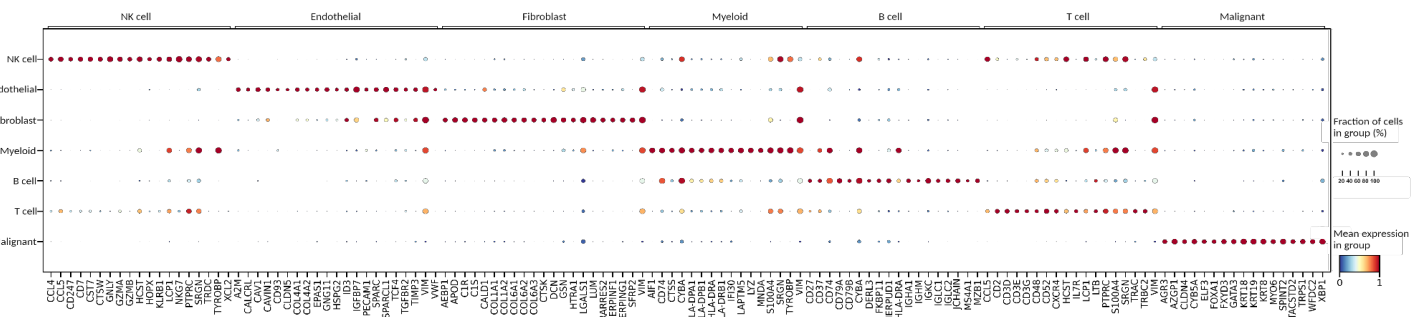
b



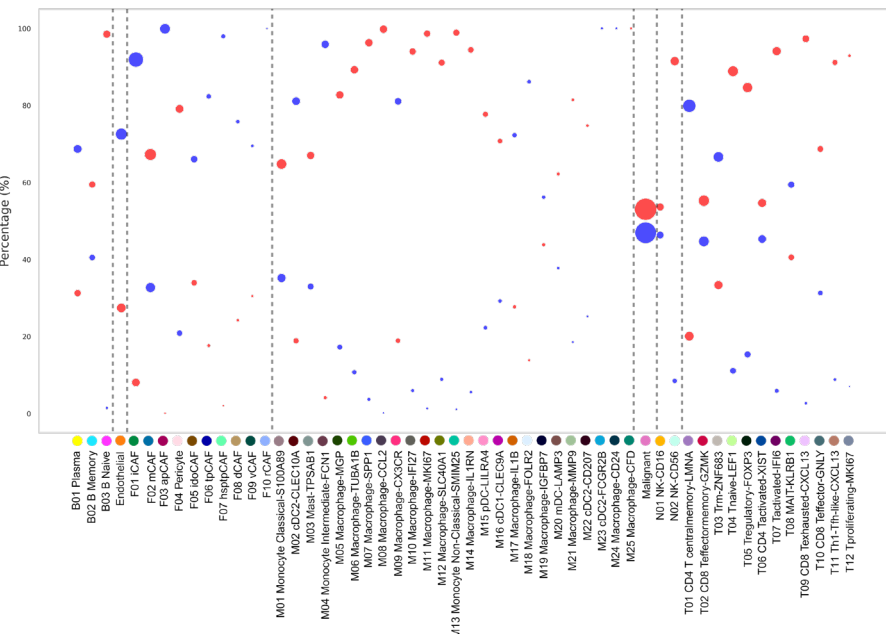
c



d



e



f

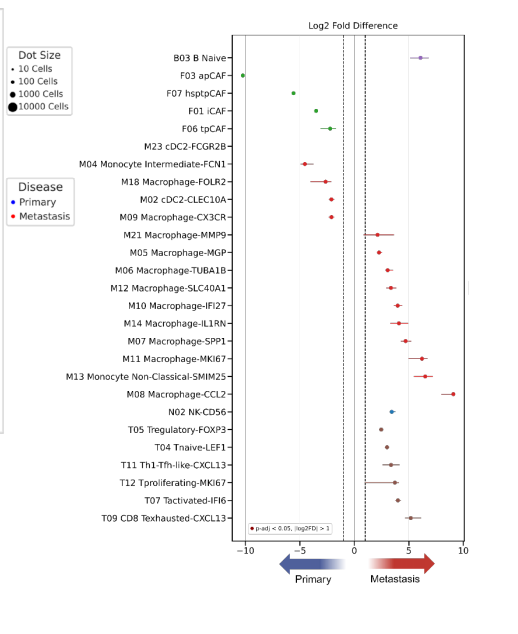


Fig.1 Single-Cell Landscape of Primary and Metastatic ER+ Breast Cancer: Characterization and Cellular Subtype Differences
a. Sample collection workflow and data analysis overview. (Created in BioRender. [BioRender.com/y20n356](https://www.biorender.com/y20n356)) **b.** UMAP visualization of the unified cell map, consisting of 7 major cell types colored by clusters. **c.** UMAP visualization of the unified cell map, sub-segmenting by 54 minor cell types uniquely colored within the 7 major cell types. **d.** Dot plot representation of differentially expressed genes across the 7 major cell type clusters. **e.** Relative percentage of 54 minor cell types by metastatic status and corresponding cell number indicated by dot size. **f.** Significant changes in minor cell type proportions between primary and metastatic breast cancer

Fig.1 Single-Cell Landscape of Primary and Metastatic ER+ Breast Cancer: Characterization and Cellular Subtype Differences

- a.** Sample collection workflow and data analysis overview. (Created in BioRender. BioRender.com/y20n356)
- b.** UMAP visualization of the unified cell map, consisting of 7 major cell types colored by clusters.
- c.** UMAP visualization of the unified cell map, sub-segmenting by 54 minor cell types uniquely colored within the 7 major cell types.
- d.** Dot plot representation of differentially expressed genes across the 7 major cell type clusters.
- e.** Relative percentage of 54 minor cell types by metastatic status and corresponding cell number indicated by dot size.
- f.** Significant changes in minor cell type proportions between primary and metastatic breast cancer

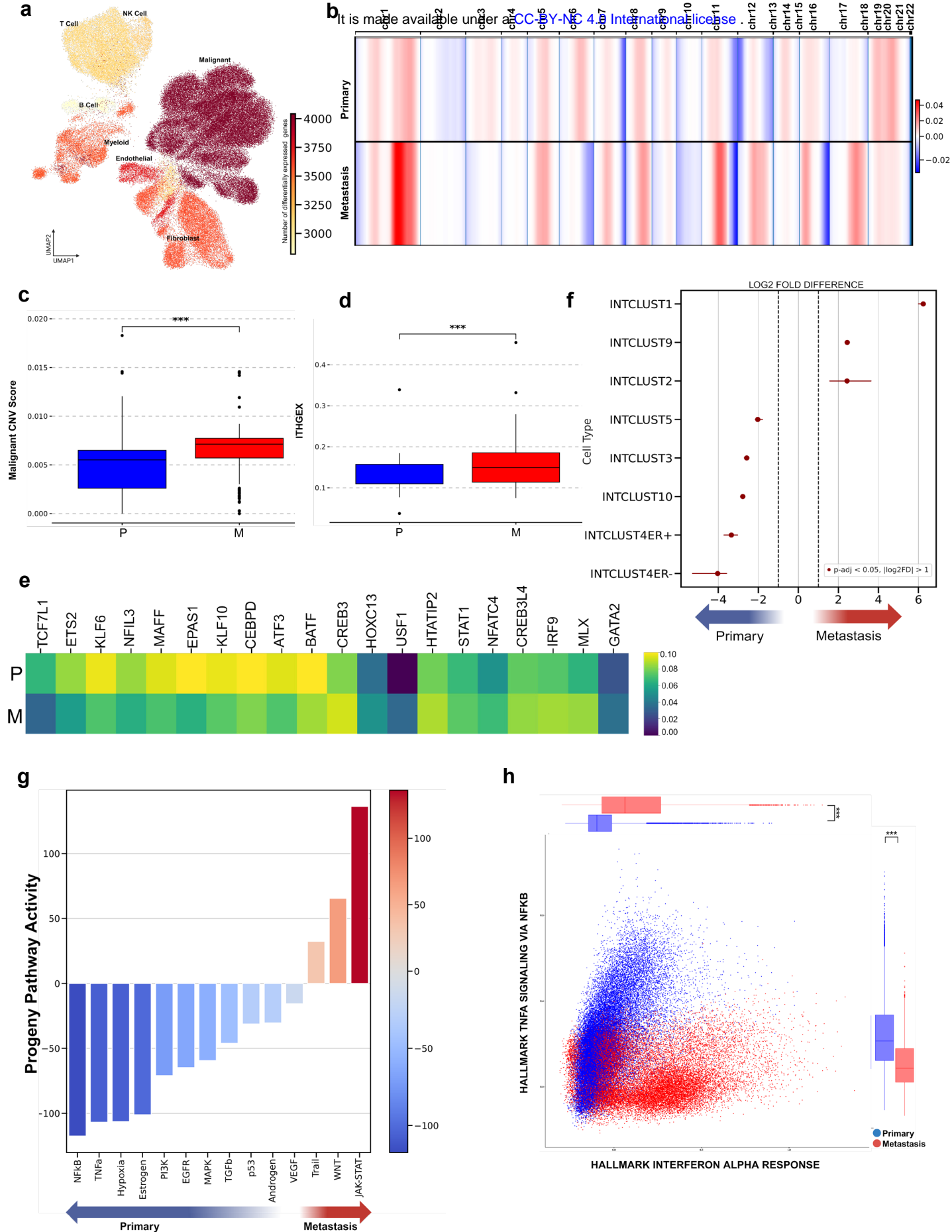


Fig.2 Genomic and Phenotypic Alterations in Malignant Cells of Primary vs. Metastatic Breast Cancer: Insights from CNV Analysis, Gene Regulatory Networks, and Integrative Clusters

