The Contribution of Partial and Full Epithelial-to-Mesenchymal Transition to Breast Cancer Progression

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Fabiana Maria Lüönd

aus Sattel, Schwyz

Basel, 2020

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von:

Prof. Dr. Gerhard Christofori
Prof. Dr. Verdon Taylor

Basel, den 17.12.2019

Prof. Dr. Martin Spiess
Dekan

Summary

Breast cancer is a leading cause of mortality in women worldwide. The lethality associated with the disease is primarily a consequence of the systemic spread of cancer cells and the formation of metastasis, particularly in the lung, liver, brain and bones. Multiple cellular and molecular mechanisms may contribute to the complex process of metastasis formation. In order to devise clinically effective strategies to combat metastatic disease, these need to be better understood.

Epithelial-to-mesenchymal transition (EMT) is a cellular program of trans-differentiation which is critical for embryonic development and wound healing but which may also contribute to the dissemination of cancer cells. During an EMT, epithelial cancer cells lose their cell-cell adhesion, dedifferentiate, and acquire a migratory and invasive phenotype. Experimental induction of an EMT is sufficient for metastasis formation. However, whether a spontaneous EMT is required for metastasis in an unperturbed system *in vivo* is a major focus of current research. Within recent years, it has become clear that EMT is not a binary switch but covers a spectrum of intermediate EMT hybrid states which differ in their functional characteristics and metastatic potential. However, due to the transient and reversible nature of the process, the extent to which cancer cells spontaneously undergo a partial or full EMT *in vivo* and the functional consequences regarding metastasis remain unknown.

With my PhD work, I have aimed at assessing the contribution of partial and full EMT to breast cancer progression and metastasis. To this end, I have established two novel color-switching lineage tracing models in transgenic mice. Based on an irreversible switch from mCherry to GFP expression, this model allows to visualize and track cancer cells that have undergone a partial or full EMT, even if they re-differentiate by undergoing a mesenchymal-to-epithelial transition (MET). I show that cancer cells mostly transition between epithelial/mesenchymal hybrid states but rarely undergo a full EMT. Furthermore, cells which have undergone a partial EMT are highly enriched in lung metastases compared to primary tumors. In particular, metastasis with a mosaic composition of mCherry and GFP positive cells are observed, pointing towards a collective dissemination of cells that have undergone a partial EMT together with epithelial cancer cells. In contrast, cells that have undergone a full EMT retain a more quiescent mesenchymal phenotype and do not colonize the lung. In conclusion, these

data suggest that although a full EMT may not be required, a partial EMT contributes to experimental breast cancer metastasis.

In addition, I have further characterized the mammary gland-specific flippase driver-line which we have generated for the lineage tracing experiments. These mice may serve as a versatile tool for studying mammary gland biology and breast carcinogenesis.

In summary, my PhD work provides novel insights into the dynamics of EMT and MET *in vivo*, as well as the contribution of partial and full EMT to breast cancer metastasis. Furthermore, our newly established mouse models offer novel opportunities to study the contribution of partial and full EMT towards distinct aspects of breast cancer progression *in vivo*.

Table of Contents

List of Abbreviations

α-SMA	α-smooth muscle actin
CK	Cytokeratin
cKO	Conditional knockout
CSC	Cancer stem cell
CTC	Circulating tumor cell
$CreER^{T2}$	Cre fused to mutated ligand-binding domain of the estrogen receptor
E-cadherin	Epithelial cadherin
ECM	Extracellular matrix
E/M	Epithelial/mesenchymal
EMT	Epithelial-to-mesenchymal transition
EpCAM	Epithelial cellular adhesion molecule
FACS	Fluorescence-activated cell sorting
Flp	Flippase
Flpo	Mouse codon usage-optimized Flp
Fsp1	Fibroblast specific protein 1
MET	Mesenchymal-to-epithelial transition
MMTV	Mouse mammary tumor virus
PyMT	Polyoma middle T
Tam	Tamoxifen
TGFβ	Transforming growth factor beta

1 General Introduction

1.1 Epithelial-to-Mesenchymal Transition

Epithelial-to-mesenchymal transition (EMT) is a transient and reversible program of dedifferentiation during which epithelial cells gradually lose their apical-basal polarity and cell-cell adhesion to acquire a motile and invasive mesenchymal phenotype [1]. EMT and the reverse process, a mesenchymal-to-epithelial transition (MET) are fundamental for embryonic development during which highly regulated rounds of dedifferentiation, cell migration and differentiation are required for tissue and organ formation. In the adult, EMT is activated as a physiological response to tissue injury and inflammation and contributes to wound healing [2]. However, aberrant activation of an EMT also contributes to disease. In a context of chronic inflammation with repeating tissue injury and repair, EMT contributes to organ fibrosis [3]. Finally, EMT may contribute to malignant transformation and cancer progression. EMT promotes cell survival and confers epithelial cancer cells with increased migratory and invasive capabilities and stem cell-like properties. This is most notably associated with metastasis and therapy resistance [4].

1.1.1 The Phenotypic Changes During an EMT

Epithelial and mesenchymal cells differ dramatically in morphology and functionality. Epithelial cells are organized in cell layers with apical-basal polarity and strong adherence to the basement membrane and to their neighboring cells. The migratory potential of epithelial cells is low. In contrast, the mesenchymal single cell state has much weaker adhesive properties and greatly increased migratory and invasive capabilities [5]. Therefore, a complete transition from an epithelial to mesenchymal phenotype (or vice versa) must involve drastic structural remodeling (Figure 1).

