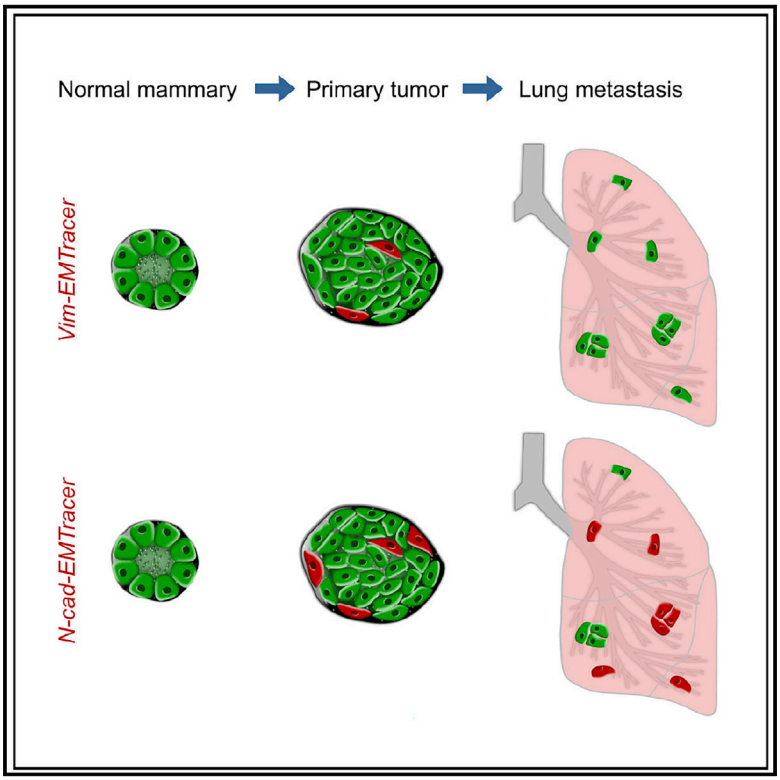
Article

Developmental Cell

Genetic Fate Mapping of Transient Cell Fate Reveals N-Cadherin Activity and Function in Tumor Metastasis

Graphical Abstract



Authors

Yan Li, Zan Lv, Shaohua Zhang, ..., Jing Feng, M. Angela Nieto, Bin Zhou

Correspondence

zhoubin@sibs.ac.cn

In Brief

Li, Lv et al. establish a genetic system to seamlessly record transient cell fate *in vivo*. Using this approach, they study epithelial-to-mesenchymal transition (EMT) gene activity from the local primary tumor to a distant metastatic site showing that N-cadherin but not vimentin is activated and functionally required during breast-to-lung tumor metastasis.

Highlights

- Generation of a genetic system for recording transient cell fate *in vivo*
- Genetic tracing of EMT gene activity during breast-to-lung tumor metastasis
- Vimentin is not involved during mammary tumor metastasis in MMTV-PyMT model
- N-cadherin is activated and functionally required during tumor metastasis

Li et al., 2020, Developmental Cell *54*, 593–607
September 14, 2020
https://doi.org/10.1016/j.devcel.2020.06.021

CellPress

Developmental Cell

CellPress

Article

Genetic Fate Mapping of Transient Cell Fate Reveals N-Cadherin Activity and Function in Tumor Metastasis

Yan Li,[1,11] Zan Lv,[1,11] Shaohua Zhang,[1] Zhuo Wang,[2] Lingjuan He,[1] Muxue Tang,[1] Wenjuan Pu,[1] Huan Zhao,[1] Zhenqian Zhang,[1] Qihui Shi,[2] Dongqing Cai,[3] Mingfu Wu,[4] Guohong Hu,[5] Kathy O. Lui,[6] Jing Feng,[7] M. Angela Nieto,[8] and Bin Zhou[1,3,9,10,12,*]

[1]State Key Laboratory of Cell Biology, Shanghai Institute of Biochemistry and Cell Biology, Center for Excellence in Molecular Cell Science, Chinese Academy of Sciences, University of Chinese Academy of Sciences, Shanghai 200031, China
[2]Shanghai Key Laboratory of Medical Epigenetics, Institutes of Biomedical Sciences, Fudan University, 201100, Shanghai, China
[3]Key Laboratory of Regenerative Medicine of Ministry of Education, Institute of Aging and Regenerative Medicine, Jinan University, Guangzhou 510632, China
[4]Center for Cardiovascular Sciences, Albany Medical College, Albany, NY 12208, USA
[5]Shanghai Institute of Nutrition and Health, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai 200031, China
[6]Department of Chemical Pathology, Li Ka Shing Institute of Health Sciences, The Chinese University of Hong Kong, Prince of Wales Hospital, Shatin, Hong Kong SAR 999077, China
[7]Laboratory Medicine, Southern Medical University Affiliated Fengxian Hospital, Shanghai 201400, China
[8]Institute de Neurociencias CSIC-UMH, Avda. Ramon y Cajal s/n, 03550 San Juan de Alicante, Spain
[9]School of Life Science and Technology, ShanghaiTech University, Shanghai 201210, China
[10]School of Life Science, Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Hangzhou 310024, China
[11]These authors contributed equally
[12]Lead Contact
*Correspondence: zhoubin@sibs.ac.cn
https://doi.org/10.1016/j.devcel.2020.06.021

SUMMARY

Genetic lineage tracing unravels cell fate and plasticity in development, tissue homeostasis, and diseases. However, it remains technically challenging to trace temporary or transient cell fate, such as epithelial-to-mesenchymal transition (EMT) in tumor metastasis. Here, we generated a genetic fate-mapping system for temporally seamless tracing of transient cell fate. Highlighting its immediate application, we used it to study EMT gene activity from the local primary tumor to a distant metastatic site *in vivo*. In a spontaneous breast-to-lung metastasis model, we found that primary tumor cells activated vimentin and N-cadherin *in situ*, but only N-cadherin was activated and functionally required during metastasis. Tumor cells that have ever expressed N-cadherin constituted the majority of metastases in lungs, and functional deletion of N-cad significantly reduced metastasis. The seamless genetic recording system described here provides an alternative way for understanding transient cell fate and plasticity in biological processes.

INTRODUCTION

Cre-loxP technology has been widely used for genetically tracing cell lineages in development, tissue homeostasis, and diseases (Kretzschmar and Watt, 2012). The temporally controlled genetic tracing utilizes tamoxifen-induced CreER-loxP recombination, such that the Cre-expressing cells and their descendants are genetically labeled after tamoxifen treatment (Tian et al., 2015). Importantly, the selected temporal window of tamoxifen treatment depends on the critical information of CreER expression, which is determined by the gene activity indicative of specific cell fate. However, cell fate transition *in vivo* could be dynamic and transient, which is manifested by transient gene expression (Süel et al., 2006; Maamar et al., 2007; Weinberger et al., 2008) and remains technically hard to capture in real time. For example, epithelial-to-mesenchymal transition (EMT) has been proposed as an essential step for tumor metastasis, but the EMT program and genes are usually transiently expressed in tumor cells (Nieto et al., 2016; Cano et al., 2000; Yang et al., 2004; Mani et al., 2008; Ye et al., 2015). Due to its transient nature, tracking EMT in tumor metastasis by conventional Cre-loxP remains technically challenging. Whether EMT is involved in the tumor progression to the metastatic state remains as a significant yet highly contentious subject in the field for the past decade (Krebs et al., 2017; Xu et al., 2017; Brabletz et al., 2018; Ledford, 2011; Fischer et al., 2015; Zheng et al., 2015; Maheswaran and

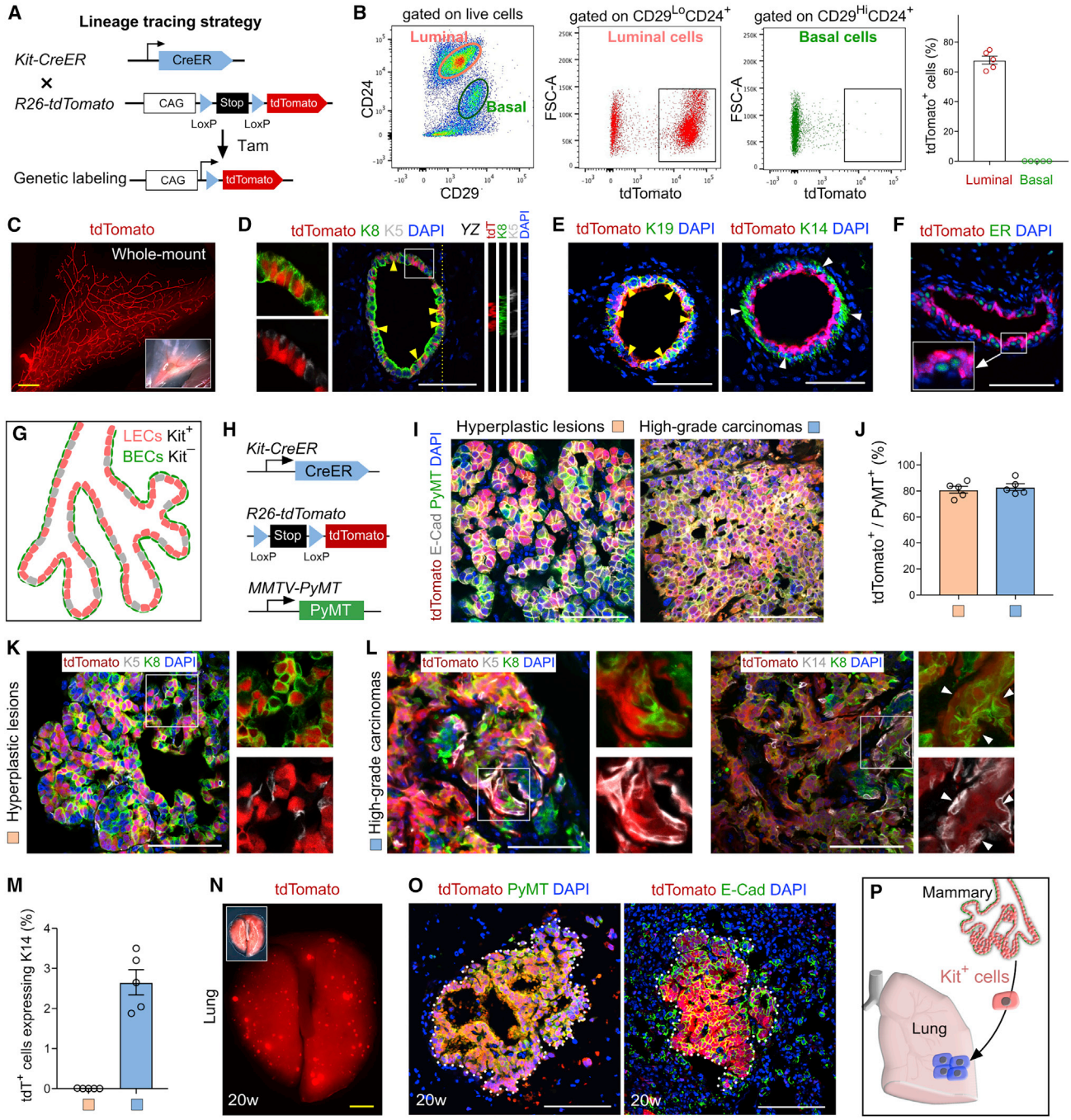


Figure 1. Kit$^+$ Luminal Epithelial Cells Contribute to Mammary Tumor Growth and Metastasis in Lung

(A) Schematic figure showing experimental strategy for tracing Kit$^+$ cells. Tamoxifen (Tam) was induced at 6–7 weeks old.

(B) Flow cytometric analysis of luminal or basal epithelial cells labeled by *Kit-CreER*. Data are mean ± SEM; n = 5.

(C) Whole-mount fluorescence of adult mammary gland. Insert indicates bright-field image.

(D–F) Immunostaining for tdTomato, basal cell markers K5 and K14, luminal cell markers K8 and K19 (D and E), endogenous estrogen receptor, ER (F) on *Kit-CreER;R26-tdTomato* mammary sections. *YZ* indicates signals from dotted lines on z stack images in (D). Yellow arrowheads, tdTomato$^+$ luminal cells; white arrowheads, tdTomato$^-$ basal cells.

(G) Cartoon image showing Kit$^+$ luminal epithelial cells (LECs) and Kit$^-$ basal epithelial cells (BECs).

(H) Generation of tumor lineage tracing model using *Kit-CreER;R26-tdTomato;MMTV-PyMT* mouse.

(I) Immunostaining for tdTomato, E-Cad, and PyMT on hyperplastic and high-grade mammary tissues.

(J) Quantification of the percentage of PyMT$^+$ tumor cells expressing tdTomato.

(K and L) Immunostaining for tdTomato, K8, K5, or K14 on mammary tissues. Arrowheads, K14$^+$tdTomato$^+$ cells. Hyperplastic lesions in (K) and high-grade carcinomas in (L).

(legend continued on next page)

Haber, 2015; Ye et al., 2017; Aiello et al., 2017), and there is a lack of clear *in vivo* genetic fate-mapping evidence for EMT.

In this study, we generated a genetic approach to seamlessly record transient cell fate. Using dual recombinases-mediated genetic lineage tracing, we examined whether EMT genes, such as Vimentin and N-cadherin, were activated and functionally required from breast-to-lung tumor metastasis. The fate-mapping results revealed that Vimentin was not activated during tumor metastasis, nor was it functionally required for tumor EMT. Instead, tumor cells that have ever expressed N-cadherin constitute the majority of tumor metastases in lung, with increased expression of other EMT genes and cell invasion capability. Specific knockout of N-cadherin in mammary epithelial cells resulted in a significant reduction of tumor metastasis. Our study provided genetic evidence of the activity and function of EMT gene N-cadherin in tumor metastasis.

RESULTS

Establishment of Epithelial Cell Tracing in Mammary Tumors and Metastasis

A recent study demonstrated that the c-KIT tyrosine kinase receptor (encoded by the Kit gene) was expressed in mammary luminal epithelial cells (Lim et al., 2009). To genetically trace these Kit$^+$ mammary luminal epithelial cells *in vivo*, we generated a *Kit-CreER* mouse strain and crossed this strain with loxP reporter *R26-tdTomato* mice to track Kit expression in normal mammary glands (Figure 1A). Tamoxifen was induced when these mice were 6–7 weeks old, and tissues were collected 2 days after tamoxifen treatment. Flow cytometric analysis showed that Kit was mainly expressed in mammary luminal epithelial cells and not in basal epithelial cells (Figure 1B). Whole-mount fluorescence images of mammary gland showed the classic branching patterns of epithelial cells labeled by *Kit-CreER* (Figure 1C). Immunostaining for tdTomato, luminal epithelial cell markers K8 and K19, or basal epithelial cell markers K5 and K14, on tissue sections showed that tdTomato$^+$ cells were luminal but not basal epithelial cells (Figures 1D and 1E). Of note, these tdTomato$^+$ cells were mainly endogenous estrogen receptor (ER)-negative luminal epithelial cells (Figure 1F). Taken together, the above results demonstrated the specific expression of Kit in luminal epithelial cells, and *Kit-CreER* could be used for genetic tracing of Kit$^+$ mammary luminal epithelial cells (Figure 1G).

To trace Kit$^+$ luminal epithelial cells during mammary tumor development and metastasis, we crossed *Kit-CreER;R26-tdTomato* mice with *MMTV-PyMT* (*M*ouse *M*ammary *T*umor *V*irus-*P*olyoma virus *M*iddle *T* antigen) mice, which form spontaneous breast adenocarcinomas (Lin et al., 2003; Guy et al., 1992) that resemble the human luminal subtype and have high lung metastasis penetrance (Figure 1H). Tamoxifen was used to induce triple-positive mice at 6–7 weeks old, and tissue samples were collected at the early (8–12 weeks old) or late (18–24 weeks old) stages of tumor formation. In the hyperplastic mammary lesions at the early stage, we found that Kit$^+$ luminal cells constituted the majority of PyMT$^+$ tumor cells and these tdTomato$^+$ tumor cells were E-Cad$^+$ (Figures 1I and 1J). These Kit-derived cells expressed luminal cell marker K8, but not basal cell markers K5, indicating that these cancer cells kept their luminal epithelial phenotype at an early stage of tumor formation (Figure 1K). In the high-grade carcinoma at a later stage, the majority of PyMT$^+$ tumor cells were tdTomato$^+$ (Figures 1I and 1J). Of note, a subset of tdTomato$^+$ cancer cells expressed basal cell markers K5 and K14, and quantification data showed that 2.65% ± 0.32% of these tdTomato$^+$ cells expressed K14 (Figures 1L and 1M), indicating that luminal tumor cells could express basal epithelial genes in high-grade carcinoma. In the lung tissue at the late stage of tumor formation, we observed many tdTomato$^+$ lung metastases (Figure 1N), and immunostaining data showed that these tdTomato$^+$ cells in metastatic lung colonies were PyMT$^+$ (Figure 1O). The original Kit$^+$ cells residing in the lung are endothelial cells (Liu et al., 2015). Notably, these tdTomato$^+$ metastases expressed the epithelial cell markers E-Cad (Figure 1O), suggesting the epithelial phenotype of the cells in lung metastases. Thus, Kit$^+$ mammary luminal epithelial cell contributes to tumor cells in primary tumor and lung metastasis (Figure 1P). In this study, we used these mouse strains to explore EMT activity during mammary tumor development and metastasis.

A Genetic System for Fate Mapping of EMT by Temporally Seamless Tracing

Conventional lineage tracing is based on an inducible Cre-loxP system, in which tamoxifen treatment leads to CreER translocation into the nucleus and results in permanent genetic labeling (Figure 2AI). Since the gene activity or expression could be transient and reversible, such as in EMT genes that are transiently expressed during metastasis, tamoxifen-induced Cre-loxP system may not capture these transient EMT gene activities (Figure 2AII). Therefore, it is technically challenging to capture transient and reversible EMT gene activity using this previously used lineage tracing strategy (Figure 2A).

