Lymphatic system is the mainstream for breast cancer dissemination and metastasis revealed by single-cell lineage tracing

Kai Miao[1,2,3*†], Aiping Zhang[1,2,4†], Xiaodan Yang[1,2], Yipeng Zhang[1,2], Anqi Lin[1,2], Lijian Wang[1,2], Xin Zhang[1,2], Heng Sun[1,2], Jun Xu[1,2], Jingyao Zhang[5], Yuzhao Feng[1,2], Fangyuan Shao[1,2], Sen Guo[1,2], Zhihui Weng[1,2,3], Peng Luo[1,2], Dong Wang[6], Shuai Gao[5], Xiao-Yang Zhao[6], Xiaoling Xu[1,2,3] and Chu-Xia Deng[1,2,3*]

Abstract

Cancer metastasis is the primary cause of cancer-related death, yet the forces that drive cancer cells through various steps and different routes to distinct target organs/tissues remain elusive. In this study, we applied a barcoding system based single-cell lineage tracing approach to study the metastasis rate and route of breast cancer cells and their interactions with the tumor microenvironment (TME) during metastasis. The results indicate that only a small fraction of cells, accounting for fewer than 3% of total barcodes, can intravasate from the primary site into the blood circulation, whereas more cells disseminate through the lymphatic system to different organs. Tumor cells derived from the same progenitor cell exhibit different gene expression patterns in different soils, and the cancer cell-TME communication paradigm varies significantly between primary and metastatic tumors. Furthermore, metastable cells require a prewired particular cytokine expression ability which may be specific for lymph metastasis route although the underlying mechanism requires further investigation. In summary, leveraging a single-cell lineage tracing system, we demonstrate that the crosstalk between tumor cells and the TME is the driving force controlling the preferential metastatic fate of cancer cells through the lymphatic system.

Keywords Barcode system, Lymphatic metastasis, Lineage tracing, Tumor microenvironment, Soil-seed interaction

†Kai Miao and Aiping Zhang contributed equally to this work.

*Correspondence:
Kai Miao
kaimiao@um.edu.mo
Chu-Xia Deng
cxdeng@umac.mo
[1]Cancer Center, Faculty of Health Sciences, University of Macau, Macau SAR, China
[2]Centre for Precision Medicine Research and Training, Faculty of Health Sciences, University of Macau, Macau SAR, China
[3]MOE Frontier Science Centre for Precision Oncology, University of Macau, Macau SAR, China
[4]Division of Hepatobiliary & Pancreatic, Department of Surgery, The University of Hong Kong-Shenzhen Hospital, Shenzhen, China
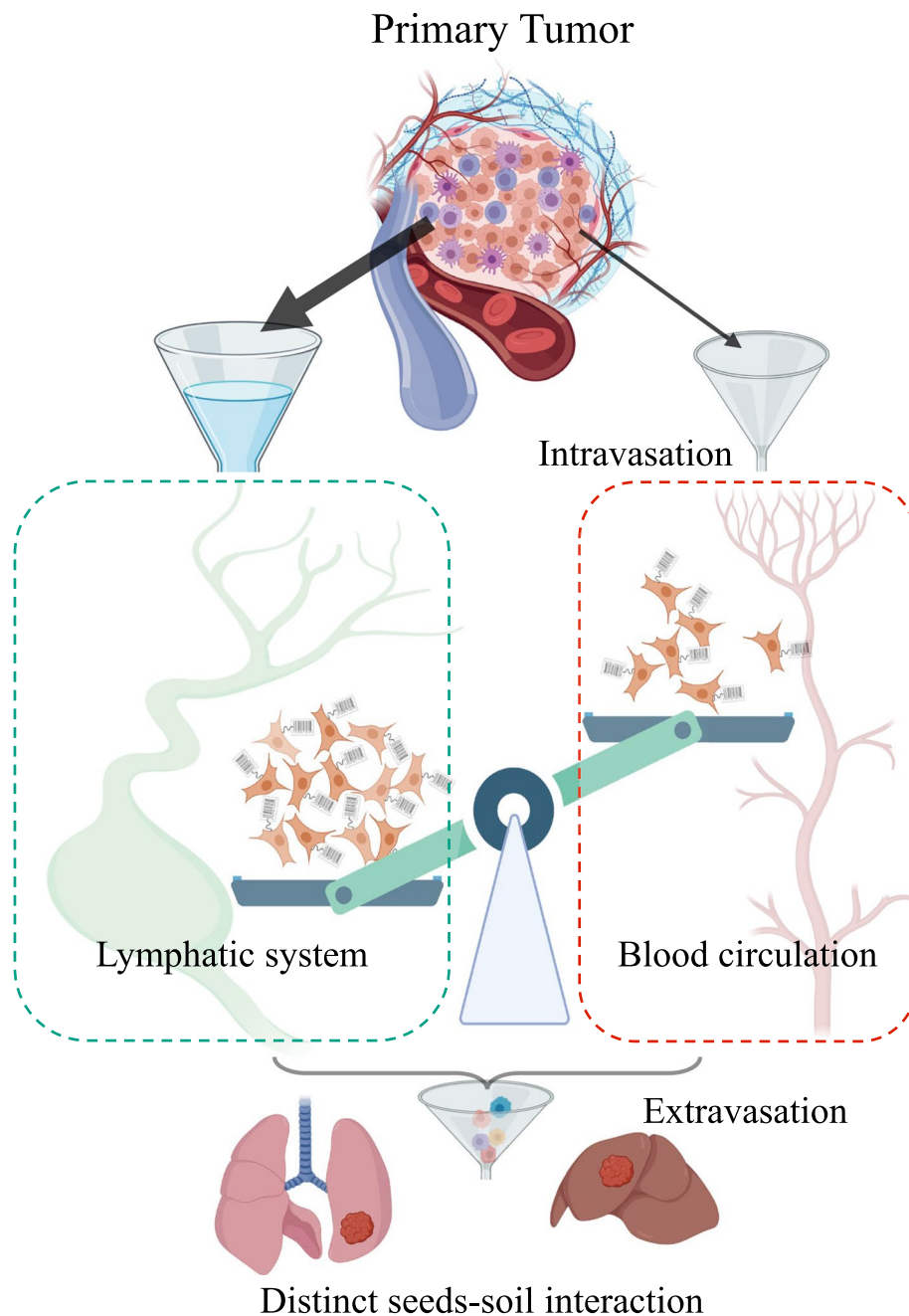[5]Key Laboratory of Animal Genetics, Breeding and Reproduction of the MARA, National Engineering Laboratory for Animal Breeding, College of Animal Science and Technology, China Agricultural University, Beijing, China
[6]Department of Developmental Biology, State Key Laboratory of Organ Failure Research, School of Basic Medical Sciences, Southern Medical University, Guangzhou, Guangdong, China.