a) Epithelial junction disassembly and loss of cell polarity: Epithelial cells are connected to their neighboring cells via tight junctions, adherence junctions and desmosomes. Furthermore, gap junctions form channels between adjacent cells and allow the exchange of small molecules. During an EMT, these types of cell-cell

junctions are disassembled and their respective junctional proteins are delocalized or degraded. Tight junctions are localized closely to the apical surface to seal adjacent cells. Claudins and occludins are part of the intercellular tight junction strands and are linked to the cytoplasmic proteins zona occludens 1 (ZO1), ZO2 and ZO3 which form a link to the actin cytoskeleton to stabilize the junctions [5, 6]. During an EMT, expression of claudins and occludins is repressed, and ZO1 translocates to the cytoplasm. Adherence junctions are mediated by cadherin proteins which form calcium-dependent homophilic interactions. Epithelial cadherin (E-cadherin), in a complex with β-catenin, α-catenin and p120 catenin, is connected to the actin cytoskeleton and is indispensable for epithelial cell adhesion [7]. The loss of E-cadherin is a hallmark of an EMT and is mediated by transcriptional repression, delocalization and degradation [8]. During later stages of an EMT, cells may express N-cadherin which contributes to mesenchymal cell motility, dissemination and invasion. This 'cadherin-switch', is another hallmark of EMT [7, 9, 10]. Finally, desmosomes consisting of desmosomal cadherin, desmogleins, desmocollins, plakoglobins and plakophilins are connected to cytokeratin intermediate filaments [5] and are disrupted during an EMT [11]. It has been proposed that a sequential disassembly of tight junctions, adherence junctions and desmosomes during an EMT progresses via intermediate states resembling an "unzipping" starting from the apical compartment [12].

Apical-basal polarity of epithelial cells is maintained by antagonistic interactions of the cell polarity complexes crumbs, partitioning-defective (Par) and scribble. Crumbs and Par associate with tight junctions and define the apical compartment while scribble depends on E-cadherin and defines the basolateral compartment [13]. Consequently, disassembly of epithelial junctions leads to disruption of cell polarity and vice versa which allows cells to adopt front-rear polarization that is critical for single cell migration.

b) Cytoskeleton remodeling and motility: Adherence junctions connected to actin form a tight "adhesion belt" linked to a cortical actin skeleton in epithelial cells. During an EMT, the actin cytoskeleton is reorganized into a network of highly dynamic and contractile actin stress-fibers that are critical for directional migration of mesenchymal cells [8]. Stress-fibers are connected to focal adhesions, which are complexes of clustered integrin receptors and associated proteins that link the cell to the ECM and serve as traction sites during cell migration. Briefly, mesenchymal cell

migration is characterized by generation of cell protrusions, adhesion sites and traction force at the leading edge followed by cell contraction in an oscillatory manner, which protrudes the cell forwards [14, 15]. Small Rho GTPase family members, particularly RhoA, Rac1 and Cdc42 are crucial for these processes. RhoA induces stress-fiber formation, while Rac1 and Cdc42 are involved in the formation of migratory membrane protrusions, such as lamellipodia and filopodia, respectively, and the organization of focal adhesion sites [16, 17]. The intermediate filament system which enhances the cells mechanical strength and mediates trafficking of organelles and membrane proteins is also remodeled during an EMT. Epithelial cytokeratin is exchanged with vimentin filaments, which is crucial for mesenchymal cell shape, motility and adhesion [5, 18].

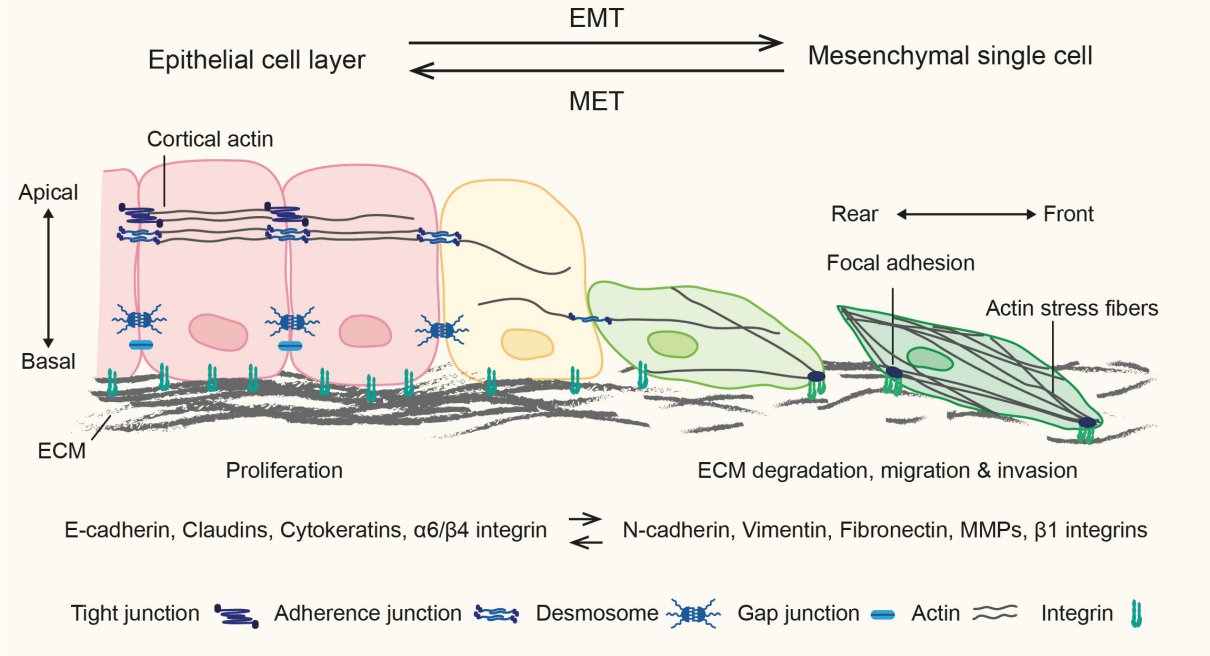


Figure 1: The phenotypic changes during an EMT. Proliferative epithelial cells are organized in cell layers characterized by strong cell adhesion mediated by tight junctions, adherence junctions and desmosomes. They display a cortical actin cytoskeleton, apical-basal polarization and adhere to the ECM via integrins. During an EMT, epithelial cell markers (E-cadherin, claudins...) are downregulated, while mesenchymal cell markers (N-cadherin, vimentin…) are expressed. Epithelial cell adhesion is lost and the actin cytoskeleton is reorganized into stress-fibers. Cells adopt front-rear polarization and attach to the ECM via dynamic focal adhesions allowing cell migration. Mesenchymal cells degrade the ECM by secretion of proteases (e.g. MMP) and invade the surrounding tissue.