To seamlessly trace and permanently record gene activity in a defined time window (i.e., from primary tumor to remote metastasis), we took advantage of two orthogonal recombination systems, Cre-loxP and Dre-rox (He et al., 2017), to monitor EMT gene activities over time. We first generated *Kit-CreER;EMTgene-LSL-Dre;NR1* triple knock-in mice, which were referred to as *EMTracer* mice throughout this study (Figure 2B). *EMTgene-LSL-Dre* was composed of a constitutive Dre recombinase driven by an EMT gene promoter, and there was a transcriptional Stop cassette (Stop) flanked by two loxP sites (LSLs) before Dre. For nested reporter 1 (*NR1*) (He et al., 2017), Cre-loxP recombination first resulted in ZsGreen labeling, and Dre-rox recombination switched the genetic labeling from ZsGreen to tdTomato

(M) Quantification of the percentage of tdTomato$^+$ cells expressing basal cell marker K14. Data are mean ± SEM; n = 5.
(N) Whole-mount fluorescence images of lung collected from *Kit-CreER;R26-tdTomato;MMTV-PyMT* mice. Insert indicates bright-field image.
(O) Immunostaining for tdTomato and PyMT and E-Cad on lung sections.
(P) Cartoon image showing Kit$^+$ cells in mammary tumor growth and subsequent metastasis to lung.
Scale bars, yellow, 2 mm; white, 100 μm.

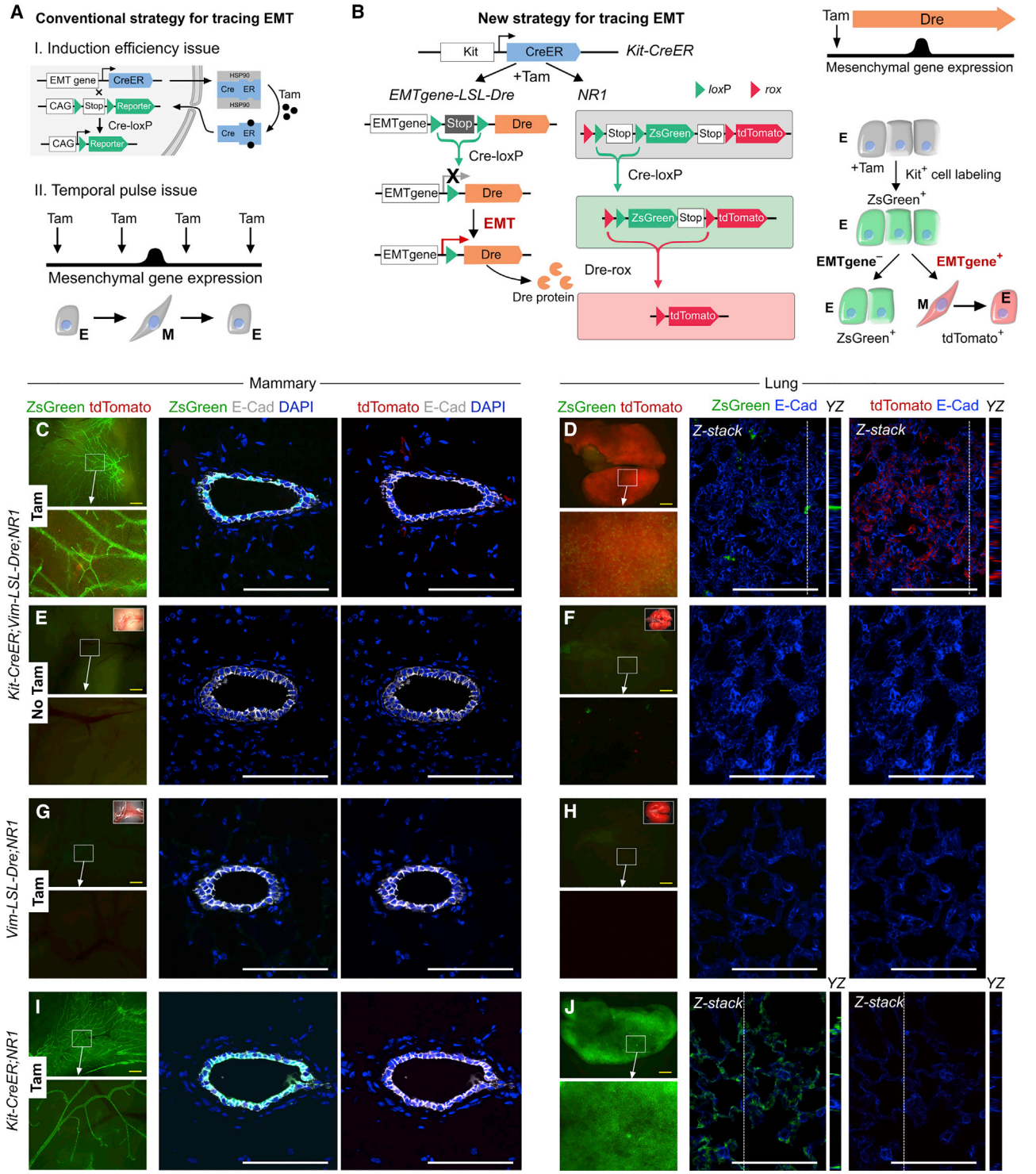


Figure 2. Generation of Dual Recombinases-Mediated Tracing System for Temporally Seamless Recording of Transient EMT Status

(A) Schematic showing strategy for Cre-loxP mediated conventional lineage tracing of transient EMT status. Labeling efficiency depends on inducible CreER activity; temporal activation of CreER by pulse induction of tamoxifen may not capture transient mesenchymal gene activity.

(B) Schematic showing strategy for Cre-loxP mediated Dre generation and subsequent Dre-rox recombination that switches ZsGreen to tdTomato on *NR1* reporter (left). After pulse Tam induction, constitutive Dre recombinase genotype is generated and driven by EMT gene promoter to monitor transient EMT gene activity (right top). After Tam, the system would switch ZsGreen to tdTomato labeling on Kit^{+} cells that have expressed EMT marker gene (Dre recombinase) (right bottom).

(legend continued on next page)

(Figure 2B). After tamoxifen treatment, *Kit-CreER* efficiently and specifically labeled luminal epithelial cells with ZsGreen by *NR1* in mammary glands (Figures 2B and S1). Meanwhile, *Kit-CreER* also removed the *Stop* in the *EMTgene-LSL-Dre* allele, placing the constitutively active Dre directly under the control of the EMT gene promoter (i.e., *EMTgene-Dre*, Figure 2B). Thereafter, Dre would be expressed when the EMT gene was activated, subsequently switching the labeling from ZsGreen to tdTomato (Figure 2B). Thus, after removal of loxP-flanked Stop cassette, the activity of the constitutively active Dre would no longer depend on tamoxifen treatment, and *EMTgene-Dre* could seamlessly monitor and record EMT gene activity, even if the EMT gene activity was transient, with the permanent genetic marker tdTomato after Dre-rox recombination (Figure 2B).

Generation and Characterization of *Vim-EMTracer*

We first established a genetic lineage tracing system for EMT characterized by the activation of vimentin (hereafter abbreviated as Vim$^+$ EMT). We generated a *Vim-CreER* knock-in mouse strain, and by labeling Vim$^+$ cells, we found that Vim was actively expressed in both fibroblasts (mesenchymal cell lineage) and endothelial cells in multiple tissues, such as the heart valve, breast, and lung (Figures S2A–S2C), which was consistent with previous reports (Kalluri and Zeisberg, 2006; Franke et al., 1979). In the mammary and lung, Vim was active mainly in CDH5$^+$ endothelial cells, and a subset of PDGFRa$^+$ fibroblasts, but not in E-Cad$^+$ epithelial cells (Figures S2D–S2F). We then generated a *Vim-LSL-Dre* knock-in mouse allele (Figure S3A) and then crossed this strain with *Kit-CreER;NR1* mice to make *Kit-CreER;Vim-LSL-Dre;NR1* triple knock-in mice to trace Vim$^+$ EMT in Kit$^+$ mammary epithelial cells (hereafter referred to as *Vim-EMTracer*, Figure S3B). As internal positive control, mesenchymal cells in the endocardial cushion were tdTomato$^+$ cells, as they were mainly derived from endocardial cells after EMT in the early developing heart (Figure S3C). In the normal mammary tissues of the adult *Vim-EMTracer* mice, 0% and 75.42% ± 3.19% of mammary luminal epithelial cells were tdTomato$^+$ and ZsGreen$^+$, respectively (Figures 2C and S3D), demonstrating that there was no Vim$^+$ EMT occurring in normal Kit$^+$ mammary epithelial cells during homeostasis. As positive control, TGF-beta addition induced a substantial number of ZsGreen$^+$ mammary tumor cells to switch into tdTomato$^+$ (Figure S3E), indicating these ZsGreen$^+$ epithelial cells were responsive to some EMT inducers. In the lungs of the same mouse, CDH5$^+$ or PECAM$^+$ endothelial cells, but not E-Cad$^+$ epithelial cells or PDGFRa$^+$ mesenchymal cells, were efficiently labeled with tdTomato by *Vim-EMTracer* (Figures 2D and S3F), indicating that there was vimentin expression in a subset of lung endothelial cells that also expressed Kit (Liu et al., 2015).

As technical controls, we performed the following three experiments. First, in *Vim-EMTracer* mice without tamoxifen induction (No Tam), no detectable fluorescence signal was found in the mammary or lung tissues (Figures 2E and 2F), indicating that there was no leakiness of *Vim-EMTracer*. Second, in *Vim-LSL-Dre;NR1* mice treated with tamoxifen, no ZsGreen or tdTomato signal was detected in the mammary or lung tissues (Figures 2G and 2H), indicating that Dre was tightly controlled by CreER-mediated recombination. Third, in *Kit-CreER;NR1* mice induced with tamoxifen, mammary luminal epithelial cells were ZsGreen$^+$tdTomato$^-$ (Figure 2I), demonstrating that Cre targeted loxP sites but not rox sites in *NR1* and confirming that Cre-loxP and Dre-rox were orthogonal systems for recombination (Anastassiadis et al., 2009; Zhang et al., 2016). In the lungs of the same mouse, Kit$^+$ endothelial cells but not E-Cad$^+$ epithelial cells were ZsGreen$^+$tdTomato$^-$ (Figure 2J). These data confirmed the establishment of *Vim-EMTracer* mice and demonstrated that there was no vimentin activity in Kit$^+$ mammary epithelial cells in normal mammary tissues.

Assessment of Vim$^+$ EMT in Tumor Growth and Metastasis

To test if a Vim$^+$ EMT program was activated during tumor growth and metastasis, we crossed *Vim-EMTracer* mice with *MMTV-PyMT* mice to generate *Vim-EMTracer;MMTV-PyMT* mice. Tamoxifen induction was performed when these mice were 6–7 weeks old, and tissue samples were collected at early (8–12 weeks old) or late (18–24 weeks old) stages of tumor formation (Figure 3A). We could readily detect ZsGreen$^+$tdTomato$^-$ tumor nodules in the mammary tissues but not in the lung metastases in the early stage of tumor development (Figure 3B). Immunostaining for PyMT, ZsGreen, and tdTomato on mammary sections from early stage of tumor formation showed that 76.22% ± 3.54% of PyMT$^+$ tumor cells were ZsGreen$^+$, and only 0.11% ± 0.049% of PyMT$^+$ cells were tdTomato$^+$ (Figures 3C and 3D), indicating that very few, if any, tumor cells expressed vimentin in the early stage of tumor progression.

In the late stage of tumor formation, mammary primary tumors exhibited robust ZsGreen and tdTomato fluorescence (Figure 3E). Immunostaining for PyMT, ZsGreen, and tdTomato in mammary sections showed that 74.42% ± 3.40% of PyMT$^+$ cells were ZsGreen$^+$, and 2.06% ± 0.48% of PyMT$^+$ cells were tdTomato$^+$ (Figures 3F and 3G), indicating that a subset of tumor cells exhibited Vim$^+$ EMT during tumor growth. Immunostaining for tdTomato, ZsGreen, and Vimentin on mammary tumors showed that as subset of tdTomato$^+$ exhibited spindle shape and expressed Vimentin in the tumor (Figure 3H). In the lung, we could readily detect ZsGreen$^+$ nodules that ranged from a small (<100 μm) to a large (>0.5 mm) size (Figure 3I). During metastasis, we focused on the Kit-derived fluorescent tumor nodules in the lung for the subsequent analysis of potential Vim$^+$ EMT. In small metastasis, all PyMT$^+$ cells were ZsGreen$^+$, and no tdTomato$^+$ tumor cells were found (Figures 3J and 3K).

(C, E, G, and I) Whole-mount fluorescence images of mammary and immunostaining for ZsGreen, tdTomato, and E-Cad on mammary gland sections from 20-week adult mouse. *EMTracer* mouse with tamoxifen treatment (C) and without tamoxifen treatment (E), *Vim-LSL-Dre;NR1* mouse (G) and *Kit CreER;NR1* mouse (I) with tamoxifen treatment. Tamoxifen or oil (No Tam) was induced at 6–7 weeks.

(D, F, H, and J) Whole-mount fluorescence images of lung and immunostaining for ZsGreen, tdTomato, and E-Cad on lung sections from same mice. *EMTracer* mouse with tamoxifen treatment (C) and without tamoxifen treatment (E), *Vim-LSL-Dre;NR1* mouse (G) and *Kit CreER;NR1* mouse (I) with tamoxifen treatment. *YZ* indicates signals from dotted lines on z stack images in (D and J).

Scale bars, yellow, 2 mm; white, 100 μm.

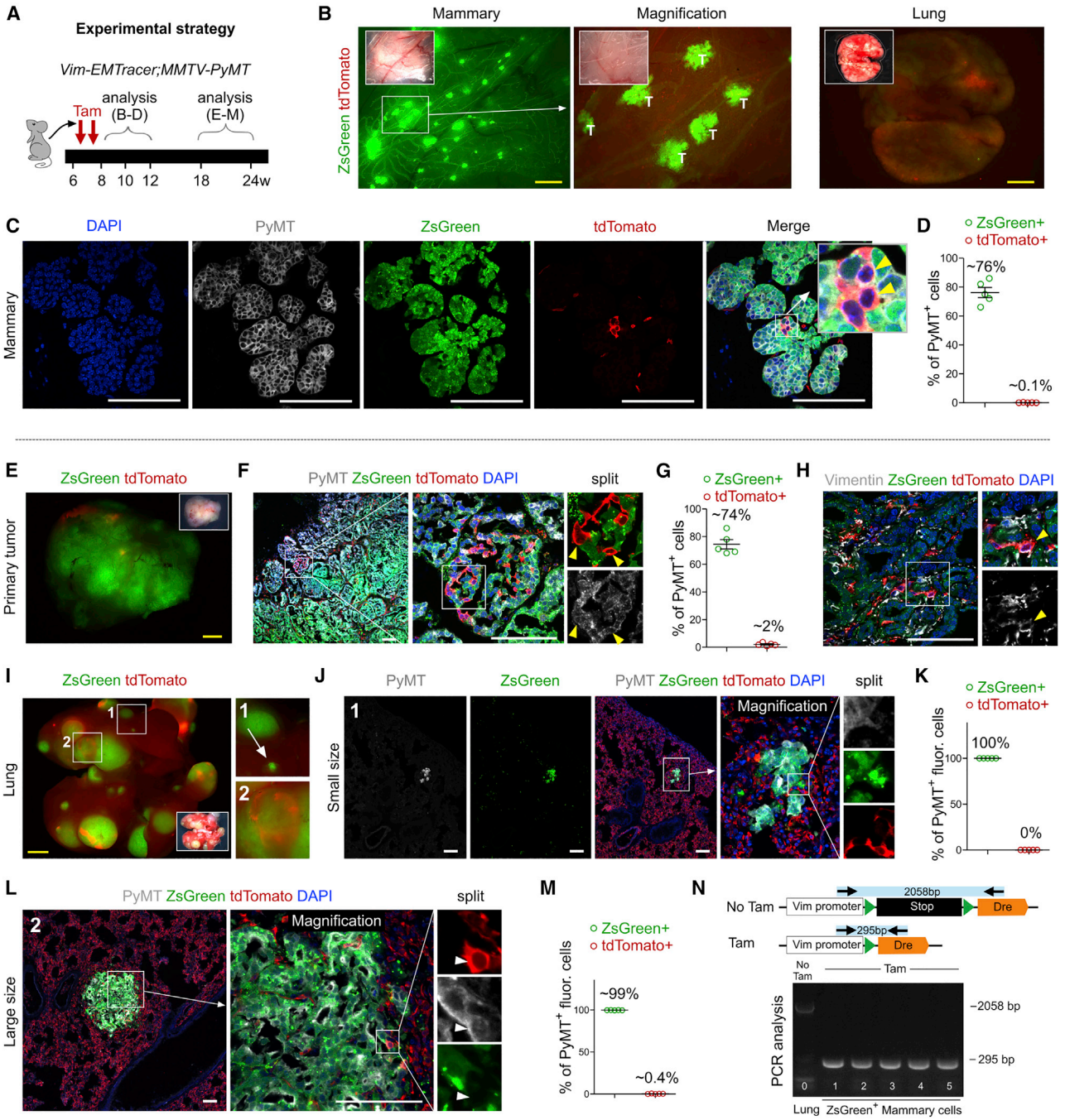


Figure 3. Analysis of Vim⁺ EMT in Primary Tumor and Lung Metastasis

(A) Schematic figure showing experimental strategy. Mammary and lung tissues were collected for analysis at early stage (B–D) and late stage (E–M).

(B) Whole-mount fluorescence image of mammary tissue and lung. T, tumor nodule.

(C) Immunostaining for PyMT, ZsGreen and tdTomato on sections from mammary tissue. Arrowheads indicate tdTomato⁺PyMT⁺ tumor cells.

(D) Quantification of percentage of ZsGreen⁺ or tdTomato⁺ cells in PyMT⁺ tumor cells. Data are mean ± SEM; n = 5.

(E) Whole-mount fluorescence image of primary tumor at late stage.

(F) Immunostaining for PyMT, ZsGreen and tdTomato on mammary tumor sections. Arrowheads, PyMT⁺tdTomato⁺ tumor cells

(G) Quantification of percentage of ZsGreen⁺ or tdTomato⁺ cells in PyMT⁺ tumor cells. Data are mean ± SEM; n = 5.

(H) Immunostaining for vimentin, ZsGreen, and tdTomato on mammary tumor sections. Arrowheads, Vimentin⁺tdTomato⁺ tumor cells.

(I) Whole-mount fluorescence image of lung at late stage. Small size lung metastasis (1, <0.1 mm, arrow) and large size lung metastasis (2, >0.5 mm).

(J and L) Immunostaining for PyMT, ZsGreen, and tdTomato on lung sections. In small size nodules, PyMT⁺ tumor cells are ZsGreen⁺tdTomato⁻ (J). In large size nodule, most PyMT⁺ tumor cells are ZsGreen⁺tdTomato⁻, while a minority of PyMT⁺ cells are ZsGreen⁻tdTomato⁺ (arrowheads, L).