Miao *et al. Molecular Cancer*   (2025) 24:75

Page 2 of 10

Graphical Abstract

Introduction

Cancer metastasis is a process in which tumor cells from the primary tumor invade through the basement membrane into blood or lymphatic vessels (intravasation), travel through the blood or lymphatic system, and colonize distant organs/tissues for further development. Metastasis is responsible for up to 70–90% of cancer-related mortality [1, 2]. Shedding of tumor cells into the vasculature and lymphatic vessels occurs simultaneously in the vast majority of cancers [1, 3, 4].

Miao *et al. Molecular Cancer*   (2025) 24:75

Page 3 of 10

Clonal origin analysis has demonstrated distinct lineage relationships between the lymphatic system and blood circulation [5, 6] indicating the complexity of tumor cell dissemination [7].

The metastasis of different tumor types demonstrates a biased distribution among distant organs. This phenomenon is called "organotropism" or "organ-specific metastasis" and is believed to be caused primarily by coordination between metastatic cancer cells ("seeds") and their organ/tissue microenvironment ("soil"), i.e., the "seed and soil" hypothesis [8, 9]. Tissue-specific gene signatures and signalling pathways in metastatic breast cancers have been identified by bulk RNA sequencing (RNA-seq) by comparing the primary tumor and distant lesions in different organs [10]. However, whether these distinct signatures represent different cell populations within the primary tumor or the same type of cell educated to exhibit different gene expression profiles remains unknown.

In this study, we used a barcoding system to track individual tumor cells to show cell fate during metastasis, including their migration ability and route. Combining this lineage tracing analysis with single-cell RNA sequencing (scRNA-seq), we demonstrated that the lymphatic system is the main route of breast cancer dissemination and metastasis with the aid of prewired particular cytokine expression, which serves as an intrinsic driver for tumor cell crosstalk with the TME. Our data also provide evidence that cancer cells derived from the same progenitor cell show different gene expression patterns and a different tumor cell-TME communication paradigm in different soils.

Results

Intravasation is the bottleneck for tumor metastasis through the blood circulation

We applied the CellTag barcoding system [11] to the mammary carcinoma cell lines 4T1 and EMT6 to trace tumor cell metastasis in mouse cancer models. Because both cell lines can spontaneously produce highly metastatic tumors in immunocompetent mice, they are ideal models to show the tumor cell dissemination route and fate for studying the intrinsic drivers and their communication patterns with stromal cells in the tumor immune microenvironment during metastasis.

First, we transfected cells with a combinatory cell indexing barcode tag fused to a GFP marker and then sorted the cells by fluorescence-activated cell sorting (FACS) to select labelled tumor cells. To ensure that each cell contained a unique barcode, we used a low multiplicity of infection (MOI) of 0.01 for infection and sorted 10,000 GFP-positive cells, a much lower number than the

number of different barcodes (19,973). Then, the GFP-positive cells were expanded in culture for 7 days prior to DNA sequencing (DNA-seq), which detected 7,665 and 7,809 unique barcodes in 4T1 and EMT6 cells, respectively. We then inoculated cells into the 4th mammary fat pad of BALB/c mice to model tumor metastasis (Fig. 1A; n = 10 mice per cell line). Mice were monitored twice a week, and blood samples were collected at 10-day intervals. A complete autopsy was performed to assess metastatic tumors when the mice were moribund (30 days for 4T1-derived tumors and 47 days for EMT6-derived tumors) (Figure S1A).

