bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Tumor-educated Gr1^{+}CD11b^{+} cells instigate breast cancer metastasis by twisting cancer cells plasticity via OSM/IL6–JAK signaling

Sanam Peyvandi[1, +, *], Manon Bulliard[1, +], Annamaria Kauzlaric[2], Oriana Coquoz[1], Yu-Ting Huang[1], Nathalie Duffey[1, +], Laetitia Gafner[1], Girieca Lorusso[1], Nadine Fournier[2], Qiang Lan[1,3,*,#], and Curzio Rüegg[1,*,#]

[1] Pathology Unit, Department of Oncology, Microbiology and Immunology (OMI), Faculty of Science and Medicine, University of Fribourg, 1700 Fribourg, Switzerland;

[2] Translational Data Science Group, Swiss Institute of Bioinformatics, Lausanne, Switzerland;

[3] Cell and Tissue Dynamics Research Program, Institute of Biotechnology, Helsinki Institute of Life Science (HiLIFE), University of Helsinki, Helsinki, Finland.

* These authors contributed equally as corresponding authors.

+ Current affiliations: S.P., Immunobiology Department, University of Lausanne (UNIL), Switzerland; M.B., Swiss Institute for Experimental Cancer Research (ISREC), Ecole Polytechnique Fédérale Lausanne (EPFL) UPCDA SV 2833 (Bâtiment SV) Lausanne, Switzerland; N.D., Department of Oncology, Centre Hospitalier Universitaire Vaudois (CHUV), University of Lausanne (UNIL), Switzerland; Y-T.H., Daiichi Sankyo Co., Ltd. Taipei City, Taiwan

Lead contacts and co-senior authors:

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

C.R. (curzio.ruegg@unifr.ch; +4126 300 8766), Department of Oncology, Microbiology, Immunology, Faculty of Science and Medicine, University of Fribourg, Chemin du Musée 18, Per 17, CH-1700 Fribourg, Switzerland

Q.L. (qiang.lan@helsinki.fi) Cell and Tissue Dynamics Research Program, Institute of Biotechnology, Helsinki Institute of Life Science (HiLIFE), Biocenter 2, P.O. Box 56, FIN-00014, University of Helsinki, Helsinki, Finland

Conflict-of-interests: The authors have no conflicts of interest to declare.

Running title: OMS/IL6 twist breast cancer plasticity to promote metastasis

Keywords: Breast cancer metastasis, chemotherapy, cancer stem cells, GR1, Oncostatin M, IL6, JAK, plasticity, heterogeneity

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

Cancer cell plasticity contributes to tumor therapy resistance and metastasis formation, which represent the main causes of cancer-related death for most cancers, including breast cancer. The tumor microenvironment drives cancer cell plasticity and metastasis and, thus, unravelling the underlying cues may provide novel effective strategies to manage metastatic disease. Here, we show that stem cell antigen-1 positive (Sca-1^{+}) murine breast cancer cells enriched during tumor progression and metastasis have higher *in vitro* cancer stem cell-like properties, enhanced *in vivo* metastatic ability, and initiate primary tumors rich in Gr1^{high}CD11b^{+}Ly6C^{low} cells. In turn, tumor-educated Gr1^{+}CD11b^{+} (Tu-Gr1^{+}CD11b^{+}) cells rapidly and transiently convert low metastatic 4T1-Sca-1^{-} cells into highly metastatic 4T1-Sca-1^{+} cells via secreted OSM and IL6. Moreover, chemotherapy-resistant and highly metastatic 4T1-derived cells maintain high Sca-1^{+} frequency through cell autonomous IL6 production. Inhibition of OSM, IL6 or JAK suppressed Tu-Gr1^{+}CD11b^{+}-induced Sca-1^{+} population enrichment *in vitro*, while JAK inhibition abrogated metastasis of chemotherapy-enriched Sca-1^{+} cells *in vivo*. Importantly, Tu-Gr1^{+}CD11b^{+} cells invoked a gene signature in tumor cells predicting shorter OS and RFS in breast cancer patients. Collectively, our data identified OSM/IL6-JAK as a clinically relevant paracrine/autocrine axis instigating breast cancer cell plasticity triggering metastasis.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Introduction

Metastasis accounts for over 90% of cancer-related death, calling for new strategies to prevent cancer cell dissemination and metastasis formation (1). Recent studies using single-cell lineage tracing and single-cell RNA sequencing (scRNA-seq) technologies have provided detailed information about intratumor heterogeneity (2, 3), whereby genetically, epigenetically and functionally diverse subpopulations of cancer cells exist within the tumor mass, spatially and temporally (4). Intratumor heterogeneity may arise by modulating cancer cell plasticity, especially of cancer stem cells (CSCs), through cell-intrinsic and -extrinsic mechanisms (5, 6).

Multiple CSC subpopulations appear to co-exist within the primary tumor mass resulting in high degree of tumor cell heterogeneity and increased aggressiveness (7) including in breast cancer (4, 8–11). In particular, CSCs can acquire metastasis initiating capacities (12) and resistance to therapy, resulting in cancer relapse (4). Moreover, non-CSCs within the tumor bulk may acquire CSC properties to repopulate the tumor(4). While CSC are defined rather functionally by their ability to initiate tumors and metastasis in low number *in vivo*, several cell surface markers associated with CSC features have been reported, including CD44, CD24, Sca-1, CD61, CD49f were used to identify breast CSCs (13, 14).

The interaction of tumor cells with the tumor microenvironment (TME) contributes to tumor cell plasticity and tumor heterogeneity (15, 16). Cells of the TME also promote tumor escape and metastasis through multiple mechanisms, including promotion of angiogenesis, cell survival, invasion, epithelial-mesenchymal transition (EMT), and immunosuppression (17–22). Recently, they have been reported to instigate expansion of CSC with metastatic ability (also known as metastasis-initiating cells) in different cancers (17, 23, 24). Thus, the TME dynamics is a key driver of cancer cell plasticity and

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

heterogeneity promoting tumor growth, progression and metastasis (4). Accurate characterization of the regulation of tumor plasticity and heterogeneity by the TME may reveal novel opportunities for developing effective anti-metastatic therapies (8, 14).

TME-derived Oncostatin M (OSM) has been shown to mediate tumor progression and CSC stemness by activating its receptor OSMR (25). OSM belongs to the IL6 family of cytokines (including IL6 itself, IL11 and LIF) (26, 27), whose members bind to dimeric receptors sharing a common subunit (gp130 or IL6ST) and activate JAK-STAT, RAS-MAPK and PI3K-AKT pathways (28, 29). Increased OSM or IL6 expression correlates with reduced survival in breast cancer patients (30, 31). OSM was shown to drive breast cancer progression and metastasis through direct effects on cancer cells, such as suppression of estrogen receptor (ER) expression (31) and promotion of EMT (25, 32), and indirect effects via TME cells, in particular the reprogramming of tumor associated macrophages and fibroblasts (33–36).

Here, by assessing the metastatic evolution of murine triple-negative breast cancer (TNBC) models *in silico* and *in vivo*, we observed that the Sca-1$^+$ tumor cell subpopulation is enriched during tumor progression. We show that tumor-educated Gr1$^+$CD11b$^+$ cells (Tu-Gr1$^+$CD11b$^+$), but not naïve Gr1$^+$CD11b$^+$ cells from spleen (Spl-Gr1$^+$CD11b$^+$) or bone marrow (BM-Gr1$^+$CD11b$^+$) in tumor-bearing mice, modulate tumor plasticity via OSM/IL6-JAK signaling by rapidly and transiently converting 4T1-Sca-1$^-$ cells into 4T1-Sca-1$^+$ cells with high metastatic capacity. Prolonged exposure of 4T1 cells to chemotherapy stably enriched for metastatic Sca-1$^+$ cells via an autocrine IL6-JAK signaling loop. A short *in vitro* treatment of these chemo-resistant cells with the JAK inhibitor Ruxolitinib, suppressed their metastatic capacity. Importantly, Tu-Gr1$^+$CD11b$^+$ invoked a gene expression signature in 4T1 cells that predicted shorter overall survival (OS) and relapse-free survival (RFS) in breast cancer patients, reinforcing the clinical relevance of these findings.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Our results reveal a novel mechanism modulating tumor plasticity and triggering the emergence of cancer cells with enhanced metastatic capacity, through paracrine (Tu-Gr1$^+$CD11b$^+$-mediated) and cell autonomous (chemotherapy-induced) OSM/IL6-JAK dependent signaling. The OSM/IL6-JAK axis may be considered as a candidate of actionable clinical targets to impinge on metastatic progression and therapy resistance.

Results

A Sca-1$^+$ tumor cell population is enriched during tumor progression and has higher *in vivo* metastatic capacity

To investigate tumor cell heterogeneity during tumor progression, we first examined the expression of the previously reported breast CSC markers CD24, CD44, CD61, Sca-1 and CD49f (13) in the publicly available RNA sequencing (RNAseq) dataset from Ross *et al.* encompassing several murine breast cancer models (37). This dataset includes data from cultured cancer cells (In_Culture), orthotopic primary tumors (OT_PT), spontaneous lung metastases (OT_LuM), and experimental lung metastases after tail vein injection (TV_LuM) (Figure 1A and Supplemental Figure 1A). *Sca-1* expression was elevated in lung metastasis in 4T1, 6DT1, Mvt1 and Met1 models compared with the respective primary tumors. Interestingly, in 4T1, 6DT1 and Mvt1 models, *Sca-1* expression was also elevated in experimental lung metastases compared with cultured cells. However, the expression of *Cd24*, *Cd44*, *Cd61* and *Cd49f* were not altered, or their expression pattern was inconsistent during progression (Supplemental Figure 1A). Thus, increased *Sca-1* expression during metastasis is consistently observed in different preclinical breast cancer models.

To investigate whether the increased *Sca-1* expression within the tumor mass was due to an increased gene expression in all cancer cells or to the enrichment of a Sca-1$^+$ population, we orthotopically injected 4T1 tumor cells and determined the frequency of

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

different cell populations present in the primary tumor and lung metastases 30 days later by flow cytometry (Figure 1B). We observed that the frequency of both Sca-1$^+$ and CD61$^+$ populations increased in lung metastases compared to primary tumors (Figure 1C). In contrast, the CD24$^+$, CD44$^+$ and CD49f$^+$ populations were not significantly altered.

The enrichment of the Sca-1$^+$ population in lung metastases prompted us to ask whether Sca-1$^+$ cells actively contribute to the metastasis. To this end, we isolated 4T1-Sca-1$^+$ and 4T1-Sca-1$^-$ cells by magnetic activated cell sorting (MACS) from parental 4T1 cells, which contains low frequency of Sca-1$^+$ population (10-15%) (Supplemental Figure 1B, C), and examined their metastatic ability *in vivo*. In the orthotopic injection model, 4T1-Sca-1$^+$ cells formed significantly more lung metastases than 4T1-Sca-1$^-$ cells, while there was no significant difference in primary tumor growth (Figure 1D-F). Upon tail vein injection, 4T1-Sca-1$^+$ cells displayed significantly greater lung colonization ability compared to 4T1-Sca-1$^-$ cells and a non-significant increase compared to parental 4T1 cells (Figure 1G-I). In addition, 4T1-Sca-1$^+$ cells showed significantly higher *in vitro* mammosphere forming efficiency than 4T1-Sca-1$^-$ cells (Supplemental Figure 2A), while *in vitro* cell growth and cell motility were comparable (Supplemental Figure 2B-C).

These results suggest that 4T1-Sca-1$^+$ and 4T1-Sca-1$^-$ cells have similar tumorigenic potential, while 4T1-Sca-1$^+$ cells have higher metastasis-initiating capacity.

The Sca-1$^+$ tumor cell population is plastic *in vitro* and *in vivo*

Growing evidence indicates that cancer cells possess plastic features, which can be modulated by both cell-intrinsic factors and microenvironmental cues (37, 38). To characterize the observed plasticity of 4T1-Sca-1$^+$ cells, we first investigated isolated 4T1-Sca-1$^+$ and 4T1-Sca-1$^-$ cells *in vitro*. The abundance of Sca-1$^+$ population in 4T1-Sca-1$^+$ cells enriched by MACS sorting (> 75%) gradually decreased to 50% after 4 days of culture

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(Supplemental Figure 2D, upper panel), while the Sca-1$^-$ cells (negatively enriched by MACS) regenerated a Sca-1$^+$ population (from less that 1% to 19%) (Supplemental Figure 2E, lower panel). Consistently, after orthotopic injection of 4T1-Sca-1$^+$ cells, the abundance of the Sca-1$^+$ population in the derived tumors decreased from 75% to about 40% after 23 days of growth, while in tumors generated from 4T1-Sca-1$^-$ cells it increased from less than 1% to 15% (Supplemental Figure 2E), similar to the frequency of Sca-1$^+$ population in tumors derived from parental 4T1 cells (Figure 1C). In addition, when tumor cells derived from primary tumors and lung metastases of parental 4T1-injected mice were cultured *ex vivo*, the abundance of the Sca-1$^+$ population significantly decreased from 20% to 4.5% and 60% to 19%, respectively (Supplemental Figure 2F).

From these observations we conclude that both Sca-1$^+$ and Sca-1$^-$ populations are highly plastic and this plasticity appears to be modulated *in vivo*.

Tumor-educated Gr1$^+$CD11b$^+$ cells expand the metastatic Sca-1$^+$ population

Immune cells in the TME are critical determinants of tumor cells functions and behaviors, including metastatic capacity (39). To collect evidence for a potential correlation between immune cells and the Sca-1$^+$ population, we characterized the inflammatory cells infiltrating the orthotopic primary tumors. We observed a significant increase of the Gr1^{high}CD11b$^+$Ly6C^{low} population and a significant decrease of the Gr1^{low}CD11b$^+$Ly6C^{high} population in tumors derived from the 4T1-Sca-1$^+$ cells, compared to tumors derived from the 4T1-Sca-1$^-$ cells (Figure 2A). Gr1^{high}CD11b$^+$Ly6C^{low} cells are immature myeloid progenitors mobilized from the bone marrow by tumor-derived signals capable of establishing an immunosuppressive environment facilitating tumor progression and metastasis (40, 41). Gr1$^+$CD11b$^+$ cells form a homogenous population in the circulation (Supplemental Figure 3A), consistent with the literature (42), while in the TME they

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

differentiated into two distinct subpopulations, Gr1^{high} and Gr1^{low}. To examine the direct contribution of the tumor-educated Gr1^{+}CD11b^{+} cells in promoting the enrichment of the Sca-1^{+} population, we isolated Gr1^{+} cells from primary tumors, bone marrow and spleen of 4T1 tumor-bearing BALB/c mice by MACS, and co-cultured them *in vitro* with parental 4T1 cells (Figure 2B). Gr1^{+} cells isolated from tumors (Tu-Gr1^{+}CD11b^{+}), but not from spleen (Spl- Gr1^{+}CD11b^{+}) or bone marrow (BM-Gr1^{+}CD11b^{+}), significantly induced the expansion of a Sca-1^{+} population in 4T1 cells (from 12.5% to 81%, $p<0.001$) in 48 hours (Figure 2C). To test whether the expansion of a Sca-1^{+} population from 4T1 cells required direct contact with Tu-Gr1^{+}CD11b^{+} or was mediated by soluble factors, we compared the induction in two different co-culture setups, either in standard wells (cell contact) or in Transwells, where Tu-Gr1^{+}CD11b^{+} and 4T1 cells were separated by a filter with 0.4 µm pores (Figure 2D). We did not observe any significant difference in the induction efficiency of Sca-1^{+} populations (measured by flow cytometry) between the two conditions and increasing the 4T1:Tu-Gr1^{+}CD11b^{+} cell ratio from 1:1 to 1:3 did not further expand the Sca-1^{+} population (Figure 2E). Furthermore, conditioned medium from co-cultured Tu-Gr1^{+}CD11b^{+} and 4T1 cells was also capable of subsequently expanding the Sca-1^{+} population from 4T1 cells alone (Supplemental Figure 3B). Interestingly, the Tu-Gr1^{+}CD11b^{+} induced Sca-1^{+} population appeared more stable in time compared to the isolated 4T1-Sca-1^{+} cells (Supplemental Figure 3C and Supplemental Figure 2E). More importantly, when Gr1^{+}CD11b^{+} cells-primed 4T1 cells were injected into the tail vein, Tu-Gr1^{+}CD11b^{+} primed ones showed higher lung colonization capacity compared to Spl-Gr1^{+}CD11b^{+} primed one (Figure 2F-H).

