

## Differential Contributions of Pre- and Post-EMT Tumor Cells in Breast Cancer Metastasis

Ana Rita Lourenco<sup>1,2</sup>, Yi Ban<sup>1,2</sup>, Michael J. Crowley<sup>1,2,3</sup>, Sharrell B. Lee<sup>1,2</sup>, Divya Ramchandani<sup>1,2</sup>, Wei Du<sup>4</sup>, Olivier Elemento<sup>4</sup>, Jason T. George<sup>5,6</sup>, Mohit Kumar Jolly<sup>5,7</sup>, Herbert Levine<sup>5</sup>, Jianting Sheng<sup>8,9</sup>, Stephen T. Wong<sup>8,9,10,11</sup>, Nasser K. Altorki<sup>1,2</sup>, and Dingcheng Gao<sup>1,2,12</sup>

### ABSTRACT

Metastases are responsible for the majority of breast cancer-associated deaths. The contribution of epithelial-to-mesenchymal transition (EMT) in the establishment of metastases is still controversial. To obtain *in vivo* evidence of EMT in metastasis, we established an EMT lineage tracing (Tri-PyMT) model, in which tumor cells undergoing EMT would irreversibly switch their fluorescent marker from RFP<sup>+</sup> to GFP<sup>+</sup> due to mesenchymal-specific Cre expression. Surprisingly, we found that lung metastases were predominantly derived from the epithelial compartment of breast tumors. However, concerns were raised on the fidelity and sensitivity of RFP-to-GFP switch of this model in reporting EMT of metastatic tumor cells. Here, we evaluated Tri-PyMT cells at the single-cell level using single-cell RNA-sequencing and found that the Tri-PyMT cells exhibited a spectrum of EMT phenotypes, with EMT-related genes concomitantly expressed with the activation of GFP. The fluorescent color switch in these cells precisely marked an unequivocal

change in EMT status, defining the pre-EMT and post-EMT compartments within the tumor. Consistently, the pre-EMT cells played dominant roles in metastasis, while the post-EMT cells were supportive in promoting tumor invasion and angiogenesis. Importantly, the post-EMT (GFP<sup>+</sup>) cells in the Tri-PyMT model were not permanently committed to the mesenchymal phenotype; they were still capable of reverting to the epithelial phenotype and giving rise to secondary tumors, suggesting their persistent EMT plasticity. Our study addressed major concerns with the Tri-PyMT EMT lineage tracing model, which provides us with a powerful tool to investigate the dynamic EMT process in tumor biology.

**Significance:** These findings confirm the fidelity and sensitivity of the EMT lineage tracing (Tri-PyMT) model and highlight the differential contributions of pre- and post-EMT tumor cells in breast cancer metastasis.

See related commentary by Bunz, p. 153

### Introduction

Metastasis represents the primary cause of mortality in patients with cancer. Therefore, exploring the mechanisms of metastasis stand as a major task in cancer research. Epithelial-to-mesenchymal transition (EMT), originally characterized in embryo development as a process of cellular phenotypic transdifferentiation from stationary epithelial cells

to motile mesenchymal cells, is also hijacked by tumor cells (1, 2). Through EMT, epithelial tumor cells lose their polarity and tight connection with neighboring cells, gain the ability to migrate and invade, exhibit resistance to apoptosis, and retrieve stemness properties (3–5). EMT has been enthusiastically proposed as an essential step for metastasis, in that EMT-associated features adequately meet the requirements for metastasis formation.

However, the mesenchymal phenotype was rarely observed in secondary tumors. Metastatic lesions usually resemble the epithelial phenotype of primary tumors (6–8). This observation has been tentatively explained by the dynamic nature of EMT. When the mesenchymal tumor cells had seeded at metastatic sites, they regained the epithelial features by undergoing a mesenchymal-to-epithelial transition (MET). To obtain *in vivo* evidence of the reversible EMT in metastasis, we established an EMT lineage tracing model in a multiple-transgenic mouse (*MMTV-PyMT/Fsp1-Cre/Rosa26<sup>mT/mG</sup>*, Tri-PyMT; ref. 9). In this model, breast tumor cells that underwent EMT would irreversibly switch their fluorescent marker from RFP<sup>+</sup> to GFP<sup>+</sup> due to the mesenchymal-specific Cre expression. Surprisingly, we found that metastatic lesions did not convert their fluorescent marker (9). Lung metastases were predominantly composed of RFP<sup>+</sup> cells exhibiting epithelial phenotypes. Although EMT tumor cells (GFP<sup>+</sup>) were detected in the primary tumor, circulation, and metastatic lungs, they were significantly outnumbered by their epithelial precursors (RFP<sup>+</sup>). Similar observations were obtained by using MMTV-Her2-driven breast tumors, or vimentin-driven Cre-mediated EMT lineage tracing models (9). These findings challenge the concept that EMT is required for metastasis, and have aroused vigorous discussion about its true contributions to metastasis (6, 10–12).

One of the major concerns with the Tri-PyMT model is that the expression of GFP only indicates the complete EMT. The partial-

<sup>1</sup>Department of Cardiothoracic Surgery, Weill Cornell Medicine, New York, New York. <sup>2</sup>Neuberger Berman Lung Cancer Center, Weill Cornell Medicine, New York, New York. <sup>3</sup>Weill Cornell Graduate School of Medical Sciences, Weill Cornell Medicine, New York, New York. <sup>4</sup>Department of Physiology and Biophysics, Weill Cornell Medicine, New York, New York. <sup>5</sup>Center for Theoretical Biological Physics, Department of Bioengineering, Rice University, Houston, Texas. <sup>6</sup>Medical Scientist Training Program, Baylor College of Medicine, Houston, Texas. <sup>7</sup>Centre for BioSystems Science and Engineering, Indian Institute of Science, Bangalore, India. <sup>8</sup>Bioinformatics and Biostatistics Cores, Houston Methodist Cancer Center, Houston Methodist Hospital, Houston Texas. <sup>9</sup>TT & WF Chao Center for BRAIN, Houston Methodist Hospital, Houston Texas. <sup>10</sup>Department of Radiology, Weill Cornell Medicine, New York, New York. <sup>11</sup>Methodist Cancer Center, Houston Methodist Hospital, Houston, Texas. <sup>12</sup>Department of Cell and Developmental Biology, Weill Cornell Medicine, New York, New York.