a. UMAP visualization of the number of differentially expressed genes within all major cell type clusters. **b.** Chromosomal CNV alteration score by metastatic status, with copy number gain and copy number loss indicated by red and blue, respectively. **c.** CNV score comparison between primary and metastatic breast cancer malignant cells. Significance indicated as (***) $p < 0.001$. **d.** ITGEX score comparison between primary and metastatic breast cancer malignant cells. Significance indicated as (***) $p < 0.001$. **e.** Regulon activity scores by metastatic status within malignant cells. **f.** Significant (Log2 FC) integrative cluster differences between primary and metastatic malignant cells ($p\text{-adj} < 0.05$). **g.** Differential progeny pathway activity scores within malignant cells across metastatic status. **h.** Cellular states plot showing hallmark gene signature activity for Interferon Alpha response versus TNFa signaling via NFKB in malignant cell populations.

Fig.2 Genomic and Phenotypic Alterations in Malignant Cells of Primary vs. Metastatic Breast Cancer: Insights from CNV Analysis, Gene Regulatory Networks, and Integrative Clusters

- a.** UMAP visualization of the number of differentially expressed genes within all major cell type clusters.
- b.** Chromosomal CNV alternation score by metastatic status, with copy number gain and copy number loss indicated by red and blue, respectively.
- c.** CNV score comparison between primary and metastatic breast cancer malignant cells. Significance indicated as (***, $p < 0.001$).
- d.** ITGEX score comparison between primary and metastatic breast cancer malignant cells. Significance indicated as (***, $p < 0.001$).
- e.** Regulon activity scores by metastatic status within malignant cells.
- f.** Significant ($\text{Log}_2 \text{FC}$) integrative cluster differences between primary and metastatic malignant cells ($p\text{-adj} < 0.05$).
- g.** Differential progeny pathway activity scores within malignant cells across metastatic status.
- h.** Cellular states plot showing hallmark gene signature activity for Interferon Alpha response versus TNF- α signaling via NF κ B in malignant cell populations.

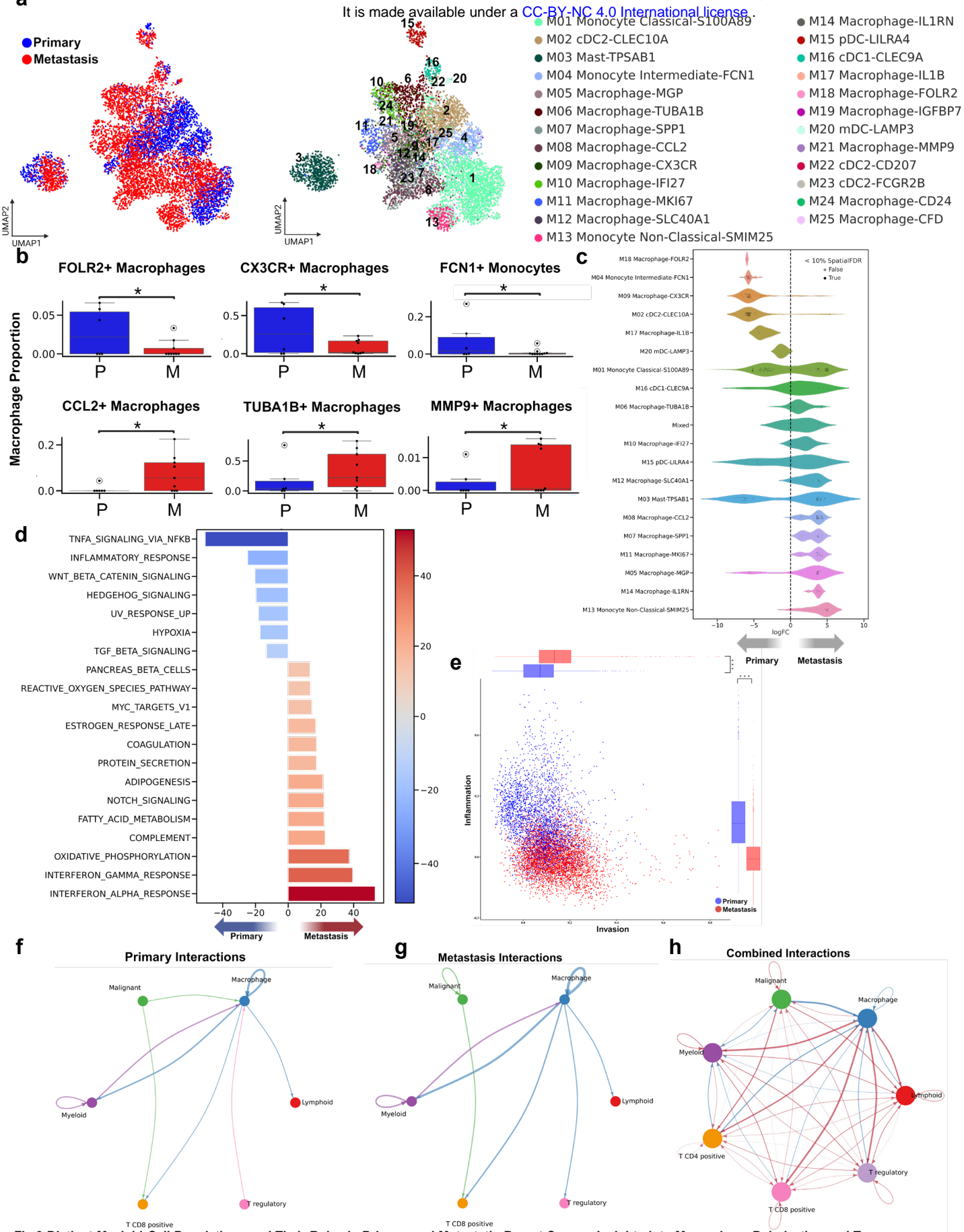


Fig.3 Distinct Myeloid Cell Populations and Their Roles in Primary and Metastatic Breast Cancer: Insights into Macrophage Polarization and Tumor Microenvironment Interactions
a. UMAP visualizations of 7651 myeloid cells. Cells are colored by tissue class of origin (left) and grouped by one of 25 identified sub-types (right), including 6 Dendritic, 3 Monocytic, 1 Mast, and 15 Macrophage lineages. **b.** Bar plots showing the proportions of cells from each tissue class of origin (P=Primary, M=Metastatic) for macrophage subtypes with a significant difference in proportion. (permutation test, $p < 0.05$, $\log_2FC > 2$) **c.** The beeswarm plot illustrates the differential abundance analysis of myeloid cell subtypes, revealing distinct variations in cell abundance across neighborhoods, stratified by tissue classes of origin in primary and metastatic breast cancer. **d.** Differential Hallmark gene signature pathway activity scores within malignant cells across metastatic status. **e.** Cellular states plot showing CANCERSEA gene signature activity for Inflammation versus Invasion in myeloid cell populations. **f.** Circle plot displaying the top 15% of primary tumor microenvironment cell-cell interactions. Arrow colors indicate sender and thickness represents interaction strength. **g.** Circle plot displaying the top 15% of metastatic tumor microenvironment cell-cell interactions. **h.** Comparative circle plot visualizing differential cell-cell communication between primary and metastatic tumor microenvironments, where red edges indicate increased interactions in metastatic tumors and blue edges denote increased interactions in primary tumors

Fig.3 Distinct Myeloid Cell Populations and Their Roles in Primary and Metastatic Breast Cancer: Insights into Macrophage Polarization and Tumor Microenvironment Interactions

- a.** UMAP visualizations of 7651 myeloid cells. Cells are colored by tissue class of origin (left) and grouped by one of 25 identified sub-types (right), including 6 Dendritic, 3 Monocytic, 1 Mast, and 15 Macrophage lineages.
- b.** Bar plots showing the proportions of cells from each tissue class of origin (P=Primary, M=Metastatic) for macrophage subtypes with a significant difference in proportion. (permutation test, $p < 0.05$, $\log_2FC > 2$)
- c.** The beeswarm plot illustrates the differential abundance analysis of myeloid cell subtypes, revealing distinct variations in cell abundance across neighborhoods, stratified by tissue classes of origin in primary and metastatic breast cancer.
- d.** Differential Hallmark gene signature pathway activity scores within malignant cells across metastatic status.
- e.** Cellular states plot showing CANCERSEA gene signature activity for Inflammation versus Invasion in myeloid cell populations.
- f.** Circle plot displaying the top 15% of primary tumor microenvironment cell-cell interactions. Arrow colors indicate sender and thickness represents interaction strength.
- g.** Circle plot displaying the top 15% of metastatic tumor microenvironment cell-cell interactions. Arrow colors indicate sender and thickness represents interaction strength.
- h.** Comparative circle plot visualizing differential cell-cell communication between primary and metastatic tumor microenvironments, where red edges indicate increased interactions in metastatic tumors and blue edges denote increased interactions in primary tumors