These results imply that tumor-educated Gr1^{+}CD11b^{+} cells induce the emergence of a highly metastatic Sca-1^{+} population through secreted factors.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Tu-Gr1$^+$CD11b$^+$-induced and tumor-inherent Sca-1$^+$ populations display distinct gene expression profiles

To unravel the molecular basis for the metastatic capacity of the inherent 4T1- Sca-1$^+$ population and the Tu-Gr1$^+$CD11b$^+$ induced Sca-1$^+$ population, we first performed transcriptomic profiling of 4T1-Sca-1$^+$ and 4T1-Sca-1$^-$ cells isolated from the parental 4T1 line. Pathway enrichment analysis showed that 4T1-Sca-1$^+$ and 4T1-Sca-1$^-$ cells expressed different genes associated with distinct signaling pathways (Figure 3A). The top 200 significantly upregulated and downregulated genes were extracted as Sca1 Positive and Sca1 Negative signatures, respectively (Supplemental Table 1). Next, we performed transcriptomic profiling of Tu-Gr1$^+$CD11b$^+$-primed 4T1, Spl-Gr1$^+$CD11b$^+$-primed 4T1 and parental 4T1 cells. Pathway enrichment analysis revealed that Tu-Gr1$^+$CD11b$^+$ and Spl-Gr1$^+$CD11b$^+$ priming induced distinct transcriptomic alterations in 4T1 cells (Figure 3B). To focus on the transcriptomic alternations related to the Sca-1$^+$ population conversion, we compared Tu-Gr1$^+$CD11b$^+$-primed vs Spl-Gr1$^+$CD11b$^+$-primed cells. Interestingly, Tu-Gr1$^+$CD11b$^+$-primed 4T1 cells expressed both Sca-1 Positive and Sca-1 Negative signatures (Figure 3C), suggesting that Tu-Gr1$^+$CD11b$^+$ may twist tumor plasticity by converting the 4T1-Sca-1$^-$ cells into 4T1-Sca-1$^+$ cells, rather than expanding the pre-existing 4T1-Sca-1$^+$ population.

To test whether, despite the diverse transcriptional profiles across different Sca-1$^+$ populations, there could be a common molecular mechanism underlying their induction and metastatic capacity, we compared the significantly differentially expressed genes between 4T1-Sca-1$^+$ versus 4T1-Sca-1$^-$ and Tu-Gr1$^+$CD11b$^+$ primed 4T1 versus Spl-Gr1$^+$CD11b$^+$ primed 4T1 cells. Strikingly, among a total of 1118 up- and 423 down-regulated genes found when comparing the two conditions, only 56 up- and one down-regulated genes were shared (Figure 3D and Supplemental Table 1). This observation was

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

consistent with the notion that the Sca-1$^+$ population in Tu-Gr1$^+$CD11b$^+$ primed 4T1 cells was different from the inherent 4T1-Sca-1$^+$ cells. Nonetheless, the fact that the 4T1-Sca-1$^+$ cells and Tu-Gr1$^+$CD11b$^+$ primed 4T1 cells possessed similar *in vivo* metastatic capacity suggested that among these common pathways some were relevant for 4T1 metastases formation. To this end, we analyzed the publicly available scRNA-seq dataset from 4T1 primary tumors of Sebastian *et al*. (43). In this dataset, several cell types, including cancer cells, epithelial cells, fibroblasts, distinct subpopulations of myeloid cells were identified (43). To determine significant ligand-receptor interactions from the scRNA-seq data, we performed cell-cell interaction analysis with CellPhoneDB (44) by focusing on the interactions between epithelial/cancer cells and myeloid cells (Figure 3E, F). The analysis identified 160 ligand-receptor interaction pairs (Supplemental Table 2). Among those pairs, OSM receptor (OSMR) and pyrimidinergic receptor P2Y6 (P2RY6) were the only ones present inside the 56 common genes shown in Figure 3D. However, P2RY6 interacts with COPA (Coatomer Complex Subunit Alpha), a membrane protein involved in membrane traffic between endoplasmic reticulum and Golgi (45) and, thus, unlikely to mediate cell-cell contact-independent induction of the Sca-1$^+$ population. As the IL6-JAK-STAT3 signaling pathway was upregulated both in the 4T1-Sca-1$^+$ population and Tu-Gr1$^+$CD11b$^+$-primed 4T1 cells (Figure 3A, B), we next examined the expression of *Osm*, *Osmr*, *Il6st*, *Il6*, and *Il6* receptor (*Il6ra*) in the Sebastian dataset (Figure 3G). The expression of *Il6* and *Osm* was restricted to myeloid cells, with *Osm* expression being more prominent, similar to a previous report (33). *Osmr* was predominantly expressed in tumor cells, while *Il6st* and *Il6ra* were homogenously expressed in all cell types.

We then explored their expression in the 4T1-Sca-1$^+$ cells and Tu-Gr1$^+$CD11b$^+$ primed 4T1 cells. *Osm* and *Il6* expression were very low in all samples (normalized count number less than 7 on average) (Figure 3H-I), consistent with data in the Sebastian dataset

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(Figure 3G). *Osmr* and *Il6ra*, however, were highly expressed in 4T1-Sca-1$^+$ cells compared with 4T1-Sca-1$^-$ cells, while the expression of *Il6st* was abundant in both populations, though higher in 4T1-Sca-1$^+$ cells (Figure 3H). On the other hand, only *Osmr* was significantly upregulated in Tu-Gr1$^+$CD11b$^+$ primed 4T1 compared with Spl-Gr1$^+$CD11b$^+$ primed 4T1 (Figure 3I).

Taken together, these results suggest that OSM/OMSR and IL6/IL6R signaling pathways may be involved in the Tu-Gr1$^+$CD11b$^+$-mediated expansion of 4T1-Sca-1$^+$ cells.

Tu-Gr1$^+$CD11b$^+$ cells promote Sca-1$^-$ to Sca-1$^+$ population conversion

The above results strongly implied that Tu-Gr1$^+$CD11b$^+$ cells convert Sca-1$^-$ population into the Sca-1$^+$ one. To further test this hypothesis, we compared the cell population dynamics in cultured cells and the orthotopic primary tumors by analyzing publicly available scRNA-seq datasets. By integrating scRNA-seq data from 3D cultured 4T1 cells (GSM4812003) (46) and tumor cells isolated from orthotopically fat pad-injected 4T1 primary tumor (PT) (GSM3502134) (47) (Figure 4A, B), we observed 5 clusters. Clusters 0, 1, 2, 4 were predominant in cultured tumor cells, while cluster 3 was predominant in primary tumors. Single-cell trajectories analysis confirmed that cluster 3 was at the end of the transformation process (Figure 4C, D). The population dynamics also showed that the fraction of cells in clusters 1, 2 and 4 decreased during the transformation, the one in cluster 0 only minimally increased, while the fraction in cluster 3 massively increased (Figure 4E). Importantly, very few cultured 4T1 cells expressed *Sca-1* while it was abundantly expressed in the majority of cells in the primary tumor (Supplemental Figure 4A, upper panel). Consistently, the fraction of cells expressing *Osmr* was higher in the primary tumor compared to cultured cells (Supplemental Figure 4A, lower panel). Similar observations were obtained when analyzing scRNA-seq data from the ER$^+$ human breast

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

cancer model MCF-7. After integrating data from cultured MCF-7 cells (GSM4681765) and tumor cells that were isolated from MCF-7 intraductal injected mammary gland (GSM5904917) (48), 6 clusters were identified (Figure 4F), with clusters 1 and 3 predominant in cultured MCF-7 cells, while clusters 2 and 4 were predominant in primary tumors (Figure 4G). Further analysis showed that clusters 2 and 4 expanded during *in vitro* to *in vivo* tumor cell transformation and represented nearly 50% of the *in vivo* primary tumor cells (Figure 4H-J). Although there is no human homolog of *Sca1* gene, *OSMR* expressing cells were increased upon tumor implantation, especially in clusters 2 and 1 (Supplemental Figure 4B).

To further investigate the signals involved in this transformation, we performed GSAE analysis for cluster 3 in 4T1 cells and clusters 2 and 4 in MCF-7 cells, respectively (Figure 4K-N). By comparing the Hallmark gene signatures, IL6-JAK-STAT3 signature was significantly upregulated in both cell populations (Figure 4K, M). Interestingly, Sca1 Positive and Sca1 Negative signatures were both upregulated (Figure 4L, N), which is consistent with our *ex vivo* induction experiment (Figure 3C). To validate the involvement of Tu-Gr1^{+}CD11b^{+} during the cell population transformation, we extracted the top 50 upregulated genes (Supplemental Table 3) identified by comparing the Tu-Gr1^{+}CD11b^{+} with the Sp-Gr1^{+}CD11b^{+}-stimulated 4T1 cells as Tu-Gr1^{+}CD11b^{+} induced signature. Both cell populations predominant in the primary tumor in both 4T1 and MCF-7 models, upregulated the Tu-Gr1^{+}CD11b^{+} induced signature (Figure 4L, N).

These data, together with our *in vivo* observations (Figure 1A-C) and *ex vivo* coculture experiments (Figure 2B-E and Figure 3C), indicate that Tu-Gr1^{+}CD11b^{+} convert the Sca-1^{-} population to Sca-1^{+} population, likely, via OSM/IL6 signaling pathway.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

OSM/IL6-JAK pathway mediates Tu-Gr1$^+$CD11b$^+$-induced Sca-1$^+$ population enrichment

To experimentally interrogate the role of OSM/IL6 in modulating the Sca-1$^+$ population, we first measured *Osm* and *Il6* mRNA expression in Spl-Gr1$^+$CD11b$^+$ and Tu-Gr1$^+$CD11b$^+$. Indeed, both *Osm* and *Il6* mRNA levels were significantly elevated in Tu-Gr1$^+$CD11b$^+$ (Figure 5A). To functionally validate the role of OSM/IL6 in the generation of Sca-1$^+$ population, we treated 4T1 cells directly with recombinant OSM and IL6 proteins *in vitro* and measured the effect on the Sca-1$^+$ population. After 2 days of treatment, both OSM and IL6 significantly increased the frequency of Sca-1$^+$ cells from 16% to 38.8% and 26.5%, respectively (Figure 5B). Conversely, blocking OSM and IL6 activities using anti- OSM or -IL6 neutralizing antibodies significantly reduced the emergence of Sca-1$^+$ population induced by Tu-Gr1$^+$CD11b$^+$ conditioned medium (Figure 5C). The combination of anti-OSM and -IL6 antibodies did not have additive effects suggesting that OSM and IL6 both contribute in promoting Sca-1$^+$ population by sharing the same signaling cascades.

OSMR and IL6R signal by activating the intracellular Janus tyrosine kinase (JAK) (49). To explore the involvement of the JAK pathway in the emergence of the Sca-1$^+$ population, we treated 4T1 cells with the JAK inhibitor (Ruxolitinib) during exposure to recombinant OSM and IL6. Ruxolitinib treatment prevented the emergence of the Sca-1$^+$ population in response to recombinant OSM and IL6 (Figure 5D).

From these results, we conclude that OSM/IL6-JAK pathway mediates Tu-Gr1$^+$CD11b$^+$ -induced Sca-1$^+$ population enrichment.

Tu-Gr1$^+$CD11b$^+$-induced Sca-1$^+$ population and 4T1-inherent Sca-1$^+$ population have distinct CSC and EMT gene expression profiles

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

OSM/IL6-JAK signaling has been reported to support tumor progression by promoting a CSC phenotype and epithelial-mesenchymal plasticity (25, 36, 50). To further characterize CSC and EMT features in 4T1-Sca-1$^+$ cells and Tu-Gr1$^+$CD11b$^+$ induced Sca-1$^+$ population, we interrogated our RNAseq data set for the expression of 17 stem cell and EMT markers (Supplemental Figure 5). The stem cell markers *Oct4* (*Pou5f1*), *Sox2* and *Nanog* were undetectable or very low in all or some samples. 4T1-Sca-1$^+$ cells had higher expression of *Aldh1a1*, *Aldh3a1* and *Podxl* but lower expression of *Klf4* and *Sox9* compared to 4T1-Sca-1$^-$ cells. There was no difference in the expression of *Abcg2* and *Has2*. Tu-Gr1$^+$CD11b$^+$ primed 4T1 cells had higher expression of *Klf4* and *Has2*, lower expression of *Aldh1a1*, *Aldh3a1* and *Sox9*, and similar expression of *Podxl* when compared with Spl-Gr1$^+$CD11b$^+$ primed 4T1. Among them, only *Has2* expression was specifically elevated in Tu-Gr1$^+$CD11b$^+$ primed 4T1 compared with control 4T1 and Spl-Gr1$^+$CD11b$^+$-primed 4T1 (Supplemental Figure 5A). On the other hand, 4T1-Sca-1$^+$ cells had lower expression of *Cdh1* and higher expression of *Snail1*, *Twist1*, *Vim* and *Foxc1* which support an EMT status, although *Zeb1* expression was reduced (Supplemental Figure 5B). Globally, the expression of most of the EMT genes were similar between Tu-Gr1$^+$CD11b$^+$ and Spl-Gr1$^+$CD11b$^+$ primed 4T1 cells, except for *Snail2* and *Vim*, whose expression was suppressed in Tu-Gr1$^+$CD11b$^+$ primed 4T1 (Supplemental Figure 5B).

Taken together, these results indicate that the Tu-Gr1$^+$CD11b$^+$-induced Sca-1$^+$ population and inherent Sca-1$^+$ population have different CSC and EMT transcriptional profiles, reinforcing the notion that, Tu-Gr1$^+$CD11b$^+$-induced Sca-1$^+$ population are not just enriched tumor-inherent Sca-1$^+$ population

Chemotherapy enriches a Sca-1$^+$ population with CSC features

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

CSC and cancer cell plasticity contribute to drug resistance in various tumor types, including breast cancer (51–55). The above results, including the significantly elevated expression of *Aldh3a1* (a marker of drug resistance) in the 4T1-Sca-1$^+$ population prompted us to investigate the resistance of this population to chemotherapy. To this end, we treated 4T1 cells for 48 hours *in vitro* with methotrexate (MTX) and doxorubicin (Dox), two widely used chemotherapy drugs, including in breast cancer. The 48 hours treatment of either drug increased the frequency of Sca-1$^+$ population in 4T1 cells (Figure 6A). Next, we mimicked a clinically relevant situation of cancer cells escaping chemotherapy by exposing 4T1 cells to 28 nM MTX, a concentration slightly higher than the IC_{50} concentration of the drug, for up to 3 weeks and recovered the surviving cells by switching to normal medium (Figure 6B). The selected cell line, named MR13, was highly enriched in Sca-1$^+$ cells (>60%) (Figure 6C). Compared to parental 4T1 tumor cells, MR13 cells exhibited a higher mammosphere forming efficiency (Figure 6D), lower *in vitro* proliferative capacity (Figure 6E) and increased cell mobility (Figure 6F), which were consistent with CSC-like properties. When tested in a 48-hours cytotoxicity assay, MR13 cells were more resistant against MTX compared to parental 4T1 (Supplemental Figure 6A).

MR13-derived tumors are highly metastatic and rich in Gr1^{high}CD11b$^+$ Ly6C^{low} cells

To characterize the *in vivo* behavior of MR13 cells, we orthotopically implanted them into BALB/c mice and monitored the progression. MR13 cells formed smaller primary tumors relative to parental 4T1 cells, that were more metastatic to the lung (Figure 6G-I) and enriched in Gr1^{high}CD11b$^+$ Ly6C^{low} cells compared to 4T1 tumors (Figure 6J), similarly to 4T1-Sca-1$^+$-derived tumors (Figure 2A). Strikingly, we observed metastases in the heart (Supplemental Figure 6B), which we never observed with the parental 4T1 cells. MR13 cells retained a large fraction of the Sca-1$^+$ population *in vitro*, even when cultured in the

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

absence of MTX (Figure 6C) and upon *in vivo* expansion (Figure 6K). Such stability of Sca-1 population contrasted with 4T1-Sca-1$^+$ cells and Tu-Gr1$^+$CD11b$^+$-primed 4T1 cells (Supplemental Figure 2E, G) that, upon isolation or induction reverted to a Sca-1$^-$ phenotype, suggesting that MR13 line was capable of self-sustain its own Sca-1$^+$ population.

Taken together, chemotherapy-selected MR13 cells share some similar *in vivo* characteristics as 4T1-Sca-1$^+$ cells, while they are capable of self-sustaining high Sca-1$^+$ abundancy both *in vitro* and *in vivo.*

IL6/IL6R-JAK autocrine signaling maintains Sca-1 positivity and metastatic capacity in MR13 cells

To better understand the chemotherapy-induced alterations in MR13 cells, we performed transcriptomic analyses comparing MR13 and parental 4T1 cells. Pathway enrichment analysis showed that the IL6-JAK-STAT3 signature was also elevated in MR13 cells (Figure 7A). Importantly, MR13 gene expression significantly positively correlated with the Sca1 Positive signature. At the same time, it negatively correlated with the Sca1 Negative signature (Figure 7B). This observation suggested that chemotherapy enriched for the inherent 4T1-Sca-1$^+$ population, rather than converting plastic Sca-1$^-$ cells, as observed upon Tu-Gr1$^+$CD11b$^+$-stimulation. In addition, *Osmr*, *Il6*, *Il6ra* and *Il6st* were all overexpressed in MR13 cells compared to parental 4T1 cells, while the *Osm* expression was not altered (Figure 7C). To functionally validate the IL6-JAK signaling pathway in Sca-1$^+$ population maintenance, we treated MR13 cells with Ruxolitinib *in vitro* for 48 hours. The treatment significantly decreased the fraction of Sca-1$^+$ cells (Figure 7D). Importantly, *in vitro* treatment of MR13 cells with Ruxolitinib for 3 days nearly completely abolished their lung metastatic capacity upon tail vein injection (Figure 7E-F).