**Note:** Supplementary data for this article are available at Cancer Research Online (<http://cancerres.aacrjournals.org/>).

A.R. Lourenco and Y. Ban contributed equally to this article.

**Corresponding Author:** Dingcheng Gao, Weill Cornell Medicine, 1300 York Avenue, 525 East, 68th Street, New York, NY 10065. Phone: 212-746-9450; Fax: 212-746-9393; E-mail: dig2009@med.cornell.edu

Cancer Res 2020;80:163–9

doi: 10.1158/0008-5472.CAN-19-1427

©2019 American Association for Cancer Research.

Lourenco et al.

EMT, a concept describing tumor cells exhibiting both epithelial and mesenchymal features (10, 12, 13), could possibly fail to launch the Fsp1-Cre-mediated fluorescence switch. Here, we reevaluated the efficiency and fidelity of the Tri-PyMT model by performing single-cell RNA-sequencing (scRNA-seq). We aimed to clarify the advantages and disadvantages of the RFP<sup>+</sup> and GFP<sup>+</sup> Tri-PyMT cells in the metastatic cascade.

## Materials and Methods

### Animals

CB-17 SCID mice (Charles River Laboratories) were used in the orthotopic and metastatic Tri-PyMT models. Animal works were approved and conducted following the guidance of the Institutional Animal Care and Use Committee at Weill Cornell Medicine (New York, NY).

### Tri-PyMT cells

Tri-PyMT cells were derived from primary Tri-PyMT tumors (9). Cell authentication was performed by RT-PCR of PyMT antigen expression (9). Cells were cultured in DMEM with 10% FBS, 2 mmol/L glutamine, and antibiotics. Experiments were performed with cells of 5th–10th passages from the primary culture. The *Mycoplasma*-free culture condition was confirmed by MycoAlert PLUS Kit (Lonza).

### Flow cytometry

Single-cell suspensions were prepared by digesting lung or tumor tissues with Collagenase A and DNase I (Roche Applied Science) in HBSS at 37°C for 30 minutes. GFP<sup>+</sup> and RFP<sup>+</sup> cells were detected via LSRII Flow Cytometer (BD Biosciences).

### RNA-seq analysis

Total RNA extracted from RFP<sup>+</sup> and GFP<sup>+</sup> Tri-PyMT cells with the RNeasy Kit (Qiagen). RNA-Seq libraries were constructed and sequenced following standard protocols (Illumina). RNA-seq data were analyzed with Cufflinks and Cuffdiff2 packages. GSEA was performed following descriptions at <http://www.broad.mit.edu/gsea>.

### scRNA-seq analysis

Single-cell suspension was prepared following a standard protocol of Drop-seq sample preparation at the Genomics and Epigenetics Core Facility at Weill Cornell Medicine. The Drop-seq libraries were prepared and sequenced on HiSeq 2500 (Illumina). For *in vivo* samples, RFP<sup>+</sup> and GFP<sup>+</sup> cells were FACS-sorted from Tri-PyMT primary tumors and remixed at 1:1. The sequencing library was prepared following 10X Genomics protocol and sequenced on HiSeq 4000 (Illumina).

The Drop-seq data analyses were performed with the Seurat R package (14). The data quality was controlled by the total number of genes (200–5,000 genes), unique molecular identifiers (>200), and the percentage of mitochondria gene (<2%). The mapping of RFP and GFP sequences were used to identify RFP<sup>+</sup> and GFP<sup>+</sup> cells, respectively. After filtering, 871 RFP cells and 3,357 GFP cells were preserved for further analyses. The top 10 principal components (PC) were selected for tSNE visualization. The Wilcoxon rank sum test in the Seurat package was employed for differential expression analysis.

### EMT score calculation of single cells

Using the identified signature genes in RFP<sup>+</sup> and GFP<sup>+</sup> cells, we normalized their expression matrix, and calculated the sum of expression values across all signature genes (unweighted) for each cell, and added 1,000 to ensure positive values. The ratio of the sum

of mesenchymal to that of epithelial genes was defined as the EMT score, which was further employed to build a binary classifier and ROC curve.

### RT-PCR analysis

Total RNA was extracted using the RNeasy Kit (Qiagen), and converted to cDNA using qScript<sup>TM</sup>\_cDNA\_SuperMix (Quanta Biosciences). PCR was performed with primers and iQTM SYBR Green master mix on a CFX96 System (Bio-Rad).

Gapdh: GGTCTCAGTGTAGCCCAAG, AATGTGTCCGTCG-TGGATCT  
 E-cad: ACACCGATGGTGAGGGTACACAGG, GCCGCCACA-CACAGCATAGTCTC  
 Vimentin: TGACCTCTCTGAGGCTGCCAACC, TTCCATC-TCACGCATCTGGCGCTC  
 Snai1: ACTGGTGAGAAGCCATTCTCCT, CTGGCACTGG-TATCT-CTTCACA  
 Fn1: CGAAGAGCCCTTACAGTTCCA, ATCTGTAGGCTGG-TTCAGGC  
 Col18a: GCAGTGCCATTCCAAGTTCTC, AACATTCTCTG-GGAAGTCTGGT  
 Mmp14: TTGTCTTCAAGGAGCGATGGT, AGGGAGGCTT-CGTCAAACAC  
 Tgfb: ACGTCACTGGAGTTGTACGG, GGGGCTGATCCCG-TTGATT  
 Ccl2: CACTCACCTGCTGCTACTCA, GCTTGGTGACAAAA-ACTACAGC  
 Cxcl12: CTTCAGATTGTTGCACGGCTG, CTCGGGGTCT-ACTGGAAAG  
 Il1b: TGCCACCTTTTGACAGTGATG, ATGTGCTGCTGCG-AGATTTG  
 Il6: AGACAAAGCCAGAGTCCCTCAG, TTAGGAGAGCATT-GGAAATTGG  
 Vegfc: CTTGTCTCTGGCGTGTCC, TTCAAAGCCTTGA-CCTCGCC  
 Vegfd: GCCTGGGACAGAAGACCACT, GCAGCAGCTCTC-CAGACTTT  
 Fgf2: GGCTGCTGGCTTCTAAGTGT, TCTGTCCAGGTCCT-GTTTTG  
 Angpt1: TTCCAGAACACGACGGGAAC, TAATTCTCAAGT-TTTTGCAGCCAC  
 Pdgfa: GGAGGAGACAGATGTGAGGTG, GGAGGAGAAACA-AAGACCGCA

### Endothelial cell proliferation assay

Mouse endothelial cells (2H11, ATCC) were seeded in 96-well plates ( $2 \times 10^3$  cells/well) in 2% FBS medium overnight, and then stimulated with supernatant collected from RFP<sup>+</sup> or GFP<sup>+</sup> Tri-PyMT cells for 3 days. Cell proliferation was measured with the CellTiter-Glo Luminescent Kit (Promega).