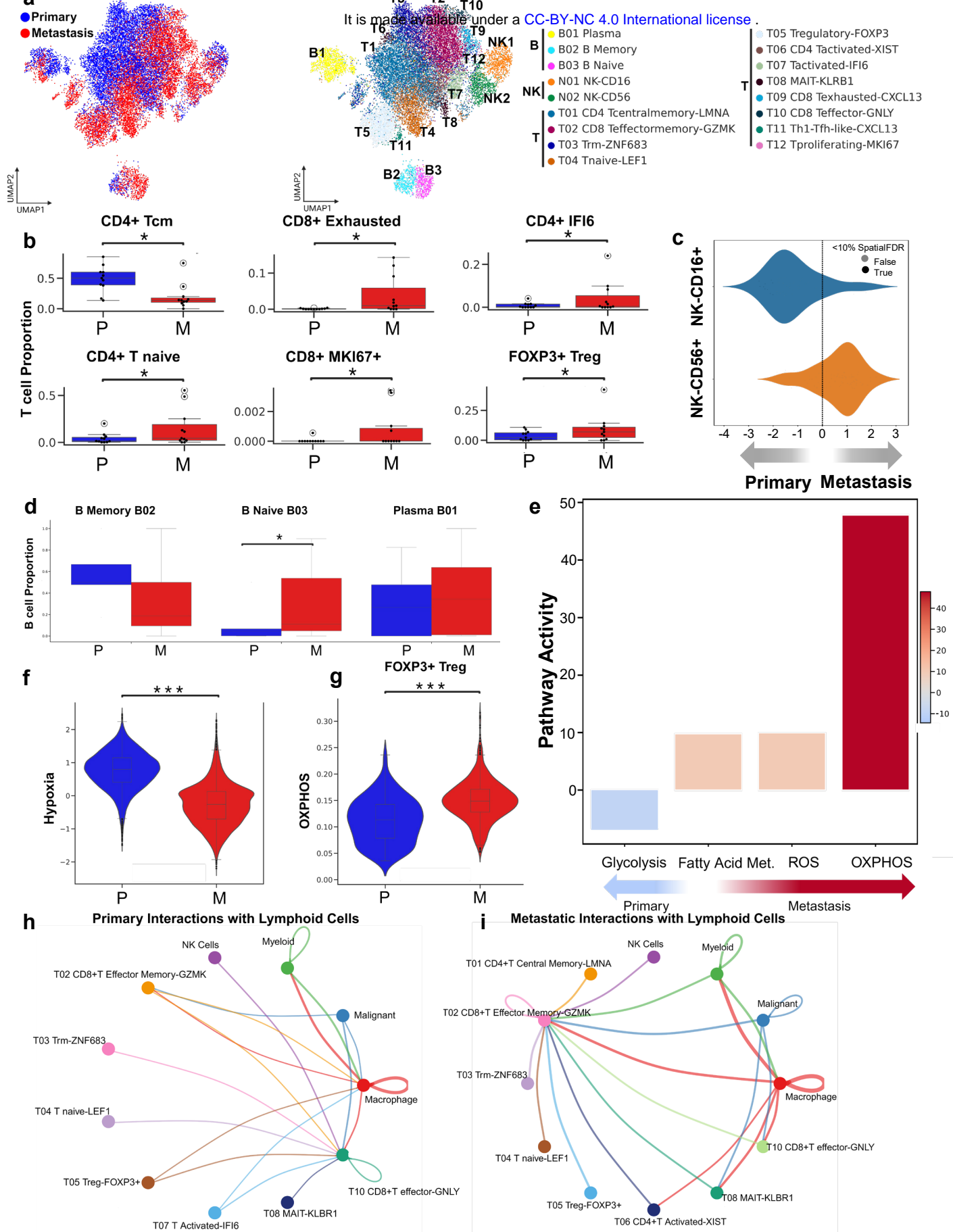


Fig.4 Lymphoid Cell Landscape and Functional Shifts in Primary vs. Metastatic Breast Cancer Tumors

a. UMAP visualizations of 19104 lymphoid cells. Cells are colored by tissue class of origin and grouped by one of 17 identified subtypes on the right, including 3 B cell subtypes, 2 NK cell subtypes, and 12 T cell subtypes. **b.** Bar plots showing the proportions of cells from each tissue class of origin (P=Primary, M=Metastatic) for T cell subtypes with a significant difference in proportion. (permutation test, $p < 0.05$, $\log_2FC > 2$) **c.** The beeswarm plot illustrating the differential abundance analysis of NK cells, uncovering variations in cell abundance across neighborhoods, with a focus on specific NK cell subtypes, categorized by tissue classes of origin in primary and metastatic breast cancer. **d.** Bar plots showing the proportions of cells from each tissue class of origin (P=Primary, M=Metastatic) for B cell subtypes with a significant difference in proportion. (permutation test, $p < 0.05$, $\log_2FC > 2$) **e.** Differential Hallmark gene signature pathway activity scores within T cells across metastatic status. **f.** Hypoxia score comparison between primary and metastatic breast cancer T cells. Significance indicated as (***, $p < 0.001$). **g.** OXPPOS score comparison between primary and metastatic breast cancer T cells. Significance indicated as (***, $p < 0.001$). **h.** Circle plot showing the top 10% of predominant cell-cell interactions in the primary tumor microenvironment, with arrow colors indicating the source (sender) of the interaction, and arrow thickness representing the strength of interaction. **i.** Circle plot showing the top 10% of predominant cell-cell interactions in the metastatic tumor microenvironment, with arrow colors indicating the source (sender) of the interaction, and arrow thickness representing the strength of interaction.

Fig.4 Lymphoid Cell Landscape and Functional Shifts in Primary vs. Metastatic Breast Cancer Tumors"

- a.** UMAP visualizations of 19104 lymphoid cells. Cells are colored by tissue class of origin on the left and grouped by one of 17 identified subtypes on the right, including 3 B cell subtypes, 2 NK cell subtypes, and 12 T cell subtypes.
- b.** Bar plots showing the proportions of cells from each tissue class of origin (P=Primary, M=Metastatic) for T cell subtypes with a significant difference in proportion. (permutation test, $p < 0.05$, $\log_2FC > 2$)
- c.** The beeswarm plot illustrating the differential abundance analysis of NK cells, uncovering variations in cell abundance across neighborhoods, with a focus on specific NK cell subtypes, categorized by tissue classes of origin in primary and metastatic breast cancer.
- d.** Bar plots showing the proportions of cells from each tissue class of origin (P=Primary, M=Metastatic) for B cell subtypes with a significant difference in proportion. (permutation test, $p < 0.05$, $\log_2FC > 2$)
- e.** Differential Hallmark gene signature pathway activity scores within T cells across metastatic status.
- f.** Hypoxia score comparison between primary and metastatic breast cancer T cells. Significance indicated as (***, $p < 0.001$).
- g.** OXPHOS score comparison between primary and metastatic breast cancer T cells. Significance indicated as (***, $p < 0.001$).
- h.** Circle plot showing the top 10% of predominant cell-cell interactions in the primary tumor microenvironment, with arrow colors indicating the source (sender) of the interaction, and arrow thickness representing the strength of interaction.
- i.** Circle plot showing the top 10% of predominant cell-cell interactions in the metastatic tumor microenvironment, with arrow colors indicating the source (sender) of the interaction, and arrow thickness representing the strength of interaction.