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Altogether, these data suggest that MR13 cells sustain the metastatic Sca-1$^+$ population through cell-autonomous activation of the IL6-JAK signaling pathway.

Tu-Gr1$^+$CD11b$^+$ invoked tumor cell signature predicts shorter overall and relapse-free survival in breast cancer patients

To evaluate the clinical relevance of the crosstalk between Tu-Gr1$^+$CD11b$^+$ and tumor cells, we tested whether the Tu-Gr1$^+$CD11b$^+$-induced 4T1 signature could predict cancer progression in patients. To this end, we interrogated the METABRIC dataset (56) with the Tu-Gr1$^+$CD11b$^+$-induced signature. Thirty-two human orthologue genes (Supplemental Table 3) in the murine 50 genes signature were present in the METABRIC dataset. Patients with the higher expression level of the signature had shorter overall survival (OS; $p = 0.0056$) and relapse-free survival (RFS; $p = 0.032$) (Figure 8A, B). Notably, OSM expression positively correlated with the signature in all patients (Supplemental Figure 7A), suggesting OSM do contribute to altered signature expression in patients. Of the thirty-two genes, five (*MX1, IRF7, OAS1, CMPK2, ISG15*) were commonly discriminant for a shorter OS and RFS (Figure 8C, D), and taken together, they further enhanced the predictive power (OS: $p=0.00055$; RFS: $p=0.00069$) (Figure 8E, F).

These data indicate that a small set of genes issued from the murine tumor cell signature invoked by Tu-Gr1$^+$CD11b$^+$ cells can predict a shorter OS and RFS in breast cancer patients, thereby reinforcing the clinical significance of the proposed model.

Discussion

Metastatic disease and therapy resistance are the leading causes of breast cancer mortality, calling for novel approaches to effectively prevent and cure for metastasis and therapy resistance. This is particularly relevant for TNBC, where in spite of the recent

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

encouraging results with targeted therapies, such as PARP inhibitors for tumors with germline BRCA mutations, or checkpoint inhibitors for PD-L1$^+$ tumors, management of metastatic disease remains challenging (57–60). It has been proposed that CSCs present in the primary tumor are responsible for tumor persistence, metastasis, and therapy resistance (52, 61). Enrichment and differentiation of CSC contribute to tumor heterogeneity (17, 19–22). Importantly, CSC features can be intrinsic or plastic and be modulated by cues from the TME (62–64).

Here, we have interrogated the contribution of the Sca-1$^+$ population to breast cancer metastasis and its modulation by the TME. We reported that tumor-educated Gr1$^+$CD11b$^+$ cells (Tu-Gr1$^+$CD11b$^+$) instigate cancer metastasis by twisting cancer cell plasticity and enriching for a Sca-1$^+$ population with enhanced metastatic capacity. We identified OSM/IL6-JAK as a paracrine communication axis between Tu-Gr1$^+$CD11b$^+$ and breast cancer cells and as an autocrine loop in chemotherapy-resistant tumor cells, promoting tumor heterogeneity, CSC features and metastatic capacity. Importantly, breast cancer patients expressing high levels of the human orthologues of the gene expression signatures invoked by Tu-Gr1$^+$CD11b$^+$ have significantly shorter OS and RFS, reinforcing the clinical significance of our findings. While some of these elements have been reported individually before, our results extend these observations by providing an integrative view of paracrine (Tu-Gr1$^+$CD11b$^+$-induced) and autocrine (chemotherapy-induced) communication in regulating tumor heterogeneity, cancer cell plasticity, metastasis, and resistance to chemotherapy.

A main observation stemming from this study is that Sca-1$^+$ population exists under three different conditions: as inherent 4T1-Sca-1$^+$ population, upon exposure to Tu-Gr1$^+$CD11b$^+$ cells and upon chemotherapy treatment (MR13). All these populations had higher metastatic abilities compared to their counterpart controls. GSEA analysis

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

suggested that the Tu-Gr1$^+$CD11b$^+$-induced Sca-1$^+$ population was likely to convert from the Sca-1$^-$ population (Figure 3C and Figure 4), while the Sca-1$^+$ population surviving chemotherapy (MR13) appeared to be enriched from the inherent Sca-1$^+$ population (Figure 7B). In addition, the different gene expression signature of Tu-Gr1$^+$CD11b$^+$-induced Sca-1$^+$ population relative to the inherent Sca-1$^+$ cells suggest a remarkable functional plasticity of these cells. This plasticity is further supported by comparing single cell gene expression of murine 4T1 and human MCF-7 breast cancer cells, using publicly available scRNA-seq datasets, before and after *in vivo* growth (Figure 4). *In vivo*, tumor cells undergo a transformation enriching for Sca-1 (in the mouse), OSMR expressing cells, Sca1 Positive, Sca1 Negative and Tu-Gr1$^+$CD11b$^+$-induced signatures, and IL6-JAK signaling pathway from precursor cells (Sca-1$^-$ in the mouse) (Figure 4). Gong *et al.* have reported that sorted Sca-1$^-$ 4T1 cells could be transiently transformed into a Sca-1$^+$ population by radiotherapy (65). An analogous observation was reported in colorectal cancer, where selective ablation of LGR5$^+$ CSCs in organoids leads to initial tumor regression, followed by regrowth driven by LGR5$^+$ CSCs reemerging from the LGR5$^-$ population (66). Taken together, by demonstrating that tumor recruited and educated Gr1$^+$CD11b$^+$ cells contribute to such plasticity by inducing the conversion of low metastatic Sca-1- population into a highly metastatic Sca-1$^+$ population, our observations further consolidate the notion that cancer consists of a heterogenous and plastic tumor mass, including highly metastatic cell populations. Dedicated time course scRNA-seq analyses together with lineage tracing experiments may help to further characterize the detailed origin, development, fate, and function of these Sca-1$^+$ populations during cancer progression.

Recruitment and accumulation of Gr1$^+$CD11b$^+$ cells in the TME, particularly through the chemokines CCL2, CXCL1 and CXCL2, or IL-33, is considered a critical step for their

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

contribution to tumor progression and metastasis (67–70). Consistent with these observations, tumors derived from sorted 4T1-Sca-1^{+} cells, or MR13 cells that are intrinsically enriched for Sca-1^{+} cells, have a higher content of Gr1^{high}CD11b^{+}Ly6C^{low} cells compared to tumors derived from 4T1-Sca-1^{-} or parental 4T1 cells, respectively (Figure 2A and 5K). Beyond recruitment, tumor-mediated education of Gr1^{+}CD11b^{+} cells, appears to be necessary to gain higher metastatic activity: Only Gr1^{+}CD11b^{+} cells recovered from primary tumors (but not from spleen or bone marrow) induced Sca-1 positivity and enhanced metastatic ability in 4T1 cells (Figure 2C, E). The role of Gr1^{+}CD11b^{+} cells in promoting metastasis has been mainly attributed to promotion of angiogenesis, EMT and immunosuppression (17, 68). Peng *et al.*, reported the ability of those cells to endow CSC-like features to breast cancer cells but their metastatic capacity was not interrogated (17). Our observations suggest a self-sustaining positive feedback mechanism between highly metastatic cancer cells (Sca-1^{+} population) and Gr1^{+}CD11b^{+} cells: inherent Sca1^{+} CSC-like, metastatic, cells promote recruitment and local education of Gr1^{+}CD11b^{+} cells, which in turn promote tumor heterogeneity, cancer cell plasticity and metastatic capacity by converting low-metastatic Sca1^{-} cells into additional high-metastatic Sca1^{+} cells (Figure 9). Recently, it was reported that neutrophils escorting blood circulating tumor cells (CTCs) expands the metastatic potential of CTCs (71). While this effect was attributed to the promotion of cell cycle progression of CTCs through direct contact with the neutrophils, in light of our findings, one may also consider the possibility that clustered neutrophils may also promote the expansion of a CSCs-like phenotype with higher metastatic capacity. Interestingly, OSM was reported to be expressed by neutrophils cocultured with breast cancer cells (36) and to promote phenotypic changes associated with mesenchymal and stem cell-like differentiation in breast cancer (36, 72). Together with our observation, these findings further reinforce the notion that boosting the *Osm* expression in Gr1^{+}CD11b^{+} cells

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

is part of their educating program prompted by the tumor. One outstanding question raised by these observations is by which mechanisms and pathways cancer cells, in particular Sca-1$^+$ ones, educate Gr1$^+$CD11b$^+$ cells to acquire cancer plasticity-promoting activity.

Besides directly activating tumor cells, OSM has also been shown to remodel macrophages and fibroblasts of the TME (33, 36). Araujo *et al.* recently reported that OSM derived from tumor-infiltrating myeloid cells reprogram fibroblasts to secrete VEGF and the chemokines CXCL1 and CXCL16, resulting in enhanced myeloid cell recruitment and breast cancer progression (33). Here we extend these observations by demonstrating that tumor-educated Gr1$^+$CD11b$^+$-derived OSM/IL6 twist cancer cell plasticity by promoting a rapid but reversible conversion of Sca-1$^-$ cells into more metastatic Sca-1$^+$ cells (Figure 5B). We further broaden the implications of the OSM/IL6-JAK axis, by demonstrating that MR13 cells that have escaped chemotherapy hijacked this paracrine mechanism in a cell autonomous manner by elevating the *Il6* expression (Figure 9). Significantly, a short *in vitro* treatment with Ruxolitinib effectively abrogated their metastatic capacity (Figure 7E, F).

While OSM/OSMR is the only interaction pair identified by the cell-cell interaction analysis (Figure 3F, G), it is possible that other molecular mechanisms may also play roles in modulating Sca-1$^+$ cell plasticity and metastasis. One candidate is the IL6/IL6R communication axis, which was less prominent in the cell-cell communication analysis (Figure 3F, G), but highly expressed in our experimental data in Tu-Gr1$^+$CD11b$^+$ compared with Spl-Gr1$^+$CD11b$^+$ (Figure 5A) and in chemotherapy-resistant MR13 cells compared with 4T1 (Figure 7C). Moreover, the neutralization of IL6 also suppressed the Tu-Gr1$^+$CD11b$^+$ conditioned medium induced Sca-1$^+$ cell enrichment, much alike OSM inhibition (Figure 5C).

Importantly, we demonstrated that a human orthologue signature of the 4T1 gene expression signature invoked by Tu-Gr1$^+$CD11b$^+$ can predict a significantly shorter OS and

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

RFS in breast cancer patients (Figure 8). This finding strengthens the clinical significance of the observed crosstalk between Tu-Gr1$^+$CD11b$^+$ and tumor cells. Strikingly, the five genes that significantly contribute to the discriminatory power of the signature are genes related to native or viral immunity or regulated by interferon. While expression of interferons and interferon response genes in breast cancer has been mainly associated with tumor suppression and improved survival (73), there is evidence also correlating interferon responses with tumor promotion, therapy resistance and reduced survival (74). As JAKs/STATS are activated by both IFN and OSM/IL6 receptors, it is conceivable that OSM/IL6 only activates a subset of the IFN-induced genes with tumor-promoting activity, as the case for Mx1 (75). Consistent with our findings, JAK/STAT signaling has recently been shown to initiate the lineage plasticity in prostate cancer as well as to promote lineage plasticity-driven targeted therapy resistance in a stem-like subpopulation of prostate cancer (76, 77). On the other hand, Aouad *et al.* showed that epithelial-mesenchymal plasticity is essential for the generation of a dormant cell state of ER$^+$ breast cancer during progression, and the activation of IL6-JAK-STAT signaling triggers tumor cell awakening and recurrence (48).

One crucial question raised by these results is whether the OSM/IL6-JAK pathway is a potential actionable clinical target to impinge on metastatic progression and therapy resistance. In particular, the observation that a short *in vitro* treatment of the highly metastatic MR13 cells profoundly suppressed their metastatic capacity *in vivo* (Figure 7E, F), suggests potential long-lasting effects consistent with an adjuvant effect. JAK inhibitors are being tested in clinical trials in breast cancer. A phase I study combining Ruxolitinib with paclitaxel in HER2-negative metastatic breast cancer showed good tolerability and evidence of activity (78). A phase I/II trial of Ruxolitinib in combination with trastuzumab in metastatic HER2 positive breast cancer and a phase II study combining Ruxolitinib with

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

capecitabine in advanced HER2$^-$ breast cancer, however, did not improve progression-free survival (79, 80). The absence of benefits in these studies in advanced breast cancer and our reported mechanistic observations on metastatic progression, raise the question of whether the JAK inhibitors should be considered in adjuvant setting in high-risk patients, to prevent progression to metastases, rather than treating patients already bearing metastases.

In conclusion, we reported here that a subpopulation of tumor cells within the tumor mass educates requited Gr1$^+$CD11b$^+$ cells to convert a low metastatic subpopulation into highly metastatic one, through the OSM/IL6-JAK signaling axis (Figure 9). The clinical relevance of this observation is supported by human transcriptomic data. This process is hijacked by tumor cells that survived chemotherapy and evolved toward a highly metastatic phenotype via cell autonomous IL6-JAK signaling. Importantly, a short *in vitro* treatment with a clinically approved JAK inhibitor, Ruxolitinib, suppresses their metastatic capacity these cells *in vivo.* These results should stimulate considering testing JAK inhibitors in the adjuvant setting in TNBC breast cancer patients at high-risk for metastatic progression.

Methods

Cell culture. The 4T1 murine breast cancer cell line was kindly provided by Dr Fred R. Miller (Michigan Cancer Foundation, Detroit, MI, USA). 4T1 cells were cultured in high glucose DMEM supplemented with 10% heat-inactivated FBS, 1% penicillin–streptomycin (P/S, from Gibco) and 1% Non-Essential Amino Acid (Gibco).

Tumor models. 4T1, MR13, sorted 4T1-Sca-1$^+$ and 4T1-Sca-1$^-$ (5x10^4 cells in 50µl PBS/10% of 8.1 mg/ml Matrigel Matrix, were injected in the fourth right mammary gland of

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

mice. Prior to surgery, ketamine (1.5 mg/kg) and xylazine (150 mg/kg) (both from Graeub) were injected intra-peritoneally to anesthetize the animals. Immune cell populations were analyzed at different time points post-tumor cell injection. Tumor length and width were measured twice a week with caliper and used to calculate tumor volume by the following equation: volume = (length x width^2) x π/6. Tumors were collected and weighted at necropsy. For the intravenous injections, 2x10^5 sorted 4T1-Sca-1$^+$ and 4T1- Sca-1$^-$ tumor cells resuspended in a volume of 50 µl of PBS were injected into the mice tail vein. Lung metastases were quantified 10 days post-injection. At each indicated time point, mice were sacrificed according to defined ethical criteria and were killed by CO_2 inhalation followed by neck dislocation or terminal bleeding. All animal procedures were performed in accordance with the Swiss legislations on animal experimentation and approved by the Cantonal Veterinary Service of the Cantons Vaud and Fribourg for experiments in Lausanne and Fribourg (VD_1486.2; 2017_34_FR; 2017_34_FR, 2014_58_FR; 2011-33-FR).

Reagents and chemicals. Growth factor reduced Matrigel Matrix (MG) was obtained from Becton Dickinson (BD Biosciences). Collagenase I was purchased from Worthington and DNAse I from Roche. Bovine serum albumin (BSA), crystal violet (CV) and paraformaldehyde (PFA) were obtained from Sigma-Aldrich. Drugs, inhibitors and cytokines: Doxorubicin and Methotrexate (generously provided by the Department of Oncology, University Hospital, University of Lausanne, Lausanne, Switzerland), Ruxolitinib (JAK inhibitor, Cat N°7064, Biotechne), anti-mouse Oncostatin M (R&D systems), anti-mouse IL-6 (BioxCell), recombinant mouse Oncostatin M and IL-6 (Biolegend, Cat #: 762802 & 575702 respectively).