### Orthotopic breast tumor model

RFP<sup>+</sup> Tri-PyMT cells were FACS-sorted and injected ( $5 \times 10^5$  cells/mouse) into the mammary fat pad (#4, right) of 8-week-old female SCID mice. Primary tumors were removed when tumor sizes reach approximately 1.5 cm in diameter. Lung metastasis was analyzed at 2–4 weeks after primary tumor removal.

### Tissue processing, immunofluorescence, and microscopy

The tumor and lung tissues were fixed in 4% paraformaldehyde overnight, followed by desiccation in 30% sucrose for 2 days. Serial

sections (10–20  $\mu\text{m}$ ) were prepared from optimal cutting temperature embedded blocks. Hematoxylin and eosin and immunofluorescent staining were performed following standard protocols. Primary antibodies include E-cadherin (DECMA-1, BioLegend), vimentin (sc-7557, Santa Cruz Biotechnology), and CD31 (MEC13.3, BioLegend). Fluorescent images were obtained using a Zeiss Fluorescent Microscope (Axiovert 200M), fitted with an apotome and an HRM camera.

### Statistical analysis

Experiment results were expressed as mean  $\pm$  SD. Data distribution in groups and significance between groups was analyzed by using the Mann–Whitney T test in GraphPad Prism software.  $P < 0.05$  was considered significant.

## Results

### The fluorescence switch in Tri-PyMT cells precisely reports a specific EMT program on the single-cell level

Tri-PyMT cells were derived from primary tumors of an *MMTV-PyMT/Fsp1-Cre/Rosa26<sup>mt/mG</sup>* transgenic mouse. The cells switch their fluorescence from RFP<sup>+</sup> to GFP<sup>+</sup> in culture with 10% FBS (Fig. 1A), concomitant with changing in morphologies, altering the expression of key EMT markers (including Ecad, Occl, Fn1, Vim, Snail, and Zeb1/2), and gaining mobility and resistance to apoptosis (9). Bulk RNA-seq of purified RFP<sup>+</sup> and GFP<sup>+</sup> cells revealed significant upregulation of “Hallmark Epithelial\_Mesenchymal\_Transition,” “Angiogenesis,” and “Hypoxia” pathways in the GFP<sup>+</sup> cells with the gene set enrichment analysis (GSEA; Supplementary Fig. S1).

To evaluate the single-cell EMT statuses, we performed scRNA-seq with Tri-PyMT cells. Interestingly, RFP<sup>+</sup> and GFP<sup>+</sup> cells were preferentially clustered on opposite sides in the dimensionality reduction tSNE plot (Fig. 1B). The adjacent and continuous localization of the two clusters suggested a gradient differentiation from RFP<sup>+</sup> to GFP<sup>+</sup> according to their overall transcriptome. Projecting the expression of EMT markers on the tSNE plot revealed that the epithelial (i.e., Epcam and Krt18) and mesenchymal markers (i.e., Vim and S100a4) were largely confined to RFP<sup>+</sup> and GFP<sup>+</sup> subpopulations, respectively (Fig. 1C).

We next performed differential expression analyses of RFP and GFP cells. With criteria of log fold change  $\geq 0.25$  and  $P < 0.05$ , 241 genes in RFP<sup>+</sup> cells and 324 genes in GFP<sup>+</sup> cells were differentially expressed (Supplementary Table 1). Many well-characterized epithelial genes (i.e., *Epcam*, *Cdh1*, *Krt14*, *Cldn7*, and *Jup*), and mesenchymal genes (i.e., *Vim*, *Fn1*, *S100a4*, *Spp1*, *Col8a1*, *Cxcl12*, *Prrx1*, and *Zeb1/2*) were included in the signature. When compared with published EMT hallmark gene sets, the EMT signature in Tri-PyMT cells overlapped with 61 of 200 genes in the “Hallmark\_EMT” of MSigDB database, and with 50 of 359 genes of the dbEMT gene sets (15). Of note, only 39 overlapping genes exist between these two databases. These results suggested that a specific gene set was activated during the RFP-to-GFP transition in Tri-PyMT cells.

To assess the efficiency of the Tri-PyMT model, we quantified the EMT status of individual cells. Basically, we calculated the sum expression level of identified mesenchymal and epithelial genes of each cell. The ratio of mesenchymal and epithelial sums was defined as the EMT score. Accordingly, epithelial-like cells were expected with low EMT scores, while mesenchymal-like cells would display high scores. Density plots of RFP<sup>+</sup> and GFP<sup>+</sup> cells showed a well-defined separation with minor overlaps (Fig. 1D, left). To evaluate the performance of the fluorescence switch in predicting the EMT status of single cells, we used their EMT scores to build a binary classifier.

ROC curve revealed a high specificity and sensitivity, with an accuracy of up to 97% (Fig. 1D, right).

Together, the scRNA-seq analyses of Tri-PyMT cells revealed a continuous EMT spectrum and a specific EMT signature of this model. The fluorescence switch reported the EMT statuses of single cells with high sensitivity and specificity. Therefore, in the Tri-PyMT model, RFP<sup>+</sup> and GFP<sup>+</sup> cells would well represent the “pre-EMT” and “post-EMT” subpopulations, respectively.