(legend continued on next page)

In large metastasis, while the majority of PyMT$^+$ cells were ZsGreen$^+$, a very small number of PyMT$^+$ cells in the periphery of nodules were tdTomato$^+$Zsgreen$^-$ (Figure 3L). These tdTomato$^+$ tumor cells constituted 0.42% ± 0.083% of PyMT$^+$ cells in the lung metastases (Figure 3M), indicating that very few, if any, tumor cells exhibited vimentin activity after tumor colonization.

To demonstrate that these tdTomato$^+$ cells had undergone Vim$^+$ EMT during tumor growth and expansion after but not during metastatic colonization, we also collected lungs before the mice were 18 weeks old, when there were only small metastatic colonies (Figure S4A). We found that all tumor metastatic cells in the lung were ZsGreen$^+$tdTomato$^-$ (Figures S4B–S4E). Given that the tumor cells in small nodules were tdTomato$^-$ before they grew into larger nodules, and considering that very few tdTomato$^+$ tumor cells were detected in only the large nodules, usually in the periphery of the nodule (Figures 3L and 3M), it was likely that ZsGreen$^+$ metastatic cells activated *in situ* vimentin expression following their colonization, thus, switching from ZsGreen$^+$ to tdTomato$^+$ labeling in the tumor metastases.

While Vim is not robustly activated during tumor metastasis, we asked if Vim could be functionally required for tumor metastasis. To address this possibility, we crossed *Kit-CreER;Vim-LSL-Dre;NR1;MMTV-PyMT* (heterozygous Vim control mice) with *Vim-LSL-Dre* to generate homozygous null allele for Vim. In *Kit-CreER;Vim-LSL-Dre*; *Vim-LSL-Dre;NR1;MMTV-PyMT* (homozygous Vim mutant mice), we could not detect any Vim expression in mammary tumors collected from Vim null mice, while Vim was readily detected in the mammary tumors collected from littermate control mice (Figure 4A). Similarly, we treated mutant mice with tamoxifen at 6 weeks, and collected mammary tissues at 18–24 weeks (Figure 4B), and could observe formation of large mammary tumors exhibiting ZsGreen and tdTomato (Figure 4C). Immunostaining for ZsGreen, tdTomato, and PyMT on mutant tumor sections revealed that a subset of PyMT$^+$ tumor cells had switched from ZsGreen to tdTomato in the primary tumors (Figure 4D). In the lung, many small and large tumor metastases were detected, and virtually all of them were ZsGreen$^+$ (Figure 4E), indicating that lack of Vim did not significantly inhibit tumor metastasis, nor was the Vim gene promoter activated during tumor metastasis. Similar to the control mice, a subset of tdTomato$^+$PyMT$^+$ cells was detected in the peripheral of some tumor colonies in the lung (Figure 4F), indicating Vim gene promoter was activated during the growth and expansion of metastases.

The lack of detectable vimentin activity during the metastasis process was likely not due to inefficient *Stop* removal from the *Vim-LSL-Dre* allele by *Kit-CreER*, as vimentin activity could be readily detected in Kit$^+$ endothelial cells (Liu et al., 2015) throughout the lung (tdTomato$^+$, Figures 3J and 3L), demonstrating that the *Stop* was efficiently removed from the *Vim-LSL-Dre* allele by *Kit-CreER* driver. In addition, we performed PCR analysis of the *Stop* DNA cassette in ZsGreen$^+$ cells isolated from mammary tumors. ZsGreen$^+$ cells had *Stop*-removed DNA signatures in the *Vim-LSL-Dre* allele (Figure 3N), indicating that the *Kit-CreER*-mediated *Stop* removal in the *Vim-LSL-Dre* allele was efficient in tumor cells. Furthermore, all the different types of internal technical controls confirmed the fidelity of the EMT genetic tracing in this tumor model (Figure S5). Taken together, these data demonstrated that there was a lack of Vim activation during the progression to tumor metastasis, indicating that breast tumors utilize Vim-independent EMT or an EMT-independent mechanism to metastasize.

Generation and Characterization of *N-Cad-EMTracer*

We next employed another commonly used EMT marker, N-cadherin, to generate a lineage tracing system for EMT characterized by the activation of N-cadherin (N-cad$^+$ EMT) during tumor growth and metastasis. We generated the *N-cad-LSL-Dre* knock-in mouse strain (Figure S6A) and crossed it with *Kit-CreER;NR1* mice to make the *Kit-CreER;N-cad-LSL-Dre;NR1* triple knock-in mouse for the lineage tracing of N-cad$^+$ EMT during tumor growth and metastasis (*N-cad-EMTracer*, Figure S6B). As internal control, we could detect tdTomato$^+$ mesenchymal cells in the endocardial cushion of developing heart (Figure S6C), demonstrating that endocardial EMT could be genetically traced by *N-cad-EMTracer*. Whole-mount fluorescence and immunostaining of mammary tissues derived from *N-cad-EMTracer* mice showed that the mammary epithelial cells were ZsGreen^{+-} tdTomato$^-$ (Figures 5A, 5B and S6D–S6G), indicating that there was no N-cadherin activity in Kit$^+$ mammary epithelial cells in the normal mammary gland during homeostasis. Moreover, no tdTomato$^+$ cells were detected in the lung (Figure 5A). Immunostaining for ZsGreen, tdTomato, and E-Cad on lung sections verified that there were no tdTomato$^+$ epithelial cells in the lung (Figure 5B), demonstrating that the N-cad$^+$ EMT program was not active in normal mammary and lung tissues.

Activity and Requirement of N-cad$^+$ EMT for Tumor Metastasis

We then crossed *N-cad-EMTracer* mice with *MMTV-PyMT* mice to generate *N-cad-EMTracer;MMTV-PyMT* mice to study N-cad$^+$ EMT in a tumor model. After tamoxifen induction when the mice were 6–7 weeks old, we collected mammary tissues at early (8–12 weeks old) or late (18–24 weeks old) stages of tumor formation for analysis. ZsGreen$^+$tdTomato$^-$ tumor nodules were readily detected at the early stage of tumor development in the mammary tissues (Figure 5C). Immunostaining for PyMT, ZsGreen, and tdTomato on mammary sections showed that 77.67% ± 2.63% and 0.22% ± 0.035% of PyMT$^+$ cells in hyperplastic lesions were ZsGreen$^+$ and tdTomato$^+$, respectively, while there were no tdTomato$^+$ epithelial cells in the morphologically normal glands of the same mammary tissues (Figures 5D and 5E). These tdTomato$^+$ tumor cells still expressed epithelial cell marker E-Cad (Figure 5F). Of note, N-cad could be detected in subset of tdTomato$^+$ tumor cells (Figure 5G). At the late stage

(K and M) Quantification of the percentage of ZsGreen$^+$ or tdTomato$^+$ cells in PyMT$^+$ fluorescent tumor cells in small (K) or large (M) size lung metastases. Data are mean ± SEM; n = 5.

(N) PCR analysis of genomic DNA from ZsGreen$^+$ cells of tumor. Samples 1–5 were from mice with Tam treatment; Sample 0 was lung tissue from mice without tamoxifen treatment.

Scale bars, yellow, 2 mm; white, 100 μm.

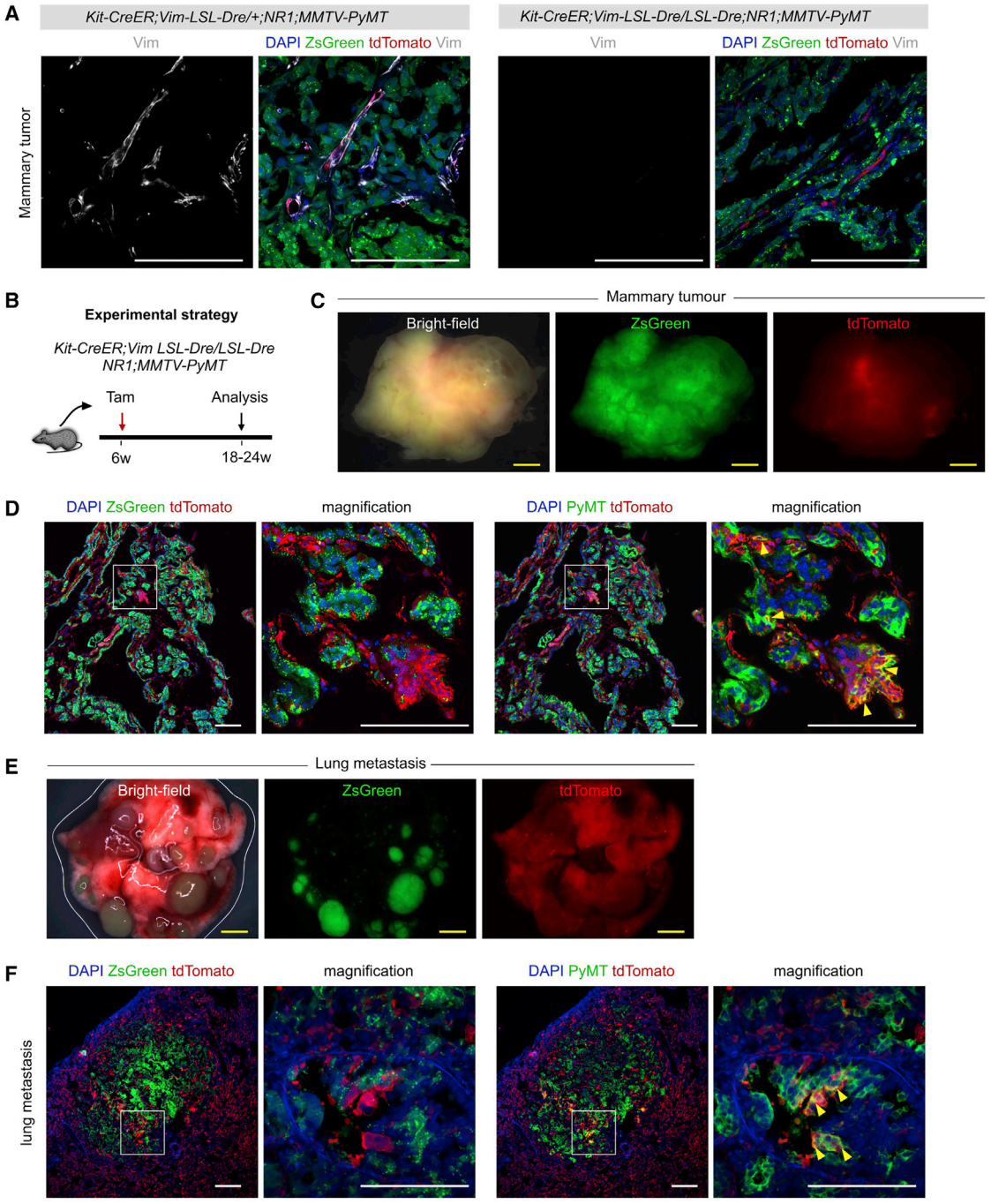


Figure 4. Analysis of Vim$^+$ EMT in Primary Tumor and Lung Metastasis of Mice with Functional Loss of Vim

(A) Immunostaining for Vim, ZsGreen, and tdTomato on mammary tumor sections.

(B) Schematic figure showing experimental strategy for analysis of mammary tissues and lungs from Vim knockout mice.

(C) Whole-mount fluorescence and bright-field images of mammary tumors.

(D) Immunostaining for ZsGreen, tdTomato, and PyMT on mammary tumor sections. Yellow arrowheads, PyMT$^+$tdTomato$^+$ cells.

(E) Whole-mount fluorescence images of lungs with tumor metastases. Most tumor metastases are ZsGreen$^+$ but not tdTomato$^+$.

(F) Immunostaining for ZsGreen, tdTomato, and PyMT on lung sections. Yellow arrowheads indicate very few PyMT$^+$tdTomato$^+$ tumor cells in the large tumor metastases in the lung.

Scale bars, yellow, 2 mm; white, 100 μm.

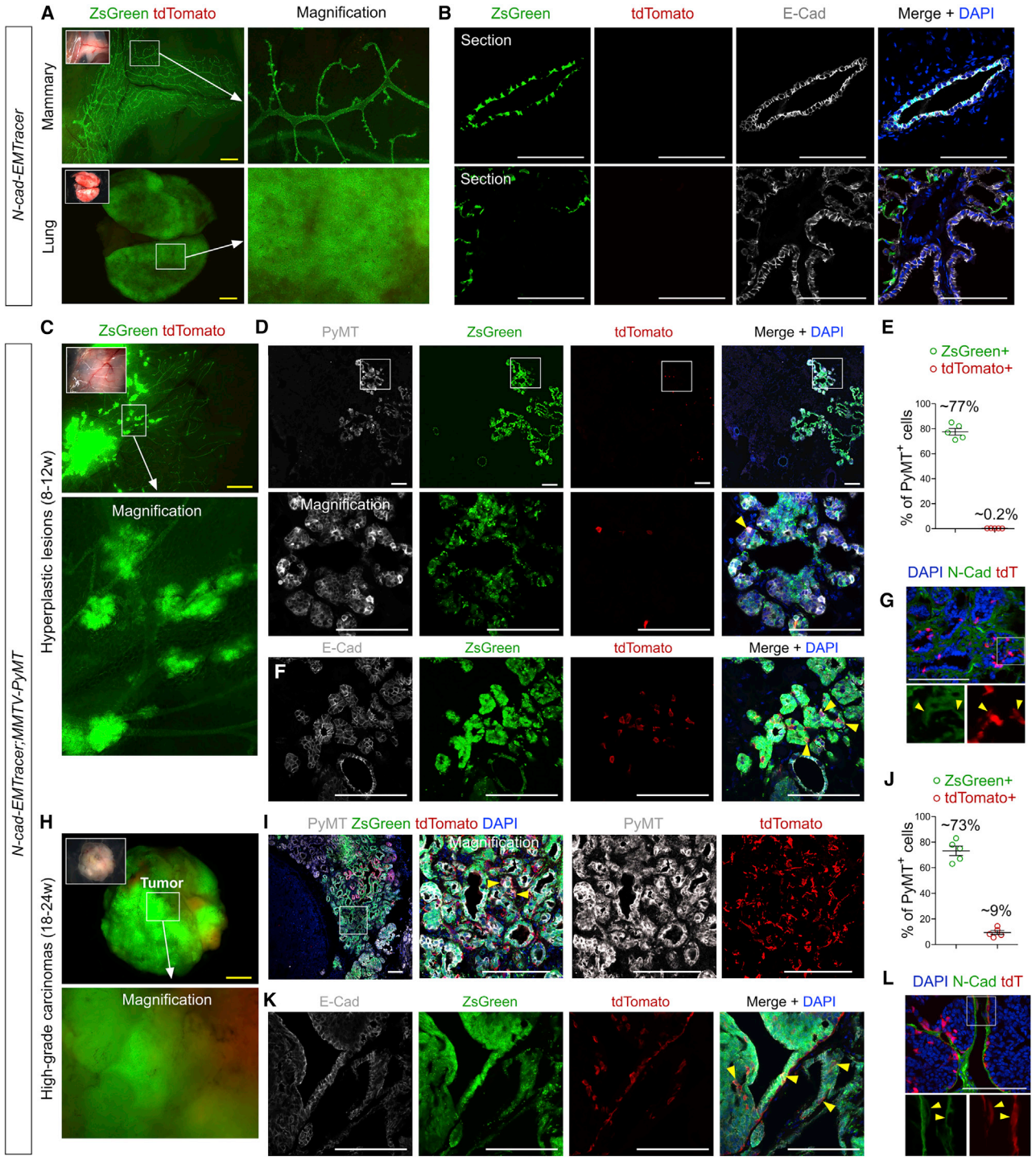


Figure 5. N-cad⁺ EMT Program in Primary Mammary Tumor

(A) Whole-mount fluorescence images of mammary and lung of *N-cad-EMTracer* mouse. Inserts are bright-field images. Tamoxifen was induced at 7 weeks and tissues were collected at 18 weeks.

(B) Immunostaining for ZsGreen, tdTomato, and E-Cad on mammary and lung sections of *N-cad-EMTracer* mouse. No tdTomato⁺ cells were detected in mammary and lung epithelial cells.

(C and H) Whole-mount fluorescence images of mammary tissue collected at early (8–12 weeks) (C) or late (18–24 weeks) (H) stages from *N-cad-EMTracer;MMTV-PyMT* mouse. Inserts are bright-field images.

(legend continued on next page)

of tumor formation, when high-grade carcinomas had formed, the tumor nodules were ZsGreen$^+$ and tdTomato$^+$ (Figure 5H). Immunostaining for PyMT, ZsGreen, and tdTomato in mammary sections showed that 73.38% ± 3.76% and 9.46% ± 1.57% of PyMT$^+$ cells were ZsGreen$^+$ and tdTomato$^+$, respectively (Figures 5I and 5J). These tdTomato$^+$ tumor cells also expressed epithelial cell marker E-Cad (Figure 5K), and a subset of tdTomato$^+$ tumor cells expressed N-cad (Figure 5L). The appearance of tdTomato$^+$ tumor cells at the early stage of tumor formation and the increase in their number of these cells at the late stage demonstrated that the N-cad$^+$ EMT program occurred and increased in frequency over a wide time span during primary tumor growth and expansion based on a genetic reporter (tdTomato).

To study tumor metastasis, we collected the lungs of *N-cad-EMTracer;MMTV-PyMT* mice at the late stages of tumor development to examine N-cad$^+$ EMT (Figure 6A). Whole-mount fluorescence imaging of the lung showed that most tumors were tdTomato$^+$, while a subset of tumors were ZsGreen$^+$ (Figure 6B). These data indicated that most lung metastases (tdTomato$^+$) had activated EMT gene N-cad, while ZsGreen$^+$ tumor cells either employed N-cad-independent EMT or an EMT-independent mechanism to metastasize. These tdTomato$^+$ cells were derived from lung metastases labeled by the EMT tracing system, as the lungs collected from *N-cad-EMTracer* siblings remained tumor free and tdTomato$^-$ with the same tamoxifen treatment regime (Figure 6C). In the lung tissues, we focused on Kit-derived fluorescent tumor metastases for the subsequent analysis of N-cad$^+$ EMT. Staining for PyMT, ZsGreen, and tdTomato in lung sections showed ZsGreen$^+$tdTomato$^-$, tdTomato$^+$ZsGreen$^-$, or ZsGreen$^+$tdTomato$^+$ hybrid clones that constituted 21.69% ± 1.72%, 66.28% ± 2.81%, and 12.03% ± 2.61% of the fluorescent tumor colonies in the lung, respectively (Figures 6D and 6E). We also found that the majority of small tumor colonies in the lung were tdTomato$^+$ZsGreen$^-$ (Figure 6F), demonstrating that tdTomato$^+$ tumor cells were the primary components of lung metastases during the initial colonization. Of note, we did not detect N-cad expression in the tdTomato$^+$ tumor cells colonizing the lung (Figure 6G), suggesting N-cad was less likely to be activated during colonization. Internal technical controls validated the fidelity of the EMT fate-mapping results (Figure S7). Taken together, these data indicated that there was N-cadherin activity and N-cad$^+$ EMT during tumor growth and metastasis.