c) Extracellular matrix (ECM) remodeling and invasion: The ECM is mainly composed of proteoglycans and fibrous proteins, such as collagens, fibronectins and

laminins, which form the interstitial matrix and basement membrane. The ECM is constantly remodeled and cues from the ECM are important regulators of many cellular processes, such as proliferation and migration [19]. The adhesion to and signal transduction from the ECM to epithelial cells is mediated by integrins which are heterodimeric type-I transmembrane proteins consisting of an α and β-chain. The combination of α and β-chains determines the specificity for a certain ECM component. During an EMT, epithelial integrins are downregulated and mesenchymal integrins are newly expressed. These mesenchymal integrins may critically contribute to EMT and cell migration [8]. For example, epithelial α6/β4 integrins which connect to laminin are downregulated during an EMT [20] and replaced by β1 integrins, such as α5/β1. Mesenchymal α5/β1 integrins connect to fibronectin and are involved in dynamic focal adhesion assembly and cell motility [21, 22]. EMT also changes the ECM composition by inducing the secretion of ECM components, including fibronectin, and the expression of proteases, particularly matrix metalloproteinases (MMPs). MMPs facilitate invasion by degrading ECM components. Additionally, MMPs cleave cell surface receptors and release growth factors that have been trapped in the ECM, thereby altering cell signaling. Thus, MMPs may act as facilitators of an EMT [23].

Taken together, the structural rearrangements during an EMT eventually lead to the emergence of elongated, spindle-shaped migratory cells. These cells have newly gained the capabilities required for ECM degradation and local invasion.

1.1.2 Induction and Molecular Regulation of EMT/MET Programs

A variety of signaling pathways can trigger an EMT, including TGFβ/Smad, MAPK, PI3K, Wnt/β-catenin, Notch, JAK/STAT, NFκB and Hippo. Contextual signals determine the activity and crosstalk of these pathways which control cell survival, proliferation and the phenotypic changes observed during an EMT [5]. In cancer, many EMT-inducing growth factors, such as TGFβ, HGF, EGF, FGF, PDGF and IGF, are secreted by tumor associated stromal cells, e.g. by cancer associated fibroblasts. An inflammatory tumor microenvironment further contributes to EMT due to secretion of cytokines, particularly TNFα, IL6 and IL8, by immune cells. Furthermore, hypoxia is a potent EMT-inducing stimulus [24, 25], and cues from the ECM such as matrix-stiffness

may also trigger an EMT [26]. Besides signals from the stroma, also cell and tissue architecture may control an EMT. It has recently been shown that apical-basal cell polarity prevents EMT, while disruption of cellular polarity induces an EMT [27, 28]. Finally, loss of E-cadherin, e.g. by mutation, can promote an EMT [29-31]. In fact, E-cadherin expression or function is frequently lost in carcinomas, particularly in invasive lobular breast cancer [32]. However, the sole loss of E-cadherin is not sufficient to induce a full EMT and its overexpression does not induce a MET [33-35]. This indicates that additional modifications are required for the execution of a complete cellular reprogramming.

Independent of the inducing stimuli, EMT programs are regulated at the transcriptional, post-transcriptional and post-translational levels. This involves transcriptional activation/repression, epigenetic modifications, mRNA degradation, alternative splicing and protein degradation or relocalization. Numerous regulatory mechanisms cooperate in a dynamic network and in a context dependent manner (Figure 2).

a) Transcriptional regulation: Completion of a full EMT involves profound transcriptional changes. These are largely mediated by a core regulatory network of EMT-inducing transcription factors of the Snail (SNAI1, SNAI2), zinc-finger E-box-binding (Zeb: ZEB1, ZEB2) and basic helix-loop-helix (bHLH: TWIST1, TWIST2, E47, Id HLHs) families which are upregulated early during an EMT. These transcription factors act as repressors of epithelial genes, such as E-cadherin, claudins and cytokeratins, and as activators of mesenchymal genes, such as N-cadherin, Vimentin and MMPs [36]. Snail and Zeb are particularly strong transcriptional repressors of epithelial genes, while Twist1/2 are activators of the mesenchymal transcription program. Yet all of them share a set of common target genes and may regulate the expression of each other. SNAI1 is a common target-gene of various EMT-inducing stimuli and induces ZEB1 and TWIST1, consistent with their frequently observed sequential activation [37, 38]. The most significant target of transcriptional repression during an EMT is E-cadherin. Snail, Zeb and Twist in association with distinct sets of histone-modifiers repress E-cadherin via binding to E-box sequences in the proximal promoter [5]. Hence, progressive epigenetic remodeling may lead to long-term silencing and robust repression of E-cadherin, while cells acquire a stable

mesenchymal phenotype [39]. The expression of Snail, Zeb and Twist family members is tightly controlled. In epithelial cells their expression is repressed by “epithelial” transcription factors ELF5 [40], GRHL2 [41] and OVOL1/2 [42]. These transcription factors also act as drivers of a MET. In addition to Snail, Zeb and Twist, several other transcription factors contribute to different aspects of EMT and MET. These include amongst others Prrx1 [43], Sox4 [44], Sox9 [45, 46], Klf4 [47], Foxc2 [48], Foxf2 [49], Ap-1 [50], Yap/Taz/Tead [51-53], Smad/Hmga2 [54, 55] and β-catenin/Tcf/Lef [56, 57]. Many of their distinct functional contributions and complex interactions are just being unraveled.