### The differential contributions of pre-EMT and post-EMT cells in tumor progression

To further evaluate the roles of pre-EMT and post-EMT tumor cells in metastasis, we established orthotopic model by implanting sorted RFP<sup>+</sup> Tri-PyMT cells in the mammary fat pad of mice. Histologic (Fig. 2A) and flow cytometry (Fig. 2B) analyses revealed that the primary tumors comprised mostly RFP<sup>+</sup> cells at 4 weeks after implantation. Interestingly, a relatively higher percentage of GFP<sup>+</sup> was observed in the early stage tumors (~15% at 1 week; Supplementary Fig. S2A and S2B). This percentage decreased to 1%–2% at 4 weeks, suggesting that post-EMT tumor cells may play a more important role in the early stage of tumor progression. Of note, post-EMT tumor cells were also detected in the blood of tumor-bearing mouse, which comprised approximately 10% of total circulating tumor cells (Supplementary Fig. S2C).

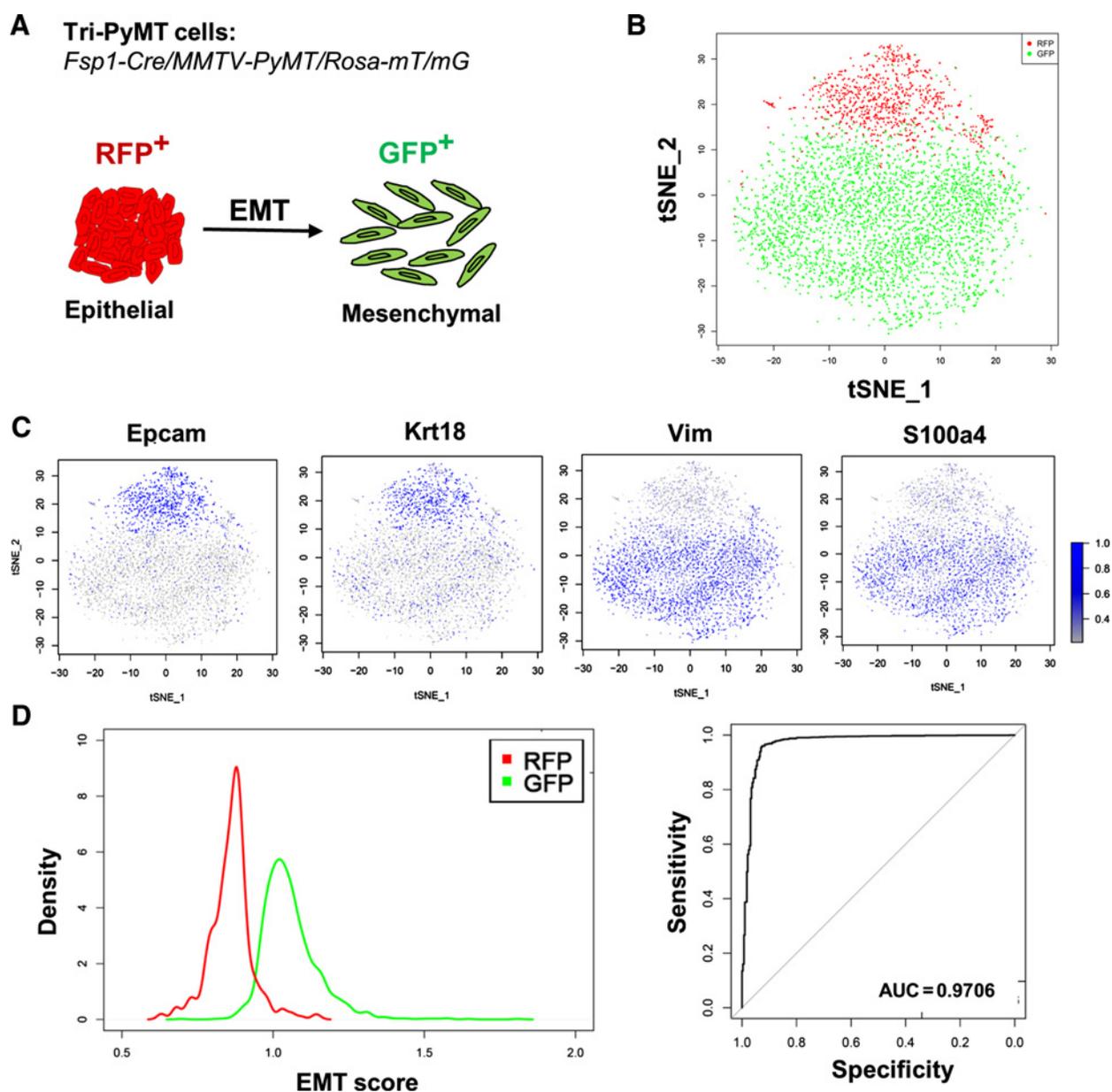
To further evaluate the fidelity of Tri-PyMT model *in vivo*, we performed scRNA-seq with sorted RFP<sup>+</sup> and GFP<sup>+</sup> cells from the orthotopic primary tumors. Two major clusters representing epithelial and mesenchymal phenotypes were detected with dimensionality reduction tSNE analysis (Supplementary Fig. S3). Consistently, RFP and GFP expression were closely associated with the epithelial (*Ecad*, *Epcam*, and *Krt18*) and mesenchymal (*Vim*, *Fn1*, *s100a4*, *Snail*, and *Zeb1*) markers, respectively. These results strongly suggested that the Tri-PyMT model reported the transition of epithelial-to-mesenchymal phenotype with high fidelity.

In the primary tumor, post-EMT (GFP<sup>+</sup>) cells were detected in close proximity to the necrotic region (Supplementary Fig. S4) and at the edge of pre-EMT (RFP<sup>+</sup>) cell clusters, directly connecting with tumor stroma (Fig. 2C). Immunofluorescence staining of CD31 (endothelial cell marker) revealed that the post-EMT (GFP<sup>+</sup>) cells were more preferentially adjacent to tumor vasculatures than pre-EMT (RFP<sup>+</sup>) cells (Fig. 2C). Consistent with the upregulated “Angiogenesis” pathway in GSEA (Supplementary Fig. S1), RT-PCR analyses of proangiogenic factors confirmed that the expression of *Vegfd*, *Pdgfa*, *Angpt1*, *IL6*, *Tgfb*, *Cxcl12*, and *Fn1* were upregulated in GFP<sup>+</sup> cells when compared with that in RFP<sup>+</sup> cells (Supplementary Fig. S5A). Moreover, supernatant collected from GFP<sup>+</sup> cells significantly enhance the proliferation of mouse endothelial (2H11) cells (Supplementary Fig. S5B). These results suggested that post-EMT cells account for a minority of primary tumor cells; however, they may play supportive roles in tumor angiogenesis.