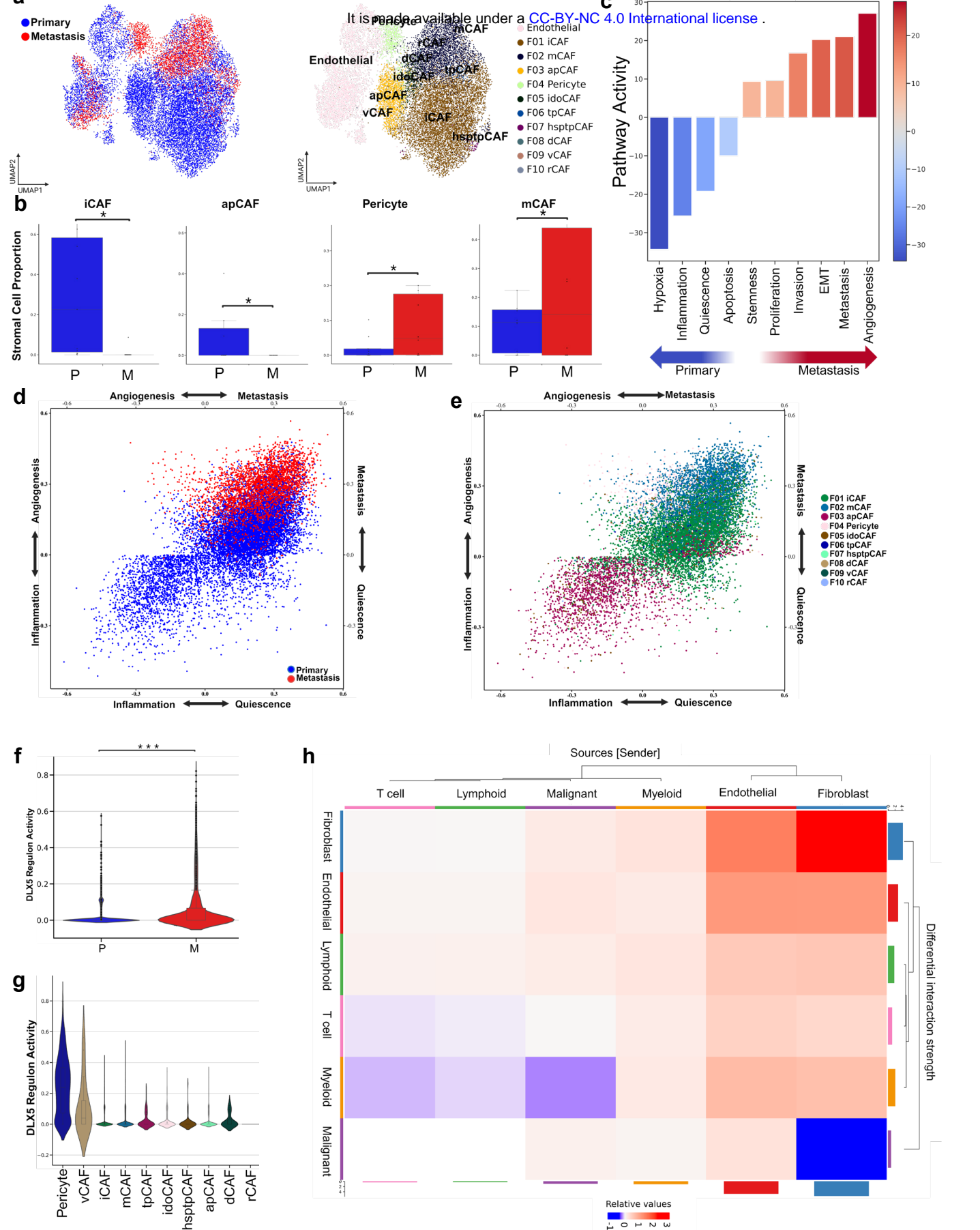
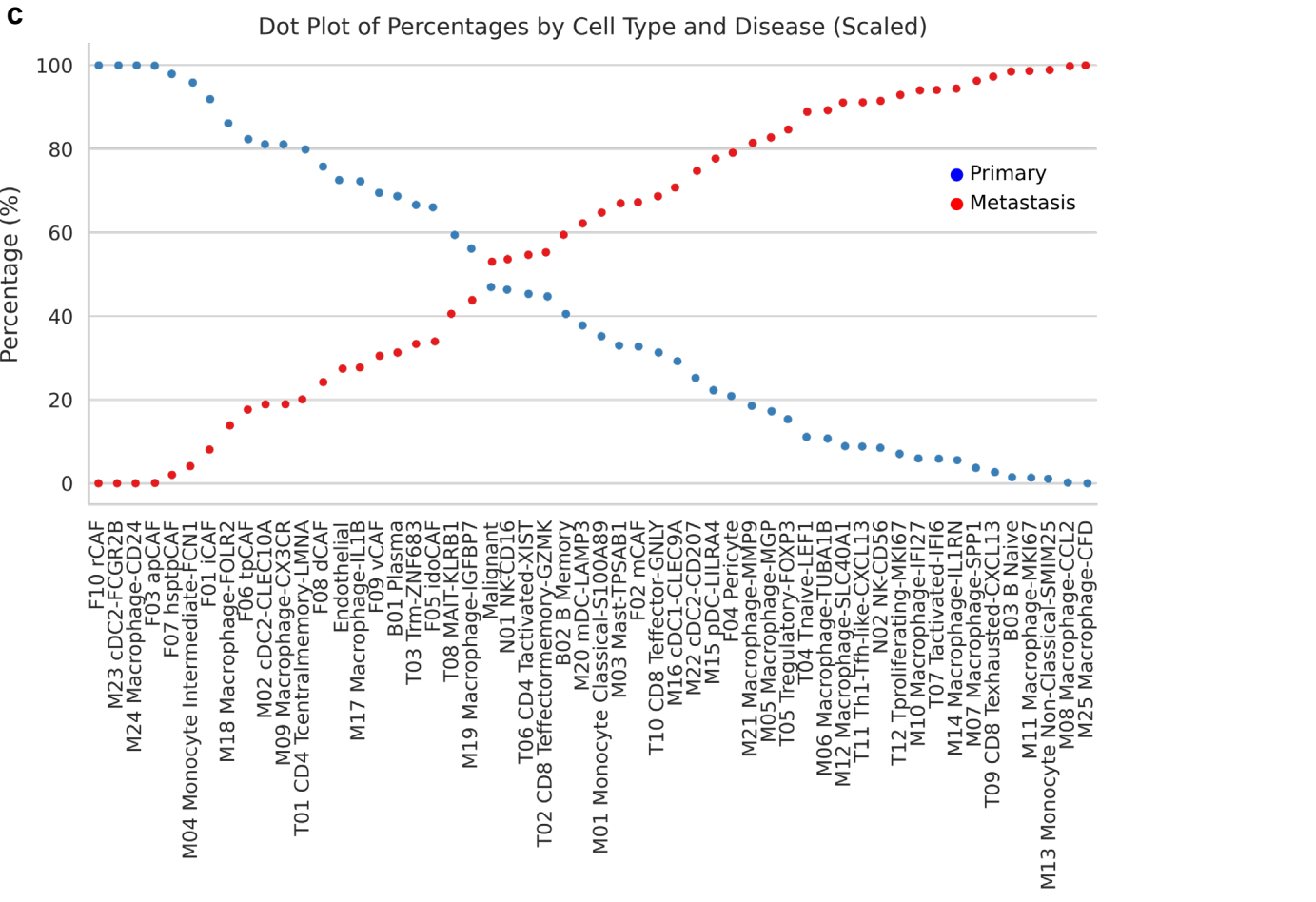
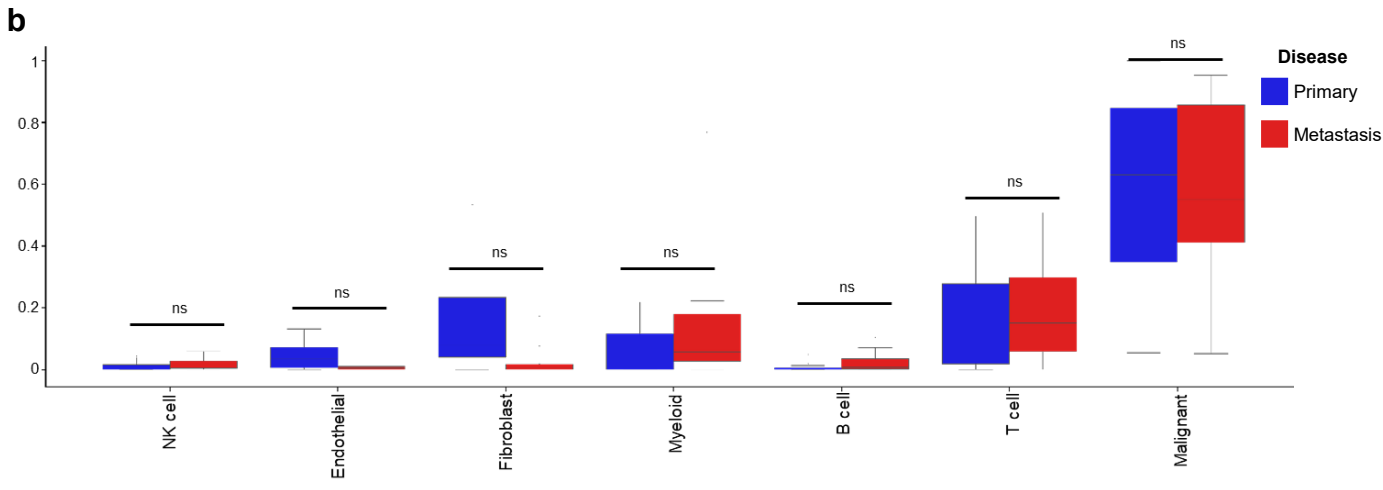
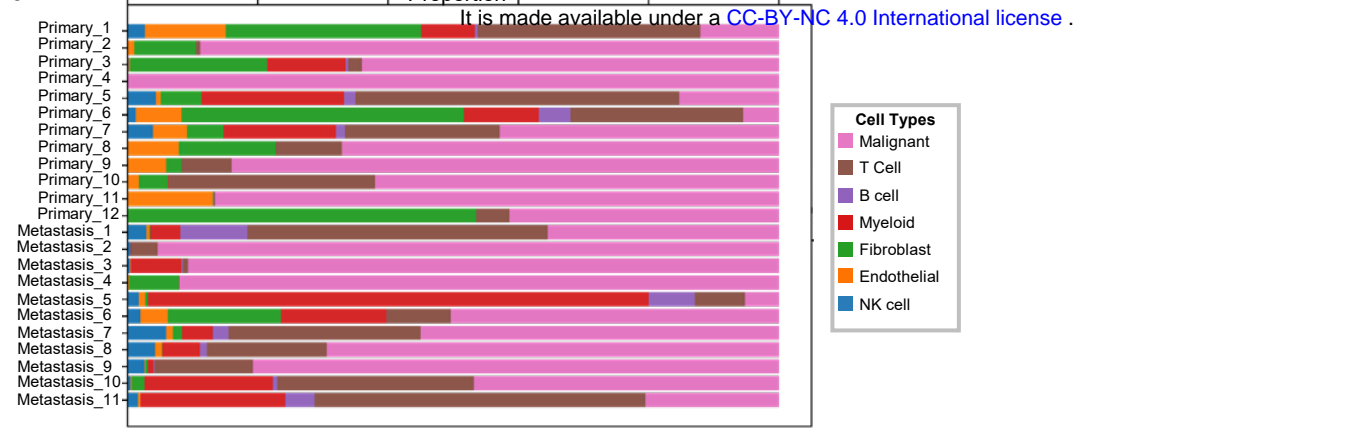