b) post-transcriptional regulation: Several micro RNAs (miRNAs) are known to target key-components involved in EMT [58]. The master-regulators Snail and Zeb are engaged in double-negative feedback loops with members of the miR34 and miR200 families, respectively. miR34 targets Snail by whom it is transcriptionally repressed [59] and Zeb represses miR200 family members [60, 61] of which some target Zeb [62-64]. These regulatory feedback-loops are critical for fine-tuning EMT and MET dynamics. Similarly to miRNAs, certain long non-coding RNAs (lncRNAs) may regulate EMT [65]. Finally, alternative splicing contributes to EMT by generating epithelial or mesenchymal protein isoforms [66, 67].

c) post-translational regulation: Many EMT inducing stimuli and their downstream signaling pathways converge on EMT programs via post-translational modifications. The stability and localization of structural proteins and EMT transcription factors are frequently regulated by phosphorylation and ubiquitylation. Snail, when phosphorylated by GSK3β, a downstream target of Wnt and PI3K signaling, is primed for ubiquitylation and degradation [68]. Phosphorylation of Snail by PAK1 or LATS2, however, increases its nuclear retention to promote an EMT [5, 69, 70]. Similarly, Twist is stabilized upon phosphorylation by MAPK [71, 72]. Another prominent target of post-translational modification is E-cadherin. Degradation of E-cadherin and dissolution of adherence junctions can for example be triggered Src-mediated phosphorylation which marks it for ubiquitylation [73, 74], by direct cleavage by MMPs [75], or by the disruption of the E-cadherin/catenin complex due to destabilizing phosphorylation of β-catenin by mediated by FAK or EGFR [76, 77].

As described above, there is extensive crosstalk between signaling pathways and amongst the multiple layers of EMT regulatory networks. Usually, activation of one pathway is not enough to induce an EMT (Figure 2).

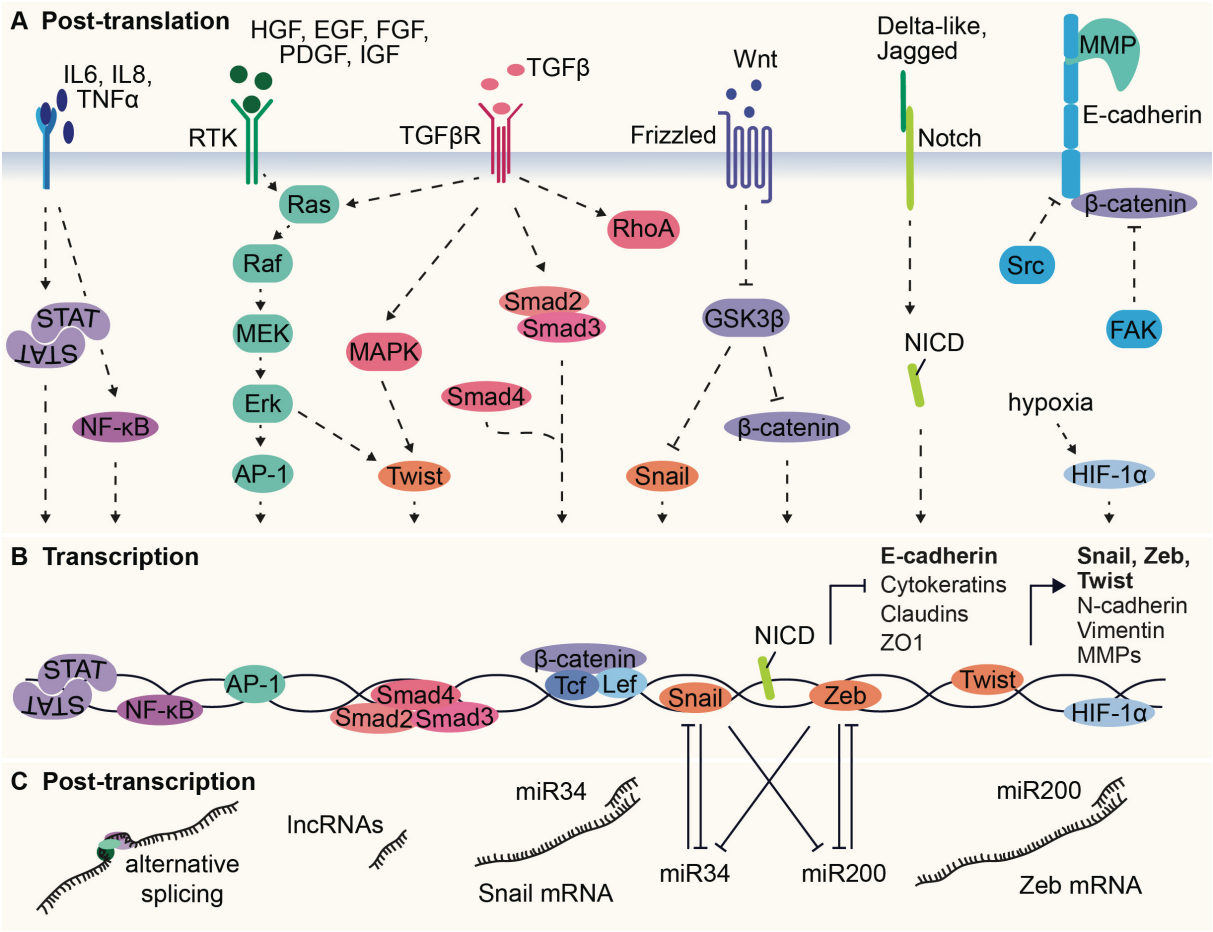


Figure 2: Molecular regulation of EMT. EMT is regulated on the transcriptional, post-transcriptional and post-translational level. **A)** Ligands secreted by stromal cells activate signaling pathways which trigger an EMT mostly via activation of EMT-inducing transcriptional regulators. TGFβ, the most prominent inducer of an EMT, acts via canonical Smad2/3 signaling and several non-canonical pathways e.g. MAPK which promotes cell survival and stabilization of Twist. Activation of RhoA by TGFβ induces cytoskeleton remodeling. Canonical Wnt signaling induces an EMT by stabilization of β-catenin and Snail. Dissociation of adherence junction is promoted by destabilizing phosphorylation of E-cadherin and β-catenin or cleavage of E-cadherin by MMPs. **B)** EMT is regulated on the transcriptional level by repression of epithelial-specific genes, activation of EMT-inducing transcription factors Snail, Zeb, Twist and expression of mesenchymal-specific genes. **C)** On the post-transcriptional level, alternative splicing and non-coding RNAs contribute to EMT. miRNAs of the miR34 and miR200 families are engaged in double-negative feedback loops with Snai1 or Zeb1, respectively, thereby being involved in the core regulatory network governing EMT.