Consistent with previous studies, lung metastases were mostly derived from the pre-EMT, but not the post-EMT compartment (Fig. 2D). RFP-dominated metastases were observed in all 69 lung nodules from 18 animals bearing Tri-PyMT tumors, suggesting a robust advantage of the pre-EMT in lung metastasis formation.

The important role of the pre-EMT tumor cells in metastasis formation was also confirmed in *MMTV-PyMT/Vim-CreERT/Rosa26<sup>mt/mG</sup>* model (9), in which the fluorescent marker switch was driven by vimentin promoter. Both primary tumors and lung metastases were mainly composed of RFP<sup>+</sup> cells (Supplementary Fig. S6), indicating the dominant role of pre-EMT cells in metastasis.

Lourenco et al.

**Figure 1.**

The Tri-PyMT model precisely reports the EMT status of Tri-PyMT cells on the single-cell level. **A**, A scheme of the Tri-PyMT EMT lineage tracing model. Epithelial tumor cells will switch their expression of fluorescent marker from RFP<sup>+</sup> to GFP<sup>+</sup> due to the activation of Fsp1 promoter during EMT. **B**, The tSNE visualization of RFP<sup>+</sup> (red) and GFP<sup>+</sup> (green) Tri-PyMT cells with the top 10 PCs of their scRNA-seq data. **C**, The relative expression of epithelial markers (Epcam and Krt18) and mesenchymal markers (Vim and S100a4) are projected to individual cells. **D**, Plots show the distribution of individual RFP<sup>+</sup> (red) and GFP<sup>+</sup> (green) cells according to their EMT scores (left), and the quantification of the sensitivity and specificity (GFP<sup>+</sup> vs. RFP<sup>+</sup>) with ROC curve (right). AUC = 0.9706.

### Pre-EMT tumor cells possess advantages in lung colonization

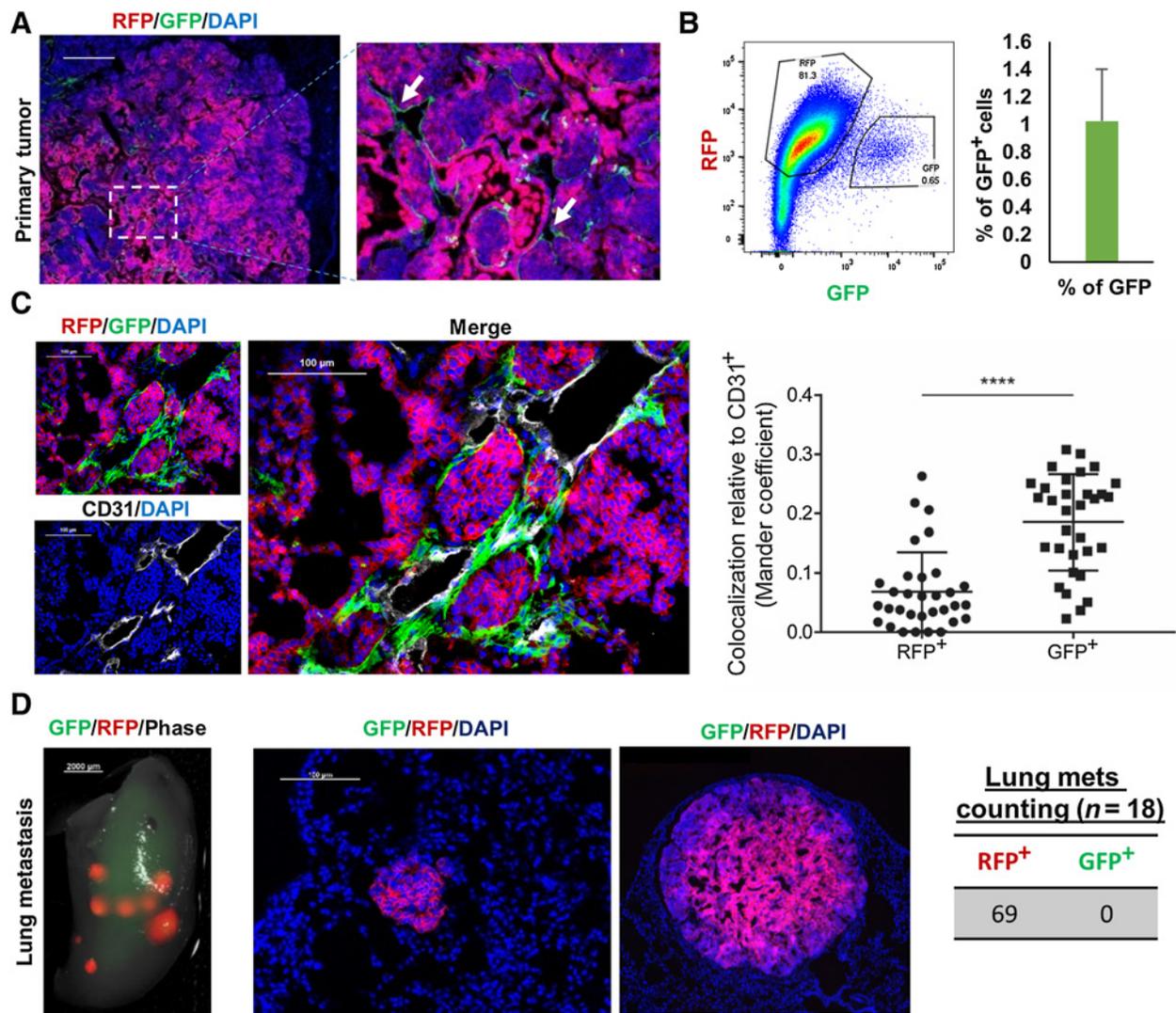
The predominant ratio of pre-EMT to post-EMT cells in the primary tumor may not be sufficient to explain the outgrowth of RFP<sup>+</sup> metastases in the lung. We further asked whether there were differences between pre-EMT and post-EMT cells in seeding and colonizing at the metastatic site. A comparative metastasis assay was performed with sorted RFP<sup>+</sup> and GFP<sup>+</sup> Tri-PyMT cells ( $2.5 \times 10^5$  cells of each) from culture via tail vein injection. Lung-seeding tumor cells were quantified via flow cytometry at various time points (Fig. 3A). We found that both RFP<sup>+</sup> and GFP<sup>+</sup> tumor cells seeded the lung with

similar low efficiencies on second day ( $0.78\% \pm 0.15\%$  and  $0.53\% \pm 0.15\%$ , respectively; Fig 3B). However, a significantly higher number of RFP<sup>+</sup> than GFP<sup>+</sup> cells were detected in the lung on 20th day after injection (Fig. 3B). These results suggest that both pre-EMT and post-EMT cells possess similar seeding capacities, whereas the pre-EMT cells have advantages in the metastasis outgrowth.