Fig.5 Stromal Cell Remodeling and Fibroblast Subtype Dynamics in Metastatic vs. Primary Breast Cancer TME

a. UMAP visualizations of 17339 stromal cells. Cells are colored by tissue class of origin on the left and grouped by one of 11 identified subtypes on the right, including 10 Fibroblast and an endothelial cell subtype. **b.** Bar plots showing the proportions of cells from each tissue class of origin (P=Primary, M=Metastatic) for Fibroblast cell subtypes with a significant difference in proportion. (permutation test, $p < 0.05$, $\log_2FC > 2$) **c.** Differential CANCERSEA gene signature pathway activity scores within fibroblast cells across metastatic status. **d.** Cellular states plot to show CANCERSEA gene signature activity for Angiogenesis, Metastasis, Inflammation, and Quiescence across the metastatic status within fibroblast cells. **e.** Cellular states plot to show CANCERSEA gene signature activity for Angiogenesis, Metastasis, Inflammation, and Quiescence across the multiple fibroblast subtypes. **f.** DLX5 Regulon Activity score comparison between primary and metastatic breast cancer stromal cells. Significance indicated as (***, $p < 0.001$). **g.** DLX5 Regulon Activity score comparison between primary and metastatic breast cancer stromal cells. Significance indicated as (***, $p < 0.001$). **h.** Heatmap comparing the differential number of interactions and interaction strength between primary and metastatic tumor microenvironments. The top-colored bar plot indicates incoming signaling, while the right-colored bar plot represents outgoing signaling for each cell type. Red indicates increased signaling in metastatic tumors, whereas blue represents increased signaling in primary tumors. Bar height reflects the magnitude of change in interaction number or strength between the two conditions.

Fig.5 Stromal Cell Remodeling and Fibroblast Subtype Dynamics in Metastatic vs. Primary Breast Cancer TME

- a.** UMAP visualizations of 17339 stromal cells. Cells are colored by tissue class of origin on the left and grouped by one of 11 identified subtypes on the right, including 10 Fibroblast and an endothelial cell subtype.
- b.** Bar plots showing the proportions of cells from each tissue class of origin (P=Primary, M=Metastatic) for Fibroblast cell subtypes with a significant difference in proportion. (permutation test, $p < 0.05$, $\log_2FC > 2$)
- c.** Differential CANCERSEA gene signature pathway activity scores within fibroblast cells across metastatic status.
- d.** Cellular states plot to show CANCERSEA gene signature activity for Angiogenesis, Metastasis, Inflammation, and Quiescence across the metastatic status within fibroblast cells.
- e.** Cellular states plot to show CANCERSEA gene signature activity for Angiogenesis, Metastasis, Inflammation, and Quiescence across the multiple fibroblast subtypes.
- f.** DLX5 Regulon Activity score comparison between primary and metastatic breast cancer stromal cells. Significance indicated as (***, $p < 0.001$).
- g.** DLX5 Regulon Activity score comparison between primary and metastatic breast cancer stromal cells. Significance indicated as (***, $p < 0.001$).
- h.** Heatmap comparing the differential number of interactions and interaction strength between primary and metastatic tumor microenvironments. The top-colored bar plot indicates incoming signaling, while the right-colored bar plot represents outgoing signaling for each cell type. Red indicates increased signaling in metastatic tumors, whereas blue represents increased signaling in primary tumors. Bar height reflects the magnitude of change in interaction number or strength between the two conditions.

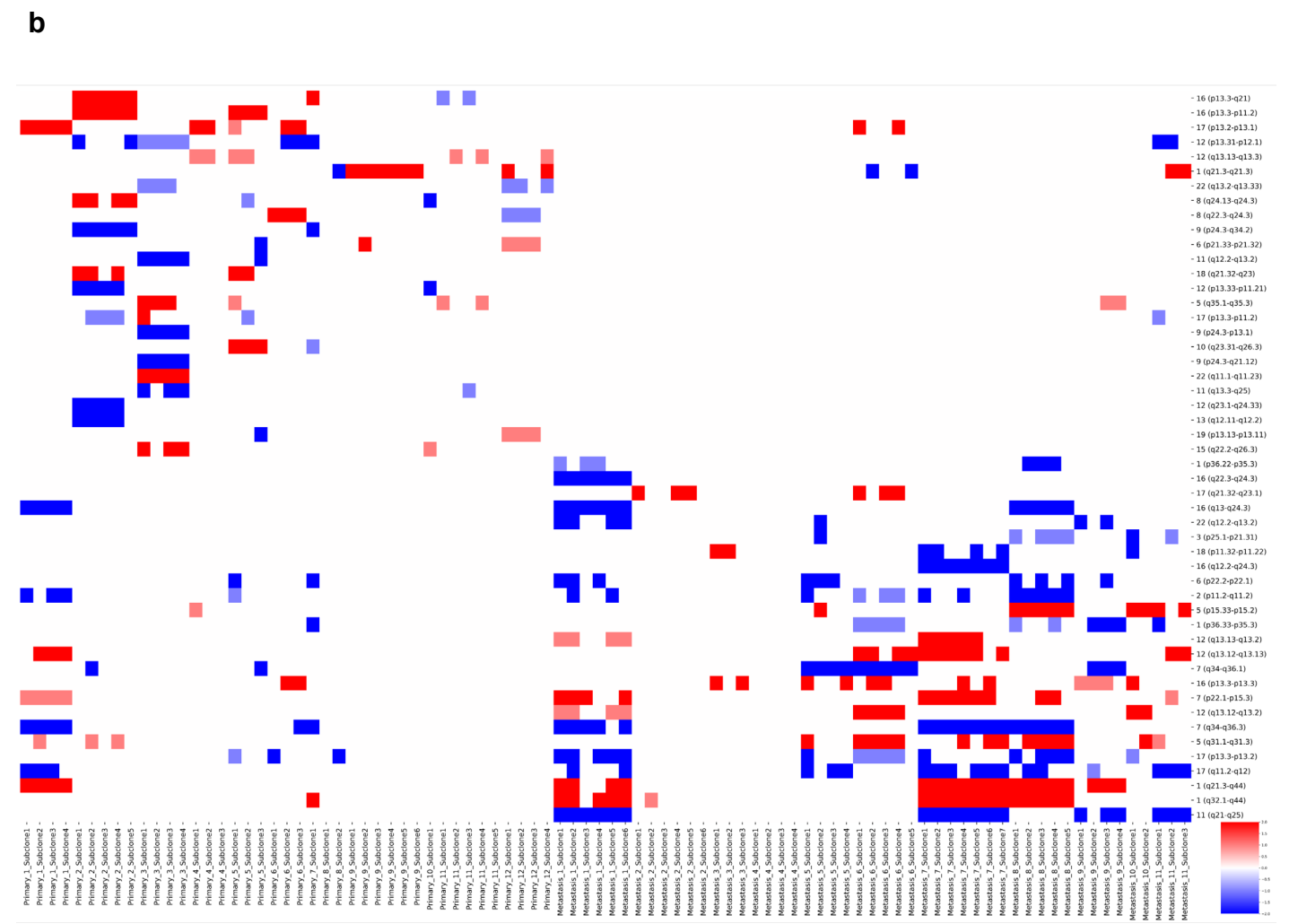
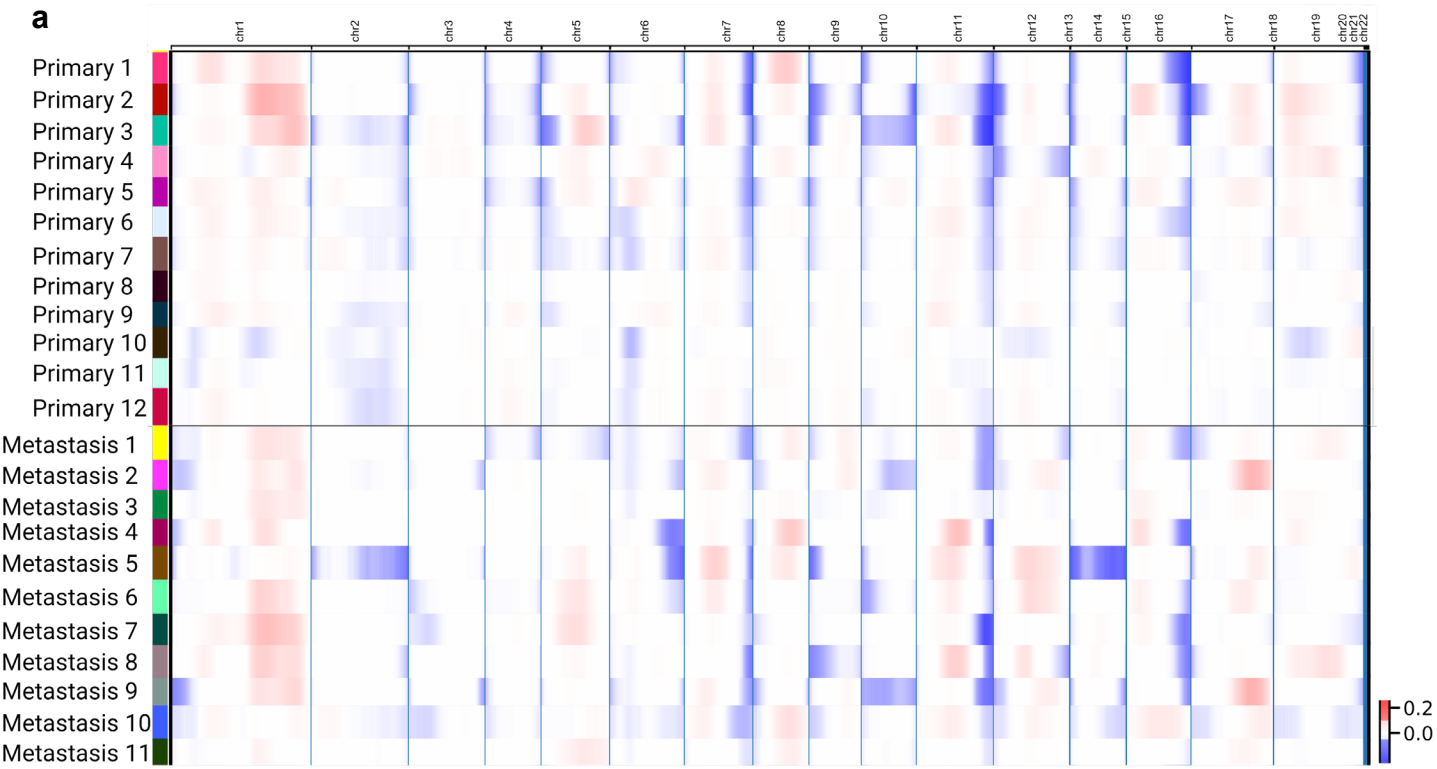