1.1.3 TGFβ and EMT in Cancer

TGFβ is the most potent inducer of EMT in cancer and contributes to various aspects of tumor progression. TGFβ activates the canonical TGFβ/Smad signal-transduction as well as several non-canonical pathways (Figure 2). Briefly, binding of TGFβ1, 2 or 3 to the TGFβ receptor (TGFβR) causes receptor autophosphorylation, followed by the recruitment and subsequent phosphorylation of Smad2/3. The Smad2/3 transcription factors associate with Smad4 and translocate to the nucleus to collaborate with other transcription factors, such as HMGA2 to induce an EMT. TGFβR also activates Smad-independent, non-canonical signaling-pathways, including p38, Jnk and Erk-MAPK pathways as well as PI3K/mTor signaling, which promotes cell survival and proliferation. Furthermore, TGFβ activates RhoA which leads to actin cytoskeleton remodeling and formation of lamellipodia and filopodia. Phosphorylation of Par6, a component of cell polarity complexes, by TGFβR causes loss of cell polarity and disassembly of cell adhesion complexes [6, 78]. Hence, TGFβ directly impinges on core transcriptional circuits regulating EMT and contributes to the phenotypic remodeling observed during the EMT on the post-translational level.

However, the cellular response to TGFβ is context-dependent, and TGFβ has paradoxical roles in cancer. In fact, TGFβ has strong cytostatic and apoptotic effects and thus acts as a tumor suppressor. In epithelial cells, Smad3/4 in association with C/EBPβ induces expression of the Cdk inhibitors p21 and p15INK4b and represses c-MYC, thereby causing G1 arrest. Once cancer cells inactivate or escape from these tumor-suppressive signals, TGFβ promotes tumor progression [78]. This may be achieved by loss of function mutations in components of the canonical TGFβ-pathway as frequently observed in aggressive carcinomas. TGFβRI and TGFβRII mutations frequently occur in colon cancer and Smad4 is mutated in 50% of pancreatic and 30% of metastatic colon cancers [79]. However, in most cases, only the tumor-suppressive branch is circumvented. In metastatic breast cancer, the dominant-negative C/EBPβ isoform LIP which inhibits the transcriptionally active isoform LAP has been found overexpressed [78, 80, 81]. Importantly, Snail1 has been observed to block TGFβ-induced apoptosis while also promoting EMT [82].

It has been assumed that upon evasion from TGFβ-mediated growth inhibition, cancer cells may undergo a TGFβ-induced EMT which further promotes cell-survival

and tumor progression. Interestingly, not all cell lines which have overcome TGFβ-mediated growth inhibition may undergo a TGFβ-induced EMT *in vitro*, indicating that cell intrinsic properties determine the ability of cells to undergo an EMT [83]. Furthermore, abrogation of the major branch of TGFβ signaling by loss of Smad4 does not seem reconcilable with the tumor-promoting effects associated with TGFβ and EMT. Interestingly, David et al. revealed that in a certain context, TGFβ could still induce an EMT in Smad4 mutant murine pancreatic ductal adenocarcinoma (PDAC) cells. This EMT was mediated via Smad2/3-dependent but Smad4-independent activation of Sox4. In Smad4 wildtype PDAC cells, however, TGFβ-dependent activation of Snail induced a conversion of Sox4 from a pro-tumorigenic to a pro-apoptotic factor, resulting in the induction of a lethal EMT. These data indicate that EMT does not always promote cell survival [84]. Taken together, this highlights once more the context-dependent functions of TGFβ signaling and EMT programs in cancer.

Besides its effects on cancer cells, TGFβ also exerts pro- and anti-tumorigenic effects on the stroma. TGFβ has been linked to immunosuppression and angiogenesis and increased production of TGFβ correlates with poor prognosis. However, the manifold consequences of TGFβ signaling, EMT and underlying mechanisms of TGFβ's paradoxical function in cancer progression are still not fully understood [79].

1.1.4 Molecular Basis Underlying EMT Hybrid States

Due to the variety of environmental stimuli which precipitate an EMT and its highly dynamic regulatory networks, EMT programs differ vastly depending on the environmental context [1]. Importantly, EMT is not an all or nothing response but involves sequential transcriptional reprograming and remodeling of the cell's architecture. Cells may undergo only a partial EMT and MET as opposed to acquiring a completely mesenchymal phenotype [85-91]. Consistent with this, cancer cells in an epithelial/mesenchymal (E/M) hybrid state that have retained part of their epithelial characteristics and simultaneously express certain mesenchymal markers have been observed *in vitro* and *in vivo* (reviewed in [91]).

To date, little is known about the spectrum of intermediate EMT states, the stability of these transition states and their functional characteristics. Based on

mathematical modelling and simulations focused on the core EMT regulatory feedback-loops of Snail/miR-34 and Zeb/miR-200, several groups predicted a two-step progression of EMT via three distinct cell phenotypes: epithelial, intermediate EMT and mesenchymal [85, 88, 92, 93]. Modeling based on a slightly extended core network including the mutually inhibitory feedback-loop between Ovol2 and Zeb1 predicted an additional intermediate step [89]. However, these studies were limited to the most basic core regulatory mechanisms with vastly reduced complexity. A comprehensive study by Meyer-Schaller et. al. which integrated transcriptional profiles of EMT-kinetics and functional perturbation of a network of 46 transcription factors and 13 miRNAs essential for EMT has shed light on the hierarchical organization and functional interactions of multiple EMT regulatory networks during a TGFβ-induced EMT *in vitro*. Multiple transition states (at least 3-4 hybrid states) with distinct gene expression profiles and gene regulatory networks could be identified [86].