All recipient mice exhibited severe metastasis in distant organs, especially the lung and liver (Figure S1B). DNA-seq detected 4,741 unique barcodes for 4T1 cells and 5,227 unique barcodes for EMT6 cells, on average, from each primary tumor (Table S1, Table S2). Notably, the diversity of the barcodes identified in the blood and distant organs decreased dramatically to approximately 1.04–2.46% of that in the primary tumour (Fig. 1B). A similar pattern was observed in EMT6 tumor models (Figure S1C). Based on the barcode system, we were able to trace metastasis fate of tumor cells in each of all animals. Using a tumor derived from 4T1 cells as an example (mouse #540), which contains 4,456 different barcodes (i.e. different types of individual cells), we found very small fraction of cells (68, 38 and 39 barcodes in the blood collected at D10, D20 and D30) could escape from the primary tumor into the blood (Figure S1D). Further analysis detected a total of 86 distinct barcodes with 26 common and varying number of barcodes specific for each time point in this tumor (Figure S1E). Next, we traced the fate of these circulating tumor cells (CTCs) by studying their presence in the metastasis tumors, and found that 33 out of 68 (48.5%) CTCs at day 10; 24 out of 38 (63%) CTCs at day 20; and 25 out of 38 (66%) CTCs at day 30 could eventually colonize in the lung and/or liver (Figure S1F). Analysis of all 10 tumors derived from both 4T1 and EMT6 models revealed a similar pattern, i.e. only small fractions, on average < 3%, of all barcodes were detected in the blood at these time points (Fig. 1B, Figure S1G, H). These data suggest that the step of cancer cell intravasation from the primary tumor is a major bottleneck for tumor metastasis. To verify this hypothesis, we directly injected the same amount of tumor cells into a blood vessel through the tail vein. By DNA-seq, we found that the average number of barcodes in lung metastatic tumors was 1,839, approximately 15 times higher than that (121 CellTags) detected in the lungs of mice implanted with the same number of cancer cells in the mammary fat pad (Fig. 1C, Figure S1I, J, Table S3).
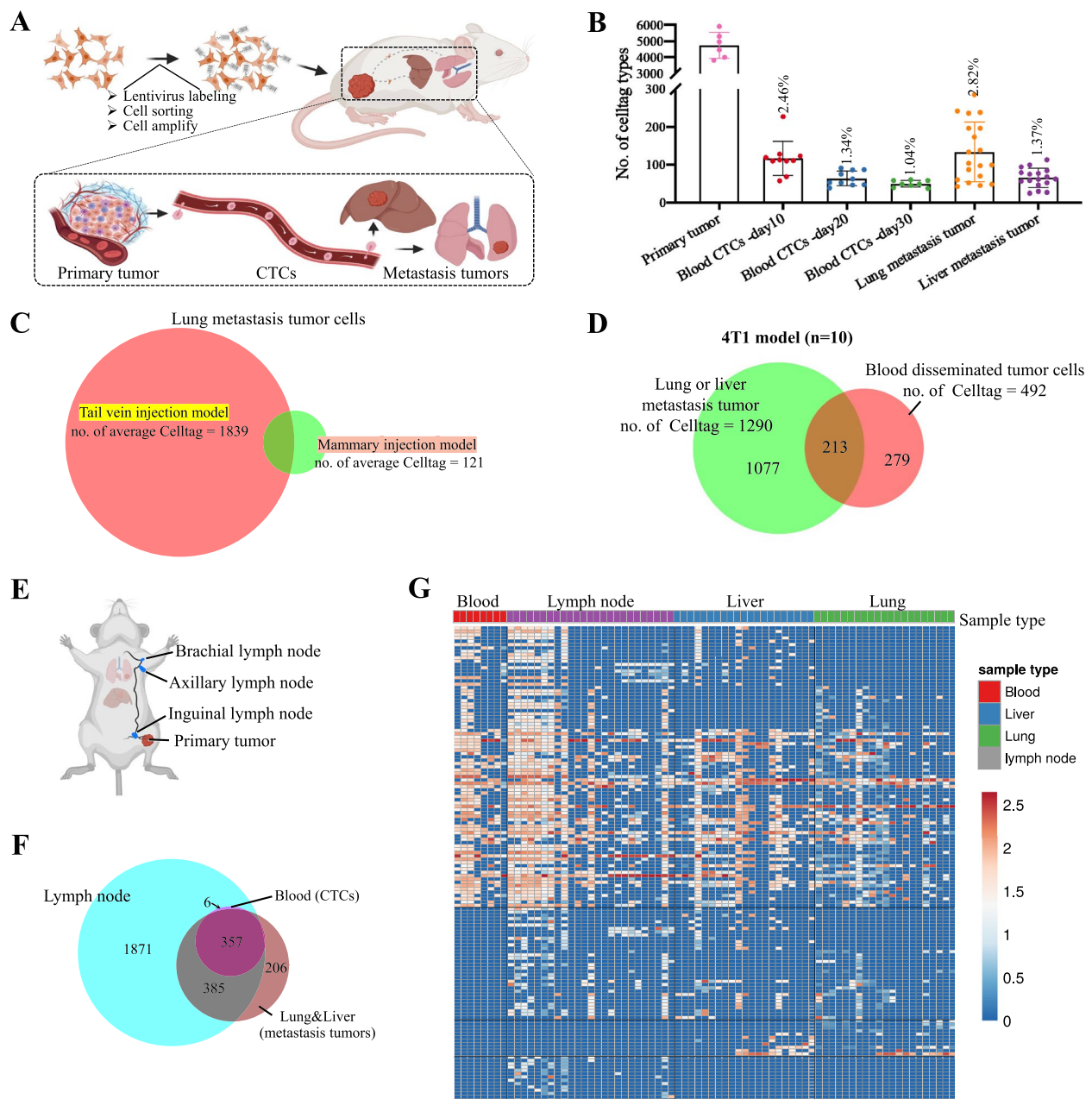
Miao *et al. Molecular Cancer*   (2025) 24:75

Page 4 of 10

Fig. 1 Lineage tracing of tumor cell metastatic rate and route. **A** Schematic illustration of the strategy for tracing tumor cell migration fate using single-cell barcoding integrated with scRNA-seq (*n* = 10). **B** CellTag compositions analysis between the primary tumor, blood CTCs collected from different days and liver/lung metastasis tumors (Liver MT, Lung MT), which were collected at the end timepoint. **C** CellTag barcode types identified in lung and liver metastatic tumors were analysed in the tail vein injection model and compared with those in the mammary fat pad inoculation model. **D** Summary of CellTag barcode types were compared between different organs/tissues and blood disseminated tumor cells in the 4T1 metastasis model (*n* = 10 mice). **E** Workflow of the sample collection strategy to demonstrate the tumor cell lymphatic and hematogenous migration routes. **F** The CellTag barcode types were compared in different organs/tissues in the 4T1 metastasis model on day 30 (*n* = 7 mice). **G** Heatmap presentation of the CellTag composition in the blood, lymph nodes, and distant organs in the 4T1 cell metastasis model. CellTags were defined based on their presence in at least 5% of tumors or more than 3 samples