Supp Fig.1
a. Stacked barplot showing the proportion of seven major cell types across individual samples. **b.** Boxplots illustrating changes in the proportions of seven major cell types between primary and metastatic breast cancer samples. (permutation test, $p < 0.05$, $\log_2FC > 2$) **c.** Dot plot illustrating the percentage of each cell subtype from primary versus metastatic tumors, highlighting each cell subtype's distribution and comparative abundance across the two conditions.

Supp Fig.1

- a.** Stacked barplot showing the proportion of seven major cell types across individual samples.
- b.** Boxplots illustrating changes in the proportions of seven major cell types between primary and metastatic breast cancer samples. (permutation test, $p < 0.05$, $\log_2FC > 2$)
- c.** Dot plot illustrating the percentage of each cell subtype from primary versus metastatic tumors, highlighting each cell subtype's distribution and comparative abundance across the two conditions.

It is made available under a [CC-BY-NC 4.0 International license](https://creativecommons.org/licenses/by-nc/4.0/).

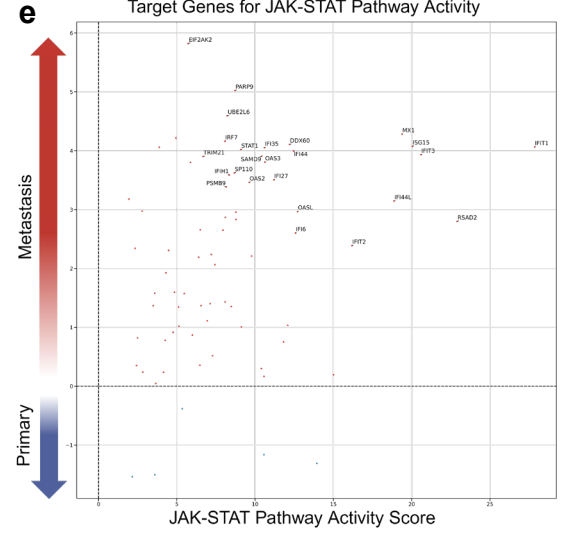
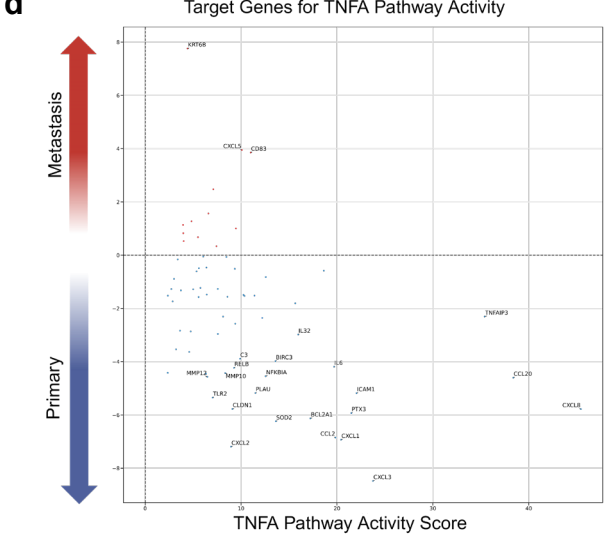
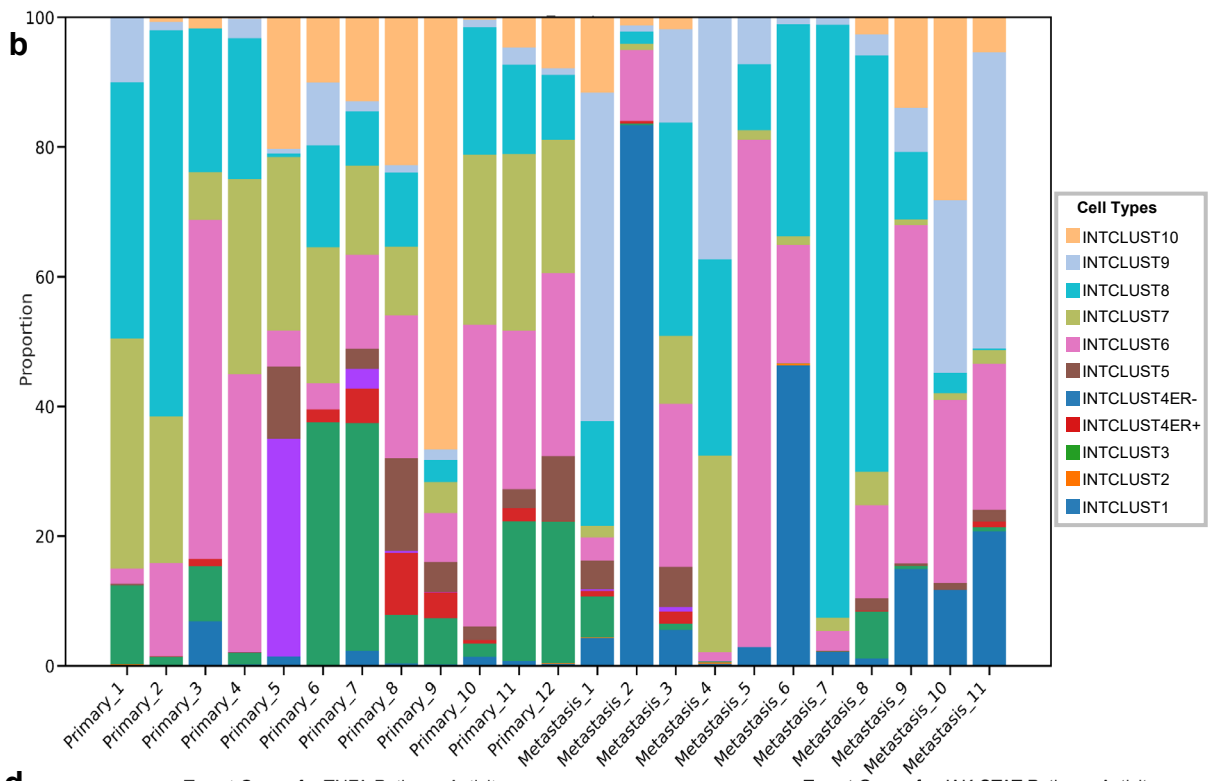
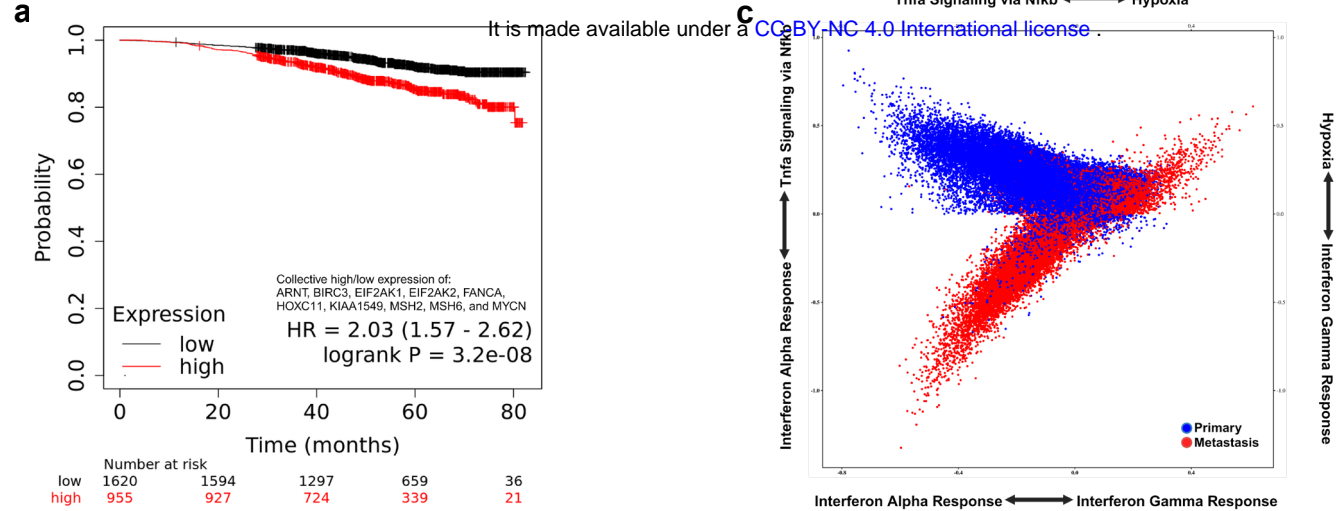