Finally, Pastushenko and colleagues could recently prove the existence of distinct intermediate hybrid states in a transgenic mouse model of skin squamous cell carcinoma *in vivo*. Based on the differential expression of four cell-surface markers (EpCAM, CD106, CD51 and CD61) and combinations thereof, the authors could identify 7 distinct cell states along an epithelial-mesenchymal gradient ($EpCAM^+$ epithelial plus 6 $EpCAM^-$ hybrid/mesenchymal populations). These exhibited distinct functional characteristics in regard of migration, invasion, metastasis and stemness and were associated with transition state-specific changes in the chromatin and transcriptional landscape [87].

Taken together, EMT and MET allow cells to transition between a spectrum of phenotypic hybrid states, therefore constituting programs of cellular plasticity rather than acting as binary phenotypic switches [94]. Yet, as discussed in the following chapters, little is known about the contribution of different EMT states to distinct aspects of tumor progression.

1.2 The Relevance of EMT in Tumor Progression

Cancer is a leading cause of mortality worldwide with currently 18.1 million new cases and 9.6 million cancer deaths per year. Carcinomas, which are of epithelial origin, are the most common type of cancer. In women, breast cancer is the most prevalent malignancy and the leading cause of cancer-related death [95]. The high lethality of carcinomas is usually a consequence of metastasis and therapy resistance, which have both been linked to EMT. Historically, EMT has been studied by experimental manipulation of the process and shown to contribute to almost all aspects of tumor progression [94, 96, 97]. However, less is known about the prevalence of spontaneous EMT in different cancer types *in vivo* and the functional relevance of EMT for certain aspects of malignancy.

1.2.1 Evidence for and Prevalence of EMT in Carcinomas

a) detection of EMT in human samples: To detect EMT in human tumors, scientists have relied on the expression of markers specific to the epithelial or mesenchymal phenotype. (Co-)stainings of epithelial and mesenchymal markers have correlated EMT signatures with poor prognosis for prostate, lung, liver and breast cancer already 10 years ago [98-101]. Most notably, co-expression of both epithelial and mesenchymal-specific markers by tumor cells, which is indicative of an E/M hybrid state, correlates with poor prognosis in breast cancer [102, 103]. More recently, the existence of cells in an intermediate EMT state could be proven by single cell-RNA-sequencing analysis [104].

b) detection of EMT facilitated by lineage tracing in murine models: Visualization of EMT based on EMT marker expression has certain constrains. Most importantly, cancer cells which have acquired a fully mesenchymal phenotype cannot be distinguished from cancer-associated fibroblast based on cell morphology or commonly used EMT-markers. Hence, co-staining with oncogenes or other cancer cell-specific genes is required, which is frequently not feasible. To circumvent this challenge, transgenic lineage-tracing models for fluorescent labeling of cancer cells have been generated to study EMT in breast, pancreatic and colorectal cancer as well

as skin squamous-cell cell carcinoma [87, 105-111] (Figure 3). Using a Pdx1-Cre, KRasG12D, p53 cKO-driven mouse model of PDAC in combination with a Rosa26-lox-stop-lox-YFP reporter for labeling of cancer cells, Rhim et al. observed up to 42% of cancer cells in an EMT state based on either loss of E-cadherin or expression of mesenchymal markers Zeb1 or fibroblast specific protein (Fsp). Interestingly, only 10% of cancer cells co-expressed E-cadherin together with either one of the mesenchymal markers which is indicative of vast heterogeneity of cells residing in an EMT hybrid state [107].

However, these lineage tracing approaches are limited to a “snap-shot” of cells undergoing an EMT at the time of sacrifice and do not include cells that have previously completed an EMT and MET. Hence, novel lineage tracing systems relying on mesenchymal-specific markers have been created to capture all cancer cells that have ever undergone an EMT as well as the progeny thereof [105, 112-116] (Figure 3). To shed light on the extent of EMT occurring in three different breast cancer mouse models, Trimboli and colleagues used the mesenchymal-specific Fsp-Cre in combination with a Rosa26-lox-stop-lox-lacZ reporter. In Myc-driven models, approximately 50% of mice presented with cancer cells that had undergone an EMT identified by means of partially epithelial morphology and lacZ expression. In contrast, hardly any EMT could be observed in mouse mammary tumor virus-polyoma middle T (MMTV-PyMT) and MMTV-Neu-driven breast tumors, indicating that EMT might be associated with distinct oncogenic events [105]. However, the single use of distinct mesenchymal markers to detect EMT has some limitations. Again, cancer cells may not be clearly distinguished from stromal cells and particularly Fsp is highly expressed in fibroblasts. Hence, all cancer cells which completely lack any traces of epithelial morphology are likely dismissed as stromal cells. Furthermore, mesenchymal cell markers are differentially expressed between distinct EMT states. Hence, markers such as Fsp might not capture all cells undergoing an EMT.

Based on the conflicting data originating from previous work, next generation lineage tracing strategies allowing cancer cell-specific EMT tracing in combination with a variety of markers for different EMT transition states are required for identifying and characterizing EMT *in vivo*.