### Post-EMT tumor cells formed secondary tumors through MET

Next, we asked whether the post-EMT cells were capable to form secondary tumors. GFP<sup>+</sup> Tri-PyMT cells were sorted from culture

## Contribution of EMT Mechanism in Breast Tumor Metastasis

**Figure 2.**

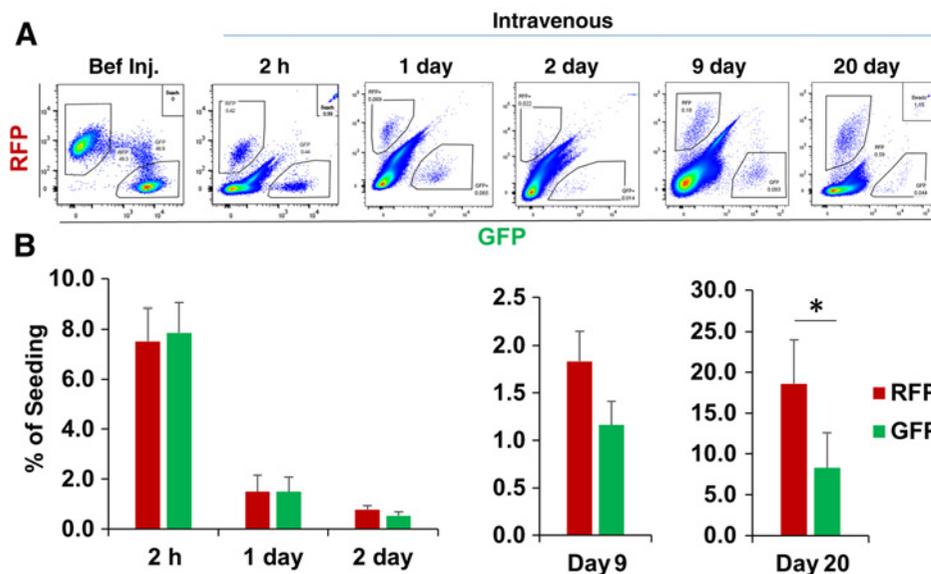
The differential contributions of pre-EMT (RFP<sup>+</sup>) and post-EMT (GFP<sup>+</sup>) cells in the Tri-PyMT tumor progression. **A**, Fluorescent images of Tri-PyMT primary tumor. Arrows, post-EMT cells (GFP<sup>+</sup>). **B**, Flow cytometry plot and quantification of GFP<sup>+</sup> tumor cells ( $n = 5$ ). **C**, Immunofluorescent image and quantification plot showing the colocalization of EMT (GFP<sup>+</sup>) tumor cells and tumor vasculature (CD31<sup>+</sup>). **D**, Fluorescent images showing the predominant role of RFP<sup>+</sup> cells in lung metastases ( $n = 18$ ).

and injected into the mammary fat pad of animals. We found that they were capable, however, formed secondary tumors at a slower growth rate as compared with their RFP<sup>+</sup> counterparts (Fig. 4A). At the 4th week after inoculation, the GFP<sup>+</sup> tumors were approximately 50% smaller than the RFP<sup>+</sup> ones. Immune staining showed that the GFP<sup>+</sup> cells regained epithelial phenotypes (Ecad<sup>+</sup>/Vim<sup>-</sup>, Fig. 4B). Given the mesenchymal features of GFP<sup>+</sup> cells in culture, these results suggested that the post-EMT tumor cells underwent “MET” to form tumors. Importantly, similar numbers of lung metastases were detected in animals bearing either GFP<sup>+</sup> or RFP<sup>+</sup> tumors (Fig. 4C). The GFP<sup>+</sup> metastatic lesions also exhibited epithelial phenotypes (Ecad<sup>+</sup>/Vim<sup>-</sup>, Fig. 4D). These results indicated that the post-EMT cells were not permanently committed to the mesenchymal lineage. Instead, they still possessed EMT plasticity to regain epithelial phenotype for initiating secondary tumors and metastases.

## Discussion

The unexpected observations that EMT is not required for tumor metastasis with Tri-PyMT model urged us to carefully evaluate the fidelity of this EMT lineage tracing model. We performed the scRNA-seq of Tri-PyMT cells from both culture and primary tumors. A specific EMT signature was identified as comparing the RFP<sup>+</sup> with GFP<sup>+</sup> cells on the single-cell level. This signature contains many well-characterized EMT markers such as Vim, Fn1, S100a4, Prrx1, and Zeb1/2. However, its overlap with the published EMT gene sets (such as the Hallmark\_EMT in MSigDB and dbEMT) was limited, suggesting the diversity of EMT program in different tumors. We preferred to describe the fluorescence switch of Tri-PyMT cells as a specific EMT program rather than a partial or hybrid EMT status, due to the lack of standardized criteria of these statuses (13). Indeed, the limited overlaps of the published EMT

Lourenco et al.