Supp Fig.2

a. Heatmap showing CNV profiles for malignant cells in individual primary and metastatic tumor samples, with T cells as a reference. **b.** Heatmap depicting CNV profiles from all tumor subpopulations per patient, comparing the top 25 copy number alterations across chromosomal arms between primary and metastatic breast cancer subclones.

Supp Fig.2

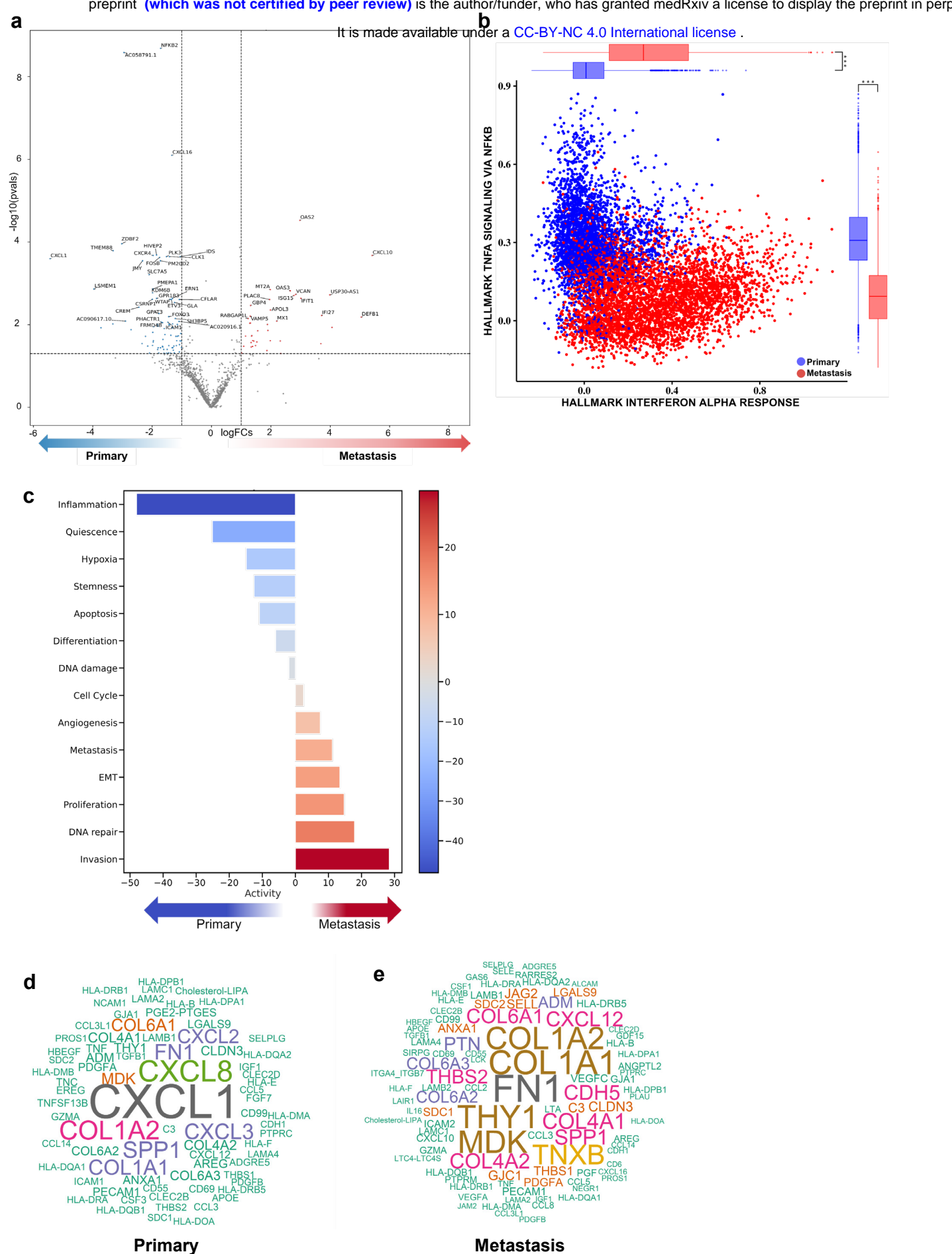
- a.** Heatmap showing CNV profiles for malignant cells in individual primary and metastatic tumor samples, with T cells as a reference.
- b.** Heatmap depicting CNV profiles from all tumor subpopulations per patient, comparing the top 25 copy number alterations across chromosomal arms between primary and metastatic breast cancer subclones.



Supp Fig.3
a. Kaplan-Meier survival curves for estrogen receptor-positive tumors among 2,575 patients, comparing survival outcomes associated with high versus low expression of the genes ARNT, BIRC3, EIF2AK1, EIF2AK2, FANCA, HOXC11, KIAA1549, MSH2, MSH6, and MYCN. **b.** Stacked bar plot showing the proportion of Integrative Clusters (IntClust) across individual samples, based on classification using the top 200 differentially expressed genes from the METABRIC study. **c.** Cellular states plot to show Hallmark gene signature activity for Hypoxia, Interferon Gamma Response, Interferon Alpha Response, and TNFA Signaling via NFKB across the metastatic status within malignant cells. **d.** Plot of the top 25 target genes affected by TNFA pathway activity, highlighting their weight and statistical significance in primary versus metastatic malignant cells based on PROGENY pathway activity. **e.** Plot of the top 25 target genes affected by JAK-STAT pathway activity, highlighting their weight and statistical significance in primary versus metastatic malignant cells based on PROGENY pathway activity.

Supp Fig.3

- a.** Kaplan-Meier survival curves for estrogen receptor-positive tumors among 2,575 patients, comparing survival outcomes associated with high versus low expression of the genes ARNT, BIRC3, EIF2AK1, EIF2AK2, FANCA, HOXC11, KIAA1549, MSH2, MSH6, and MYCN.
- b.** Stacked bar plot showing the proportion of Integrative Clusters (IntClust) across individual samples, based on classification using the top 200 differentially expressed genes from the METABRIC study.
- c.** Cellular states plot to show Hallmark gene signature activity for Hypoxia, Interferon Gamma Response, Interferon Alpha Response, and TNF- α Signaling via NFkB across the metastatic status within malignant cells.
- d.** Plot of the top 25 target genes affected by TNF- α pathway activity, highlighting their weight and statistical significance in primary versus metastatic malignant cells based on PROGENY pathway activity.
- e.** Plot of the top 25 target genes affected by JAK-STAT pathway activity, highlighting their weight and statistical significance in primary versus metastatic malignant cells based on PROGENY pathway activity.