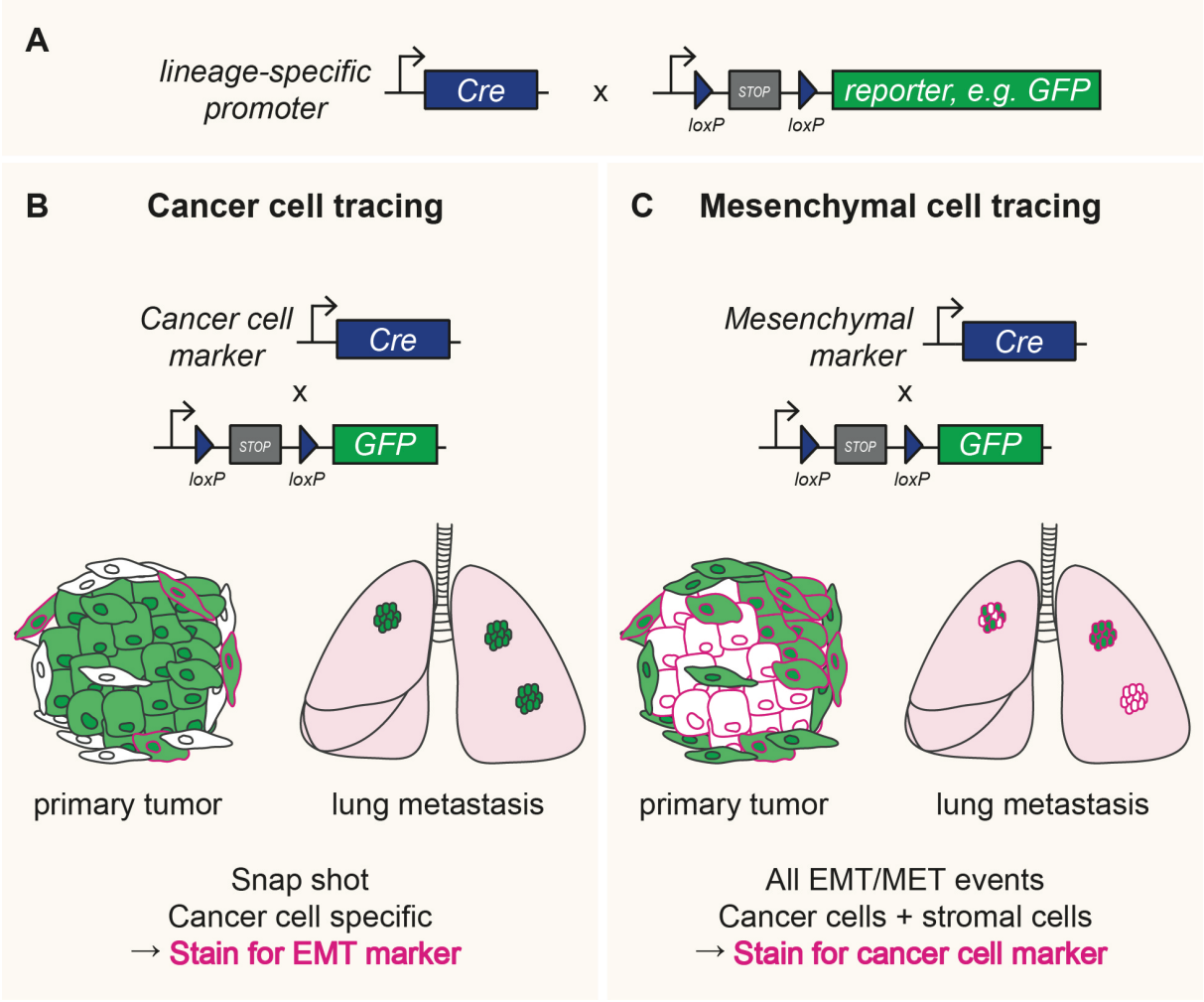


Figure 3: Lineage tracing strategies to visualize EMT in transgenic mice. A Basic lineage tracing approach: A promoter specifically active in the cell lineage of interest drives the expression of a recombinase (e.g. Cre) to recombine a conditional reporter allele (e.g. loxP-stop-loxP-GFP). **B)** Cancer cell-specific lineage tracing: A cancer-cell-specific Cre allele in combination with a floxed GFP reporter allele results in irreversible labelling of cancer cells with GFP. Staining of EMT markers allows identification of tumor cells in an EMT state but not detection of cells which have previously undergone EMT and MET, e.g. in metastasis. **C)** Mesenchymal-specific lineage tracing leading to irreversible GFP expression not only by all cancer cells which have ever undergone an EMT but also by stromal cells. This allows to capture all EMT/MET events and visualization of the contribution of EMT to metastasis. Staining for a cancer cell-specific marker is required to distinguish mesenchymal cancer cells from stromal cells.

c) EMT across different carcinomas: Cancer is a heterogeneous group of diseases and its manifestation greatly depends on the tissue of origin. Not surprisingly, the extent to which an EMT occurs varies between cancers of different tissues [117]. Cancers of the same tissues are extremely heterogeneous as well, which is in part consequence of distinct mutations underlying the individual tumor as well as the characteristics of the tumor cell of origin. As mentioned above, much higher frequencies of EMT were observed in Myc-driven compared to PyMT or Neu-driven mammary tumors, indicating that certain mutations may be genetic predispositions for

EMT [105]. Yet, the epigenetic landscape of the cell of origin seems to be important as well [111, 118, 119]. Latil and colleagues elegantly demonstrated that the cancer cell of origin predisposes cells towards undergoing an EMT using mouse models of skin squamous-cell carcinoma. When oncogenic KRas and deletion of p53 were targeted either to cells of the interfollicular epidermis (IFE) or hair follicle (HF) stem-cells, tumors arising from HF stem-cells frequently exhibited features of EMT and were highly metastatic, while transformation of IFE cells gave rise to well differentiated tumors. Transcriptional and epigenomic profiling revealed that the chromatin landscape of HF stem-cells primes cancer cells towards undergoing an EMT [111]. Similarly, introduction of the oncogenic PIK3CA into mammary epithelial cells of basal or luminal lineages induces the emergence of a variety of distinct breast-cancer subtypes. Interestingly, PIK3CA-expressing luminal lineage cells gave rise to a much higher number of aggressive carcinomas, including metaplastic carcinoma with mesenchymal features [118, 119]. These data indicate that complex mutational events as well as the cancer cell of origin contribute to the highly variable extent of EMT observed in human and murine carcinomas.

d) EMT in breast cancer subtypes: Human breast cancers are a diverse group of tumors. This heterogeneity is a hurdle for therapy and is associated with a wide range of clinical outcomes. Several classification strategies are currently in use to predict prognosis and determine the optimal therapy on the individual's level.