To further understand the function of N-cadherin during lung metastasis, we specifically knocked out the N-cadherin gene in Kit$^+$ cell lineages by crossing *Kit-CreER;N-cad-LSL-Dre;NR1;MMTV-PyMT* mice with mice with a floxed N-cad allele, *N-cad-flox* (Kostetskii et al., 2005; Figure 6H). The N-cad allele would be deleted in Kit$^+$ cells in *Kit-CreER;N-cad-LSL-Dre;N-cad-flox;NR1;MMTV-PyMT* mice (mutant). Western immunoblotting data showed that the N-cadherin protein was significantly reduced in the primary tumors of mutant mice compared with that in littermate control mice *Kit-CreER;MMTV-PyMT* (Figure 6I). We could readily detect primary tumor formation in the mammary tissues of N-cad mutant mice (Figure 6J). Notably, the number of lung metastases in mutant mice was significantly lower than that in control mice (Figures 6K and 6L). While the majority of tumor metastases in lung were tdTomato in the control mice, the very few tumor metastases in mutant mice were ZsGreen$^+$ but not tdTomato$^+$ in the mutant (Figure 6K), which was confirmed by sectional immunostaining on lung tissues (Figure 6M). In the mutant mice, while a subset of tumor cells expressed tdTomato in the primary tumors, cancer cells in the lung metastases were ZsGreen$^+$ (Figure 6N). These tdTomato$^+$ tumor cells in the mammary tumor also expressed epithelial cell marker E-Cad (Figure 6O). Taken together, the EMT gene N-cad was involved during tumor metastasis, and its activity played a critical role in tumor metastasis.

Finally, we isolated tdTomato$^+$ tumor cells from mammary tissues by FACS and compared the EMT genes with ZsGreen$^+$ tumor cells by qRT-PCR analysis (Figure 7A). While epithelial cell marker E-Cad was expressed in tdTomato tumor cells, their expression level was significantly lower than that of ZsGreen$^+$ tumor cells (Figure 7B). Notably, *N-cad* and *Vim* were significantly higher in tdTomato$^+$ tumor cells than ZsGreen$^+$ tumor cells (Figure 7B). Furthermore, expression of EMT transcriptional factors *Zeb1*, *Twist*, and *Snail* was also significantly higher in tdTomato$^+$ cells than those of ZsGreen$^+$ tumor cells (Figure 7B). A transwell assay revealed increased invasion of isolated tdTomato$^+$ tumor cells compared with ZsGreen$^+$ tumor cells (Figure 7C). Additionally, immunostaining for Zeb1 and tdTomato on mammary tumor sections revealed that a subset of tdTomato$^+$ tumor cells expressed Zeb1 (Figure 7D). To examine if N-cad gene was activated before extravasation or during colonization into lung, we used a microwell chip platform to examine the circulating tumor cells of *N-cad-EMTracer;MMTV-PyMT* mice. For tumor cells in circulation, we could find some tdTomato$^+$ EpCAM$^+$ cells, while ZsGreen$^+$ cells were largely non-tumor cells, as they were hematopoietic cells labeled by *Kit-CreER* (Figures 7E–7G). We also transplanted the isolated ZsGreen$^+$ cells into littermate with no fluorescence reporter and examined whether the colonization into lung would activate N-cad gene (Figure 7H). We found that these lung metastases were ZsGreen$^+$ (Figure 7I), indicating that EMT gene N-cad was less likely to be activated during extravasation and lung colonization.

DISCUSSION

Transient cell fate switch is manifested by transient gene expression (Süel et al., 2006; Maamar et al., 2007; Weinberger et al., 2008). Genetic tracing of transient gene expression *in vivo* remains technically challenging. The gene of interest could often be transiently activated in the embryonic stage or earlier stages

(D and I) Immunostaining for ZsGreen, tdTomato, and PyMT on mammary tissue sections collected at early (D) or late (I) stage. Yellow arrowheads, tdTomato$^+$PyMT$^+$ cells.
(E and J) Quantification of the percentage of ZsGreen$^+$ and tdTomato$^+$ cells among PyMT$^+$ cells of early (E) or late (J) stage. Data are mean ± SEM; n = 5.
(F and K) Immunostaining for tdTomato, ZsGreen, and E-Cad on tissue sections of early (F) or late (K) stage. Yellow arrowheads, tdTomato$^+$E-Cad$^+$ cells.
(G and L) Immunostaining for N-Cad and tdTomato (tdT) on tissue sections of early (G) or late (L) stage. Yellow arrowheads, N-Cad$^+$tdT$^+$ cells.
Scale bars, yellow, 2 mm; white, 100 μm.

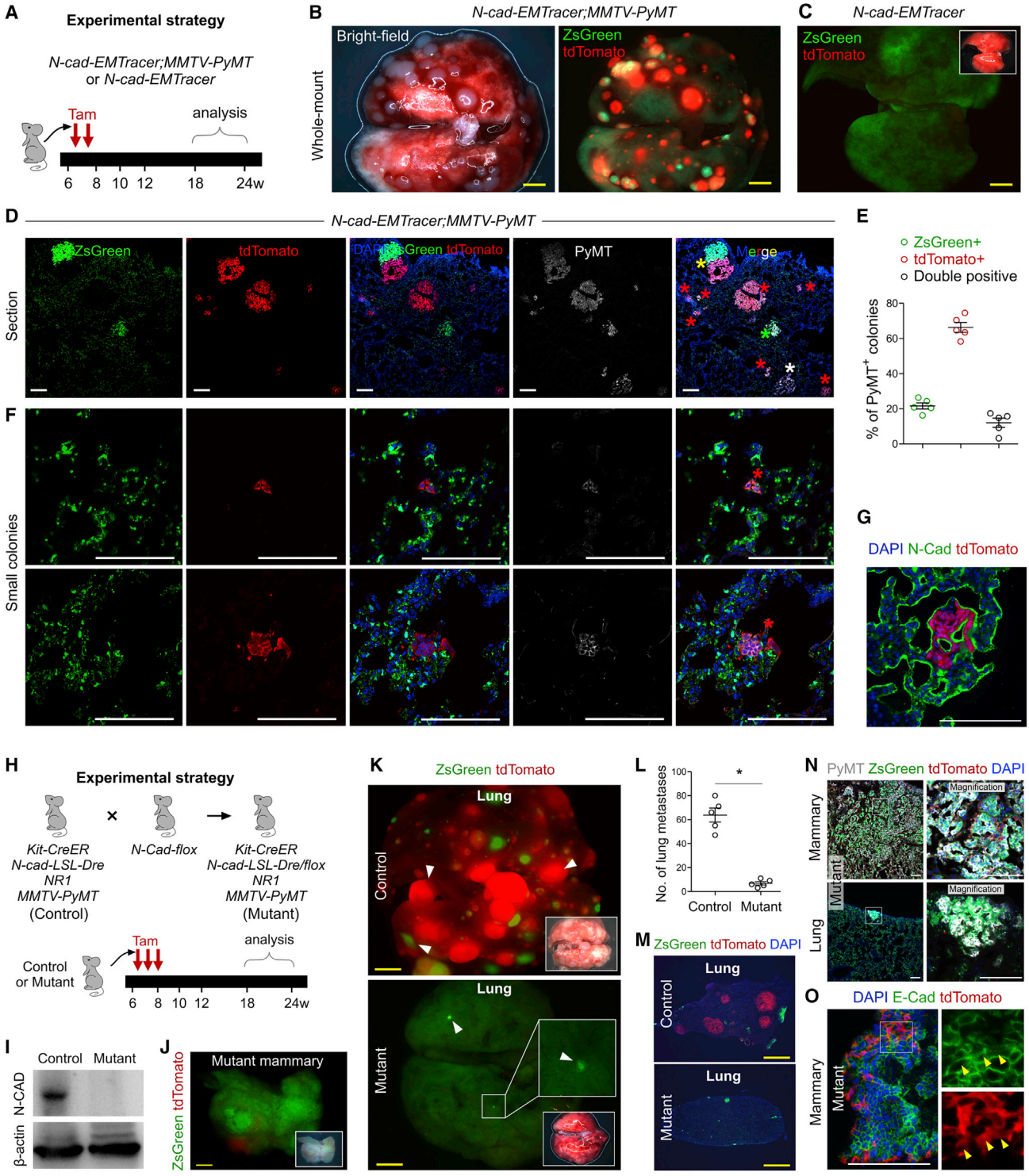


Figure 6. N-cad$^+$ EMT Activity Is Detected in Lung Metastasis with Functional Requirement of N-cad

(A) Schematic figure showing experimental strategy.

(B and C) Whole-mount fluorescence images of lungs collected from *N-cad-EMTracer;MMTV-PyMT* (B) and *N-cad-EMTracer* (C) mice.

(D and F) Immunostaining for ZsGreen, tdTomato and PyMT on lung sections shows ZsGreen$^+$ colony (green asterisk), tdTomato$^+$ colony (red asterisk), ZsGreen$^+$tdTomato$^+$ colony (yellow asterisk), and ZsGreen$^-$tdTomato$^-$ colony (white asterisk). Large size lung metastasis (D) and small size lung metastasis (F).

(E) Quantification of the percentage of ZsGreen$^+$tdTomato$^-$, ZsGreen$^-$tdTomato$^+$, or ZsGreen$^+$tdTomato$^+$ colony number among all fluorescent colonies.

(G) Immunostaining for tdTomato and N-Cad on lung sections.

(H) Schematic figure showing experimental strategy for knockout N-cadherin allele by *Kit-CreER* in tumor model.

(legend continued on next page)

before the preferred experiment time point. Constitutive recombinase directly driven by one gene promoter could seamlessly trace this gene expression from earliest stages, such as embryonic stage, and this loses the resolution to distinguish the exact time window from other stages when this gene is activated. The induced seamlessly genetic tracing system reported here not only maintains the advantage of temporal control of initial recombination by tamoxifen-induced Cre, but also relays this inducible recombination to prime constitutive Dre system for potential secondary recombination, based on the transient gene expression related to specific cell fate transition. This genetic system permits seamless fate mapping of transient cell fate transitions and could be broadly used in the fields of development, diseases, and tissue regeneration.

In this study, we found the *Kit-CreER* targeted the majority of PyMT$^+$ tumor cells in the mammary tissue. We also noticed that a small number of PyMT$^+$ tumor cells remained tdTomato$^-$. This could be due to the labeling efficiency of *Kit-CreER*, as it may not label all Kit$^+$ luminal epithelial cells. An alternative explanation is that some Kit$^-$ mammary epithelial cells may also contribute to PyMT$^+$ tumor cells. As the majority of tumor cells and their metastases were genetically labeled and traced by our *Kit-CreER*, the event of EMT gene activity could be investigated in these *Kit-CreER* targeted tumor cells during tumor growth and metastasis.

Tumor cells may transiently activate an EMT program to acquire a mesenchymal phenotype (Thiery et al., 2009; Nieto et al., 2016). The reversible activation of EMT also suggests that there is temporally transient expression of EMT genes during tumor growth, metastasis, and disease progression (Yu et al., 2013). This study provides genetic evidence of the expression of EMT gene during breast tumor metastasis and indicates that different EMT gene activity or programs can be used in different tumor types during cancer cell dissemination. In this study, we first generated a genetic system to seamlessly trace gene activity *in vivo*. By using this system, we explored EMT gene activity in tumor growth and metastasis. In the *MMTV-PyMT* mammary tumor model, we found that the EMT gene N-cad was activated in and required for tumor metastasis, while the gene activity of vimentin was not detected during tumor metastasis. Our results suggest that tumor cells may utilize a distinct EMT gene program to metastasize and colonize distant organs. Of note, N-cad gene deletion significantly reduced tumor metastasis in the *MMTV-PyMT* model. In many types of solid tumors, the aberrant expression of N-cadherin is a hallmark of EMT, resulting in the acquisition of an aggressive tumor phenotype (Gravdal et al., 2007; Araki et al., 2011). N-cadherin becomes upregulated in invasive tumor cells and induces heightened motility, invasion, and metastasis of breast cancer cells (Nieman et al., 1999; Hazan et al., 2004). In *MMTV-PyMT* mouse model of mammary tumorigenesis, N-cadherin enhanced metastasis by increasing ERK (extracellular signal-regulated kinase) activation and matrix metallopeptidase-9 expression, and attenuating the Akt pathway to increase cell motility (Hulit et al., 2007; Chung et al., 2013). Consistently, in our study, N-cad$^+$ tumor cells exhibited higher EMT gene expression and stronger migratory capacity, compared with N-cad$^-$ tumor cells.

A previous report using an inducible lineage tracing system for vimentin, *Vim-CreER*, did not find cells that had activated vimentin in the metastatic outgrowths and concluded that EMT was not involved in breast tumor metastasis formation (Fischer et al., 2015). These negative data could have been due to either a lack of the Vim gene activity or the insensitivity of the tamoxifen-induced Cre-loxP labeling approach for monitoring vimentin activity during tumor metastasis. In this study, the temporally seamless lineage tracing strategy showed that the Vim$^+$ EMT program was involved in *MMTV-PyMT* tumor growth but not in metastasis, representing a convincing genetic study of Vim activity. Of note, we did observe some tumor cells expressed tdTomato in the peripheral of large-size colonies in the lung, indicating activation of Vim during metastases growth and expansion. Whether this Vim activation was associated with chemo-resistance remains to be determined in future. Additionally, we also found that N-cad$^+$ tumor cells expressed higher Vim compared with N-cad$^-$ cells by qRT-PCR analysis. It is possible that N-cad$^+$ cells are a heterogenous population, and a subset of N-cad$^+$ cells that metastasize to lungs may not express vimentin. As different tumors could utilize different EMT gene programs to metastasize, it is possible that other types of tumors may utilize the Vim$^+$ EMT program for metastasis, which merits further studies using different tumor models and also adding additional EMT markers for tracing in future.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead Contact
  - Materials Availability
  - Data and Code Availability
- EXPERIMENTAL MODELS AND SUBJECT DETAILS
  - Mice Generation and Breeding
- METHOD DETAILS
  - Genomic PCR
  - Tamoxifen Treatment
  - Whole-mount Fluorescence Microscopy
  - Immunostaining and Confocal Fluorescence Microscopy
  - Mammary Cell Preparation and Flow Cytometry
  - Tumor Cells Isolation and Flow Cytometry
  - Transplantation Experiment
  - TGF-beta Induction

(I) Western immunoblotting of N-cadherin protein collected from mutant or littermate *Kit-CreER;MMTV-PyMT* control mice.
(J and K) Whole-mount fluorescence view of mammary tissues (J) or lungs (K). Insert, bright-field image. Arrowheads, fluorescent$^+$ tumors.
(L) Quantification of the lung metastases (nodules). Data are mean ± SEM; n = 5; *p < 0.05.
(M–O) Immunostaining for tdTomato and ZsGreen (M), or PyMT (N), or E-Cad (O) on lung or mammary tissue sections. Arrowheads, E-Cad$^+$tdTomato$^+$ cells. Scale bars, yellow, 2 mm; white, 100 μm.

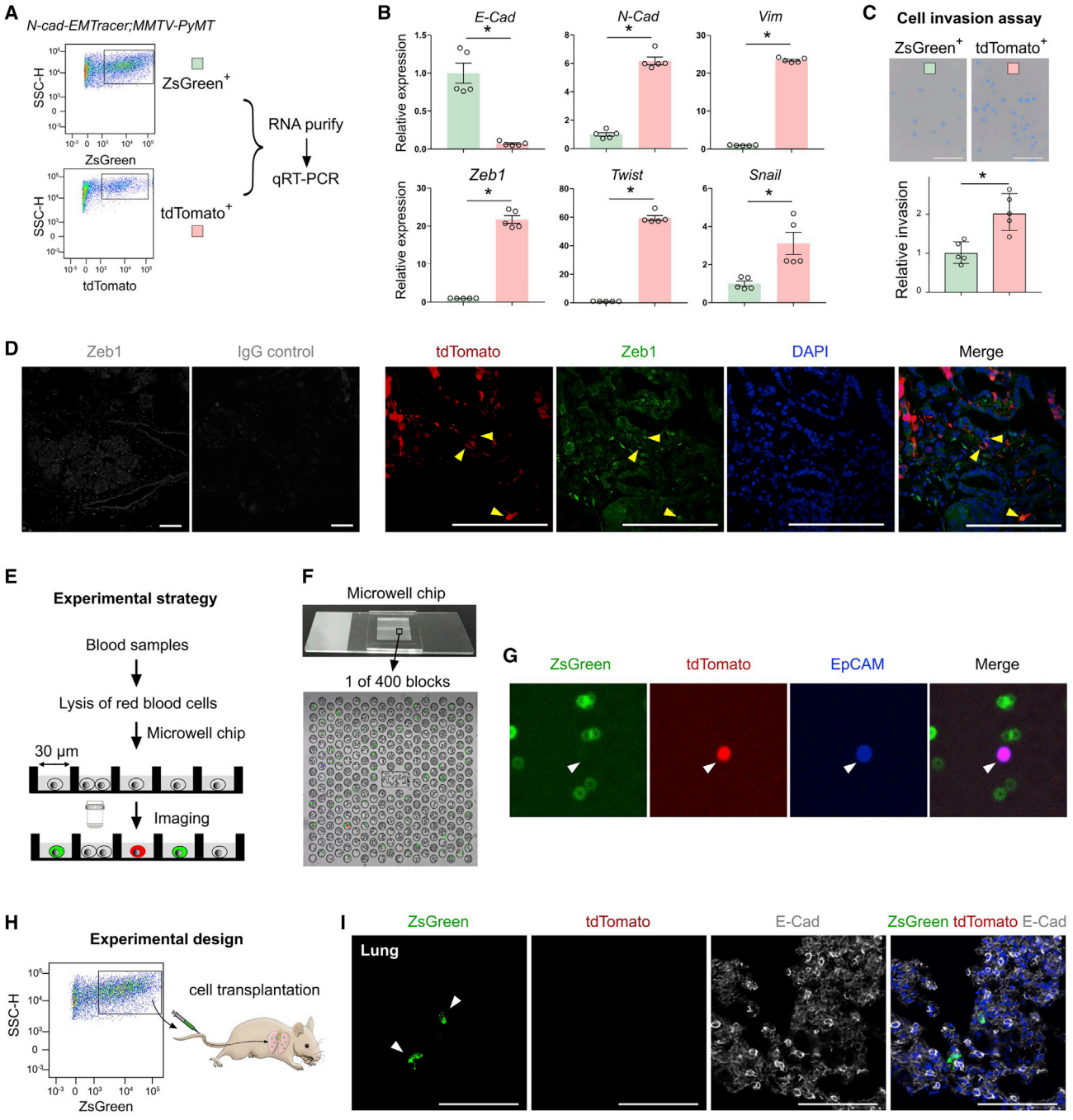


Figure 7. Gene Profile of N-Cad$^+$ EMT Cells and Analysis of Circulating Tumor Cells

(A) FACS isolation of ZsGreen$^+$ and tdTomato$^+$ cells from *N-cad-EMTracer;MMTV-PyMT* mammary tumors.