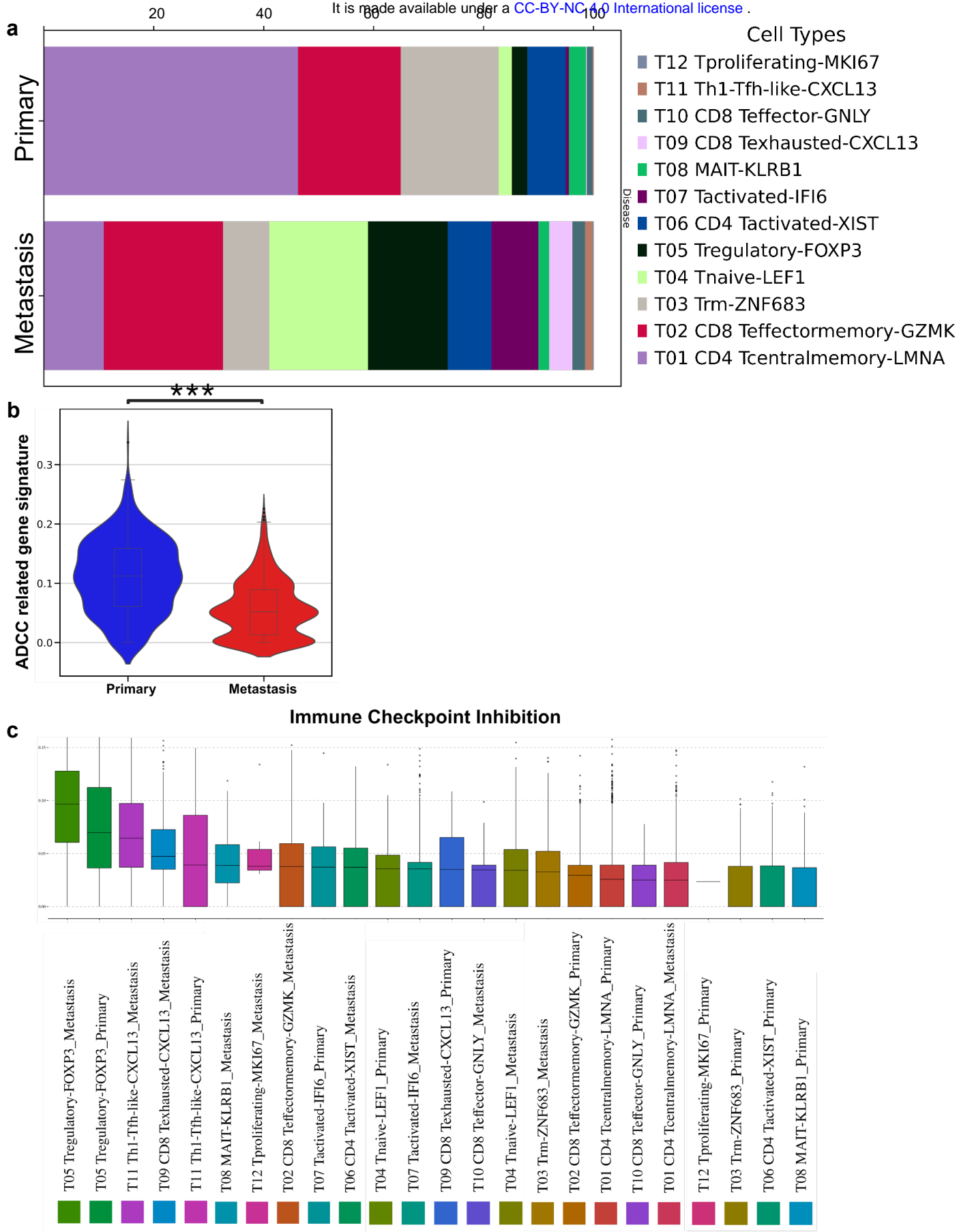


Supp Fig.4

a. Volcano plot illustrating the differential expression of genes in myeloid cells when comparing primary versus metastatic tumors. b. Cellular states plot showing Hallmark gene signature activity for Interferon Alpha Response versus TNFA Signaling via NFKB in myeloid cell populations. c. Differential CANCERSEA gene signature pathway activity scores within myeloid cells across metastatic status. d. Word cloud plot illustrating the enriched ligands in primary breast tumor microenvironment compared to metastatic cancer e. Word cloud plot illustrating the enriched ligands in metastatic breast tumor microenvironment compared to primary breast cancer

Supp Fig.4

- a.** Volcano plot illustrating the differential expression of genes in myeloid cells when comparing primary versus metastatic tumors.
- b.** Cellular states plot showing Hallmark gene signature activity for Interferon Alpha Response versus TNF- α Signaling via NF κ B in myeloid cell populations.
- c.** Differential CANCERSEA gene signature pathway activity scores within myeloid cells across metastatic status.
- d.** Word cloud plot illustrating the enriched ligands in primary breast tumor microenvironment compared to metastatic cancer
- e.** Word cloud plot illustrating the enriched ligands in metastatic breast tumor microenvironment compared to primary breast cancer

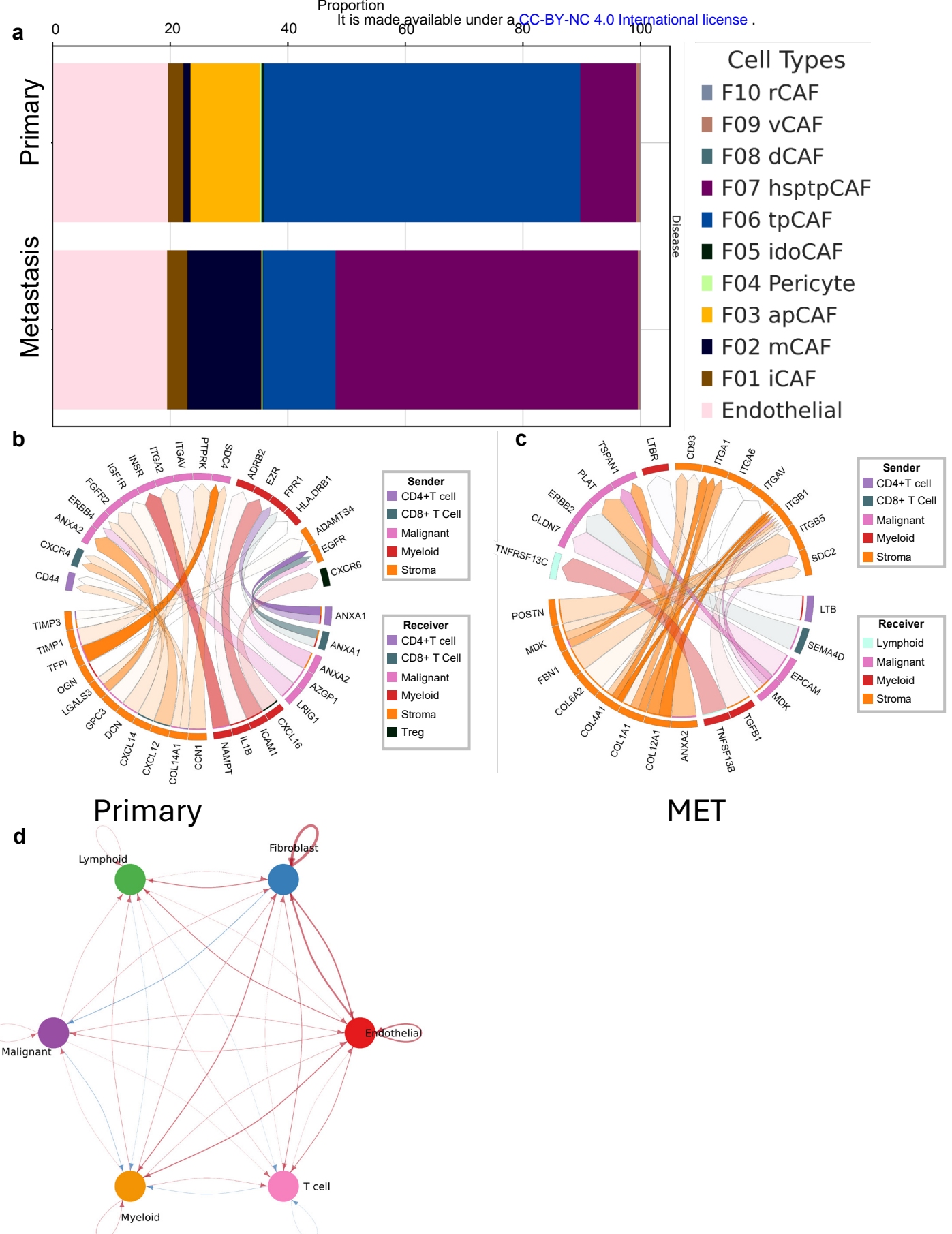


Supp Fig.5

a. Stacked bar plot showing the proportion of T cell subtypes across the metastatic status. **b.** Antibody-dependent cellular cytotoxicity gene signature comparison between primary and metastatic breast cancer NK cells. Significance indicated as (***, $p < 0.001$). **c.** Immune checkpoint inhibition levels compared across different cell types in primary and metastatic breast cancers

Supp Fig.5

- a.** Stacked bar plot showing the proportion of T cell subtypes across the metastatic status.
- b.** Antibody-dependent cellular cytotoxicity gene signature comparison between primary and metastatic breast cancer NK cells. Significance indicated as (***, $p < 0.001$).
- c.** Immune checkpoint inhibition levels compared across different cell types in primary and metastatic breast cancers



Supp Fig.6
a. Stacked bar plot showing the proportion of stromal cell subtypes across the metastatic status. **b.** Visualization of the top 25 ligand-receptor interactions between different cell types in the primary tumor microenvironment, as analyzed using the MultiNicheNet. **c.** Visualization of the top 25 ligand-receptor interactions between different cell types in the metastatic tumor microenvironment, as analyzed using the MultiNicheNet. **d.** Comparative circle plot visualizing differential cell-cell communication between primary and metastatic tumor microenvironments, where red edges indicate increased interactions in metastatic tumors and blue edges denote increased interactions in primary tumors

Supp Fig.6

- a.** Stacked bar plot showing the proportion of stromal cell subtypes across the metastatic status.
- b.** Visualization of the top 25 ligand-receptor interactions between different cell types in the primary tumor microenvironment, as analyzed using the MultiNicheNet.
- c.** Visualization of the top 25 ligand-receptor interactions between different cell types in the metastatic tumor microenvironment, as analyzed using the MultiNicheNet.
- d.** Comparative circle plot visualizing differential cell-cell communication between primary and metastatic tumor microenvironments, where red edges indicate increased interactions in metastatic tumors and blue edges denote increased interactions in primary tumors