The lymphatic system is the main route of tumor dissemination

Next, we compared the barcode types in the lung and liver with those in the blood, and the data revealed there are a total of 1290 different barcodes in the lung and liver, however, as much as 83% (1077/1290) of them were not
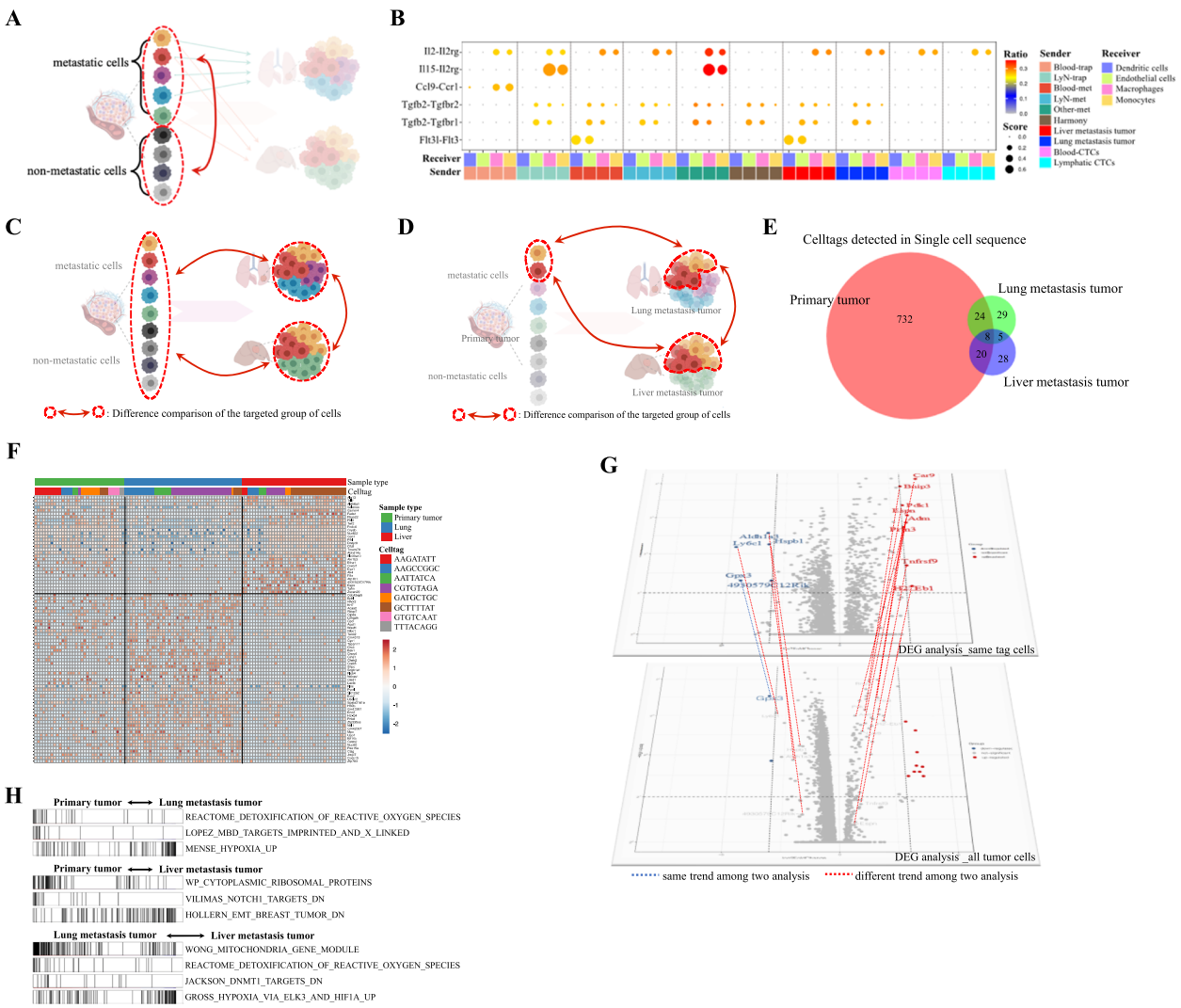
Fig. 2 scRNA-seq identifies the intrinsic features correlated with distinct metastatic fates. **A** Schematic illustration of the strategy for identifying the distinct features between metastatic cells and nonmetastatic cells using single-cell barcoding integrated with scRNA-seq. **B** Ligand–receptor pairs in tumor-–stromal interactions in primary tumors; only the differential pairs among a variety of tumor cells were used in the plot. The colour and size of the circles indicate the significance and the enrichment score, respectively. **C** Schematic showing the traditional bulk strategy for gene expression comparison analysis among primary tumor and metastasis tumors. All cells in the different organs were used for the following analysis. **D** Schematic showing the strategy for comparative analysis of gene expression. Cells with the same tags in both primary tumors and metastatic tumors were used for the following analysis. **E** Venn diagram showing the CellTags detected by the scRNA-seq of primary, lung metastatic, and liver metastatic tumors. **F** Heatmap showing that the same cells labelled with identical tags were educated for differential regulation of their gene expression pattern in distinct organs during metastasis. **G** Differentially expressed genes identified by the traditional bulk RNA analysis approach and the identical tag approach were compared. **H** GSEA of the differentially expressed genes in identical tagged cells during metastasis from the primary tumor to the lung and from the primary tumor to the liver, as well as distinct metastasis organs

detected in the blood (Fig. 1D, Figure S2A). The analysis of tumors derived from EMT6 also revealed a similar pattern, i.e., about 71% (155/221) of barcodes detected in the lung and liver metastatic tumors were not found in the blood (Figure S2B, C). The detailed analysis and comparison of barcodes regarding their appearance and abundance in the blood at all three time points from all the mice, in the lung and liver metastatic tumors revealed

similar patterns (Figure S2D, Table S4). Altogether, these data suggest that while CTCs were definitely appeared in the metastatic tumors, large fraction of metastatic cells might disseminate through other routes.

To identify the tumor cell migration route, we specifically checked the CellTag information of the lymphatic-disseminated tumor cells in the inguinal, axillary, and brachial lymph nodes (Fig. 1E, Table S5). To further show

Miao *et al. Molecular Cancer*   (2025) 24:75

Page 6 of 10

the cell migration route, we first defined lung and liver metastatic cells based on the barcode information in the lung and liver on day 30. Then, the barcode composition was compared with that of lymph node-disseminated tumor cells collected at different time points. As the Venn diagrams show, all the tumor cells in the blood circulation were found in the lymphatic system, and more metastatic tumor cells were detected in lymph nodes than in the blood (Fig. 1F). This result may indicate that tumor cells have a stronger ability to intravasate into the lymphatic system than into the blood circulation [7, 12–14], while during lymphatic circulation, the majority of tumor cells gradually die of ferroptosis or other mechanisms during the journey before colonizing remote tissues/organs to establish metastatic tumors [15, 16]. Further global cell fate comparison defined based on 7 replicates also indicated that lymphatic-disseminated tumor cells exhibited the highest diversity, covering almost all blood circulating tumor cells (CTCs), and formed metastatic tumors in distinct organs (Fig. 1G). These data indicate that the lymphatic system is the route of tumor metastasis, which has been widely reported in clinical [17].