(B) qRT-PCR of RNA expression of epithelial cell marker *E-Cad*; EMT markers N-cad and Vim; EMT transcriptional factors Zeb1, Twist, and Snail in ZsGreen$^+$ cells or tdTomato$^+$ cells. Gene expression in ZsGreen$^+$ cells is set as 1. Data are mean ± SEM; n = 5; *p < 0.05.

(C) Cell invasion assay using transwell shows more invasion in tdTomato$^+$ cells than ZsGreen$^+$ cells. Number of relative cell invasion in ZsGreen$^+$ cells is set as 1. Data are mean ± SEM; n = 5; *p < 0.05.

(D) Immunostaining for Zeb1 and IgG control (left panel) shows specific Zeb1 staining on mammary tumor sections. Immunostaining for tdTomato and Zeb1 on mammary tumor sections (right panel) shows a subset of tdTomato$^+$ tumor cells expressing Zeb1 (yellow arrowheads).

(E) Schematic figure showing experimental strategy for searching circulating tdTomato$^+$ tumor cells by high-throughput screening.

(F) Image showing a microwell chip platform that accommodates 400 numbered blocks with over 100,000 addressable microwells.

(G) Fluorescence images of ZsGreen$^+$ or tdTomato$^+$ cells in microwells of the chip. Arrowhead indicates tdTomato$^+$ EpCAM$^+$circulating tumor cell that has undergone N-cad$^+$ EMT.

(legend continued on next page)

- Transwell Assay
- PCR Analysis of DNA for Recombination Efficiency
- RNA Extraction and Quantitative RT-PCR
- Western Blot
- Microwell Chip Assay

● QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.devcel.2020.06.021.

ACKNOWLEDGMENTS

We thank Shanghai Biomodel Organism Co., Ltd. for mouse generation; and Daming Gao, Jing Yang, Yibin Kang, and Robert Weinberg for valuable suggestions. This study was supported by the National Key Research & Development Program of China (grant nos. 2019YFA011040, 2019YFA080200, 2018YFA010810, 2018YFA0107900, 2017YFC1001303, and 2016YFC1300600), Strategic Priority Research Program of the Chinese Academy of Sciences (CAS, grant nos. XDB19000000 and XDA16010507), National Science Foundation of China (grant nos. 31730112, 91849202, 31625019, 31730112 and 31900625), Key Project of Frontier Sciences of CAS (grant no. QYZDB-SSW-SMC003), International Cooperation Fund of CAS, Shanghai Science and Technology Commission (grant nos. 19JC1415700, 19YF1455300, 19ZR1479800, and 20QC1401000), China Postdoctoral Science Foundation (grant nos. 2018M640430 and 2019M660100), National Postdoctoral Program for Innovative Talents (grant nos. BX20180338 and BX20190343), the Pearl River Talent Recruitment Program of Guagdong Province (grant no. 2017ZT07S347), Royal Society-Newton Advanced Fellowship, and Sanofi-SIBS Fellowship.

AUTHOR CONTRIBUTIONS

Y.L., Z.L., and B.Z. designed the study, performed experiments, and analyzed the data; S.Z., L.H., M.T., W.P., H.Z., and Z.Z. bred the mice, performed experiments, or provided valuable comments; Z.W. and Q.S. performed microwell chip analysis; J.F., D.C., and K.L. provided valuable comments and edited the manuscript; W.M and H.G provided *N-cad-flox* and *MMTV-PyMT* mouse lines. M.A.N. contributed to interpreting the data and writing the manuscript; B.Z. supervised the study, analyzed the data, and wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing financial interests.

Received: December 9, 2019
Revised: April 26, 2020
Accepted: June 16, 2020
Published: July 14, 2020

REFERENCES

Aiello, N.M., Brabletz, T., Kang, Y., Nieto, M.A., Weinberg, R.A., and Stanger, B.Z. (2017). Upholding a role for EMT in pancreatic cancer metastasis. Nature *547*, E7–E8.

Anastassiadis, K., Fu, J., Patsch, C., Hu, S., Weidlich, S., Duerschke, K., Buchholz, F., Edenhofer, F., and Stewart, A.F. (2009). Dre recombinase, like Cre, is a highly efficient site-specific recombinase in E. coli, mammalian cells and mice. Dis. Model Mech. *2*, 508–515.

Araki, K., Shimura, T., Suzuki, H., Tsutsumi, S., Wada, W., Yajima, T., Kobayahi, T., Kubo, N., and Kuwano, H. (2011). E/N-cadherin switch mediates cancer progression via TGF-β-induced epithelial-to-mesenchymal transition in extrahepatic cholangiocarcinoma. Br. J. Cancer *105*, 1885–1893.

Brabletz, T., Kalluri, R., Nieto, M.A., and Weinberg, R.A. (2018). EMT in cancer. Nat. Rev. Cancer *18*, 128–134.

Cano, A., Pérez-Moreno, M.A., Rodrigo, I., Locascio, A., Blanco, M.J., del Barrio, M.G., Portillo, F., and Nieto, M.A. (2000). The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. Nat. Cell Biol. *2*, 76–83.

Chung, S., Yao, J., Suyama, K., Bajaj, S., Qian, X., Loudig, O.D., Eugenin, E.A., Phillips, G.R., and Hazan, R.B. (2013). N-cadherin regulates mammary tumor cell migration through Akt3 suppression. Oncogene *32*, 422–430.

Fischer, K.R., Durrans, A., Lee, S., Sheng, J., Li, F., Wong, S.T., Choi, H., El Rayes, T., Ryu, S., Troeger, J., et al. (2015). Epithelial-to-mesenchymal transition is not required for lung metastasis but contributes to chemoresistance. Nature *527*, 472–476.

Franke, W.W., Schmid, E., Osborn, M., and Weber, K. (1979). Intermediate-sized filaments of human endothelial cells. J. Cell Biol. *81*, 570–580.

Gravdal, K., Halvorsen, O.J., Haukaas, S.A., and Akslen, L.A. (2007). A switch from E-cadherin to N-cadherin expression indicates epithelial to mesenchymal transition and is of strong and independent importance for the progress of prostate cancer. Clin. Cancer Res. *13*, 7003–7011.

Guy, C.T., Cardiff, R.D., and Muller, W.J. (1992). Induction of mammary tumors by expression of polyomavirus middle T oncogene: a transgenic mouse model for metastatic disease. Mol. Cell. Biol. *12*, 954–961.

Hazan, R.B., Qiao, R., Keren, R., Badano, I., and Suyama, K. (2004). Cadherin switch in tumor progression. Ann. N. Y. Acad. Sci. *1014*, 155–163.

He, L., Li, Y., Li, Y., Pu, W., Huang, X., Tian, X., Wang, Y., Zhang, H., Liu, Q., Zhang, L., et al. (2017). Enhancing the precision of genetic lineage tracing using dual recombinases. Nat. Med. *23*, 1488–1498.

He, L., Liu, Q., Hu, T., Huang, X., Zhang, H., Tian, X., Yan, Y., Wang, L., Huang, Y., Miquerol, L., et al. (2016). Genetic lineage tracing discloses arteriogenesis as the main mechanism for collateral growth in the mouse heart. Cardiovasc. Res. *109*, 419–430.

Hulit, J., Suyama, K., Chung, S., Keren, R., Agiostratidou, G., Shan, W., Dong, X., Williams, T.M., Lisanti, M.P., Knudsen, K., and Hazan, R.B. (2007). N-cadherin signaling potentiates mammary tumor metastasis via enhanced extracellular signal-regulated kinase activation. Cancer Res. *67*, 3106–3116.

Kalluri, R., and Zeisberg, M. (2006). Fibroblasts in cancer. Nat. Rev. Cancer *6*, 392–401.

Kostetskii, I., Li, J., Xiong, Y., Zhou, R., Ferrari, V., Patel, V., Molkentin, J., and Radice, G. (2005). Induced deletion of the N-cadherin gene in the heart leads to dissolution of the intercalated disc structure. Cir Res, 346–354.

Krebs, A.M., Mitschke, J., Lasierra Losada, M., Schmalhofer, O., Boerries, M., Busch, H., Boettcher, M., Mougiakakos, D., Reichardt, W., Bronsert, P., et al. (2017). The EMT-activator Zeb1 is a key factor for cell plasticity and promotes metastasis in pancreatic cancer. Nat. Cell Biol. *19*, 518–529.

Kretzschmar, K., and Watt, F.M. (2012). Lineage tracing. Cell *148*, 33–45.

Ledford, H. (2011). Cancer theory faces doubts. Nature *472*, 273.

Lim, E., Vaillant, F., Wu, D., Forrest, N.C., Pal, B., Hart, A.H., Asselin-Labat, M.L., Gyorki, D.E., Ward, T., Partanen, A., et al. (2009). Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. Nat. Med. *15*, 907–913.

Lin, E.Y., Jones, J.G., Li, P., Zhu, L., Whitney, K.D., Muller, W.J., and Pollard, J.W. (2003). Progression to malignancy in the polyoma middle T oncoprotein mouse breast cancer model provides a reliable model for human diseases. Am. J. Pathol. *163*, 2113–2126.

(H) Schematic figure showing experimental design of ZsGreen$^+$ cells transplantation into littermate mice (no fluorescence reporter).
(I) Immunostaining for ZsGreen, tdTomato, and E-Cad on lung sections of mice transplanted with ZsGreen$^+$ tumor cells. Arrowheads indicate ZsGreen$^+$tdTomato$^-$ tumor metastases in lung.
Scale bars, 100 μm. Each image is representative of five individual biological samples

Liu, Q., Huang, X., Zhang, H., Tian, X., He, L., Yang, R., Yan, Y., Wang, Q.D., Gillich, A., and Zhou, B. (2015). c-kit(+) cells adopt vascular endothelial but not epithelial cell fates during lung maintenance and repair. Nat. Med. *21*, 866–868.

Maamar, H., Raj, A., and Dubnau, D. (2007). Noise in gene expression determines cell fate in Bacillus subtilis. Science *317*, 526–529.

Madisen, L., Zwingman, T.A., Sunkin, S.M., Oh, S.W., Zariwala, H.A., Gu, H., Ng, L.L., Palmiter, R.D., Hawrylycz, M.J., Jones, A.R., et al. (2010). A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. Nat. Neurosci. *13*, 133–140.

Maheswaran, S., and Haber, D.A. (2015). Cell fate: transition loses its invasive edge. Nature *527*, 452–453.

Mani, S.A., Guo, W., Liao, M.J., Eaton, E.N., Ayyanan, A., Zhou, A.Y., Brooks, M., Reinhard, F., Zhang, C.C., Shipitsin, M., et al. (2008). The epithelial-mesenchymal transition generates cells with properties of stem cells. Cell *133*, 704–715.

Nieman, M.T., Prudoff, R.S., Johnson, K.R., and Wheelock, M.J. (1999). N-cadherin promotes motility in human breast cancer cells regardless of their E-cadherin expression. J. Cell Biol. *147*, 631–644.

Nieto, M.A., Huang, R.Y., Jackson, R.A., and Thiery, J.P. (2016). EMT: 2016. Cell *166*, 21–45.

Süel, G.M., Garcia-Ojalvo, J., Liberman, L.M., and Elowitz, M.B. (2006). An excitable gene regulatory circuit induces transient cellular differentiation. Nature *440*, 545–550.

Tang, Y., Wang, Z., Li, Z., Kim, J., Deng, Y., Li, Y., Heath, J.R., Wei, W., Lu, S., and Shi, Q. (2017). High-throughput screening of rare metabolically active tumor cells in pleural effusion and peripheral blood of lung cancer patients. Proc. Natl. Acad. Sci. USA *114*, 2544–2549.

Thiery, J.P., Acloque, H., Huang, R.Y., and Nieto, M.A. (2009). Epithelial-mesenchymal transitions in development and disease. Cell *139*, 871–890.

Tian, X., Pu, W.T., and Zhou, B. (2015). Cellular origin and developmental program of coronary angiogenesis. Circ. Res. *116*, 515–530.

Wang, D., Cai, C., Dong, X., Yu, Q.C., Zhang, X.O., Yang, L., and Zeng, Y.A. (2015). Identification of multipotent mammary stem cells by protein C receptor expression. Nature *517*, 81–84.

Wang, Y., Nakayama, M., Pitulescu, M.E., Schmidt, T.S., Bochenek, M.L., Sakakibara, A., Adams, S., Davy, A., Deutsch, U., Lüthi, U., et al. (2010). Ephrin-B2 controls VEGF-induced angiogenesis and lymphangiogenesis. Nature *465*, 483–486.

Weinberger, L.S., Dar, R.D., and Simpson, M.L. (2008). Transient-mediated fate determination in a transcriptional circuit of HIV. Nat. Genet. *40*, 466–470.

Xu, Y., Lee, D.K., Feng, Z., Xu, Y., Bu, W., Li, Y., Liao, L., and Xu, J. (2017). Breast tumor cell-specific knockout of Twist1 inhibits cancer cell plasticity, dissemination, and lung metastasis in mice. Proc. Natl. Acad. Sci. USA *114*, 11494–11499.

Yang, J., Mani, S.A., Donaher, J.L., Ramaswamy, S., Itzykson, R.A., Come, C., Savagner, P., Gitelman, I., Richardson, A., and Weinberg, R.A. (2004). Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. Cell *117*, 927–939.

Ye, X., Brabletz, T., Kang, Y., Longmore, G.D., Nieto, M.A., Stanger, B.Z., Yang, J., and Weinberg, R.A. (2017). Upholding a role for EMT in breast cancer metastasis. Nature *547*, E1–E3.

Ye, X., Tam, W.L., Shibue, T., Kaygusuz, Y., Reinhardt, F., Ng Eaton, E., and Weinberg, R.A. (2015). Distinct EMT programs control normal mammary stem cells and tumour-initiating cells. Nature *525*, 256–260.

Yu, M., Bardia, A., Wittner, B.S., Stott, S.L., Smas, M.E., Ting, D.T., Isakoff, S.J., Ciciliano, J.C., Wells, M.N., Shah, A.M., et al. (2013). Circulating breast tumor cells exhibit dynamic changes in epithelial and mesenchymal composition. Science *339*, 580–584.

Zhang, H., Pu, W., Tian, X., Huang, X., He, L., Liu, Q., Li, Y., Zhang, L., He, L., Liu, K., et al. (2016). Genetic lineage tracing identifies endocardial origin of liver vasculature. Nat. Genet. *48*, 537–543.

Zheng, X., Carstens, J.L., Kim, J., Scheible, M., Kaye, J., Sugimoto, H., Wu, C.C., LeBleu, V.S., and Kalluri, R. (2015). Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer. Nature *527*, 525–530.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
tdTomato	ChromoTek	Cat# 600-401-379;RRID:AB_2209751
ZsGreen	Clontech	Cat# 632474; RRID:AB_2491179
CDH5	R&D	Cat# AF1002;RRID:AB_2077789
PyMT	Abcam	Cat# ab15085;RRID:AB_301631
PECAM	BD PHharmingen	Cat# 553370;RRID:AB_394816
PDGFRa	R&D	Cat# AF1062;RRID:AB_2236897
ESR	Abcam	Cat# ab32063;RRID:AB_732249
EpCAM	Abcam	Cat# ab92382;RRID:AB_30163
E-Cadherin	R & D	Cat# AF748;RRID:AB_355568
N-Cadherin	Thermo Fisher Scientific	Cat# 33-3900;RRID:AB_2313779
N-Cadherin	Santa Cruz	Cat# sc-31030;RRID:AB_2077520
β-actin	Proteintech	Cat# 60008-1-Ig;RRID:AB_2289225
K5	Covance	Cat# PRB-160p;RRID:AB_10063444
K14	Covance	Cat# PRB-155P-100;RRID:AB_292096
K8	DSHB	Cat# TROMA-I;RRID:AB_531826
K19	DSHB	Cat# TROMA-III;RRID:AB_2133570
CD24	eBioscience	Cat# 17-0291;RRID:AB_1210793
CD29	eBioscience	Cat# 13-0291;RRID:AB_1518777
CD31	eBioscience	Cat# 25-0311;RRID:AB_2734962
CD45	eBioscience	Cat# 25-0451;RRID:AB_469625
TER119	eBioscience	Cat# 14-5921-85;RRID:AB_467728
Chemicals, Peptides, and Recombinant Proteins		
Tamoxifen	Sigma-Aidrich	Cat# T5648
Paraformaldehyde (PFA)	Sigma-Aidrich	Cat# P6148-1KG
Sucrose	Sigma-Aidrich	Cat# V900116
O.C.T.	Sakura	Cat# 4583
Donkey serum	JIR	Cat# 017-000-121
Fc block solution	BD PHharmingen	Cat# 14-0161-82
Triton X-100	Sigma-Aidrich	Cat# Sigma-X-100
FBS	ThermoFisher	Cat# 10099141c
PBS	invitrogen	Cat# C10010500BT
HBSS	invitrogen	Cat# 14175103
DMEM	Hyclone	Cat# 11965092
DMEM/F12	Thermo Scientific	Cat# 11320082
RPMI 1640	invitrogen	Cat# 22400089
Collagenase III	Worthington	Cat# LS004183
Elastase	Worthington	Cat# LS002279
Collagenase A	Roche	Cat# 10103586001
DNase I	Worthington	Cat# LS002139
RBC	eBioscience	Cat# 00-4333-57
Trypsin-EDTA	invitrogen	Cat# 25200072
Violet dye	Invitrogen	Cat# L34955
DAPI	Sigma	Cat# D21490
TGF-beta	Novus Biologicals	Cat# NBP2-51925