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Antibodies. The following anti-mouse antibodies were used following manufacturer's instructions: anti-CD16/CD32 Fc blocking antibody (BD Biosciences), anti-CD24-FITC (clone M1/69, eBioscience), anti-CD29-PE and-PE-Cy5 (clone HMβ1-1, BioLegend), anti-Sca-1-APC (clone D7, eBioscience), anti-CD61-Alexa 647 (clone 2C9.62(HMβ3-1), BioLegend, anti-CD45-PE (clone 30-F11, BD Biosciences), anti-Gr1-eFlour450 (clone RB6-8C5,eBioscience), anti-Ly6C FITC (clone HK1.4, BioLegend), anti-Ly6G-APC (clone 1A8-Ly6g,eBioscience), anti-F4/80 PerCP/Cy5 (clone BM8, eBioscience), anti-CD11b-PE-Cy7 (clone M1/70,eBioscience), anti-CD11c-APC-eFluor780 (clone N418, eBioscience), anti-CD4-FITC (clone 6K1.5, eBioscience), anti-CD8-PE (clone 53-6.7, eBioscience), anti-B220-APC (clone RA3-6B2, eBioscience), anti-CD49b-eFlour450 (clone DX5, BioLegend), Annexin V-APC (clone B217656, BioLegend), Propidium Iodide-PerCP (Clone V13245, Life technologies).

Magnetic cell sorting (MACS). MACS separators were used for positive and negative cell selections based on manufacturer's instructions. Briefly, cells were counted and resuspend in 500 µl of MACS buffer with 10 µl of fluorescent coupled antibody of interest (APC-conjugated anti-Sca-1 and PE-conjugated anti-Gr1) per 10^7 cells added. Cells were incubated for 30 minutes in the dark at 4°C and then washed with MACS buffer. After centrifugation and resuspension in 80 µl of MACS buffer per 10^7 cells, 20 µl of anti- APC-conjugated magnetic beads (Miltenyi Biotec) per 10^7 cells was added. After a 20 minutes incubation in the dark at 4°C, cells were washed with MACS buffer. After the centrifugation they were resuspended in MACS buffer and the magnetic separation was performed using LS MACS column (maximum 10^8 labeled cells) for positive selection and LD MACS column (maximum 10^8 labeled cells) for negative selection (Miltenyi Biotec). The purity of positive subpopulation was >70% and <99% for the negative subpopulation.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Co-culture of Gr1$^+$CD11b$^+$ cells sorted from tumors or spleen of 4T1 tumor bearing mice. Gr1$^+$ cells were sorted from tumors or spleen at day 23 post-injection (see below Magnetic beads cell sorting section). Afterwards, 1.5x10^5 4T1 cells were co-cultured in 6 wells plates using Transwell plates (0.4 μm, Nunc, ThermoFisher Scientific) with Gr1$^+$CD11b$^+$ sorted cells (top) and 4T1 cells (bottom) at different 4T1: Gr1$^+$CD11b$^+$ ratios (1:1, 1:3 and 1:5). After 48 hours of co-culture, cells were analyzed for Sca-1 expression and gene expression by flow cytometry and semi-quantitative real-time qPCR as indicated.

Flow cytometry analysis on tissue samples. Mice were sacrificed at different time points for blood and tumors collection. Tumors were cut in small pieces with scissors, washed, and digested in serum free medium supplemented with Collagenase I and DNAse I (Roche). The mixture was incubated at 37 °C for 45 minutes on a shaking platform. Subsequently, serum-supplemented medium was added to neutralize the enzymatic reaction and the tissue suspensions were filtered through a 100 µm and a 70 µm sterile nylon gauzes. Upon centrifugation (5 minutes at 1400 rpm), pellets were recovered and red blood cells lysed with ACK buffer (Biolegend). The staining procedure and the flow cytometry acquisition are as described previously (81). Data acquisition was performed using the FACSCalibur (BD Biosciences) or MACSQuant flow cytometer from Miltenyi Biotec and data analyzed by FlowJo v10.0.7 (tree Stat Inc.).

***In vitro* cell proliferation assay**. Cells were collected and seeded in tissue culture 96-wells-plates (Costar) at 3’000 cells/well. Cells were grown in complete medium for 24, 36, 72 and 96 hours. At each time point cells were washed once with PBS, then fixed with 4% PFA and stained with 0.5% crystal violet solution for 0.5 hours. The stained cells were

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

gently washed with deionized water to remove the extra dye and air-dried overnight at room temperature. After solubilizing the dye with crystal violet eluting buffer (70% ethanol and 1% acetic acid), cell viability was assessed by reading the absorbance at 595 nm wavelength in a multiwell plate reader (Modulus II microplate reader, Turner Biosystems). Results were analyzed by Prism (Graph pad software, Inc.) expressed as mean values of optical density (OD) of octuplet determinations ± SEM.

***In vitro* cytotoxic assay.** Tumor cells were plated at a concentration of 3'000 cell/well into 96-wells plates. The following day, a series of concentrations of the different drugs were supplemented to the culture medium. Untreated control cells were kept in normal culture medium. Cell viability of each well was assessed with crystal violet staining 48 hours after treatment, as described above. Results were analyzed by Prism software by a non-linear regression analysis and expressed as relative cell viability compared with non-treated control. The 50% maximum inhibition concentrations (IC_{50}) were used to determine the drug-resistant ability of treated cells.

Real-time reverse transcription qPCR and primers. Changes in mRNA expression levels were determined by semi-quantitative real-time qPCR. RNA samples were obtained from adherent cells using RNeasy kit from QIAGEN according to manufacturer's instructions. From each sample, 1 µg RNA was retro-transcribed using SuperScript II Reverse Transcriptase kit (Life Technologies – Invitrogen), according to manufacturer's instructions. The reactions were performed in a StepOnePlus™ thermocycler (Applied Biosystems, Life Technologies – Invitrogen) using the KapaSYBR® FAST SYBR Green Master Mix (Kapa Biosystems). Each reaction was performed in triplicates and values were normalized to murine 36B4 housekeeping gene. The comparative C_t method was used to

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

calculate the difference of gene expression between samples. The following murine-specific primers (Microsynth AG) were used:
Sca-1 (F:5'-TCAGGAGGCAGCAGTTATTGTG-3',R: 5'-TGGCAACAGGAAGTCTTCACG-3'), 36B4(F:5'-GTGTGTCTGCAGATCGGGTAC-3',R:5'-CAGATGGATCAGCCAGGAAG-3'), OSM (F: 5'- ATGCAGACACGGCTTCTAAGA-3', R: 5'- TTGGAGCAGCCACGATTG G-3'), OSMR: (F: 5'- CATCCCGAAGCGAAGTCTTGG-3', R: 5'- GGCTGGGACAGTCCATTC TAAA-3'), IL-6: (F: 5'- TACCACTTCACAAGTCGGAGGC-3', R: 5'- CTGCAAGTGCATCAT CGTTGTTC-3').

Histopathology. Tumors and lungs were harvested at the end of the experiments, fixed in formalin and embedded in paraffin. 5 µm thick serial sections were cut from the tissue blocks. 3-4 sections taken at 100 µm distance were stained with hematoxylin and eosin (H&E) and used to assess tumor morphology and quantify lung metastasis. Slides were scanned by Nanozoomer (Hamamatsu Photonics) and metastasis were counted manually using NDP.viewer2 software (Hamamatsu Photonics). Metastatic index was calculated by normalizing the metastasis number with the volume of primary tumor.

Mammosphere-forming assay 5'000 cells/well were seeded in non-adhesive U bottom 96-wells plate in a semi-solid MEGM medium supplemented with 20 ng/ml EGF, and 20 ng/ml bFGF and heparin. Medium was gently replaced every 3 to 4 days. After 11 to 14 days culture, mammospheres of 50 to 150 µm diameter were detected under the microscope (bright field) and counted to quantify the sphere formation efficiency (SFE) as percentage of the initial number of seeded cells per well.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Bulk RNA sequencing and data analysis. Four independent sorts of 4T1-Sca-1$^+$ and 4T1-Sca-1$^-$ cells (as described in the magnetic cell sorting section) or four independent 4T1 cells primed with Tu- Gr1$^+$CD11b$^+$ or Spl-Gr1$^+$CD11b$^+$ were prepared. RNAs of these cells were isolated using the NucleoSpin RNA protocol of Macherey-Nagel (as described in real-time (RT) qPCR and primers section). Samples were normalized for 1 µg RNA in a volume of 20 µl, sequenced on the NextSeq500 sequencer using the NextSeq 500/550 HT reagent v2 kit (Illumina) at the Swiss Integrative Center for Human Health (SICHH) in Fribourg, or the Lausanne Genomics Technologies Facility (GTF, UNIL) in Lausanne, Switzerland. For data analysis, all sequencing reads were processed for quality control, removal of low quality reads, adaptor sequence and ribosomal RNA by fastqc(0.11.8) (82), multiqc (1.9) (83), Trimmomatic (0.39) (84) and SortMeRNA(2.1) (85) accordingly. The filtered reads were mapped to the reference genome (mm10) using htseq-count (0.6.1) (86) or Salmon (0.99.0) (87). The normalization of the read counts and the analysis of the differential expression between the groups of samples were performed with in R(v4.1.3), a free software environment available at https://www.r-project.org/ using packages DESeq2 (v3.15) (88). Pathway enrichment analysis was performed using packages GSVA(1.42.0) (89) and GSEABase(1.56.0) (90) with the input of DESeq2 normalized count numbers using ssgsea method comparing the Hallmark genesets from MSigDB (v7.4.1) (91) with default settings. The significant altered pathways were determined by computing moderated t-statistics and false discovery rates with the limma(3.50.3)(92) for pair-wised comparison. The heatmaps were produced with R package pheatmap(1.0.12)(93) with default settings while pathway hierarchy clustering was performed by similarity based on Euclidean distance and the ward aggregation algorithm. The Sca1 Positive and Sca1 Negative signatures were extracted from the top 200 most upregulated or downregulated genes in 4T1-Sca-1$^+$ and 4T1-Sca-1$^-$ RNAseq data, respectively, with the threshold of

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

adjusted p-value <0.05, fold change >1.5 or <-1.5 and average normalized count number >20. For Venn diagram, the genes fulfilling the threshold of adjusted p-value < 0.05, fold change > 1.5 or < -1.5 and average normalized count number >20 are compared. The figures were produced with R package venn (1.10) (94) . Further analysis and figures generation were performed in R using packages tidyverse(1.3.1)(95), ggplot2(3.3.6) (96), circlize(0.4.15) (97), biomaRt(2.50.3) (98, 99), RColorBrewer (1.1-3) (100), clusterProfiler (4.2.2) (101, 102), enrichplot (1.14.2) (103), ggpubr(v0.4.0)(104) ggbreak (0.1.0) (104).

Microarray hybridization and data analysis Experiments were performed as previously described (105). Briefly, triplicates wells of cultured 4T1 and MR13 cells were used for RNA extraction using RNeasy kit (QIAGEN). Probe synthesis and GeneChip Mouse Gene Exon 1.0 ST Array (Affymetrix Ltd) hybridization were performed at the GTF, UNIL, Lausanne, Switzerland. Microarray analyses were carried out with R. After quantification of gene expression with robust multi-array normalization (99) using the BioConductor package Affy, (http://www.bioconductor.org/) significance of differential gene expression was determined by computing moderated t-statistics and false discovery rates with the limma package (92). Annotation was based on the genome version NCBI Build 36 (Feb. 2006). The obtained p-values were corrected for multiple testing by calculating estimated false discovery rates (FDR) using the method of Benjamini-Hochberg. Heatmaps were produced by color-coding gene-wise standardized log gene expression levels (mean zero standard deviation one). Probe-sets were shown hierarchically clustered by similarity based on Euclidean distance and the ward aggregation algorithm.

Public RNAseq data analysis The RPKM normalized gene expression data in Ross dataset (GSE150928) from multiple murine models of breast cancer metastasis was

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

752 obtained from Gene Expression Omnibus (GEO) in the NCBI data repository. The data
753 were analyzed and plot with R package tidyverse(1.3.1)(95), ggplot2 (3.3.6) (96) and
754 ggpubr (v0.4.0) (106). The single-cell RNAseq data in Sebastian dataset (43) was obtained
755 from Dryad data repository (https://doi.org/10.6071/M3238R). The data were analyzed with
756 R package Seurat (3.0) (107). The tumor and different myeloid cell populations were
757 extracted for cell-cell interaction analysis using CellPhondDB (2.0) (44). The identified
758 interaction pairs were extracted and plot using circlize (0.4.15) (97) and ComplexHeatmap
759 (2.11.2) (108). The ligands and receptors annotated in the circular plot were compiled from
760 databases in CellTalkDB(109), SingleCellSignalR (110). To investigate the tumor cell
761 dynamics, datasets for 4T1 (GSE158844 and GSM3502134) and MCF-7 (GSM4681765
762 and GSM5904917) were obtained from GEO. The data were filtered and normalized
763 separately before merging with IntegrateData function included in Seurat. Cell cycle
764 regression was performed according to the standard protocol of Seurat. For MCF-7, due
765 the huge difference of the sample size, 2000 cells were randomly selected from
766 GSM4681765 data, and then merged with GSM5904917. Single-cell trajectories analysis
767 was performed with monocle 3 (111–114) with default settings and the root and start point
768 were selected manually for pseudotime caculation. The GSEA analysis of selected clusters
769 was performed with R package fgsea(115) and the gene rank was calculated with Wilcoxon
770 rank sum test and auROC analysis using wilcoxauc function included in presto package.
771 The Tu-Gr1^{+}CD11b^{+}-induced signature was extracted by comparing gene expression
772 between Tu-Gr1^{+}CD11b^{+} and Spl-Gr1^{+}CD11b^{+}-educated 4T1 cells with adjusted p-value
773 <0.05, fold change >2, and the top 50 genes were selected.
774
775 **Clinical data analysis.** To validate our finding in clinical data, the human orthologs of
776 murine Tu-Gr1^{+}CD11b^{+}-induced signature genes were used. Conversion from murine to

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

human gene symbols and Entrez IDs was performed with the biomaRt package (2.46.3) (98, 99), using the reference mart https://dec2021.archive.ensembl.org. Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) breast cancer data was downloaded from cBioPortal (116–118) in August 2022, and expression data was log2 transformed. Expression values were stratified in two groups by median values. Survival curves were generated using the ggsurvplot function from the survminer package (0.4.9)(106), and were compared between groups using a log-rank test. Survival curves were created using the survfit function from the survival package (3.2.11) (119). Cox proportional hazard regression model was performed through the coxph function of the same package.

Statistical analyses. Unless specified, the data were presented as mean ± SEM from at least 3 independent experiments, unless otherwise indicated. Statistical comparisons were performed by an unpaired Student's t test with a two-tailed distribution, one-way ANOVA analysis of variance with Bonferroni post-test or two-way ANOVA with Tukey's or Dunnett's multiple comparison test using Prism 7.0 GraphPad Software, Inc. or R package ggpubr (v0.4.0)(104).

Graphic illustrations. Illustrative schemes were created with BioRender.com.

Data and code availability. The raw and processed bulk RNAseq data used to generate figures in Figure 3 and Supplemental Figure 5 and microarray data used to generate figures in Figure 7 have been deposited in the GEO database under the access code GSEXXXXXX. The code used for the analyses is open-source and available through with R packages described in methods.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Author contributions

Conceptualization, S.P., Q.L., C.R.; Methodology, S.P., N.F., Q.L., C.R.; Software, A.K., N.F., Q.L.; Validation, S.P., M.B., A.K., Q.L.; Formal analysis, S.P., M.B., A.K., Q.L.; Investigation, S.P., M.B., A.K., O.C., Y-T.H., N.D., L.G., Q.L.; Resources, S.P., M.B., A.K., O.C., N.D., L.G., G.L.; Data Curation, S.P., M.B., A.K., N.D., Q.L., C.R.; Writing – original draft preparation, S.P., Q.L., C.R.; Writing – review and editing, S.P., M.B., A.K., Y-T.H., G.L., N.D., Q.L., C.R.; Visualization, S.P., M.B., A.K., Q.L., C.R.; Supervision, S.P., Q.L., C.R.; Project administration, Q.L., C.R.; Funding acquisition, C.R.

Acknowledgments

The authors wish to thank Sarah Cattin and Melissa Rizzo for assistance with FACS analysis, Dr. Ana-Marija Sulić (Institute of Biotechnology, HiLIFE, University of Helsinki) for the insightful discussion about scRNA-seq data analysis, Dr. Fred R. Miller (Michigan Cancer Foundation, Detroit,MI, USA) for providing 4T1 cells, Dr. Khalil Zaman (Department of Oncology, University Hospital, University of Lausanne, Lausanne, Switzerland) for providing chemotherapeutic drugs. This work was supported by grants from the Swiss National Science Foundation (31003A_179248; 310030_208136), the Swiss Cancer League (KFS 4400-02-2018) and the Medic Foundation (to C.R.).

References

1. Fares J, Fares MY, Khachfe HH, Salhab HA, Fares Y. Molecular principles of metastasis: a hallmark of cancer revisited. *Signal Transduct. Target. Ther.* 2020;5(1):1–17.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

2. Valdés-Mora F et al. Single-cell transcriptomics reveals involution mimicry during the specification of the basal breast cancer subtype. *Cell Rep.* 2021;35(2):108945.