Thus, by using the CellTag system, we successfully discovered the tumor cell metastatic route and consequently defined different types of metastatic tumor cells. Therefore, based on this lineage information, we can determine the fates of cells in primary tumors, i.e., which cells could migrate into the blood and/or lymphatic system, which cells could colonize the lung or liver via the blood circulation and which cells could disseminate into the lung or liver through the lymphatic system.

These results indicate that tumor cells may already be wired for a particular migration route before disseminating from the primary site [15, 17, 18], a finding that demonstrated heritability and reproducibility among the mouse models.

Communication interactions between tumor cells and stromal cells correlate with their metastasis route

To identify the intrinsic features that govern the metastatic fate of tumor cells, we conducted scRNA-seq on the primary tumors, and 6,066 cells passed quality control. Dimensionality reduction via uniform manifold approximation and projection (UMAP) clustering yielded 10 distinct clusters, including tumor cells, fibroblasts, macrophages, T cells, and other types of stromal cells, based on distinct marker gene expression patterns (Figure S3, S4A). Then, we focused specifically on tumor cells for dimensionality reduction to reveal the differences between different types of metastatic cells and nonmetastatic cells (Fig. 2A).

To investigate the driving force responsible for the metastatic fate of different types of tumor cells, we conducted cell–cell interaction analysis by using the CellCall program [19]. The results indicated that tumor cells communicated closely with a variety of stromal cells, especially macrophages, fibroblasts, endothelial cells, and dendritic cells (Figure S4B). Different types of tumor cells interacted differently with immune cells (Figure S4C, D). Analysis of signal transduction from tumor cells to stromal cells showed that different types of cells secreted different cytokines or chemokines into the microenvironment, which might help to direct their migration route. Specifically, most of the tumor cells, except the Blood-trap tumor cells, demonstrated a high expression level of Tgfβ2 to interact with endothelial cells and macrophages. In contrast, the Blood-trap tumor cells communicated strongly with macrophages and monocytes. The LyN-trap and Other-met tumor cells showed very similar interaction patterns, suggesting that Other-met tumor cells also disseminated through the lymphatic system but could not be detected, possibly because they were moderately difficult to detect (Fig. 2B).

Through comparative analysis of the cell–cell interactions between the abovementioned groups as well as the nonmetastasized cells, we identified several specific features that correlated with lymphatic and haematogenous dissemination and with lung or liver metastasis. Notably, the observation that the Harmony cells lacked Il2-Il2rg, Il15-Il2rg, Ccl9-Ccr1, Flt3l-Flt3–mediated communication with stromal cells indicating the "prewired" cytokine signalling played a key role in tumor metastasis (Fig. 2B, S4E).

Tumor cells demonstrate distinct properties in congenial soil

It is a well know concept that metastases develop only when the seed and soil are compatible. Samples from the same organ/tissue demonstrated comparable CellTag diversity even in different mice (Figure S5A, B), indicating that tumor cells exhibited greater similarity in more similar sample types and that highly conserved features existed in the tumor cells to modulate the migration fate, which could be identified by the dissimilar CellTags.

To further demonstrate the features of cancer cells during metastasis, we also conducted scRNA-seq on primary, lung and liver metastatic tumors to provide greater insights into the features of organ-specific tumor metastasis. A total of 23,843 cells in the 4T1 model passed quality control, of which 6,066, 8,582, and 9,195 cells were derived from the primary tumors, lung metastatic tumors, and liver metastatic tumors, respectively. Dimensionality reduction via UMAP clustering analysis yielded 10 distinct clusters, including tumor cells,

Miao *et al. Molecular Cancer*   (2025) 24:75

Page 7 of 10

fibroblasts, macrophages, T cells, and other types of stromal cells, based on distinct specific marker gene expression patterns (Figure S5C, D).

To assess the dynamic changes in tumor cells during metastasis, we conducted gene set enrichment analysis (GSEA) specifically on cancer cells to show the differences among primary tumors and metastatic tumors. First, we included all tumor cells, similar to the traditional bulk approach, for comparative analysis (Fig. 2C). After tumor cells migrated into the lung, the cytokine response, including the TGFβ signalling pathway, was upregulated, whereas the Toll-like receptor-related pathway and metabolism-regulated growth-related pathway were downregulated (Figure S5E). Tumor cells demonstrated similar gene expression patterns, except for the upregulation of Myc related pathways, after they colonized the liver (Figure S5F).

As demonstrated in our previous results, only a small portion of tumor cells could eventually colonize distant organs. To demonstrate the faithful properties of seeds in different soils, we focused specifically on cells that carried the same tag in primary tumors, lung metastatic tumors, and liver metastatic tumors (Fig. 2D). Based on the scRNA-seq data, eight types of cells fit the criteria and were therefore used for further analysis (Fig. 2E). Next, we conducted transcriptomic analysis by comparing the expression patterns of cells with the same barcodes in the primary tumor, the lung and liver metastatic tumors. Of 8 different cell tags analysed, the data revealed a site-specific expression pattern, i.e. the same types of cells elicited distinct expression patterns when they were at different sites. The expression in the lung and liver metastatic tumors was quite distinct (Fig. 2F). On the other hand, the difference was less obvious when these 8 different types of cells were compared in the same sites (lung metastasis, liver metastasis and primary tumors). These data suggest that soil play a significant role in regulating expression of cancer cells.