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Trizol	invitrogen	Cat# 15596018
HEPES	invitrogen	Cat# 15630080
SYBR Green qPCR master mix	Thermo Fisher Scientific	Cat# 4367659
RIPA lysis buffer	Beyotime	Cat# P0013B
Matrigel	Corning	Cat# 356234
Fluorescence-protecting mounting medium	Vector Lab	Cat# H-1000
Critical Commercial Assays		
PrimeScript RT kit	Takara	Cat# R045A
Cyanine 5 Amplification Reagent	PerkinElmer	Cat# NEL745001KT
Cyanine 3 Amplification Reagent	PerkinElmer	Cat# NEL744001KT
Deposited Data		
Raw data and analyzed data	This paper	N/A
Experimental Models: Organisms/Strains		
Mouse: *Vim LSL-Dre*	This paper	Shanghai Biomodel Organism Co., Ltd
Mouse: *N-cad-LSL-Dre*	This paper	Shanghai Biomodel Organism Co., Ltd
Mouse: *Kit-CreER*	Liu et al., 2015	N/A
Mouse: *Vim-CreER*	He et al., 2017	N/A
Mouse: *Cdh5-CreER*	Wang et al., 2010	N/A
Mouse: *R26-tdTomato*	Madisen et al., 2010	N/A
Mouse: *R26-RSR-tdTomato*	Zhang et al.2016	N/A
Mouse: *MMTV-PyMT*	Guy et al., 1992	N/A
Mouse: *NR1*	He et al., 2017	N/A
Mouse: *N-cad-flox*	Kostetskii et al., 2005	N/A
Oligonucleotides For qRT-PCR		
Primer1 of N-Cadherin: 5'- AGGCTTCTGGTGAAATTGCAT-3'	This paper	N/A
Primer2 of N-Cadherin: 5'- GTCCACCTTGAAATCTGCTGG-3'	This paper	N/A
Primer1 of Vimentin: 5'- CGTCCACACGCACCTACAG-3'	This paper	N/A
Primer2 of Vimentin: 5'- GGGGGATGAGGAATAGAGGCT -3'	This paper	N/A
Primer1 of E-Cad: 5'- CAGTTCCGAGGTCTACACCTT-3'	This paper	N/A
Primer2 of E-Cad: 5'- TGAATCGGGAGTCTTCCGAAAA -3'	This paper	N/A
Primer1 of Epcam: 5'- GCGGCTCAGAGAGACTGTG -3'	This paper	N/A
Primer2 of Epcam: 5'- CCAAGCATTTAGACGCCAGTTT -3'	This paper	N/A
Primer1 of Zeb: 5'- ACCGCCGTCATTTATCCTGAG -3'	This paper	N/A
Primer2 of Zeb: 5'- CATCTGGTGTTCCGTTTTCATCA -3'	This paper	N/A
Primer1 of Twist: 5'- GGACAAGCTGAGCAAGATTCA -3'	This paper	N/A
Primer2 of Twist: 5'- CGGAGAAGGCGTAGCTGAG -3'	This paper	N/A
Primer1 of Snail: 5'- CACACGCTGCCTTGTGTCT -3'	This paper	N/A
Primer2 of Snail: 5'- GGTCAGCAAAAGCACGGTT -3'	This paper	N/A

RESOURCE AVAILABILITY

Lead Contact

Requests for further information, reagent, and resource should be directed to and will be fulfilled by the Lead Contact, Bin Zhou (zhoubin@sibs.ac.cn).

Materials Availability

This study did not generate any new unique reagents.

Data and Code Availability

N/A

EXPERIMENTAL MODELS AND SUBJECT DETAILS

Mice Generation and Breeding

All mouse studies were carried out according to the Institutional Animal Care and Use Committee of the Institute of Biochemistry and Cell Biology and Institute for Nutritional Sciences, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences. The *Kit-CreER*, *R26-tdTomato*, *R26-rox-tdTomato*, *NR1*, and *MMTV-PyMT* were reported previously (Liu et al., 2015; Madisen et al., 2010; Zhang et al., 2016; He et al., 2017; Guy et al., 1992; Wang et al., 2010). The *Vim-CreER*, *Vim-LSL-Dre*, and *N-cad-LSL-Dre* knock-in mouse lines were generated by genome editing using CRISPR-Cas9 technology, as described previously (Zhang et al., 2016). In brief, a cDNA encoding $CreER^{(T2)}$ fusion protein was inserted into the translation start codon of the vimentin, or a cDNA containing loxp-stop-loxp-Dre-polyA (LSL-Dre) cassette was inserted into the translation start codon of the vimentin or N-cadherin gene. To prevent the leakiness of the stop sequence, multiple polyA sequences were inserted between two loxp loci. To strengthen the stability of the transcripts, the woodchuck hepatitis virus posttranscriptional regulatory element (WPRE) was also included following the Dre signal sequence. The *Vim-CreER*, *Vim-LSL-Dre*, and *N-cad-LSL-Dre* knock-in mouse lines were kindly generated by Shanghai Model Organism Center, Inc. (SMOC). *Kit-CreER* mice were crossed with *R26-tdTomato* and *MMTV-PyMT* mice to obtain *Kit-CreER*;*R26-loxp-tdTomato*;*MMTV-PyMT* triple positive mice for genetic lineage tracing of c-kit^+ cells in mammary tumor and lung metastases. To generate the genetic tracing system for EMT activity, we crossed *Kit-CreER* with *EMTgene-LSL-Dre* (*Vim-LSL-Dre* or *N-cad-LSL-Dre*) and *NR1* mice to acquire *Kit-CreER;EMTgene-LSL-Dre;NR1* (*EMTracer*) triple positive mice for genetic lineage tracing the epithelial-to-mesenchymal transition in homeostasis. These *EMTracer* mice were crossed with *MMTV-PyMT* for genetic lineage tracing the epithelial-to-mesenchymal transition (EMT) in tumorigenesis and metastasis. Both male and female mice were used for embryonic study. Female mice were included in this work to study the mammary tumor growth and metastasis in the adult stage. At least 5 mice were used in each experiment in this study.

METHOD DETAILS

Genomic PCR

Genomic DNA was prepared from embryonic yolk sac or tail of mice. Simply, tissues were lysed in lysis buffer (100 mM Tris HCl (pH 7.8), 5 mM EDTA, 0.2% SDS, 200 mM NaCl and 100 μg/ml proteinase K) overnight at 55°C, followed by centrifugation at maximum speed (21,130*g*) for 8 min to obtain supernatants with genomic DNA. Genomic DNA was precipitated by adding isopropanol and washed by 70% ethanol. PCR primers spanning the genomic DNA and inserted cassette were designed to test for the correct genotype in all mouse lines as described previously. PCR primers for newly generated transgenic mouse lines will be provided if requested.

Tamoxifen Treatment

To induce *Kit-CreER* activity in our mouse lines, tamoxifen (Sigma, T5648, 20 mg/ml, dissolved in corn oil) was administered by gavage at the indicated time points as described previously (Liu et al., 2015). The day of vaginal plug detection was regarded as E0.5. The dosage of tamoxifen administration was based on the development stage of mice (embryonic experiment, 0.1-0.15mg/g body weight of pregnant mouse; adult lineage tracing experiment, 4 mg every time every other day, 3 times at the age of 6-7 weeks; adult expression experiment, 2-3mg tamoxifen for once, and tissues were collect within 48 hours); for N-cadherin protein knockout in Kit^+ cells, mouse was treated with tamoxifen for at least 6 times at the age of 6∼8 weeks, 4mg every time every other day.

Whole-mount Fluorescence Microscopy

After mouse was sacrificed, the mouse skin was cut off by scissors and was fixed by needle in place. Then the mammary was exposed and whole-mount bright-field or fluorescence image was taken using a Zeiss stereo microscope (AxioZoom V16). The third and fourth mammary glands were usually tested in our experiments. Other tissues, primary tumor and lung were washed in phosphate-buffered saline (PBS) and fixed in 4% paraformaldehyde (PFA) at 4°C for 1 hour followed by three times washes with PBS. Then tissues were placed in PBS on an agar gel for whole-mount bright-field or fluorescence imaging.

Immunostaining and Confocal Fluorescence Microscopy

Immunostaining was performed as according to previously described protocols (He et al., 2016). Briefly, embryos or adult tissues were collected and washed in PBS to remove excessive blood, and then tissues were fixed in PFA at 4°C for 30 minutes to 1 hour depending on tissue size (embryo: 30 min; adult tissue: ∼1 hour). After washing in PBS for three times, tissues were dehydrated in 30% sucrose in PBS overnight at 4°C. Then, embryos or tissues were socked in OCT (Sakura) for 1 hour at 4°C. We embedded the tissue in block and froze it at -80°C. Usually, 10μm thick cryosections were collected on slides and stored at -20°C until use. For immunostaining, after air drying for ∼1 hour at room temperature, tissue section was blocked with blocking buffer (5% donkey serum and 0.1% Triton X-100 in PBS) for 30 min at room temperature and incubated in primary antibody at 4°C overnight. The next day, after washing out the primary antibodies by washing the section three times with PBS, tissue section was then developed with Alexa Fluor

fluorescent antibodies (Invitrogen) for 30 minutes at room temperature and then fully washing out the antibodies by PBS. Horseradish peroxidase (HRP)- or biotin-conjugated antibodies (Jackson ImmunoResearch) with a tyramide signal amplification kit (PerkinElmer) were used for antigens that were hard to detect by Alexa Fluor fluorescent antibodies. We mounted the tissue section with mounting medium containing the nuclear stain DAPI (Vector Lab). Primary antibodies and dilution ratios are listed below: tdTomato (Rockland, 600-401-379; 1:1,000), ZsGreen (Clontech, 632474; 1:2000), K5 (Covance, PRB-160p; 1:500), K14 (Covance, PRB-155P-100; 1:500), K8 (DSHB, Troma-I-C; 1:500), K19 (Developmental Studies Hybridoma Bank, TROMA-III; 1:500), E-Cad (R & D, AF748; 1:200), Ep-CAM (Abcam, ab92382; 1:1000), PyMT (Abcam, ab15085; 1:300), PDGFRa (R & D, AF1062; 1:500), PECAM (BD PHharmingen, 553370; 1:500), CDH5 (R & D, AF1002; 1:100), vimentin (Santa Cruz; sc-7557; 1:100), Estrogen receptor (Abcam; ab32063; 1:300), N-caherin (Thermo Fisher Scientific; 33-3900; 1:300). Images were acquired on an Olympus confocal microscope (FV1200), a Nikon confocal microscope (A1) or a Zeiss stereo microscope (AxioZoom V16). To obtain z-stack images, we scanned ~6 consecutive XY images on the Z-axis by confocal microscope. The obtained pictures were analyzed by ImageJ (NIH) software.

Mammary Cell Preparation and Flow Cytometry

To analyze the luminal or basal epithelial cells labeled by Kit-CreER, we performed flow cytometric analysis of dissociated tissue after digestion. The isolation of mammary epithelium cells was performed as previously (Wang et al., 2015). Usually, the third and fourth mammary glands of female mice were collected and were minced into small pieces. Then the minced tissues were incubated with an enzyme cocktail (RPMI 1640 with 25 mM HEPES, 5% fetal bovine serum, 1% penicillin-streptomycin-glutamine), 300 U /ml collagenase III (Worthington)) at 37°C for 90 min. The fatty waste and most digestive media were removed by centrifugation (1000 rpm, 5 min). The pellets were treated with red blood cell lysis buffer (eBioscience, 00-4333-57) for 5 minutes at room temperature to lyse the red blood cells. Then a single-cell suspension was obtained by digesting with 0.25% trypsin-EDTA (invitrogen/Gibco/MP) at 37°C for 5 min and 0.1 mg/ml DNase I (Worthington, LS002139) for 5 min at 37°C with gentle pipetting. The pellet was re-suspended by adding Hank's Balanced Salt Solution (HBSS) containing 5% FBS and then the cells were filtered through 70 μm cell strainers (BD Biosciences). The dissociated cells were centrifuged and then were re-suspended by adding 50 μl of Violet dye (Invitrogen L34955, 1:1,000 in PBS), incubated at 4°C for 30 min to gate living cells. Afterwards, cells were treated with Fc block solution (eBioscience, 14-0161-82, 1:100 in Isolation buffer) at room temperature for 5 min and then incubated with antibody at 4°C for 30 min. The FACS antibodies CD24 (eBioscience, 17-0291; 1:100), CD29 (eBioscience, 13-0291; 1:20) were used to gate mammary luminal, basal and stromal cells. CD31 (eBioscience, 25-0311; 1:40), CD45 (eBioscience, 25-0451; 1:200), and TER119 (eBioscience, 25-5921-82; 1:200) were also used to exclude the lineage positive populations. We performed flow cytometric analysis on Aria II Flow Cytometer (BD Bioscience), and the raw data was processed using Flowjo (TreeStar).

Tumor Cells Isolation and Flow Cytometry

For mammary tumor, cells were isolated as according to a previous protocols (Fischer et al., 2015). Tumor tissues were isolated from mouse and washed in PBS to remove excessive blood and waste. Then tumor tissues were minced into ~1 mm^3 blocks and digested in an enzyme cocktail (collagenase A (Roche, 10103578001), elastase (Worthington, LS002279), and DNase I (Worthington, LS002139)) in HBSS buffer at 37°C for 60 min. The cell suspension was filtered through 40 μm cell strainers (BD Biosciences) and the cells were treated with red blood cell lysis buffer (eBioscience, 00-4333-57) for 5 minutes at room temperature to lyse the red blood cells. Then cells were washed with PBS for three times and re-suspended by DMEM/F12. The dead cells were identified by adding DAPI. We sorted $ZsGreen^+$ and $tdTomato^+$ cells by Aria II cell sorter (BD Bioscience).

Transplantation Experiment

$ZsGreen^+$ and $tdTomato^+$ cells isolated from *EMTracer* tumor mouse were diluted in sterile PBS. Cells were injected into tail vein of littermate control that were reporter negative by a dose of 5x 10^5 for one mouse, the volume of injection was limited to 200μl. About one month later, lung tissues were collected for analysis.

TGF-beta Induction

To test our EMTracer system could be induced by TGF-beta in vitro cell culture, we isolated $ZsGreen^+$ cells of the mammary tumor of *EMTracer;MMTV-PyMT* mouse and cells were cultured in DMEM with 2% FBS and 2ng/ml TGF-beta (Novus Biologicals) for 5 days, then cells were imaged by a Zeiss stereo microscope (AxioZoom V16).

Transwell Assay

To test the invasion of N-cad positive tumors, $ZsGreen^+$ and $tdTomato^+$ cells isolated from *N-cad-EMTracer;MMTV-PyMT* mice. We add 300μl complete medium to each well (24-wells), then we place the inserts in wells and add each insert with 50μl matrigel at a concentration of 1 mg/ml, then incubate at 37°C for 1 hour. Cells were resuspended in serum-free media at a concentration of 2 x 10^5 cells/ml. We add 500μl to the top chamber of each well and incubated at 37°C for 24 hours. Imaging and quantification for the cells attached on each well. Images were acquired by an Olympus XI73 microecope. Each assay was repeated in 5 times.

PCR Analysis of DNA for Recombination Efficiency

We performed PCR analysis to test the efficiency of Stop sequence removal on *EMTgene-LSL-Dre* allele by *Kit-CreER* after tamoxifen administration. Briefly, $ZsGreen^+$ mammary tumor tissues were collected and the genomic DNA from $ZsGreen^+$ cells was pre-

pared for PCR. PCR primers spanning the vimentin promoter and inserted Dre cassette were designed (forward: TTGTCCAGTCCTCTGCCACTCTTG, reverse: TACTCCTTGCCGATGTTCCTCAGG). After the Stop sequence on *Vim-LSL-Dre* allele was removed by *Kit-CreER*, the band size was 295bp; if the Stop sequence was not removed, a larger size band (2057bp) would be detected.

RNA Extraction and Quantitative RT-PCR

Total RNA was extracted from ZsGreen^{+} and tdTomato^{+} cells isolated from *EMTracer* tumor mouse treated with tamoxifen. Cells were lysed with Trizol, the total RNA extraction was performed as according to the instructions of Invitrogen. The total RNA was reverse transcribed in cDNA by a PrimeScript RT kit (Takara). We used SYBR Green qPCR master mix (Applied Biosystems) to do qRT-PCR and cDNA was amplified on a StepOnePlus Real-time PCR System (Applied Biosystems). *Gapdh* was used as a internal control.

Western Blot

Mammary tumor tissues were removed from experimental and control group, mammary tumor tissue from *MMTV-PyMT* was removed for control. Tissues were washed in 1x PBS to remove massive blood. Then about 20mg tissues were lysed with 400μl RIPA lysis buffer (Beyotime, P0013B), smashed with refiner and incubated at room temperature for 30 min, followed by centrifugation at 12,000 g at 4°C for 10 min to get the protein superatant. All protein samples were mixed with 5 x loading buffer and boiled for 10∼15min. The total proteins were resolved by 10% SDS PAGE and then transferred onto the polyvinylidene fluoride (PVDF) membrane (Immobilon, Millipore) by a Mini Trans Blot system (Bio-Rad). The membranes were blocked with PBST (10 mM Tris-HCl pH 8.0, 150 mM NaCl and 0.5% vol/vol Tween-20) containing 5% BSA at room temperature for 1h. Then, membranes were incubated with primary antibodies at 4°C overnight. The next day, after washing primary antibodies with PBST for 3 times, membranes were incubated with HRP-conjugated secondary antibodies at room temperature for 1h. Finally, signals were detected by enhanced chemiluminescence (Pierce) according to the manufacturer's instructions. Antibodies and dilution ratios are listed below: anti-N-cadherin (Santa Cruz, sc-31030, 1:1000), anti-β-actin (Proteintech, 60008-1-Ig, 1:1000), HRP-anti-goat IgG (Jackson ImmunoResearch, 705-035-147; 1:5000), HRP-anti-mouse IgG (Jackson ImmunoResearch, 715-035-150; 1:5000).

Microwell Chip Assay

To identify the candidate CTCs in whole blood, we performed Microwell chip assay as previously described (Tang et al., 2017). Briefly, the mice were treated with 200μl heparin (6.25U/μl) through intraperitoneal injections. 15 minutes later, the mice were anesthetized with pentobarbital sodium at 80 mg/kg body weight intraperitoneally. We collected ∼1 ml blood through the right ventricle by a needle. Then, blood was transferred into EDTA-contained vacutainer (BD, 367841) tubes and processed within 1 hour. Then blood was treated with red blood cell lysis buffer to lyse the red blood cells and cells were counted by Countless II (Thermo), then incubated with EpCAM antibody (eBioscience, 17-5791-82; 1:200). The images were acquired on ImageXpress MicroXLS Widefield High Content Screening System (Molecular Devices) and were analyzed by ImageJ (NIH) software.