3. Yang D et al. Lineage tracing reveals the phylodynamics, plasticity, and paths of tumor evolution. *Cell* 2022;185(11):1905-1923.e25.

4. Torborg SR, Li Z, Chan JE, Tammela T. Cellular and molecular mechanisms of plasticity in cancer. *Trends Cancer* 2022;8(9):735–746.

5. Baccelli I, Trumpp A. The evolving concept of cancer and metastasis stem cells. *J. Cell Biol.* 2012;198(3):281–293.

6. Lawson DA, Kessenbrock K, Davis RT, Pervolarakis N, Werb Z. Tumour heterogeneity and metastasis at single-cell resolution. *Nat. Cell Biol.* 2018;20(12):1349–1360.

7. Celià-Terrassa T, Jolly MK. Cancer Stem Cells and Epithelial-to-Mesenchymal Transition in Cancer Metastasis. *Cold Spring Harb. Perspect. Med.* 2020;10(7):a036905.

8. Dagogo-Jack I, Shaw AT. Tumour heterogeneity and resistance to cancer therapies. *Nat. Rev. Clin. Oncol.* 2018;15(2):81–94.

9. McDonald K-A et al. Tumor Heterogeneity Correlates with Less Immune Response and Worse Survival in Breast Cancer Patients. *Ann. Surg. Oncol.* 2019;26(7):2191–2199.

10. Hata AN et al. Tumor cells can follow distinct evolutionary paths to become resistant to epidermal growth factor receptor inhibition. *Nat. Med.* 2016;22(3):262–269.

11. Ramirez M et al. Diverse drug-resistance mechanisms can emerge from drug-tolerant cancer persister cells. *Nat. Commun.* 2016;7(1):10690.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

12. Oskarsson T, Batlle E, Massagué J. Metastatic stem cells: sources, niches, and vital pathways. *Cell Stem Cell* 2014;14(3):306–321.

13. Ye F et al. CD49f Can Act as a Biomarker for Local or Distant Recurrence in Breast Cancer. *J. Breast Cancer* 2017;20(2):142–149.

14. Vaillant F et al. The Mammary Progenitor Marker CD61/β3 Integrin Identifies Cancer Stem Cells in Mouse Models of Mammary Tumorigenesis. *Cancer Res.* 2008;68(19):7711–7717.

15. Li JJ, Tsang JY, Tse GM. Tumor Microenvironment in Breast Cancer—Updates on Therapeutic Implications and Pathologic Assessment. *Cancers* 2021;13(16):4233.

16. Ge R, Wang Z, Cheng L. Tumor microenvironment heterogeneity an important mediator of prostate cancer progression and therapeutic resistance. *Npj Precis. Oncol.* 2022;6(1):1–8.

17. Peng D et al. Myeloid-Derived Suppressor Cells Endow Stem-like Qualities to Breast Cancer Cells through IL6/STAT3 and NO/NOTCH Cross-talk Signaling. *Cancer Res.* 2016;76(11):3156–3165.

18. Li K et al. Myeloid-derived suppressor cells as immunosuppressive regulators and therapeutic targets in cancer. *Signal Transduct. Target. Ther.* 2021;6(1):1–25.

19. Lorusso G, Ruegg C. The tumor microenvironment and its contribution to tumor evolution toward metastasis. *Histochem Cell Biol* 2008;130(6):1091–103.

20. Sleeman JP et al. Concepts of metastasis in flux: the stromal progression model. *Semin Cancer Biol* 2012;22(3):174–86.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

21. Cha YJ, Koo JS. Role of Tumor-Associated Myeloid Cells in Breast Cancer. *Cells* 2020;9(8). doi:10.3390/cells9081785

22. Pastaki Khoshbin A, Eskian M, Keshavarz-Fathi M, Rezaei N. Roles of Myeloid-Derived Suppressor Cells in Cancer Metastasis: Immunosuppression and Beyond. *Arch Immunol Ther Exp Warsz* 2019;67(2):89–102.

23. Panni RZ et al. Tumor-induced STAT3 activation in monocytic myeloid-derived suppressor cells enhances stemness and mesenchymal properties in human pancreatic cancer. *Cancer Immunol. Immunother.* 2014;63(5):513–528.

24. Cui TX et al. Myeloid-derived suppressor cells enhance stemness of cancer cells by inducing microRNA101 and suppressing the corepressor CtBP2. *Immunity* 2013;39(3):611–621.

25. Abaurrea A, Araujo AM, Caffarel MM. The Role of the IL-6 Cytokine Family in Epithelial–Mesenchymal Plasticity in Cancer Progression. *Int. J. Mol. Sci.* 2021;22(15):8334.

26. Zarling JM et al. Oncostatin M: a growth regulator produced by differentiated histiocytic lymphoma cells. *Proc. Natl. Acad. Sci. U. S. A.* 1986;83(24):9739–9743.

27. Felcher CM, Bogni ES, Kordon EC. IL-6 Cytokine Family: A Putative Target for Breast Cancer Prevention and Treatment. *Int J Mol Sci* 2022;23(3). doi:10.3390/ijms23031809

28. Johnson DE, O'Keefe RA, Grandis JR. Targeting the IL-6/JAK/STAT3 signalling axis in cancer. *Nat Rev Clin Oncol* 2018;15(4):234–248.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

29. Quintas-Cardama A, Verstovsek S. Molecular pathways: Jak/STAT pathway: mutations, inhibitors, and resistance. *Clin Cancer Res* 2013;19(8):1933–40.

30. Tawara K et al. HIGH expression of OSM and IL-6 are associated with decreased breast cancer survival: synergistic induction of IL-6 secretion by OSM and IL-1beta. *Oncotarget* 2019;10(21):2068–2085.

31. West NR, Murphy LC, Watson PH. Oncostatin M suppresses oestrogen receptor-α expression and is associated with poor outcome in human breast cancer. *Endocr. Relat. Cancer* 2012;19(2):181–195.

32. Guo L et al. Stat3-coordinated Lin-28-let-7-HMGA2 and miR-200-ZEB1 circuits initiate and maintain oncostatin M-driven epithelial-mesenchymal transition. *Oncogene* 2013;32(45):5272–5282.

33. Araujo AM et al. Stromal oncostatin M cytokine promotes breast cancer progression by reprogramming the tumor microenvironment. *J Clin Invest* 2022;132(7). doi:10.1172/JCI148667

34. Bryson BL, Junk DJ, Cipriano R, Jackson MW. STAT3-mediated SMAD3 activation underlies Oncostatin M-induced Senescence. *Cell Cycle* 2017;16(4):319–334.

35. Taniguchi K, Karin M. IL-6 and related cytokines as the critical lynchpins between inflammation and cancer. *Semin. Immunol.* 2014;26(1):54–74.

36. West NR, Murray JI, Watson PH. Oncostatin-M promotes phenotypic changes associated with mesenchymal and stem cell-like differentiation in breast cancer. *Oncogene* 2014;33(12):1485–1494.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

37. Ross C et al. Metastasis-Specific Gene Expression in Autochthonous and Allograft Mouse Mammary Tumor Models: Stratification and Identification of Targetable Signatures. *Mol. Cancer Res. MCR* 2020;18(9):1278–1289.

38. Fico F, Bousquenaud M, Rüegg C, Santamaria-Martínez A. Breast Cancer Stem Cells with Tumor- versus Metastasis-Initiating Capacities Are Modulated by TGFBR1 Inhibition. *Stem Cell Rep.* 2019;13(1):1–9.

39. Hanahan D. Hallmarks of Cancer: New Dimensions. *Cancer Discov.* 2022;12(1):31–46.

40. Ouzounova M et al. Monocytic and granulocytic myeloid derived suppressor cells differentially regulate spatiotemporal tumour plasticity during metastatic cascade. *Nat. Commun.* 2017;8(1):14979.

41. Veglia F, Sanseviero E, Gabrilovich DI. Myeloid-derived suppressor cells in the era of increasing myeloid cell diversity. *Nat. Rev. Immunol.* 2021;21(8):485–498.

42. Bronte V et al. Recommendations for myeloid-derived suppressor cell nomenclature and characterization standards. *Nat. Commun.* 2016;7(1):12150.

43. Sebastian A et al. Single-Cell Transcriptomic Analysis of Tumor-Derived Fibroblasts and Normal Tissue-Resident Fibroblasts Reveals Fibroblast Heterogeneity in Breast Cancer. *Cancers* 2020;12(5):1307.

44. Efremova M, Vento-Tormo M, Teichmann SA, Vento-Tormo R. CellPhoneDB: inferring cell–cell communication from combined expression of multi-subunit ligand–receptor complexes. *Nat. Protoc.* 2020;15(4):1484–1506.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

45. Brandizzi F, Barlowe C. Organization of the ER–Golgi interface for membrane traffic control. *Nat. Rev. Mol. Cell Biol.* 2013;14(6):382–392.

46. Chen K et al. Phenotypically supervised single-cell sequencing parses within-cell-type heterogeneity. *iScience* 2021;24(1):101991.

47. Yeo SK et al. Single-cell RNA-sequencing reveals distinct patterns of cell state heterogeneity in mouse models of breast cancer. *eLife* 2020;9:e58810.

48. Aouad P et al. Epithelial-mesenchymal plasticity determines estrogen receptor positive breast cancer dormancy and epithelial reconversion drives recurrence. *Nat. Commun.* 2022;13(1):4975.

49. Jin W. Role of JAK/STAT3 Signaling in the Regulation of Metastasis, the Transition of Cancer Stem Cells, and Chemoresistance of Cancer by Epithelial–Mesenchymal Transition. *Cells* 2020;9(1):217.

50. Junk DJ et al. Oncostatin M promotes cancer cell plasticity through cooperative STAT3-SMAD3 signaling. *Oncogene* 2017;36(28):4001–4013.

51. Naz F, Shi M, Sajid S, Yang Z, Yu C. Cancer stem cells: a major culprit of intra-tumor heterogeneity. *Am. J. Cancer Res.* 2021;11(12):5782–5811.

52. Zhang R, Tu J, Liu S. Novel molecular regulators of breast cancer stem cell plasticity and heterogeneity. *Semin Cancer Biol* 2022;82:11–25.

53. Qin S et al. Emerging role of tumor cell plasticity in modifying therapeutic response. *Signal Transduct. Target. Ther.* 2020;5(1):228.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

54. Kong D, Hughes CJ, Ford HL. Cellular Plasticity in Breast Cancer Progression and Therapy. *Front. Mol. Biosci.* 2020;7:72.

55. Muzio G, Maggiora M, Paiuzzi E, Oraldi M, Canuto RA. Aldehyde dehydrogenases and cell proliferation. *Free Radic. Biol. Med.* 2012;52(4):735–746.

56. Curtis C et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* 2012;486(7403):346–352.

57. Shen M et al. A review of current progress in triple-negative breast cancer therapy. *Open Med Wars* 2020;15(1):1143–1149.

58. Al-Mahmood S, Sapiezynski J, Garbuzenko OB, Minko T. Metastatic and triple-negative breast cancer: challenges and treatment options. *Drug Deliv Transl Res* 2018;8(5):1483–1507.

59. Azim HA, Ghosn M, Oualla K, Kassem L. Personalized treatment in metastatic triple-negative breast cancer: The outlook in 2020. *Breast J* 2020;26(1):69–80.

60. Vagia E, Mahalingam D, Cristofanilli M. The Landscape of Targeted Therapies in TNBC. *Cancers Basel* 2020;12(4). doi:10.3390/cancers12040916

61. Hua Z, White J, Zhou J. Cancer stem cells in TNBC. *Semin Cancer Biol* 2022;82:26–34.

62. Batlle E, Clevers H. Cancer stem cells revisited. *Nat Med* 2017;23(10):1124–1134.

63. Brooks MD, Burness ML, Wicha MS. Therapeutic Implications of Cellular Heterogeneity and Plasticity in Breast Cancer. *Cell Stem Cell* 2015;17(3):260–71.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

64. Yeo SK, Wen J, Chen S, Guan JL. Autophagy Differentially Regulates Distinct Breast Cancer Stem-like Cells in Murine Models via EGFR/Stat3 and Tgfbeta/Smad Signaling. *Cancer Res* 2016;76(11):3397–410.

65. Gong J et al. Genotoxic stress induces Sca-1-expressing metastatic mammary cancer cells. *Mol. Oncol.* 2018;12(8):1249–1263.

66. Shimokawa M et al. Visualization and targeting of LGR5+ human colon cancer stem cells. *Nature* 2017;545(7653):187–192.

67. Oo MW et al. Resident stroma-secreted chemokine CCL2 governs myeloid-derived suppressor cells in the tumor microenvironment. *JCI Insight* 2022;7(1). doi:10.1172/jci.insight.148960

68. Franklin DA et al. MEK activation modulates glycolysis and supports suppressive myeloid cells in TNBC. *JCI Insight* 2020;5(15). doi:10.1172/jci.insight.134290

69. Bullock K, Richmond A. Suppressing MDSC Recruitment to the Tumor Microenvironment by Antagonizing CXCR2 to Enhance the Efficacy of Immunotherapy. *Cancers Basel* 2021;13(24). doi:10.3390/cancers13246293

70. Jovanovic IP et al. Interleukin-33/ST2 axis promotes breast cancer growth and metastases by facilitating intratumoral accumulation of immunosuppressive and innate lymphoid cells. *Int J Cancer* 2014;134(7):1669–82.

71. Szczerba BM et al. Neutrophils escort circulating tumour cells to enable cell cycle progression. *Nature* 2019;566(7745):553–557.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

72. Doherty MR et al. The opposing effects of interferon-beta and oncostatin-M as regulators of cancer stem cell plasticity in triple-negative breast cancer. *Breast Cancer Res* 2019;21(1):54.

73. Cavalli LR, Riggins RB, Wang A, Clarke R, Haddad BR. Frequent loss of heterozygosity at the interferon regulatory factor-1 gene locus in breast cancer. *Breast Cancer Res. Treat.* 2010;121(1):227–231.

74. Weichselbaum RR et al. An interferon-related gene signature for DNA damage resistance is a predictive marker for chemotherapy and radiation for breast cancer. *Proc. Natl. Acad. Sci. U. S. A.* 2008;105(47):18490–18495.

75. Aljohani AI et al. Myxovirus resistance 1 (MX1) is an independent predictor of poor outcome in invasive breast cancer. *Breast Cancer Res. Treat.* 2020;181(3):541–551.

76. Chan JM et al. Lineage plasticity in prostate cancer depends on JAK/STAT inflammatory signaling. *Science* 2022;377(6611):1180–1191.

77. Deng S et al. Ectopic JAK–STAT activation enables the transition to a stem-like and multilineage state conferring AR-targeted therapy resistance. *Nat. Cancer* 2022;3(9):1071–1087.

78. Lynce F et al. Phase I study of JAK1/2 inhibitor ruxolitinib with weekly paclitaxel for the treatment of HER2-negative metastatic breast cancer. *Cancer Chemother Pharmacol* 2021;87(5):673–679.

79. Kearney M et al. Phase I/II trial of ruxolitinib in combination with trastuzumab in metastatic HER2 positive breast cancer. *Breast Cancer Res Treat* 2021;189(1):177–185.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

80. O'Shaughnessy J et al. A randomized, double-blind, phase 2 study of ruxolitinib or placebo in combination with capecitabine in patients with advanced HER2-negative breast cancer and elevated C-reactive protein, a marker of systemic inflammation. *Breast Cancer Res Treat* 2018;170(3):547–557.

81. Lan Q et al. Type I interferon/IRF7 axis instigates chemotherapy-induced immunological dormancy in breast cancer. *Oncogene* 2019;38(15):2814–2829.

82. FASTQC. A quality control tool for high throughput sequence data | BibSonomy [Internet]https://www.bibsonomy.org/bibtex/f230a919c34360709aa298734d63dca3. cited October 15, 2022

83. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* 2016;32(19):3047–3048.

84. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;30(15):2114–2120.

85. Kopylova E, Noé L, Touzet H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinforma. Oxf. Engl.* 2012;28(24):3211–3217.

86. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 2015;31(2):166–169.

87. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* 2017;14(4):417–419.

88. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):550.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

89. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinformatics* 2013;14(1):7.

90. Morgan M, Falcon S, Gentleman R. GSEABase: Gene set enrichment data structures and methods. R package version 1.58.0.2022;

91. Subramanian A et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci.* 2005;102(43):15545–15550.

92. Ritchie ME et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015;43(7):e47.

93. Kolde R. Pheatmap: pretty heatmaps. *R Package Version* 2012;1(2):726.

94. Dusa, Adrian. Draw Venn Diagrams R package venn2016;

95. Wickham H et al. Welcome to the Tidyverse. *J. Open Source Softw.* 2019;4(43):1686.

96. Wickham H. *ggplot2 [Internet]*. New York, NY: Springer; 2009:

97. Gu Z, Gu L, Eils R, Schlesner M, Brors B. circlize Implements and enhances circular visualization in R. *Bioinforma. Oxf. Engl.* 2014;30(19):2811–2812.