Compared with the traditional bulk approach, the CellTag-based analysis identified different marker genes during tumor cell migration (Fig. 2G, Figure S6A, B). As predicted, we identified a group of genes that were specifically differentially regulated after cell colonization of the lung and/or liver (Figure S6C). Specifically, Tufm and Atp1b1 were specifically upregulated after cells were colonized in the live, while some genes were down-regulated to adapt liver environment, such as Acot1, Orc5, Evc2 et al. There was a group of genes, including Sftpc, Bak1, Cnpy4, Gins1 et al., were activated to promote the cell growth in lung (Fig. 2F). GSEA and KEGG pathway enrichment analysis indicated that in addition to the terms identified by the bulk sequencing approach, the CellTag-based analysis approach also allowed us to

identify more specific functional terms related to tumor metastasis, i.e., reactive oxygen species detoxification, hypoxia, DNA methylation, and mitochondria terms (Fig. 2H, Figure S6D). These data indicate that identification of the cell migration fate provides detailed information on tumor metastasis.

The seed-soil communication paradigm differs among primary tumors and metastatic tumors

To further investigate cell–cell interactions and communication in the TME, we used CellCall to estimate the putative intercellular networks based on not only ligand and receptor expression but also receptor downstream transcription factors (TFs). These analyses revealed comprehensive crosstalk among tumor cells and stromal cells (Figure S7A-B). Specifically, tumor cells at the primary site demonstrated much stronger interactions with endothelial cells and fibroblasts, for example, Vegfc-Kdr/Flt4/Flt1 and Pdgfc-Pdgfrb/Pdgfra/Kdr/Flt4/Flt1 interactions (Figure S7C). Functional analysis indicated that some of these genes are mainly growth factors that consistently exhibit high expression levels in primary tumors. In addition, groups of ligand–receptor-transcription factor interaction pairs were commonly upregulated in all tumor metastasis steps, including Tslp-Il17r/Crlf2, Il11-Il16st/Il16ra, Tgfβ1/3-Tfgbe1/2, Csf3-Csf3r, and Csf1-Csf1r. These interactions were functionally enriched in cytokine-, chemokine- and interferon-gamma response-related genes, as well as the Tgfβ signalling pathway.

In addition, tumor cells exhibited distinct responses to stromal signals in different soils through differentially expressed receptors, which were also reflected by the levels of their downstream transcription factors (Figure S7D, E). For example, tumor cells highly expressed Bmpr2 after migrating into the lungs, with subsequent activation of its downstream targets, including Id1, Id2, Id4, and Smad4.

These results are consistent with those of GSEA, indicating that tumor cells have different gene expression patterns in different soils and that the cancer cell-TME communication paradigm varies significantly between primary tumors and metastatic tumors.

Discussion

Tumor cells metastasize mainly through the blood circulation and the lymphatic system. Numerous studies have been carried out to distinguish the role of both systems in the various stages of cancer metastasis [1, 5, 10, 20], yet which one is the "preferred" route for distant metastasis remains unknown. Phylogenetic analysis of colorectal and breast cancers has shown that in the majority of cases, lymphatic and distant metastases arose from independent subclones in the primary tumor; moreover,

Miao *et al. Molecular Cancer*   (2025) 24:75

Page 8 of 10

the two sites are subject to different levels of selection, indicating that the origin cells of lymphatic and distant metastases are fundamentally different [5, 6].

Cancer cells enter the bloodstream through intravasation to become CTCs [1], which have been proven to be highly correlated with the survival outcomes of breast cancer patients [21]. The lymphatic system is reported to be critical for metastasis. Lymphatic system-experienced melanoma cells were more resistant to ferroptosis and formed more metastases [16]. In primary tumor resection models, cancer cells can further disseminate from lymph nodes into blood vessels and form lung metastatic tumors [22]. Although metastatic spread is feasible via both routes, the relative importance of each process for a given cancer type and target organ is unknown owing to the lack of appropriate quantitative tools. Leveraging the CellTag system, we demonstrated that the majority of lung and liver metastatic cells can be detected in the lymphatic system instead of as blood-derived CTCs, indicating that lymphatic vessels could be the preferred route for tumor cell metastasis, at least in breast cancer.

Tumor cells demonstrate high intra- and intertumor heterogeneity [23] and exhibit multiclonal evolution during tumorigenesis, drug treatment, and metastasis [24–26]. Via the cell labelling system, we were able to perform lineage tracing to define individual cell fates. In our metastasis models, different tagged cells, even in established cell lines, consistently demonstrated migration route bias and organotropic metastasis, indicating that intrinsic mechanisms were prewired in each tumor cell before it disseminated from the primary site, controlling its heritable and reproducible migration fate in individual mice. Owing to its tremendous complexity and diversity, the immune system performs multifaceted functions in the TME, with different immune cells playing distinct or sometimes even opposing roles. Emerging evidence indicates that intrinsic tumor genetic aberrations shape the tumor milieu. Intratumor heterogeneity predisposes tumors to an elusive tumor-TME crosstalk pattern, making exploitation of the intrinsic properties of cancer cells impossible without high-throughput approaches for cell fate classification. Further scRNA seq and cell–cell interaction analyses allowed us to identify the unique features of metastatic tumor cells in primary tumors, for example, Il2-Il2rg and Flt3l-Flt3, which mediate interactions between dendritic cells and endothelial cells. Consistent with this, in our subsequent preliminary tests, we found silencing Il2 in 4T1 cells could dramatic decrease lymphatic-disseminated tumor cells. We will continually work on this to further under the mechanism of lymphatic dissemination and explore the therapeutic disruptions in our future efforts.