QUANTIFICATION AND STATISTICAL ANALYSIS

Data are presented means ± SEM. Statistical analyses were performed on the mean values of biological replicates in each group using unpaired two-tailed Student's *t* test for comparison of differences between two groups; while date from over two groups was analyzed by ANOVA with Tukey's multiple comparisons test using GraphPad Prism method for multiple comparisons. $P < 0.05$ was considered statistically significant. In each experiment, 5 biological replicates were used for quantification and statistical analysis.

Developmental Cell, Volume *54*

Supplemental Information

Genetic Fate Mapping of Transient Cell Fate Reveals N-Cadherin Activity and Function in Tumor Metastasis

Yan Li, Zan Lv, Shaohua Zhang, Zhuo Wang, Lingjuan He, Muxue Tang, Wenjuan Pu, Huan Zhao, Zhenqian Zhang, Qihui Shi, Dongqing Cai, Mingfu Wu, Guohong Hu, Kathy O. Lui, Jing Feng, M. Angela Nieto, and Bin Zhou

Developmental Cell

Supplemental Information

Genetic Fate Mapping of Transient Cell Fate Reveals N-cadherin Activity and Function in Tumor Metastasis

Yan Li, Zan Lv, Shaohua Zhang, Zhuo Wang, Lingjuan He, Muxue Tang, Wenjuan Pu, Huan Zhao, Zhenqian Zhang, Qihui Shi, Dongqing Cai, Mingfu Wu, Guohong Hu, Kathy O. Lui, Jing Feng, M. Angela Nieto, and Bin Zhou

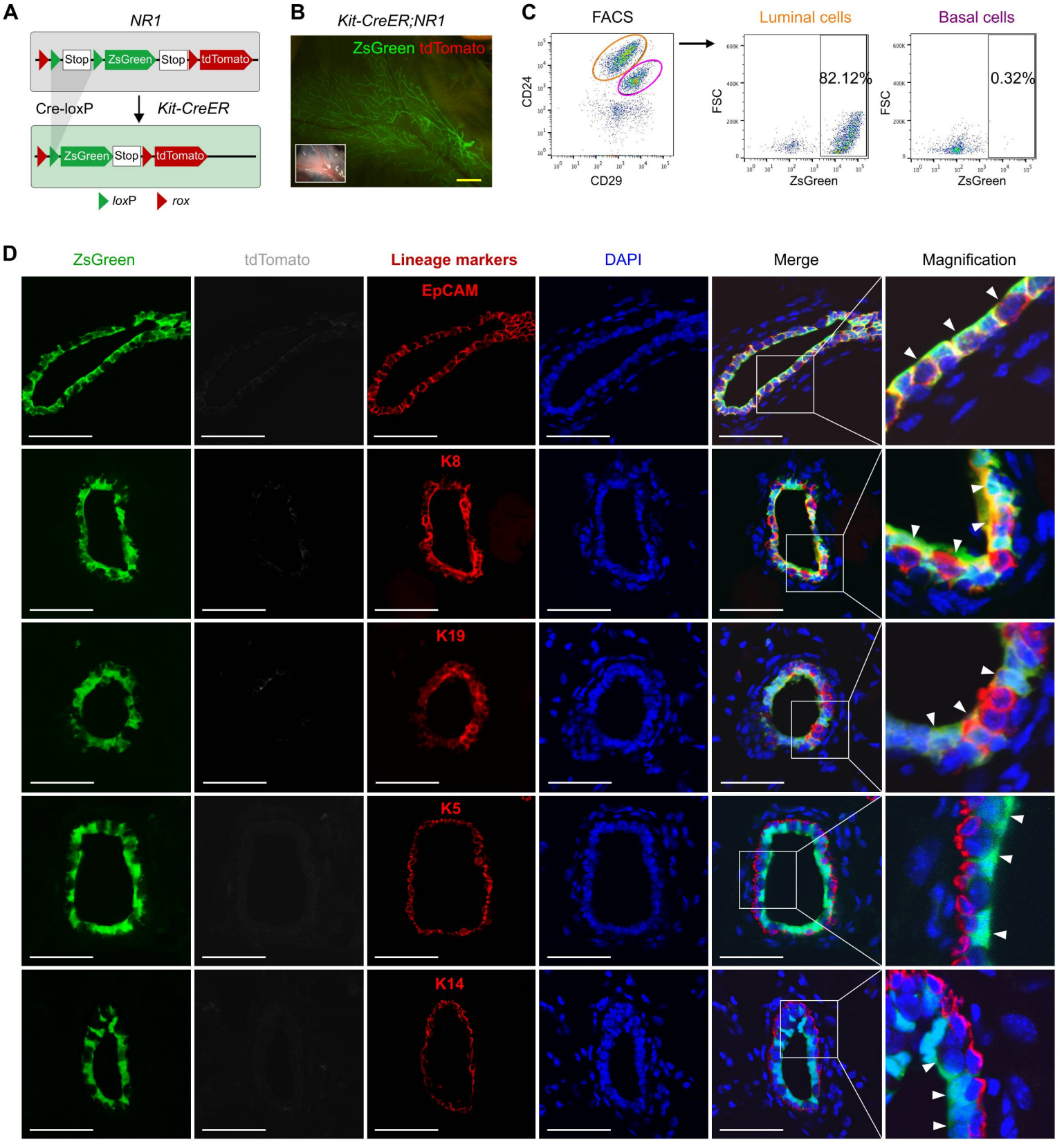


Figure S1 (Related to Figure 2). *Kit-CreER* labels luminal epithelial cells by *NR1*

(**A**) Schematic figure showing *Kit-CreER* mediated Cre-loxP recombination on *NR1* allele.

(**B**) Whole-mount fluorescence image of *Kit-CreER;NR1* mouse mammary tissue. Tamoxifen was induced at 6 weeks old and tissues were collected 2 days later.

(**C**) Flow cytometric analysis of *Kit-CreER* labeled mammary luminal or basal epithelial cells.

(**D**) Immunostaining for ZsGreen, tdTomato and epithelial cell lineage markers EpCAM, K8, K19, K5 or K14 on *Kit-CreER;NR1* mammary gland tissue sections. Arrowheads indicate ZsGreen^{+} luminal epithelial cells.

Scale bars, yellow, 2 mm; white, 100 μm. Each image is representative of 5 individual samples.

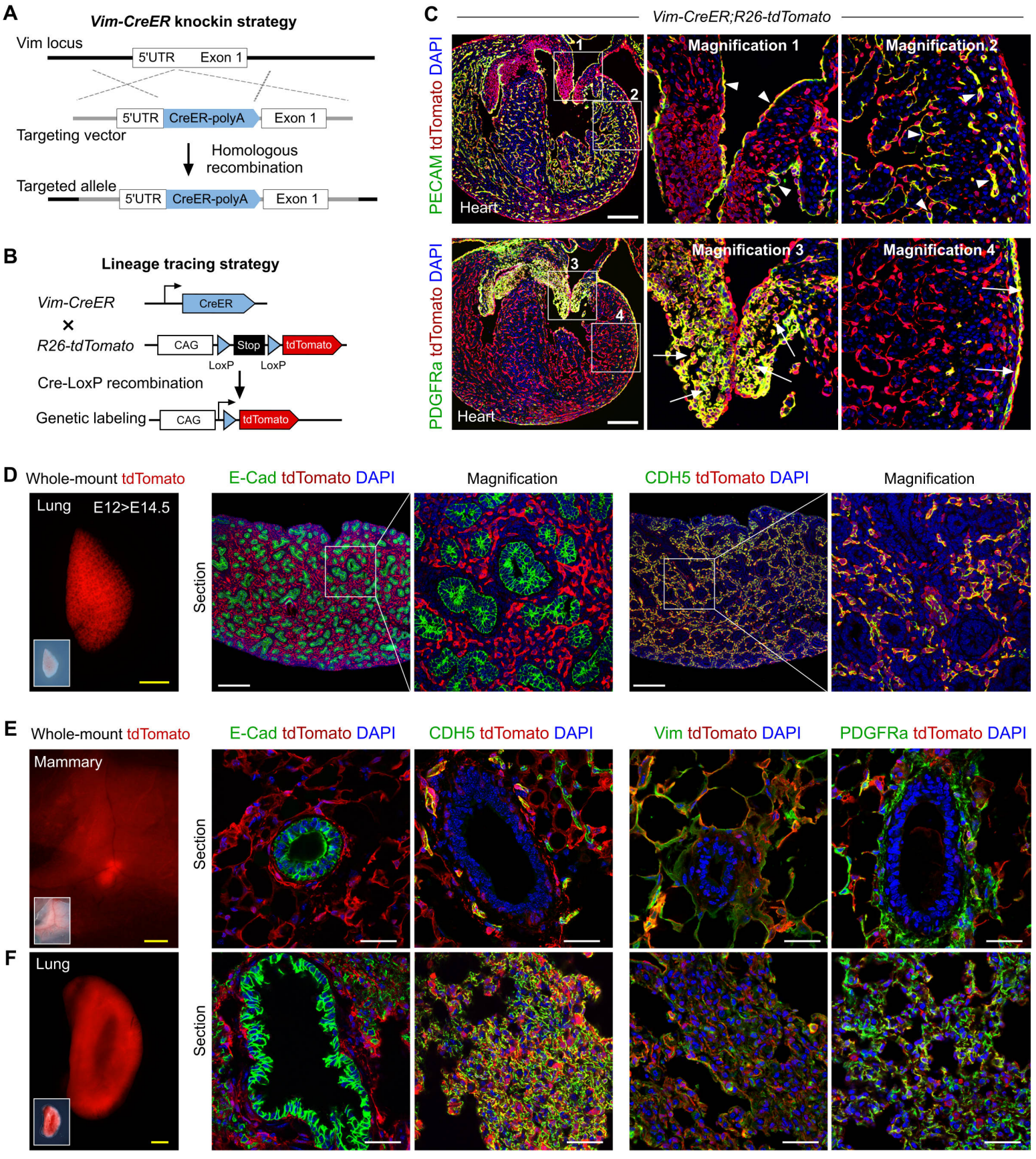


Figure S2 (Related to Figure 2). Vim⁺ cells are mainly endothelial and mesenchymal cell lineages

(**A**) Schematic figure showing generation of *Vim-CreER* by homologous recombination.

(**B**) Strategy of lineage tracing by Cre-loxP recombination.

(**C**) Immunostaining for tdTomato and PECAM or PDGFRa on E14.5 heart sections. Tamoxifen was induced 2 days before tissue collection. Vim⁺ cells are PECAM⁺ endothelial cells (arrowheads) or PDGFRa⁺ mesenchymal cells (arrows).

(**D**) Whole-mount fluorescence images for E14.5 lung (left). Immunostaining for tdTomato and E-Cad or CDH5 on lung tissue sections shows Vim⁺ cells are CDH5⁺ endothelial cells but not E-Cad⁺ epithelial cells.

(**E,F**) Whole-mount fluorescence view of adult mammary or lung tissue (left). Immunostaining for tdTomato and E-Cad, CDH5, Vim or PDGFRa on mammary (E) or lung (F) sections. In mammary and lungs, *Vim-CreER* labeled cells are CDH5⁺ endothelial cells, Vim⁺ mesenchymal cells or PDGFRa⁺ fibroblasts, and *Vim-CreER* does not label E-Cad⁺ epithelial cells.

Scale bars, yellow 1 mm; white, 100 μm. Each image is representative of 5 individual samples.

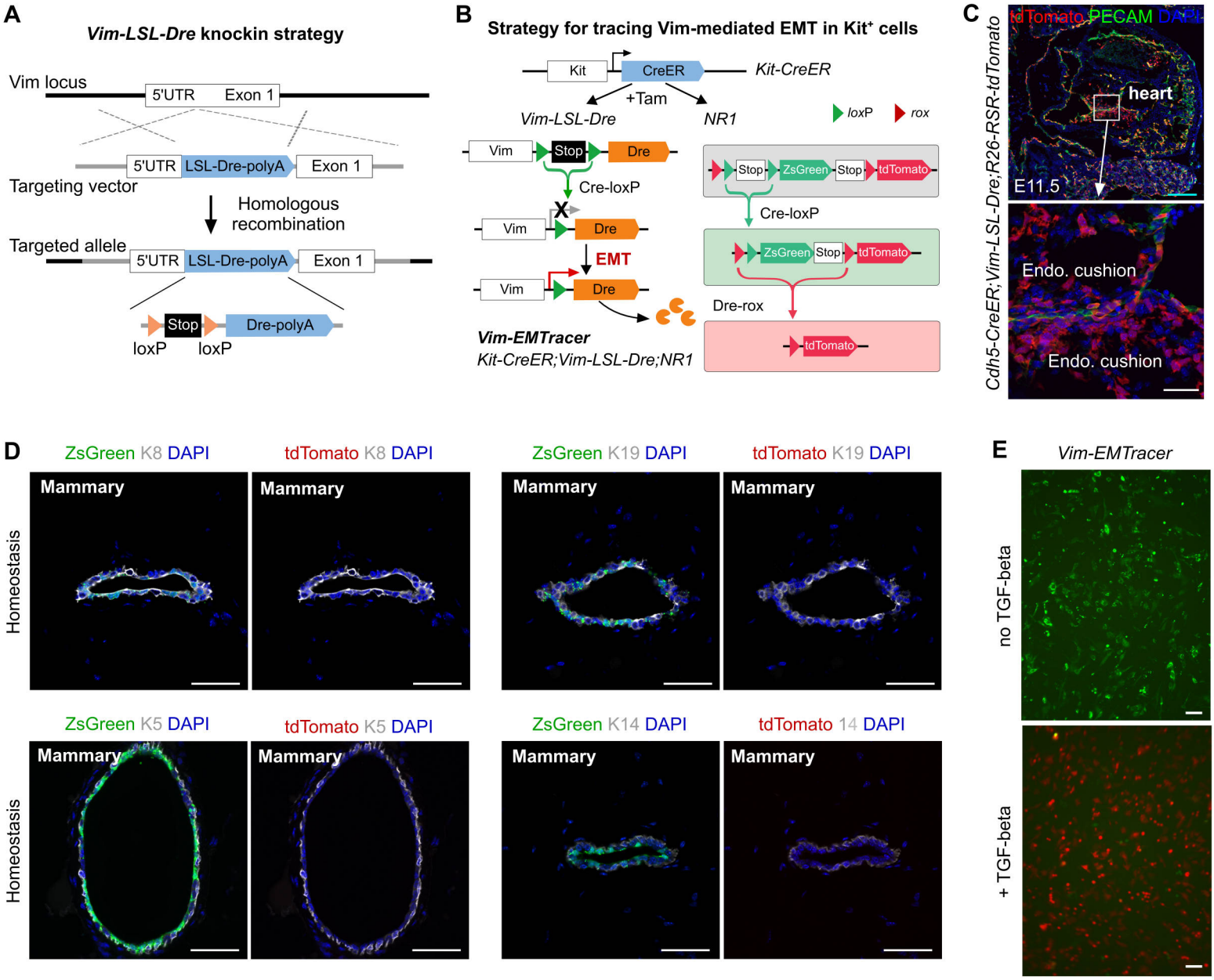


Figure S3 (Related to Figure 2). Generation of *Vim-LSL-Dre* allele for EMT tracing

(**A**) Strategy for targeting LSL-Dre into Vim gene locus by homologous recombination.

(**B**) Strategy for tracing Vim$^+$ EMT in Kit$^+$ cells by *Vim-EMTracer* (*Kit-CreER;Vim-LSL-Dre;NR1*). After tamoxifen induction, Kit$^+$ cells will be ZsGreen$^+$. Only after Vim$^+$ EMT, Dre-rox recombination leads to switch of ZsGreen into tdTomato expression in targeted cells.

(**C**) Immunostaining for tdTomato and PECAM on E11.5 *Cdh5-CreER;Vim-LSL-Dre;R26-RSR-tdTomato* embryonic section. This data is used as positive control to detect Vim$^+$ EMT in endocardial (endo.) cushion. Endocardial cells undergo EMT from E8.5 to E11.5 in the endo. cushion in developing heart. Tamoxifen was induced at E8.5, and embryos were collected at E11.5 for analysis.

(**D**) Immunostaining for ZsGreen and tdTomato with luminal cell marker K8 or K19; or with basal cell marker K5, K14 on mammary tissue sections of *Vim-EMTracer* mice. Tamoxifen was induced at 6-7 weeks old and tissues were collected at 24 weeks old for analysis.

(**E**) Positive control for EMT induced by TGF-beta. Fluorescence images of ZsGreen$^+$ cells isolated from mammary tissues of *Vim-EMTracer;MMTV-PyMT* mice. TGF-beta treatment significantly induced tdTomato expression.

(**F**) Immunostaining for CDH5, PECAM, PDGFRa, and tdTomato on lung tissue sections. Arrowheads, tdTomato$^+$CDH5$^+$ cells or tdTomato$^+$PECAM$^+$ cells.

Scale bars, blue, 500 μm; white, 50 μm. Each figure is representative of 5 individual biological samples.