98. Durinck S et al. BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinforma. Oxf. Engl.* 2005;21(16):3439–3440.

99. Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.* 2009;4(8):1184–1191.

100. Neuwirth E. RColorBrewer: ColorBrewer Palettes:5.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

101. Wu T et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innov. Camb. Mass* 2021;2(3):100141.

102. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics J. Integr. Biol.* 2012;16(5):284–287.

103. Yu G, Hu E. enrichplot: Visualization of Functional Enrichment Result [Internet]2022; doi:10.18129/B9.bioc.enrichplot

104. Xu S et al. Use ggbreak to Effectively Utilize Plotting Space to Deal With Large Datasets and Outliers. *Front. Genet.* 2021;12:774846.

105. Monnier Y et al. CYR61 and alphaVbeta5 integrin cooperate to promote invasion and metastasis of tumors growing in preirradiated stroma. *Cancer Res.* 2008;68(18):7323–7331.

106. Kassambara A. ggpubr: “ggplot2” Based Publication Ready Plots [Internet]2020;https://CRAN.R-project.org/package=ggpubr. cited October 15, 2022

107. Stuart T et al. Comprehensive Integration of Single-Cell Data. *Cell* 2019;177(7):1888-1902.e21.

108. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinforma. Oxf. Engl.* 2016;32(18):2847–2849.

109. Shao X et al. CellTalkDB: a manually curated database of ligand–receptor interactions in humans and mice. *Brief. Bioinform.* 2021;22(4):bbaa269.

110. Cabello-Aguilar S et al. SingleCellSignalR: inference of intercellular networks from single-cell transcriptomics. *Nucleic Acids Res.* 2020;48(10):e55.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted November 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

111. Trapnell C et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* 2014;32(4):381–386.

112. Qiu X et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods* 2017;14(10):979–982.

113. Cao J et al. The single-cell transcriptional landscape of mammalian organogenesis. *Nature* 2019;566(7745):496–502.

114. McInnes L, Healy J, Melville J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction [Internet]2020;http://arxiv.org/abs/1802.03426. cited October 21, 2022

115. Korotkevich G et al. Fast gene set enrichment analysis [Internet]2021;060012.

116. Gao W et al. IL20RA signaling enhances stemness and promotes the formation of an immunosuppressive microenvironment in breast cancer. *Theranostics* 2021;11(6):2564–2580.

117. Cerami E et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* 2012;2(5):401–404.

118. Ramos M et al. Multiomic Integration of Public Oncology Databases in Bioconductor. *JCO Clin. Cancer Inform.* 2020;(4):958–971.

119. Therneau TM, Grambsch PM. *Modeling Survival Data: Extending the Cox Model [Internet]*. New York, NY: Springer; 2000:

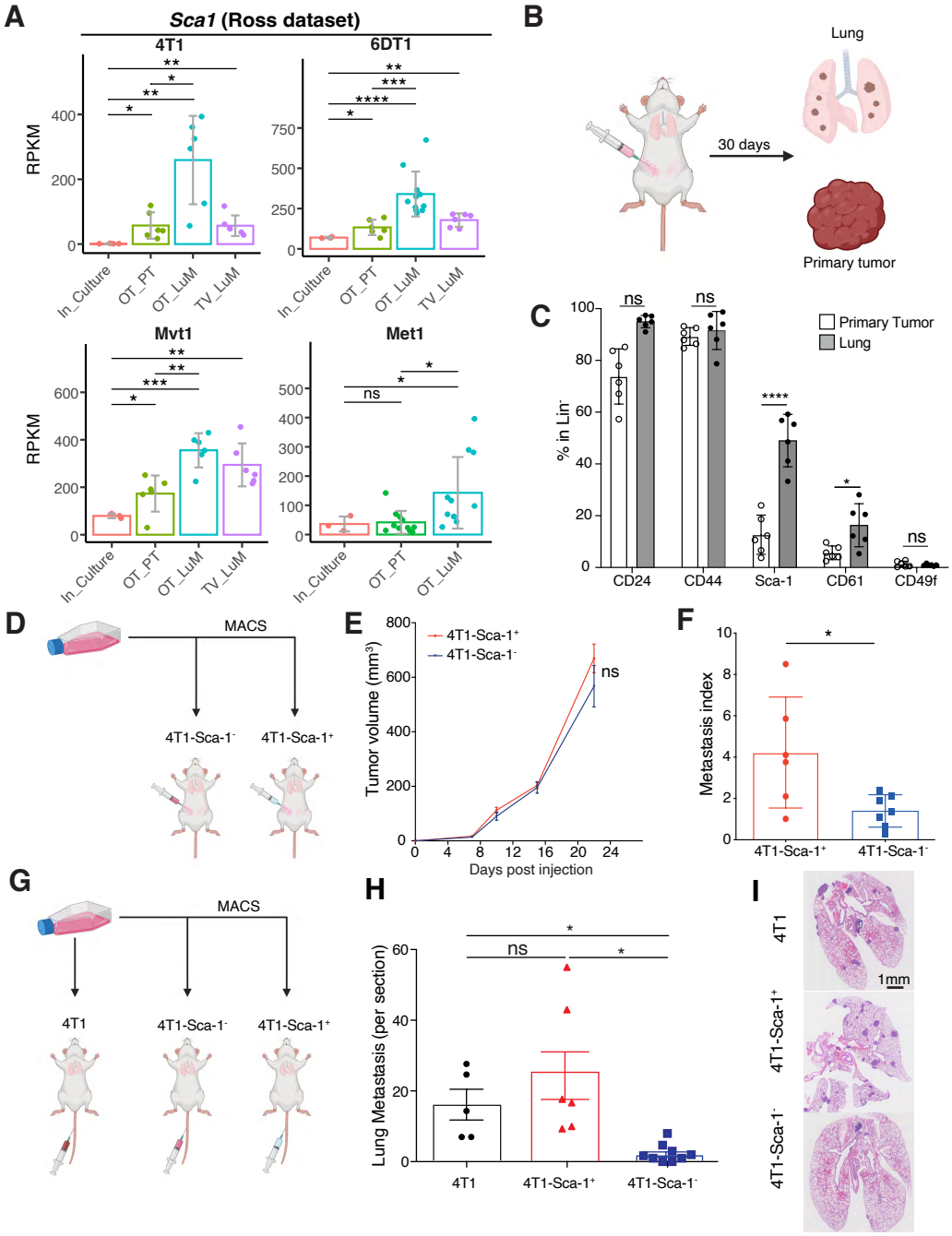

Figure 1. Sca-1+ population is enriched during in vivo metastasis across multiple breast cancer models

(**A**) *Sca-1* mRNA expression in the metastatic murine breast cancer models 4T1, 6DT1, Mvt1 and Met1, extracted from the Ross dataset. Analyzed samples consist of cultured cells (In_Culture), orthotopic injected primary tumors (OP_PT), spontaneous lung metastases (OP_LuM) and lung metastases induced by tail-vain injection (TV_LuM). Data are presented as mean of Reads Per Kilobase of transcript per Million mapped reads (RPKM) ± SD (unpaired two-tailed student's t test with Holm correction).

(**B**) Experimental set up for *in vivo* experimental validation. The 4T1 tumor cells were orthotopically injected into the 4th mammary fat pad. Thirty days later, cells from primary tumors and lungs were isolated to examine CSC marker expression by flow cytometry.

(**C**) Frequency of CSC marker expression in primary tumors and lung metastases. Results give the percentage of CD24, CD44, Sca-1, CD61 and CD49f positive cells gated in lineage negative cells ($CD45^-CD31^-TER119^-$).

(**D-F**) Experimental set up (**D**) of the in vivo experiment to assess tumor growth (e) and lung metastatic ability (metastatic index) (**F**) of 4T1-Sca-1$^+$ and 4T1-Sca-1$^-$ populations isolated from parental 4T1 cells orthotopically injected into the 4th mammary fat pad. Metastases are assessed 21 days after tumor cell injection. n=8/group, 3 independent experiments.

(**G-I**) Experimental set up (**G**) of the in vivo experiment to assess lung colonization capacity of sorted parental 4T1, 4T1-Sca-1$^+$ and 4T1-Sca-1$^-$ cells by tail-vain injection. Lung metastatic nodule numbers (**H**) and representative images (**I**) of lungs from mice 10 days post injection (n=5-6, 2 independent experiments). Scale bar=1 mm.

Except **A**, data are represented as mean values ± SEM. P values were calculated using unpaired two-tailed student's t test with Holm correction (**A**), unpaired two-tailed student's t test (**C, F**), two-way ANOVA with Tukey multiple-comparison test (**E**) or one-way ANOVA with Tukey multiple-comparison test **(H)**. *, p<0.05; **, p< 0.01; ***, p< 0.001; ****, p< 0.0001; ns, non-significant.

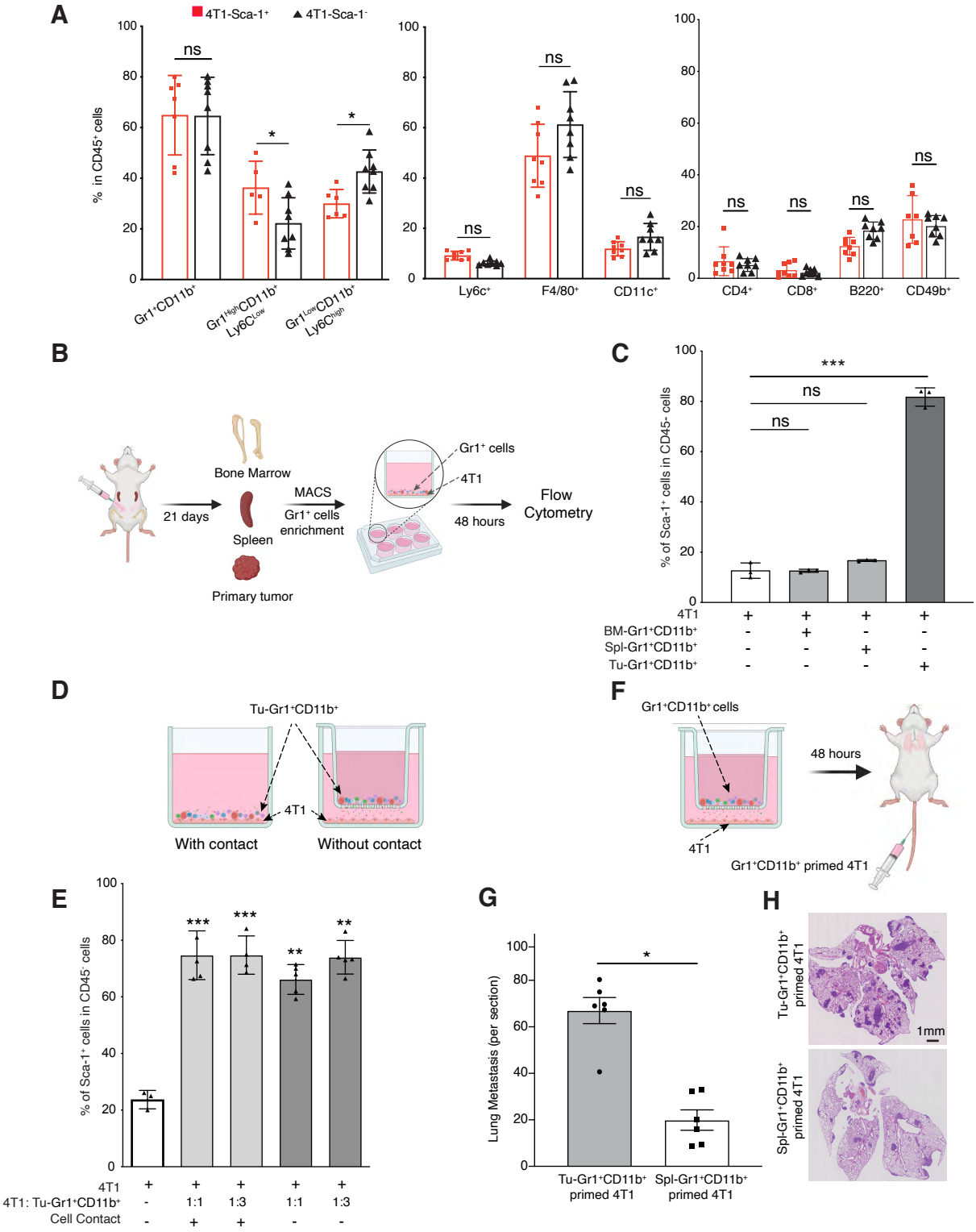

Figure 2. Sca-1 expression is modulated by the tumor microenvironment

Figure 2. Sca-1 expression is modulated by the tumor microenvironment

(A) Frequency of different immune cell populations in primary tumors of mice orthotopically injected with 4T1-Sca-1$^+$ and 4T1-Sca-1$^-$ 21 days post injection. Populations are determined in CD45 negative, viable cells (n=8 mice/Group).

(B-C) Illustrative scheme **(B)** showing the experimental design for isolating Gr1$^+$ cells from different sites of tumor-bearing mice. Twenty-one days after tumor implantation, Gr1$^+$ cells were isolated from bone marrow (BM-Gr1+CD11b+), spleen (Spl- Gr1+CD11b+) or primary tumor (Tu-Gr1$^+$CD11b$^+$) and co-cultured for 48 hours with parental 4T1 cells in vitro. Sca-1 expression in tumor cells was examined by flow cytometry **(C)**. Co-cultures conditions are indicated in the bar graph (n=3/group).

(D) Illustrative scheme of the experimental co-culture setup.

(E) MACS-enriched Gr1$^+$ cells were cocultured with 4T1 cells with or without Transwell inserts of 0.4 µm pore size. The 4T1 cells were seeded in the bottom well and Gr1$^+$CD11b$^+$ cells in the upper part of the insert. After 48 hours, 4T1 cells were examined for Sca1 expression by FACS. Co-cultures conditions are indicated in the bar graph. The ratio of tumor cells and Tu-Gr1$^+$CD11b$^+$ is varied from 1:1 to 1:3. Significant enrichment of Sca-1$^+$ population were observed in all conditions. n= 3-5/group.

(F) Illustrative scheme of the experimental metastasis setup.

(G-H) Evaluation of the metastatic capacity of Gr1$^+$CD11b$^+$-educated 4T1 cells in vivo. The 4T1 tumor cells were primed with Tu-Gr1$^+$CD11b$^+$ or Spl-Gr1$^+$CD11b$^+$ in vitro without cell-cell contact for 48 hours and injected into the tail vein of mice. Lung metastases were quantified 10 days after injection **(G)**, and representative H&E staining images of lung sections are showed **(H)**. Scale bar=1mm, n=7 mice/group.

Data are represented as mean values ± SEM from at least 3 independent experiments. P values were calculated using unpaired two-tailed student's t test (**A, G**), one-way ANOVA with Dunnett's multiple-comparison test (**C, E**). *, $p<0.05$; **, $p< 0.01$; ***, $p< 0.001$; ns, non-significant.