Besides, communication interactions between tumor cells and stromal cells correlate with their metastasis route. For instance, we observed that fibroblasts exhibited extensive communication with tumor cells, particularly through the secretion of cytokines and growth factors such as TGFβ and PDGF. These interactions are consistent with previous studies showing that cancer-associated fibroblasts (CAFs) promote tumor cell invasion, migration, and survival by remodeling the extracellular matrix (ECM) and creating a pro-metastatic niche [27, 28]. Specifically, our single-cell RNA sequencing (scRNA-seq) data reveal that fibroblasts in the primary tumor microenvironment expressed high levels of TGFβ2, which is known to enhance tumor cell invasiveness and epithelial-mesenchymal transition (EMT) [29, 30]. Additionally, fibroblasts were found to interact with tumor cells through ligand-receptor pairs such as PDGFC-PDGFRB, which are implicated in promoting tumor cell proliferation and survival [31]. These findings suggest that fibroblasts not only support tumor cell dissemination but also play a role in preparing the "soil" for metastatic colonization. Beyond fibroblasts, our data also show that tumor cells interact closely with endothelial cells, particularly through VEGF and TGFβ signaling pathways. VEGF, a well-known angiogenic factor, is highly expressed by tumor cells and interacts with endothelial cell receptors such as KDR and FLT1. This interaction is crucial for tumor cell intravasation, as VEGF signaling promotes vascular permeability, allowing tumor cells to enter the bloodstream or lymphatic system [32]. Moreover, our study highlights the role of TGFβ signaling in endothelial cell-tumor cell communication. TGFβ2, expressed by tumor cells, is found to interact with endothelial cells, potentially promoting endothelial cell activation and facilitating tumor cell extravasation at distant sites. This is consistent with previous studies showing that TGFβ signaling enhances endothelial cell permeability and supports the formation of metastatic niches [33].

Tumor cells colonizing distinct organs/tissues to form metastatic tumors usually exhibit tremendously different gene expression patterns [10]. Because only a small portion of shed tumor cells can form metastatic tumors, RNA-seq (at either the bulk or single-cell level) cannot determine whether organotropism is generated because different seeds preferentially migrate to different organs or because the same seeds are educated for differential gene expression in different soils. Therefore, this lineage tracing approach coupled with scRNA-seq analysis provided a powerful means to distinguish between these alternatives. In this regard, our data precisely elucidated that tumor cells derived from the same progenitor cell exhibit different gene expression patterns in different soils and that the cancer cell-TME

Miao *et al. Molecular Cancer*   (2025) 24:75

Page 9 of 10

communication paradigm varies significantly between primary tumors and metastatic tumors.

In summary, leveraging the CellTag system, we demonstrated the route of breast cancer metastasis and identified the potent intrinsic drivers of lymphatic dissemination (Figure S8). To strengthen the clinical utility of tumor metastasis prediction and therapy, prospective studies should exploit effective cellular markers for cellular and molecular properties and identify targets for pharmacological treatment to block or eliminate cancer cell dissemination.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12943-025-02279-w.

Additional file 1. Table S1-Celltag information in 4T1 tumor-bearing mice. Table S2-Celltag information in EMT6 tumor-bearing mice. Table S3-Celltag information in tail vein injection mice model. Table S4-Candidate celltag for tumor cell dissemination. Table S5-Lymph node Celltag information in 4T1 tumor-bearing mice. Table S6 primers for Bulk DNA barcode sequencing.

Authors' contributions
K.M. and C.D. designed this study; K.M. and A. Z. established the cell indexing system; A. Z. performed bioinformatics analyses; K.M., X. Y., Y. Z., L. W., X. Z., Y. F., F. S., and S. G. performed experiments; H. S., and J. X. conducted the cell sorting experiments and data analysis; A. L., J. Z., performed cell–cell communication analysis; Z. W., P. L., D. W., supervised the bioinformatics analysis; S. G., X. Z., X. X., C. D. supervised the experiments; and K.M., A. Z., and C.D. wrote the manuscript.

Data availability
The scRNA-seq data have been deposited in the Sequence Read Archive (SRA) database under accession code SRP330790. All other data supporting the findings of this study are included in the supplementary files.

Declarations

Ethics approval and consent to participate
All mouse experiments were performed under the ethical guidelines of the University of Macau (animal protocol number: UMAEC-015–2019, UMARE-023–2022).

Competing interests
The authors declare no competing interests.

Author details
[1]Cancer Center, Faculty of Health Sciences, University of Macau, Macau SAR, China. [2]Centre for Precision Medicine Research and Training, Faculty of Health Sciences, University of Macau, Macau SAR, China. [3]MOE Frontier Science Centre for Precision Oncology, University of Macau, Macau SAR, China. [4]Division of Hepatobiliary & Pancreatic, Department of Surgery, The University of Hong Kong-Shenzhen Hospital, Shenzhen, China. [5]Key Laboratory of Animal Genetics, Breeding and Reproduction of the MARA, National Engineering Laboratory for Animal Breeding, College of Animal Science and Technology, China Agricultural University, Beijing, China. [6]Department of Developmental Biology, State Key Laboratory of Organ Failure Research, School of Basic Medical Sciences, Southern Medical University, Guangzhou, Guangdong, China.