**Figure 3.**

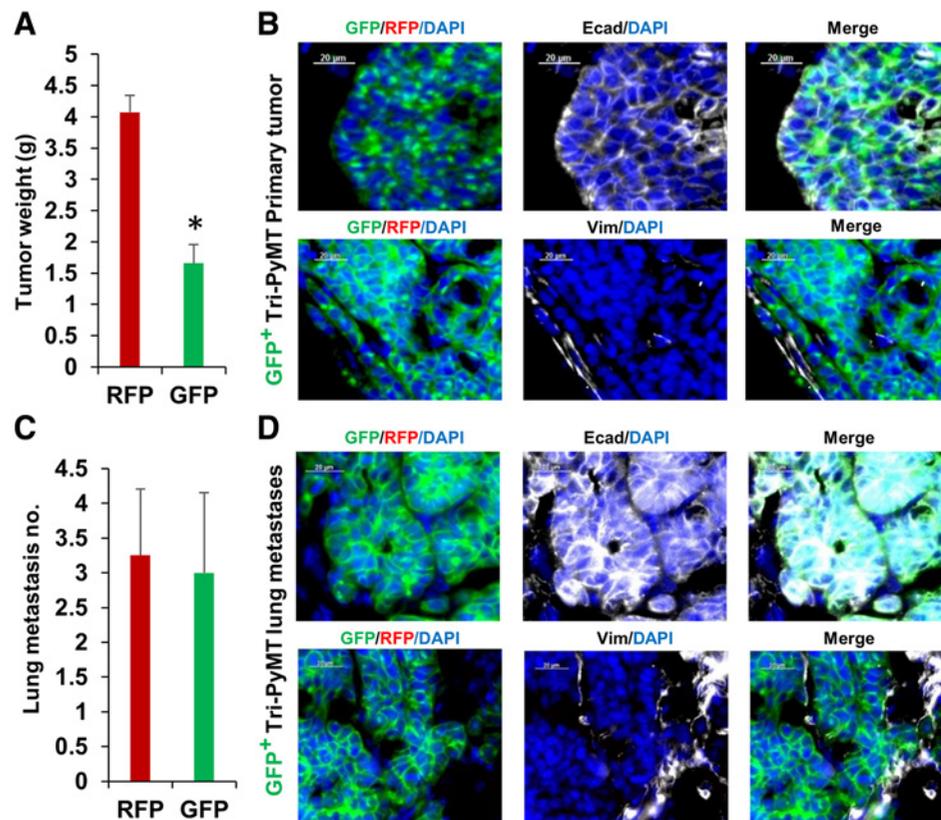
Characterizing the seeding and out-growth of pre-EMT (RFP<sup>+</sup>) and post-EMT (GFP<sup>+</sup>) tumor cells in the lung. RFP<sup>+</sup> and GFP<sup>+</sup> Tri-PyMT cells were sorted from culture, remixed at 1:1, and injected into mice through tail vein. Flow cytometry plots (**A**) and quantification (**B**) show the recovery of RFP<sup>+</sup> and GFP<sup>+</sup> cells at different times (2 hours, 1 day, 2 days, 9 days, and 20 days) after tail vein injection. \*,  $P < 0.01$ ;  $n = 3-6$  mice.

gene sets also suggest that different EMT programs are adopted by different tumors.

In the Tri-PyMT model, GFP<sup>+</sup> cells were derived from RFP<sup>+</sup> cells, sharing the same origin with their precursors. We assessed the sensitivity and specificity of Tri-PyMT model in reporting the overall EMT status at the single-cell level. An evident separation of RFP<sup>+</sup> and GFP<sup>+</sup> subpopulations was observed via an EMT-scoring assay, which

suggested the fluorescent marker switch reported a gain of mesenchymal features in Tri-PyMT cells with high accuracy.

Conflicting conclusions of EMT in metastasis are, at least in part, due to the lack of standardized EMT scoring system. Efforts have been made to describe the variety of EMT states existing in different tumor types. However, the quantified EMT statuses did not necessarily correlate with overall survival of patients with cancer (15, 16). The

**Figure 4.**

Post-EMT Tri-PyMT cells form secondary primary tumors and metastases with epithelial phenotype. **A**, Weight of GFP<sup>+</sup> and RFP<sup>+</sup> primary tumors at 4 weeks postinjection,  $n = 5$ . \*,  $P < 0.05$ . **B**, E-cadherin and vimentin staining (white pseudocolor) of GFP<sup>+</sup> Tri-PyMT primary tumors. **C**, Quantification of lung metastases in mice bearing either RFP<sup>+</sup> or GFP<sup>+</sup> primary tumor,  $n = 5$ . **D**, Fluorescent images of Ecad and Vim staining of the GFP<sup>+</sup> Tri-PyMT lung metastases.

## Contribution of EMT Mechanism in Breast Tumor Metastasis

extremely mesenchymal phenotype did not favor the metastasis development in patients with breast cancer (16). We also evaluated the EMT statuses of Tri-PyMT cells with the EMT scoring metric (16). Both RFP<sup>+</sup> and GFP<sup>+</sup> cells were categorized as E/M hybrid phenotype (Supplementary Fig. S7). Of note, such E/M hybrid statuses were referred to NCI60 cell lines, which include a wide range of fully differentiated epithelial and mesenchymal cells. This provided another clue that the EMT program could be tumor type dependent. Importantly, the Tri-PyMT model enabled us to compare tumor cells with different EMT statuses within the same tumor and assess their relative contributions to metastasis. The GFP<sup>+</sup> cells were clearly more skewed to the mesenchymal phenotypes with higher Vim/Cdh1 ratio and lower Cldn7 expression when compared with RFP<sup>+</sup> cells (Supplementary Fig. S7). The lack of GFP<sup>+</sup> metastases in Tri-PyMT model suggested that further activation of EMT programming did not grant advantages in metastasis.

With the Tri-PyMT model, we confirmed that post-EMT tumor cells were not the metastasis-initiating cells. These results do not decline the biological contributions of EMT in tumor progression. Instead, EMT is believed to endow tumor cells with many metastasis-related features including migration, invasion, and apoptosis resistance (10, 12, 17, 18). Cooperations between EMT and non-EMT tumor cells were also demonstrated (19, 20). While the EMT tumor cells invade and degrade extracellular matrix, the non-EMT tumor would follow and metastasize to the secondary organ. The EMT tumor cells could also assist non-EMT tumor cells to metastasize through a noncell autonomous activation of the GLI signal (20). In the Tri-PyMT tumors, post-EMT cells colocalize with tumor vasculature and secrete more proangiogenic factors, suggesting their supportive role in tumor angiogenesis. Nevertheless, our results do not support the hypothesis that the post-EMT tumor cells are the metastasis-initiating cells. The

biological involvement of post-EMT tumor cells in metastasis still needs further elucidation.

## Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

## Authors' Contributions

**Conception and design:** A.R. Lourenco, Y. Ban, H. Levine, D. Gao

**Development of methodology:** Y. Ban, H. Levine, S.T.C. Wong, D. Gao

**Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.):** Y. Ban, M.J. Crowley

**Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis):** A.R. Lourenco, Y. Ban, M.J. Crowley, W. Du, O. Elemento, J.T. George, M.K. Jolly, J. Sheng, S.T.C. Wong, D. Gao

**Writing, review, and/or revision of the manuscript:** A.R. Lourenco, Y. Ban, M.J. Crowley, S.B. Lee, D. Ramchandani, O. Elemento, J.T. George, N.K. Altorki, D. Gao

**Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases):** S.B. Lee, S.T.C. Wong, N.K. Altorki

**Study supervision:** N.K. Altorki

## Acknowledgments

This work was supported by NIH (5R01CA205418), the Neuberger Berman Foundation Lung Cancer Research Center, and by generous donations to the Division of Thoracic Surgery to N.K. Altorki. J. Sheng and S.T.C. Wong were supported by NIH U01CA-188388, John S Dunn Research Foundation, TT & WF Chao Foundation, and Johnston Estate Endowment.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received May 6, 2019; revised September 4, 2019; accepted October 22, 2019; published first November 8, 2019.

## References

- Thiery JP, Acloque H, Huang RY, Nieto MA. Epithelial-mesenchymal transitions in development and disease. *Cell* 2009;139:871-90.
- Lim J, Thiery JP. Epithelial-mesenchymal transitions: insights from development. *Development* 2012;139:3471-86.
- Mani SA, Guo W, Liao MJ, Eaton EN, Ayyanan A, Zhou AY, et al. The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* 2008;133:704-15.
- Ye X, Tam WL, Shibue T, Kaygusuz Y, Reinhardt F, Ng Eaton E, et al. Distinct EMT programs control normal mammary stem cells and tumour-initiating cells. *Nature* 2015;525:256-60.
- Kalluri R, Weinberg RA. The basics of epithelial-mesenchymal transition. *J Clin Invest* 2009;119:1420-8.
- Nieto MA, Huang RY, Jackson RA, Thiery JP. EMT: 2016. *Cell* 2016;166:21-45.
- Chaffer CL, San Juan BP, Lim E, Weinberg RA. EMT, cell plasticity and metastasis. *Cancer Metastasis Rev* 2016;35:645-54.
- Tsai JH, Yang J. Epithelial-mesenchymal plasticity in carcinoma metastasis. *Genes Dev* 2013;27:2192-206.
- Fischer KR, Durrans A, Lee S, Sheng J, Li F, Wong ST, et al. Epithelial-to-mesenchymal transition is not required for lung metastasis but contributes to chemoresistance. *Nature* 2015;527:472-6.
- Ye X, Brabletz T, Kang Y, Longmore GD, Nieto MA, Stanger BZ, et al. Upholding a role for EMT in breast cancer metastasis. *Nature* 2017;547:E1-E3.
- Fischer KR, Altorki NK, Mittal V, Gao D, Fischer et al. reply. *Nature* 2017;547:E5-E6.
- Yeung KT, Yang J. Epithelial-mesenchymal transition in tumor metastasis. *Mol Oncol* 2017;11:28-39.
- Grigore AD, Jolly MK, Jia D, Farach-Carson MC, Levine H. Tumor budding: the name is EMT. *Partial EMT*. *J Clin Med* 2016;5:pii: E51.
- Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* 2018;36:411-20.
- Tan TZ, Miow QH, Miki Y, Noda T, Mori S, Huang RY, et al. Epithelial-mesenchymal transition spectrum quantification and its efficacy in deciphering survival and drug responses of cancer patients. *EMBO Mol Med* 2014;6:1279-93.
- George JT, Jolly MK, Xu S, Somarelli JA, Levine H. Survival outcomes in cancer patients predicted by a partial EMT gene expression scoring metric. *Cancer Res* 2017;77:6415-28.
- Aiello NM, Maddipati R, Norgard RJ, Balli D, Li J, Yuan S, et al. EMT subtype influences epithelial plasticity and mode of cell migration. *Dev Cell* 2018;45:681-95.
- Xu Y, Lee DK, Feng Z, Bu W, Li Y, Liao L, et al. Breast tumor cell-specific knockout of Twist1 inhibits cancer cell plasticity, dissemination, and lung metastasis in mice. *Proc Natl Acad Sci U S A* 2017;114:11494-9.
- Tsuji T, Ibaragi S, Shima K, Hu MG, Katsurano M, Sasaki A, et al. Epithelial-mesenchymal transition induced by growth suppressor p12CDK2-AP1 promotes tumor cell local invasion but suppresses distant colony growth. *Cancer Res* 2008;68:10377-86.
- Neelakantan D, Zhou H, Oliphant MUJ, Zhang X, Simon LM, Henke DM, et al. EMT cells increase breast cancer metastasis via paracrine GLI activation in neighbouring tumour cells. *Nat Commun* 2017;8:15773.

# Cancer Research

The Journal of Cancer Research (1916–1930) | The American Journal of Cancer (1931–1940)

## Differential Contributions of Pre- and Post-EMT Tumor Cells in Breast Cancer Metastasis

Ana Rita Lourenco, Yi Ban, Michael J. Crowley, et al.

*Cancer Res* 2020;80:163-169. Published OnlineFirst November 8, 2019.

**Updated version** Access the most recent version of this article at:  
doi:[10.1158/0008-5472.CAN-19-1427](https://doi.org/10.1158/0008-5472.CAN-19-1427)

**Supplementary Material** Access the most recent supplemental material at:  
<http://cancerres.aacrjournals.org/content/suppl/2019/11/08/0008-5472.CAN-19-1427.DC1>

**Cited articles** This article cites 20 articles, 6 of which you can access for free at:  
<http://cancerres.aacrjournals.org/content/80/2/163.full#ref-list-1>

**E-mail alerts** [Sign up to receive free email-alerts](#) related to this article or journal.

**Reprints and Subscriptions** To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at [pubs@aacr.org](mailto:pubs@aacr.org).

**Permissions** To request permission to re-use all or part of this article, use this link  
<http://cancerres.aacrjournals.org/content/80/2/163>.  
Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.