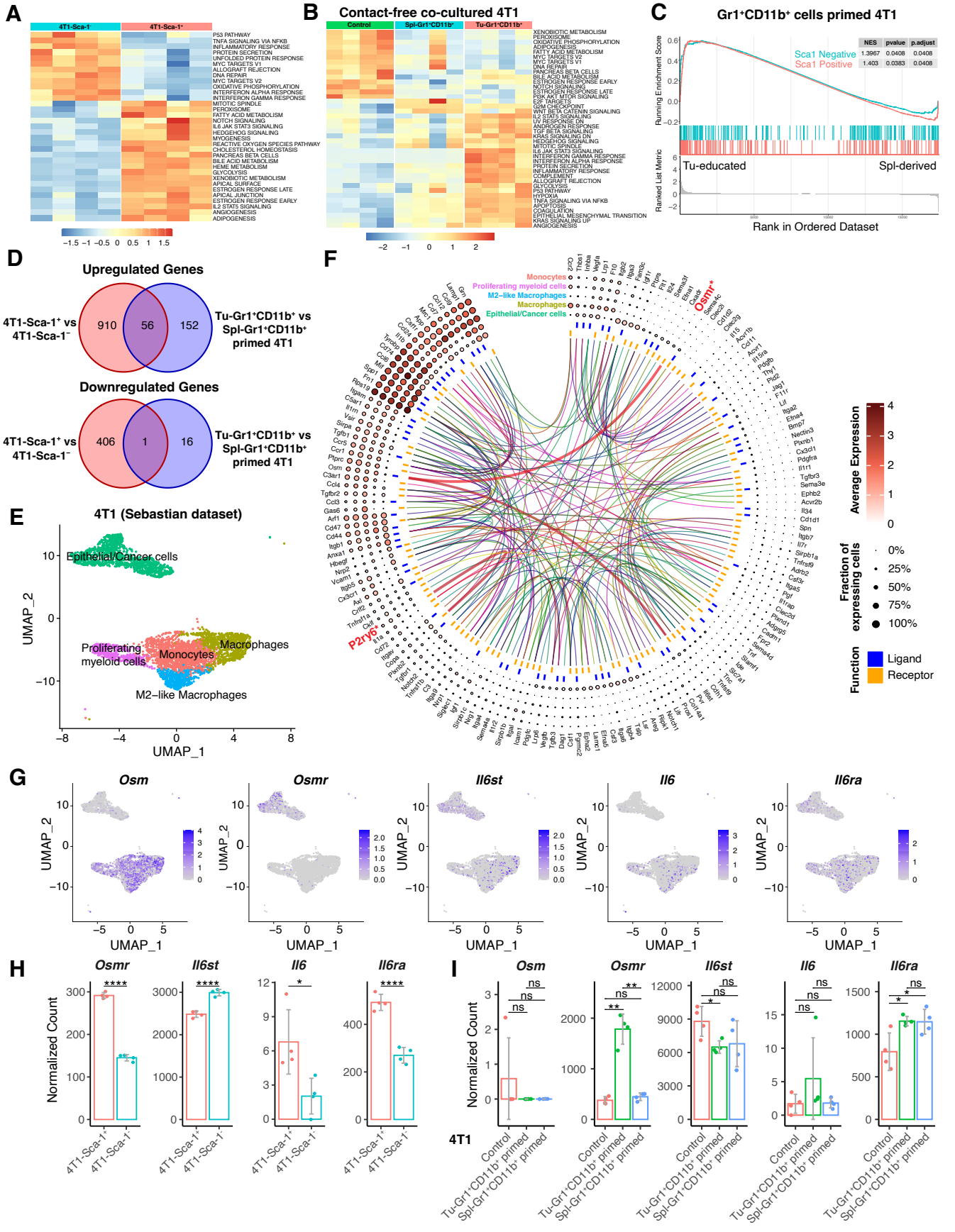

Figure 3. Transcriptomic analysis of Sca-1⁺ tumor cells

Figure 3. Transcriptomic analysis of Sca-1$^+$ tumor cells

(A) Heatmap showing the signature score of the hallmark pathways analysis in 4T1-Sca-1$^+$ and 4T1-Sca-1$^-$ population sorted from parental 4T1 cells. The colors code the expression levels relative to average levels as indicate at the bottom.

(B) Heatmap showing the signature score of the hallmarks pathway analysis in parental 4T1 (4T1), Spl-Gr1$^+$CD11b$^+$ primed 4T1 and Tu-Gr1$^+$CD11b$^+$ primed 4T1 cells. The colors code the expression levels relative to average levels as indicate at the bottom.

(C) Gene set enrichment analysis (GSEA) comparing the Tu-Gr1$^+$CD11b$^+$ and Spl-Gr1$^+$CD11b$^+$ primed 4T1 cells. GSEA shows positive correlations of both Sca1 Positive and Sca1 Negative signatures. NES, normalized enrichment score.

(D) Venn diagrams showing that 56 upregulated genes, and 1 downregulated gene are shared between endogenous and Tu-Gr1$^+$CD11b$^+$ induced Sca-1$^+$ population in 4T1 tumor cells.

(E) UMAP plot showing clusters of cancer cells and myeloid cell populations in orthotopically growing 4T1-derived primary tumors extracted from the Sebastian dataset (see Materials and Methods for details).

(F) Circos diagram showing the potential interactions between cancer cells and different myeloid cell populations determined by CellPhoneDB (see Materials and Methods for details) based on the Sebastian dataset. Only OSMR and P2RY6 are shared with the common 56 gene list showed in panel **D**.

(G) UMAP plot showing the gene expression pattern of *Osm*, *Osmr*, *Il6st*, *Il6* and *Il6ra* in different cell populations in the Sebastian dataset.

(H-I) mRNA expression of *Osm*, *Osmr*, *Il6st*, *Il6* and *Il6ra* based on RNAseq data used to generate the heatmaps shown in **A** & **B**, respectively. *Osm* expression is not detected in sorted 4T1 cells.

Data are presented as mean of normalized count ± SD. P values: *, $p<0.05$; **, $p< 0.01$; ***, $p< 0.001$; ****, $p< 0.0001$; ns, non-significant (unpaired two-tailed student's t test for **H**, and unpaired two-tailed student's t test with Holm correction for **I**).

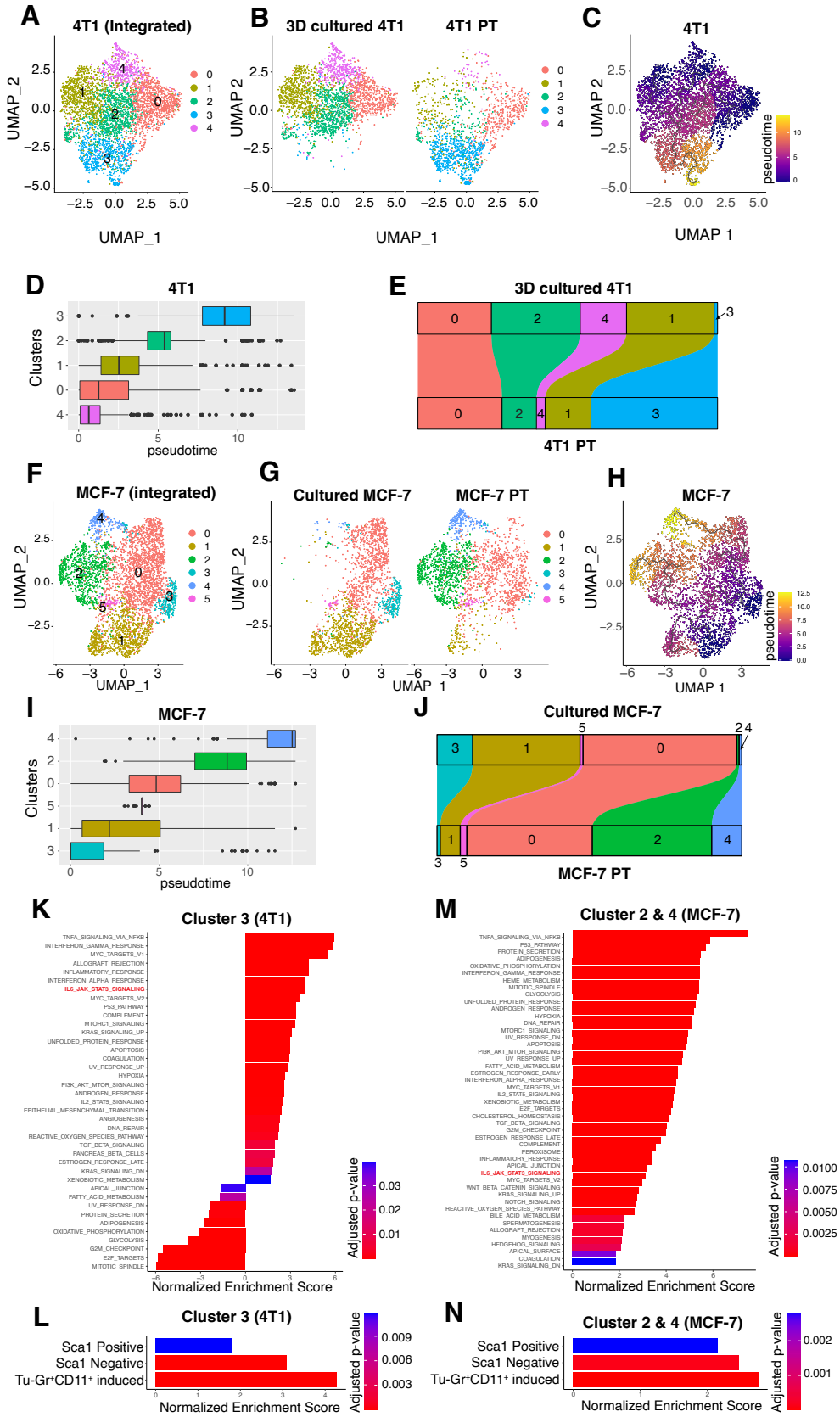

Figure 4. Transformation dynamics of tumor cell populations induced by the tumor microenvironment

Figure 4. Transformation dynamics of tumor cell populations induced by the tumor microenvironment
(A) UMAP plots showing 4T1 clusters based on integrated scRNA-seq data from 4T1 cells in 3D culture or in primary tumor.
(B) Distribution of specific clusters in 4T1 cells in 3D culture or in primary tumor (PT).
(C-D) UMAP plot **(C)** and boxplot **(D)** showing the clusters in pseudo-time during the transformation of 4T1 cells from *ex vivo* culture to *in vivo*.
(E) Sankey diagram showing the dynamic of each cluster during the transformation of 4T1 cells from *ex vivo* culture to in vivo. Cluster 3 was largely expanded *in vivo*.
(F) UMAP plots showing MCF-7 clusters based on integrated scRNA-seq data from MCF-7 cells in culture or in primary tumor.
(G) Distribution of the specific clusters in cultured MCF7 cells or in MCF7 primary tumors (PT).
(H-I) UMAP plot **(H)** and boxplot **(I)** showing the clusters in pseudo-time during the transformation of MCF-7 cells from *ex vivo* culture to *in vivo.*
(J) Sankey diagram showing the dynamic of each cluster during the transformation of MCF-7 cells from *ex vivo* culture to *in vivo*. Cluster 2 and 4 were largely expanded *in vivo*.
(K-L) GESA analysis of Hallmark gene sets **(K)** and Sca1 Positive signature, Sca1 Negative Signature and Tu-Gr1^{+}CD11b^{+} induced signature **(L)** of cluster 3 in 4T1 data. Only the signatures with adjusted p-value <0.05 were shown.
(M-N) GESA analysis of Hallmark gene sets (M) and Sca1 Positive signature, Sca1 Negative Signature and Tu-Gr1^{+}CD11b^{+} induced signature **(N)** of cells in cluster 2 or cluster 4 in MCF-7 data. Only the signatures with adjusted p-value <0.05 were shown.
Analyses are based on publicly available data (4T1: GSM4812003 and GSM3502134; MCF-7: GSM4681765 and GSM5904917).

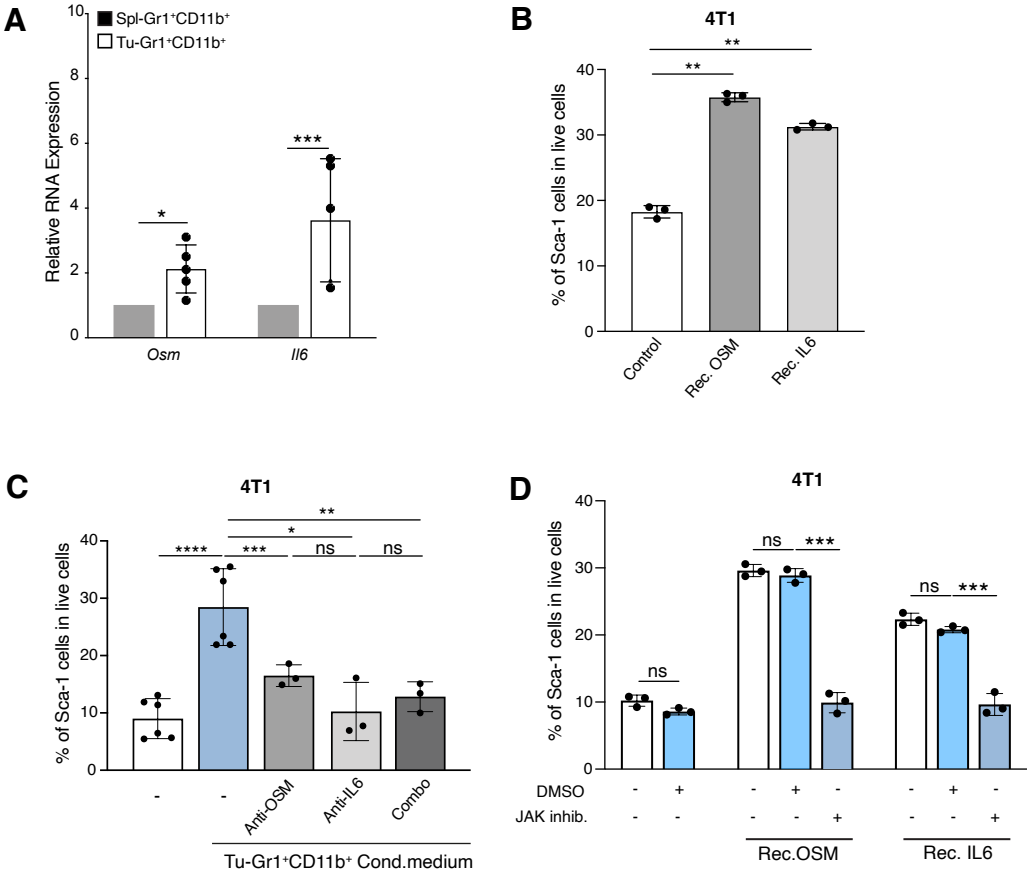

Figure 5. Sca-1⁺ population is modulated by the OSM/IL6-JAK pathway
(**A**) Quantitative PCR analysis of *Osm* and *Il6* mRNA expression in Tu-Gr1⁺CD11b⁺ and Spl- Gr1⁺CD11b⁺. Tu-Gr1⁺CD11b⁺ express significantly higher *Osm* and *Il6* levels compared with Spl-Gr1⁺CD11b⁺. n = 4-5/group.
(**B**) Fraction of 4T1-Sca1⁺ cells upon exposure to recombinant Il6 or Osm protein (1 µg/ml for 48 hours) as determined by flow cytometry. Both cytokines induced the Sca-1⁺ population in cultured 4T1 tumor cells. n = 3/group.
(**C**) Inhibition of OSM and IL6 from Tu-Gr1⁺CD11b⁺ conditioned medium with anti-OSM or anti-IL6 neutralizing antibody as indicated. Treatment with either antibody significantly suppressed the Sca-1⁺ population enrichment. n = 3-6/group.
(**D**) Treatment with the JAK inhibitor Ruxolitinib (5 µM) of cultured 4T1 cells stimulated with recombinant IL6 or OSM protein (1 µg/ml, 48 hours exposure) as indicated. Ruxolitinib inhibited Sca-1⁺ population enrichment induced by recombinant IL6 or OSM protein. n = 3/group. Data are represented as mean ± SEM from 3 independent experiments. *, p<0.05; **, p< 0.01; ***, p< 0.001; ****, p< 0.0001; ns, non-significant (unpaired two-tailed student's t test for **A**, and one-way ANOVA with Dunnett's multiple-comparison test for **B** and **D**, and one-way ANOVA with Tukey multiple-comparison test for **C**).

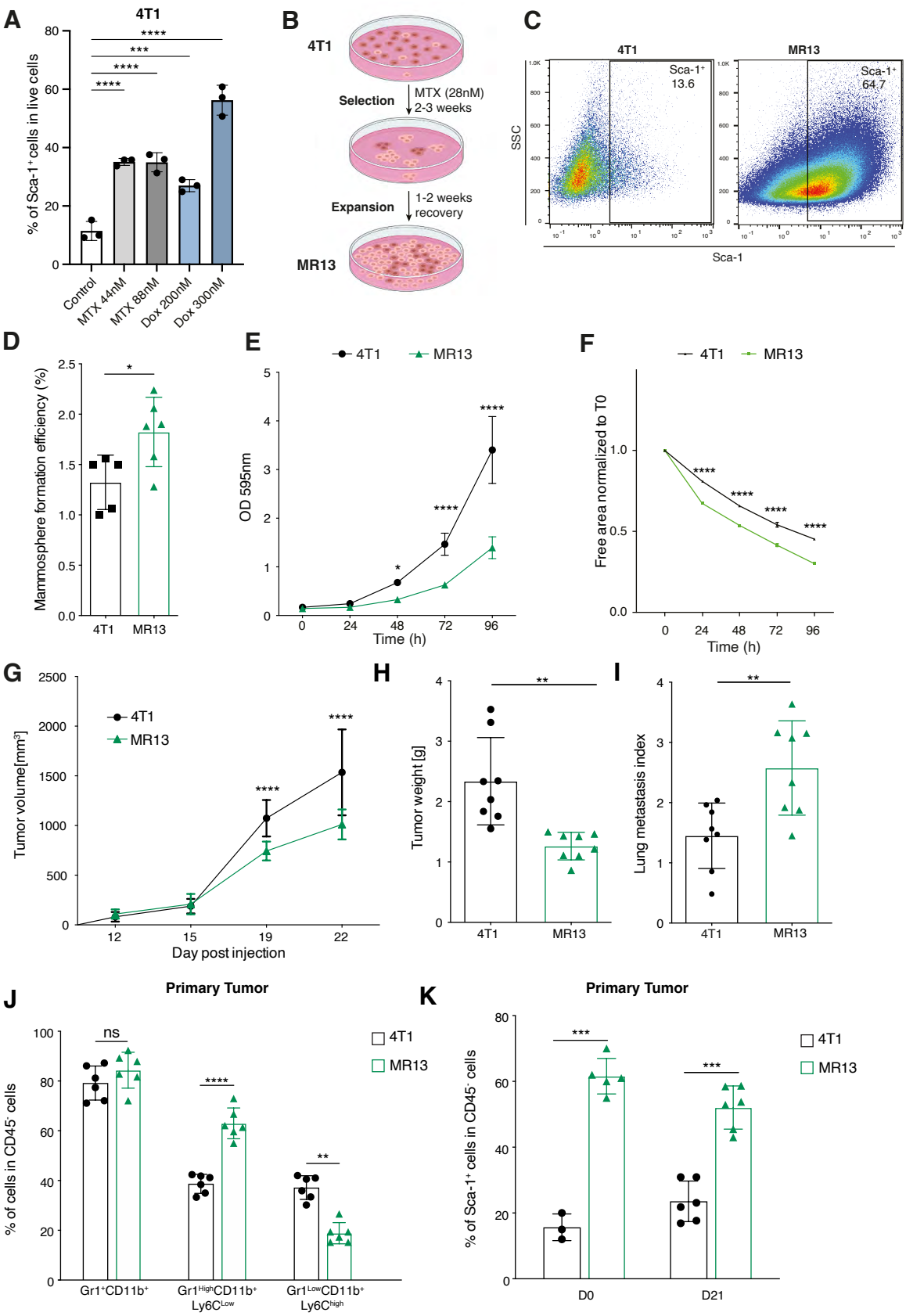

Figure 6. Long term chemotherapy treatment of 4T1 cells induces a stable Sca-1⁺ population (MR13) with higher metastatic capacity and CSC features

Figure 6. Long term chemotherapy treatment of 4T1 cells induces a stable Sca-1$^+$ population (MR13) with higher metastatic capacity and CSC features

(A) Fraction of 4T1-Sca1$^+$ cells upon short-term (48 hours) Methotrexate (44 and 88 nM) and Doxorubicin (200 and 300 nM) treatments. All treatments induced enrichment of Sca-1$^+$ population. n = 3/group.