References
1. Chaffer CL, Weinberg RA. A Perspective on Cancer Cell Metastasis. Science. 2011;331(6024):1559–64.
2. Dillekas H, Rogers MS, Straume O. Are 90% of deaths from cancer caused by metastases? Cancer Med. 2019;8(12):5574–6.
3. Lambert AW, Pattabiraman DR, Weinberg RA. Emerging Biological Principles of Metastasis. Cell. 2017;168(4):670–91.
4. Nguyen DX, Bos PD, Massague J. Metastasis: from dissemination to organ-specific colonization. Nat Rev Cancer. 2009;9(4):274–84.
5. Reiter JG, Hung WT, Lee IH, Nagpal S, Giunta P, Degner S, Liu G, Wassenaar ECE, Jeck WR, Taylor MS, et al. Lymph node metastases develop through a wider evolutionary bottleneck than distant metastases. Nat Genet. 2020;52(7):692.
6. Naxerova K, Reiter JG, Brachtel E, Lennerz JK, van de Wetering M, Rowan A, Cai T, Clevers H, Swanton C, Nowak MA, et al. Origins of lymphatic and distant metastases in human colorectal cancer. Science. 2017;357(6346):55–60.
7. Wong SY, Hynes RO. Lymphatic or hematogenous dissemination: how does a metastatic tumor cell decide? Cell Cycle. 2006;5(8):812–7.
8. Peinado H, Zhang H, Matei IR, Costa-Silva B, Hoshino A, Rodrigues G, Psaila B, Kaplan RN, Bromberg JF, Kang Y, et al. Pre-metastatic niches: organ-specific homes for metastases. Nat Rev Cancer. 2017;17(5):302–17.
9. Gao Y, Bado I, Wang H, Zhang W, Rosen JM, Zhang XH. Metastasis Organotropism: Redefining the Congenial Soil. Dev Cell. 2019;49(3):375–91.
10. Chen WJ, Hoffmann AD, Liu HP, Liu X. Organotropism: new insights into molecular mechanisms of breast cancer metastasis. Npj Precision Oncology. 2018;2:4.
11. Biddy BA, Kong WJ, Kamimoto K, Guo C, Waye SE, Sun T, Morris SA. Single-cell mapping of lineage and identity in direct reprogramming. Nature. 2018;564(7735):219–+.
12. Leong SP, Witte MH. Cancer metastasis through the lymphatic versus blood vessels. Clin Exp Metastas. 2024;41(4):387–402.
13. Miyazaki K, Wada Y, Okuno K, Murano T, Morine Y, Ikemoto T, Saito Y, Ikematsu H, Kinugasa Y, Shimada M, et al. An exosome-based liquid biopsy signature for pre-operative identification of lymph node metastasis in patients with pathological high-risk T1 colorectal cancer. Mol Cancer. 2023;22(1):2.
14. Dieterich LC, Tacconi C, Ducoli L, Detmar M. Lymphatic vessels in cancer. Physiol Rev. 2022;102(4):1837–79.
15. Lee CK, Jeong SH, Jang C, Bae H, Kim YH, Park I, Kim SK, Koh GY. Tumor metastasis to lymph nodes requires YAP-dependent metabolic adaptation. Science. 2019;363(6427):644–9.
16. Ubellacker JM, Tasdogan A, Ramesh V, Shen B, Mitchell EC, Martin-Sandoval MS, Gu Z, McCormick ML, Durham AB, Spitz DR, et al. Lymph protects metastasizing melanoma cells from ferroptosis. Nature. 2020;585(7823):113–8.
17. Massague J, Ganesh K. Metastasis-Initiating Cells and Ecosystems. Cancer Discov. 2021;11(4):971–94.
18. Esposito M, Ganesan S, Kang Y. Emerging strategies for treating metastasis. Nat Cancer. 2021;2(3):258–70.
19. Zhang Y, Liu T, Hu X, Wang M, Wang J, Zou B, Tan P, Cui T, Dou Y, Ning L, et al. Cell Call: integrating paired ligand-receptor and transcription factor activities for cell-cell communication. Nucleic Acids Res. 2021;49(15):8520–34.
20. Venet D, Fimereli D, Rothe F, Boeckx B, Maetens M, Majjaj S, Rouas G, Capra M, Bonizzi G, Contaldo F, et al. Phylogenetic reconstruction of breast cancer reveals two routes of metastatic dissemination associated with distinct clinical outcome. EBioMedicine. 2020;56: 102793.

Miao *et al. Molecular Cancer*   (2025) 24:75

Page 10 of 10

21. Rack B, Schindlbeck C, Juckstock J, Andergassen U, Hepp P, Zwingers T, et al. Circulating tumor cells predict survival in early average-to-high risk breast cancer patients. J Natl Cancer Inst. 2014;106(5):dju066:1–11.
22. Pereira ER, Kedrin D, Seano G, Gautier O, Meijer EFJ, Jones D, Chin SM, Kitahara S, Bouta EM, Chang J, et al. Lymph node metastases can invade local blood vessels, exit the node, and colonize distant organs in mice. Science. 2018;359(6382):1403–7.
23. McGranahan N, Swanton C. Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. Cell. 2017;168(4):613–28.
24. Lawson DA, Kessenbrock K, Davis RT, Pervolarakis N, Werb Z. Tumour heterogeneity and metastasis at single-cell resolution. Nat Cell Biol. 2018;20(12):1349–60.
25. Dagogo-Jack I, Shaw AT. Tumour heterogeneity and resistance to cancer therapies. Nat Rev Clin Oncol. 2018;15(2):81–94.
26. Zhang A, Miao K, Sun H, Deng CX. Tumor heterogeneity reshapes the tumor microenvironment to influence drug resistance. Int J Biol Sci. 2022;18(7):3019–33.
27. Li C, Teixeira AF, Zhu HJ, Ten Dijke P. Cancer associated-fibroblast-derived exosomes in cancer progression. Mol Cancer. 2021;20(1):154.
28. Sahai E, Astsaturov I, Cukierman E, DeNardo DG, Egeblad M, Evans RM, Fearon D, Greten FR, Hingorani SR, Hunter T, et al. A framework for advancing our understanding of cancer-associated fibroblasts. Nat Rev Cancer. 2020;20(3):174–86.
29. Pickup M, Novitskiy S, Moses HL. The roles of TGFβ in the tumour microenvironment. Nat Rev Cancer. 2013;13(11):788–99.
30. Tan HY, Wang N, Lam W, Guo W, Feng Y, Cheng YC. Targeting tumour microenvironment by tyrosine kinase inhibitor. Mol Cancer. 2018;17(1):43.
31. Östman A, Heldin CH. PDGF receptors as targets in tumor treatment. Adv Cancer Res. 2007;97:247–74.
32. Hung SK, Yu CC, Lin HY, Chiou WY, Lee MS, Lin RI, Lu MC. Targeting PADI2 as a potential therapeutic strategy against metastasis in oral cancer via suppressing EMT-mediated migration and invasion and CCL3/5-induced angiogenesis. Clin Exp Metastas. 2024;41(6):925–35.
33. Padua D, Massagué J. Roles of TGFβ in metastasis. Cell Res. 2009;19(1):89–102.

Publisher's Note

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