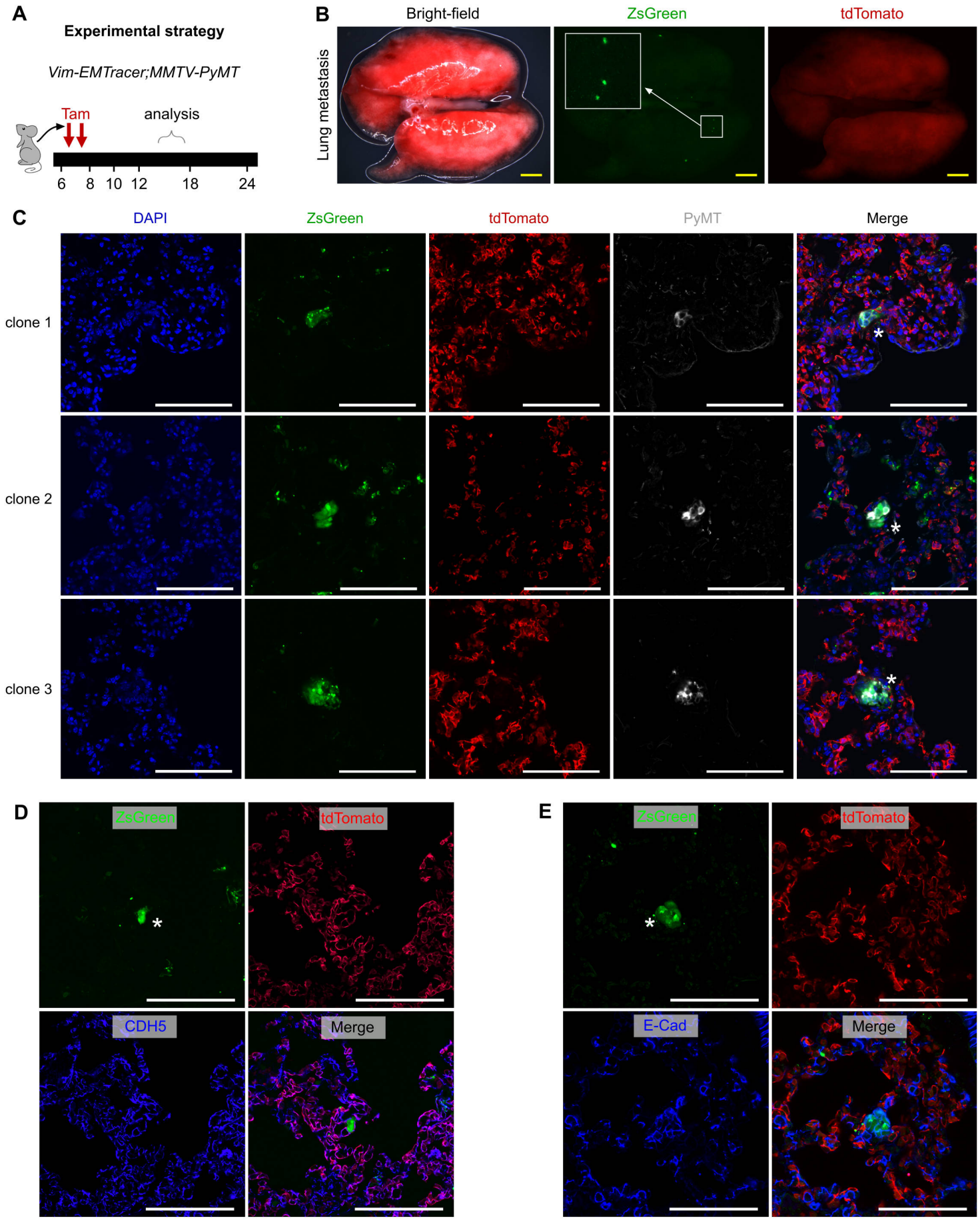


Figure S4 (Related to Figure 3). All small tumor colonies in lung were ZsGreen⁺ in *Vim-EMTracer;MMTV-PyMT* mouse at early metastasis
(**A**) Schematic figure showing experimental strategy.
(**B**) Whole-mount fluorescence images of *Vim-EMTracer;MMTV-PyMT* lung.
(**C**) Immunostaining for ZsGreen, tdTomato and PyMT on lung sections shows PyMT⁺ tumor cells (asterisks) are ZsGreen⁺tdTomato⁻.
(**D**) Immunostaining for ZsGreen, tdTomato and CDH5 on lung section shows ZsGreen⁺ tumor cells (asterisk) are CDH5⁻. As *Kit-CreER* labels lung endothelial cells that also express Vim, a subset of CDH5⁺ endothelial cells in lungs are tdTomato⁺.
(**E**) Immunostaining for ZsGreen, tdTomato and E-Cad on lung sections shows ZsGreen⁺ tumor cells (asterisk) are E-Cad⁺, and tdTomato⁺ endothelial cells are E-Cad⁻. Scale bars: yellow, 2 mm; white, 100 μm. Each image is representative of 5 individual samples.

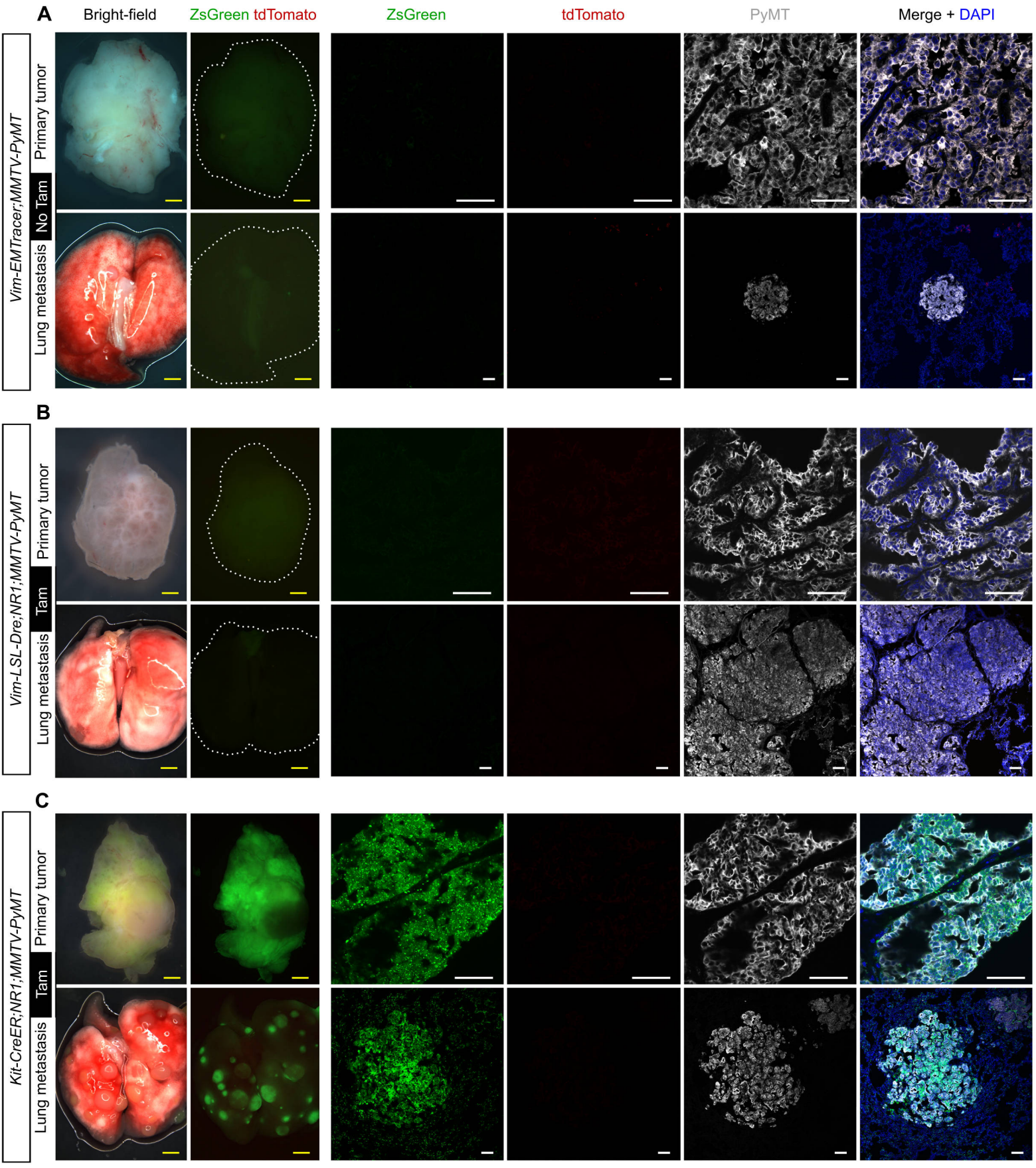


Figure S5 (Related to Figure 3). Technical controls for *Vim-EMTracer;MMTV-PyMT* mouse

(**A**) Whole-mount images of mammary (up) and lung (below) of *Vim-EMTracer;MMTV-PyMT* mouse at late stage, and immunostaining for ZsGreen, tdTomato and PyMT on mammary (up) and lung (below) sections. The mouse was not treated with tamoxifen. There is no leakiness for *Vim-EMTracer;MMTV-PyMT* without tamoxifen induction.

(**B**) Whole-mount images of mammary (up) and lung (below) of *Vim-LSL-Dre;NR1;MMTV-PyMT* mouse at late stage, and immunostaining for ZsGreen, tdTomato and PyMT on mammary (up) and lung (below) sections. The mouse was treated with tamoxifen at 6-7 weeks.

(**C**) Whole-mount images of mammary (up) and lung (below) of *Kit-CreER;NR1;MMTV-PyMT* mouse at late stage, and immunostaining for ZsGreen, tdTomato and PyMT on mammary (up) and lung (below) sections. Tamoxifen was induced at 6-7 weeks. Without *Vim-LSL-Dre* allele, Kit-derived cells are ZsGreen^{+}tdTomato^{-} in mammary tumor and lung metastasis of *Kit-CreER;NR1;MMTV-PyMT* mouse.

Scale bars, yellow, 2 mm; white, 50 μm. Each image is representative of 5 individual biological samples.

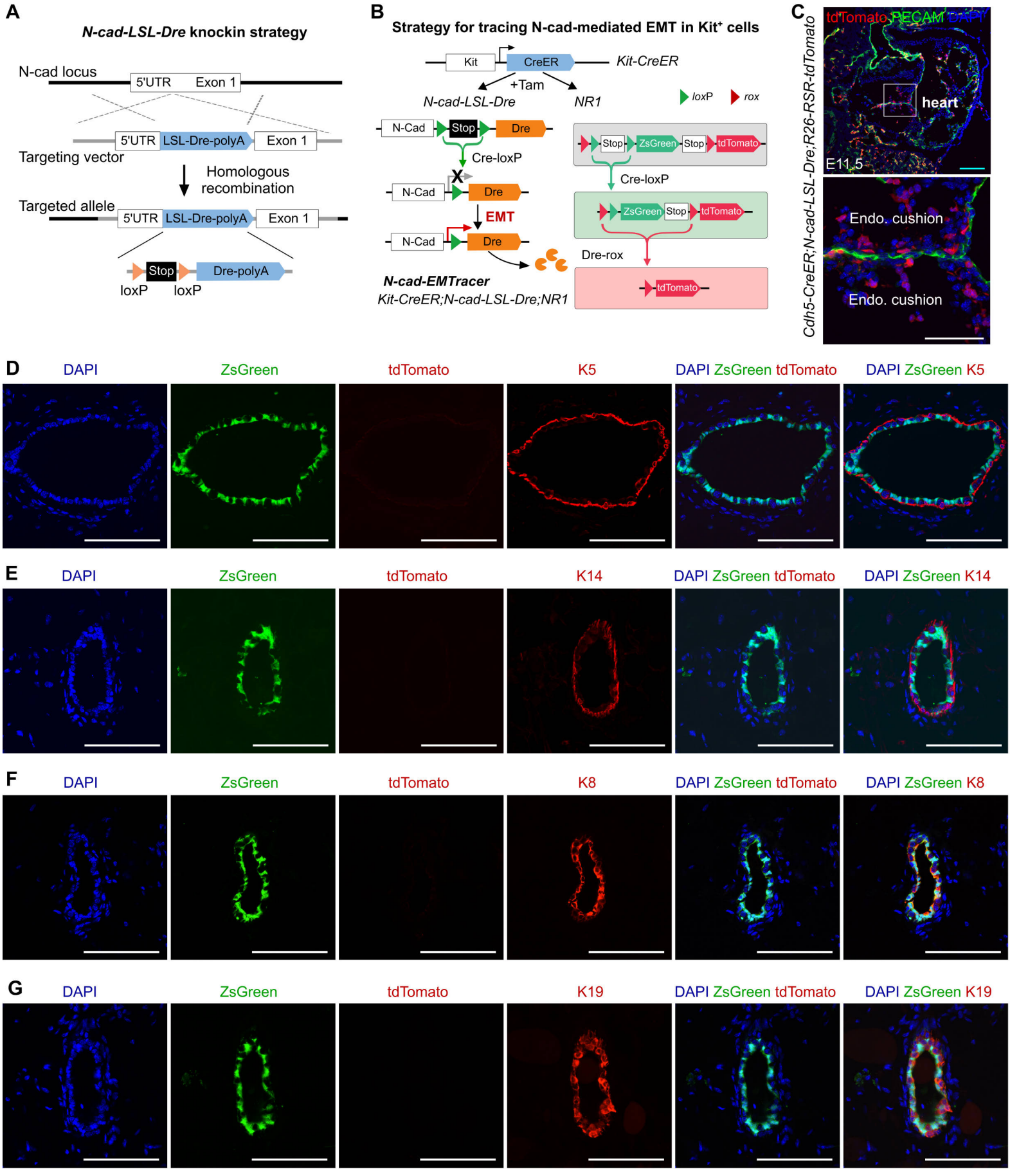


Figure S6 (Related to Figure 5). Characterization of *N-cad-EMTracer* in normal mammary

(**A**) Generation of *N-cad-LSL-Dre* allele by homologous recombination.

(**B**) Strategy for tracing N-Cad$^+$ EMT in Kit$^+$ cells by *N-Cad-EMTracer* (*Kit-CreER;N-cad-LSL-Dre;NR1*). After tamoxifen induction, Kit$^+$ cells will be ZsGreen$^+$. Only after N-Cad$^+$ EMT, Dre-rox recombination leads to switch of ZsGreen into tdTomato expression in targeted cells.

(**C**) Immunostaining for tdTomato and PECAM on E11.5 *Cdh5-CreER;N-cad-LSL-Dre;R26-RSR-tdTomato* embryonic section. This data is used as positive control to detect N-Cad$^+$ EMT in endocardial (endo.) cushion. Endocardial cells undergo EMT from E8.5 to E11.5 in the endo. cushion in developing heart. Tamoxifen was induced at E8.5, and embryos were collected at E11.5 for analysis.

(**D-G**) Immunostaining for ZsGreen, tdTomato and lineage markers K5, K14, K8 or K19 on mammary tissues of *N-cad-EMTracer* mouse. *N-cad-EMTracer* efficiently labels luminal epithelial cells by ZsGreen but not tdTomato, indicating no N-Cad$^+$ EMT program in normal mammary. Tamoxifen was induced at 6 weeks old and tissues were collected at 20 weeks old.

Scale bars, blue, 500 μm; white, 100 μm. Each image is representative of 5 individual biological samples.

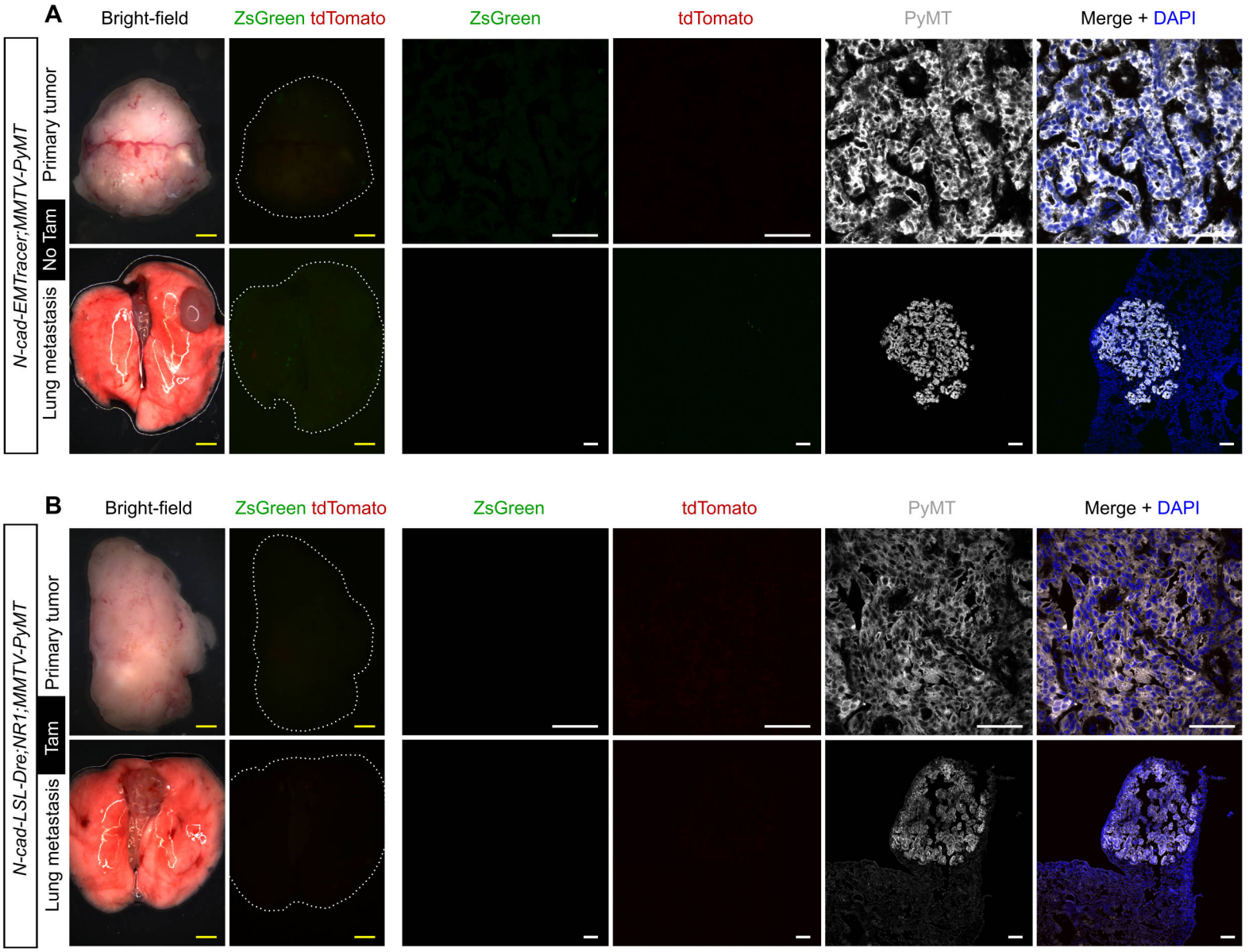


Figure S7 (Related to Figure 6). Technical controls for *N-cad-EMTracer;MMTV-PyMT* mouse

(**A**) Whole-mount images of mammary (up) and lung (below) of *N-cad-EMTracer;MMTV-PyMT* mouse at late stage, and immunostaining for ZsGreen, tdTomato and PyMT on mammary (up) and lung (below) sections. There is no leakiness for *N-cad-EMTracer;MMTV-PyMT* without tamoxifen induction. (**B**) Whole-mount images of mammary (up) and lung (below) of *N-cad-LSL-Dre;NR1;MMTV-PyMT* mouse at late stage, and immunostaining for ZsGreen, tdTomato and PyMT on mammary (up) and lung (below) sections. The mouse was treated with tamoxifen at 6-7 weeks. There is no leakiness for *N-cad-LSL-Dre;NR1;MMTV-PyMT*.

Scale bars, yellow, 2 mm; white, 50 μm. Each image is representative of 5 individual biological samples.