(B) Illustrative scheme of the experimental design to obtain chemotherapy resistant MR13 cells from 4T1.

(C) Dot plots representing Sca-1 expression vs SSC determined by flow cytometry in MR13 chemotherapy resistant cells vs parental 4T1 tumor cells in vitro.

(D) Quantification of the mammosphere forming efficiency of 4T1 and MR13 tumor cells. n = 5-6/group.

(E) Cell proliferation curve of 4T1 and MR13 tumor cells in vitro determined by crystal violet assay. The results are presented as mean of optical density (OD). n = 8/group.

(F) Cell motility of 4T1 and MR13 tumor cells determined by a scratch wound healing assay. n=5-6/group. Results are presented as cell-free area relative to the initial wound area from 3 independent experiments.

(G) Growth curves of primary tumors in BALB/c mice orthotopically injected with 4T1 and MR13 tumor cells (n=10-11/group).

(H) Tumor weight of 4T1 and MR13 tumors recovered from BALB/c mice at day 22 post injection (n=8-9/group).

(I) Lung metastasis index 23 days post injection. The number of metastatic nodules is determined by H&E staining and normalized based on the primary tumor weight (n=8-9/group).

(J) Frequency of different CD11b$^+$ myeloid cells subpopulations in primary tumors from MR13 and 4T1 injected mice determined by flow cytometry 21-days post injection (n= 6). Subpopulations are determined in CD45 positive, viable cell population.

(K) Percentage of Sca-1$^+$ tumor cells at time of injections (D0) of 4T1 and MR13 cells and in primary tumors recovered at day 21 (D21). Sca-1 expression is determined in CD45 negative, viable cell population. n = 3-6/group.

Data are represented as mean ± SEM from at least 3 independent experiments. *, $p<0.05$; **, $p< 0.01$; ***, $p< 0.001$; ****, $p< 0.0001$; ns, non-significant (one-way ANOVA with Dunnett's multiple-comparison test for **A**, unpaired two-tailed student's t test for **D**, **H-K**, and two-way ANOVA with Tukey's multiple comparison test for **E-G**).

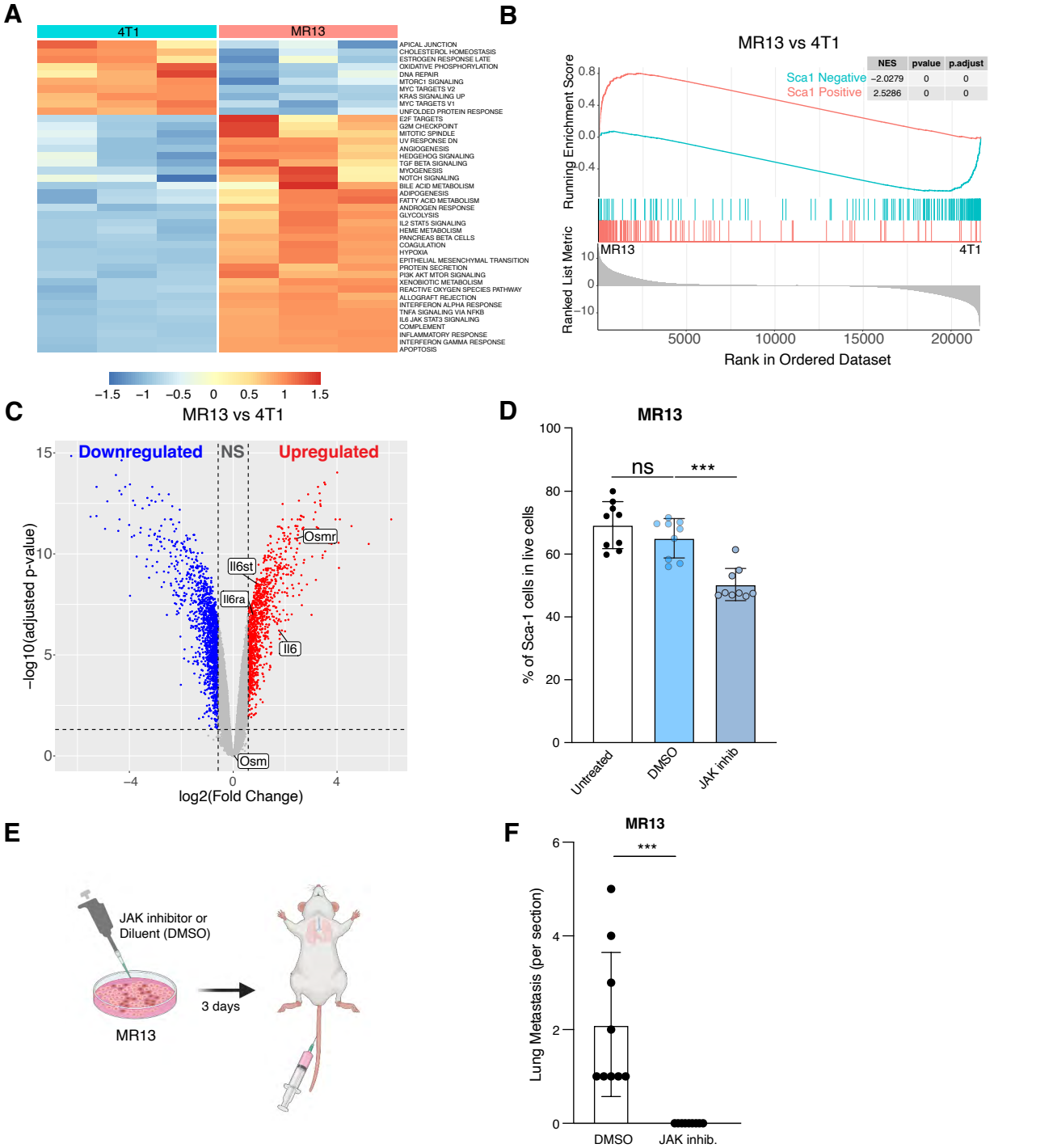

Figure 7. IL6-JAK pathway promotes Sca-1⁺ persistence and metastatic capacity in chemotherapy resistant MR13 tumor cells
(A-B) Gene expression analysis of parental 4T1 and chemotherapy resistant MR13 cells. Heat map represents the signature score of the hallmark pathways analysis. Results from 3 biological replicates are shown **(A)**. GSEA results showing that MR13 cells are positively enriched with for the Sca1 Positive signature and negatively with for the Sca1 Negative signature **(B)**.
(C) Volcano plot showing the differential expression of *Osm*, *Osmr*, *Il6st*, *Il6* and *Il6ra* mRNA in MR13 vs 4T1 tumor cells.
(D) Fraction of Sca-1⁺ population in MR13 tumor cells treated for 48 hours with Ruxolitinib (5 μM) relative to vehicle control (DMSO) treatment.
(E) Illustrative scheme of the experimental design for testing the effect of Ruxolitinib on MR13 metastatic capacity shown in F. MR13 tumor cells were treated with Ruxolitinib or DMSO in vitro for 72 hours and then injected into the mice tail vein. Lungs were examined for metastasis 10 days after tumor cell injection.
(F) Number of metastatic nodules in the lungs from mice injected with MR13 treated in vitro with Ruxolitinib or DMSO and indicated (n=9/group).
Data are represented as mean values ± SEM from 3 independent experiments. P values: ns, non-significant, ***< 0.001 (one-way ANOVA with Dunnett's multiple-comparison test for **D**, and unpaired two-tailed student's t test for **F**).

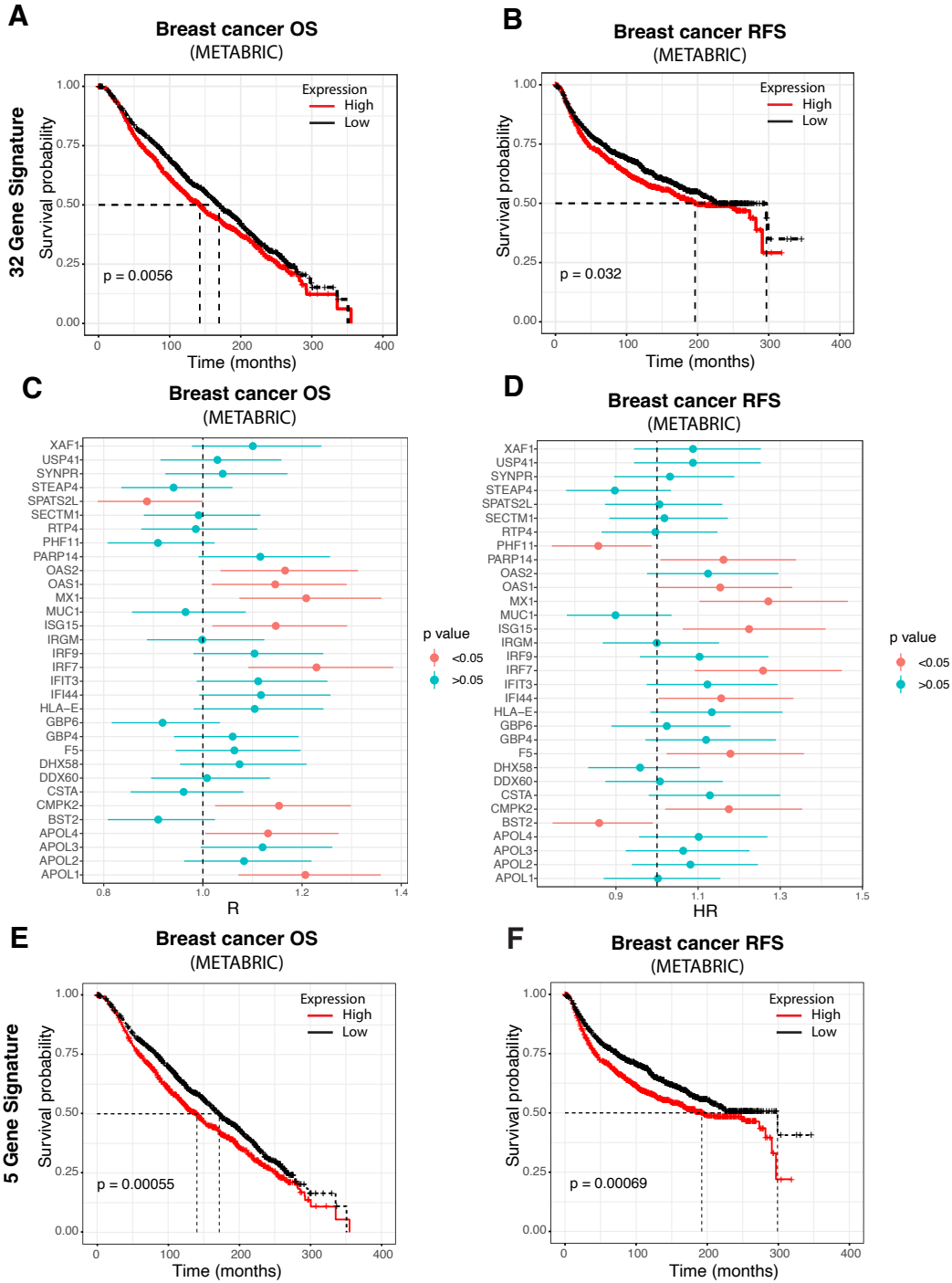

Figure 8. Tu-Gr1⁺CD11b⁺-induced tumor cell signature predicts worse outcome in breast cancer patients

(**A-B**) Kaplan-Meier curves showing overall survival (OS) (**A**) or relapse-free survival (RFS) (**B**) for breast cancer patients according to high or low expression of an orthologue 32 gene signature, based on the Tu-Gr1⁺CD11b⁺-induced 4T1 cell signature, in the METABRIC datasets.

(**C-D**) Forest plots showing the Cox proportional hazard regression (HR) for OS (**C**) and RFS (**D**) of the individual 32 orthologues of the Tu-Gr1⁺CD11b⁺-induced signature, based on gene expression in tumor samples from METABRIC dataset.

(**E-F**) Kaplan-Meier curves showing OS (**E**) and RFS (**F**) according to the reduced 5 orthologue gene signature expression in the METABRIC datasets.

The p value for **A**, **B**, **E**, **F** was calculated using the log-rank test and high and low expression levels were stratified by median values.

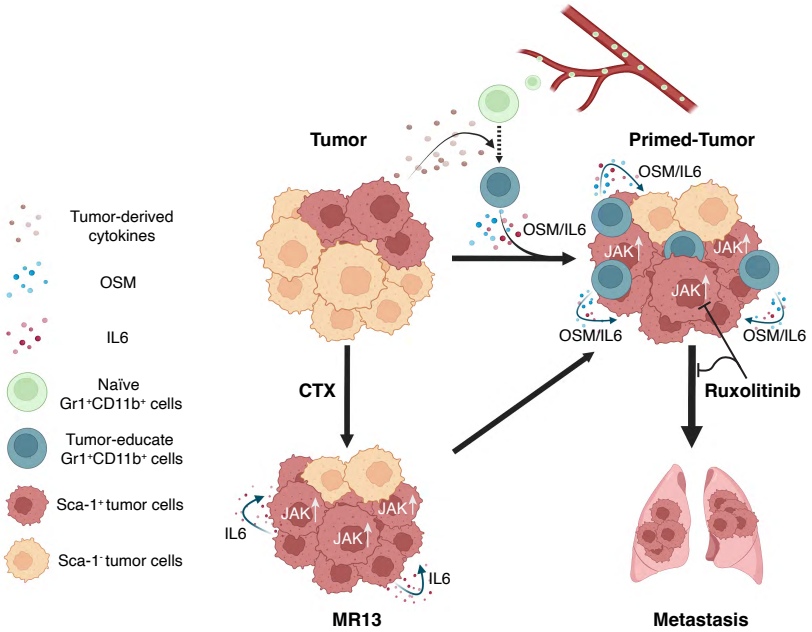

Figure 9 Illustrative scheme of the proposed model for cancer cell plasticity modulated by OSM/IL6 during tumor progression and chemotherapy

Parental tumor cells maintain a small portion of highly metastatic Sca-1$^+$ population. During tumor progression, naïve Gr1$^+$CD11b$^+$ are recruited to the TME and educated into Tu-Gr1$^+$CD11b$^+$ by tumor-derived factors. In turn, Tu-Gr1$^+$CD11b$^+$ secrete OSM and IL6 to convert Sca-1$^-$ population into a highly metastatic Sca-1$^+$ population. Chemotherapy (CTX) enriched for Sca-1$^+$ population due to its intrinsic resistance against cytotoxic treatment. Resistant cells express IL6 to maintain the high portion of Sca-1$^+$ population with high metastatic ability. JAK inhibitor Ruxolitinib suppresses the conversion to Sca-1$^+$ population and metastasis. Inhibition of OSM, IL6 and the activated downstream kinase JAK are candidate therapeutic targets to impinge on metastatic breast cancer progression during natural evolution and following therapy-resistance.