Historically, invasive breast cancer has been divided into different subtypes based on histological morphology. Invasive carcinoma of no special type - previously termed invasive ductal carcinoma - is the predominant type. Less frequent subtypes include invasive lobular carcinoma and the highly dedifferentiated, mesenchymal metaplastic carcinomas [120].

Immunopathological classification, which is based on the expression of estrogen receptor (ER) and progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2, also referred to as Neu or Erbb2), is indicative for certain therapies [121]. Overall, ER and PR-positive tumors are epithelial and well differentiated, they tend to respond to endocrine therapy and have a rather good prognosis. HER2-positive tumors frequently harbor HER2 amplifications or mutations. These tumors are less differentiated and more aggressive but can be targeted by HER2-specific antibodies or tyrosine kinase inhibitors. Finally, triple-negative breast

cancers (TNBCs) which do not express any of these three receptors are the most dedifferentiated subtype with the worst prognosis due to the lack of specific treatment.

Recent technological advances have led to the molecular characterization of breast cancers with currently six subtypes: luminal A, luminal B, HER2-enriched, basal-like, claudin-low and normal breast-like [122, 123]. The luminal A and B subtypes are mostly ER-positive and of well-differentiated, epithelial morphology. Basal-like and claudin-low signatures overlap and these tumors are frequently triple-negative. Both subtypes are highly aggressive with unfavorable prognosis. The claudin-low subtype is characterized by an EMT gene-expression signature with low expression of epithelial claudins, occludin and luminal markers cytokeratin 8/18. Compared to the basal-like subtype it has a more pronounced EMT-signature, lower expression of proliferative genes and has been associated with stem cell-like properties [124]. As expected, metaplastic carcinomas are frequently triple-negative and characterized as basal-like or claudin-low [125]. Surprisingly, invasive lobular carcinoma which is most frequently associated with the loss of E-cadherin, mostly correlates with the well-differentiated, epithelial luminal A subtype. This confirms again that the loss of E-cadherin is not enough to induce an EMT [32].

While the claudin-low and the basal-like subtype are strongly associated with EMT, the underlying reasons are not well understood [35, 126, 127]. It has been observed that in basal-like but not luminal breast cancer cell lines the Zeb1 promoter is kept in a bivalent state allowing these cells to readily undergo an EMT upon an inducing stimulus, thereby being poised towards an EMT [128]. As discussed above, the epigenetic landscape of the tumor cell of origin influences the cell-intrinsic ability to undergo an EMT. Based on correlation of breast cancer gene-expression profiles with the signatures of mammary stem cells (MaSCs) and progenitor cells, it has been proposed that oncogenic events in MaSCs might give rise to claudin-low tumors, while luminal progenitor cells might be the origin of basal-like tumors. The rather epithelial breast cancer subtypes could be more closely matched to mature luminal cells [123]. However, whether the origin of cancer really includes the stem-cell compartment is still heavily debated. Importantly, oncogenic events might have an impact on differentiation and some oncogenes may even induce multipotency [118, 119]. Indeed, the mutational landscape has been found to differ between distinct breast cancer subtypes [129-131]. Particularly, Brca1 and p53 mutations frequently co-occur and are associated with the basal-like subtype [123, 132, 133]. Whether the combination of these mutations is

causally involved in the observed EMT-phenotype is currently unknown, yet both p53 and Brca1 have been linked to EMT [134-136]. Finally, the molecular signatures frequently change during cancer progression and may be affected by therapy as well. Interestingly, the claudin-low signature was found enriched in tissue remaining after endocrine and chemotherapy [137]. Whether this is due to intrinsic resistance and stemness of a minority of cells existing pre-treatment or an adaptive response, e.g. related to an EMT during therapy remains to be determined.

Taken together, a variety of factors influences the extent and functional relevance of EMT in certain cancer subtypes. Furthermore, there is extensive evidence linking EMT with aggressive disease. Yet, as discussed within the next chapters, the functional relevance of EMT towards distinct aspects of tumor progression needs further evaluation.

1.2.2 EMT in the Early Stages of Tumorigenesis and Cancer Stem Cells

EMT has traditionally been associated with late tumor stages and metastasis and is predominantly observed at the invasive front [106, 138, 139]. However, accumulating evidence suggests that EMT might actively promote malignant transformation and tumor development. In line with this, mesenchymal cancer cells have been observed already in premalignant lesions of a transgenic PDAC mouse model, although to a lesser extent compared to later stages of disease [107]. In particular, the EMT-inducing transcription factors Snail, Zeb and Twist have intrinsic oncogenic functions and confer cells with abilities required to overcome oncogene-induced senescence and apoptosis, which is critical for tumor initiation [82, 140]. Furthermore, EMT promotes the expression of cytokines involved in immunoregulation, thereby protecting tumors from immune attacks [4, 82]. Taken together, EMT constitutes a mechanism of cell survival thereby contributing to tumorigenesis.

Stemness is critical for tumor initiation, progression and particularly metastatic outgrowth [141]. The cancer stem cell (CSC) model proposes that tumor growth and progression are driven by a subpopulation of CSCs which have the ability to self-renew and differentiate, thereby giving rise to the bulk of tumor cells. CSCs are not

necessarily a rare population and certain tumors such as claudin-low breast cancer exhibit a high degree of stemness [124]. Although it has been assumed that CSCs originate from transformed stem or progenitor cells, increasing amount of evidence supports the idea that CSCs may be derived from "regular" cancer cells by dedifferentiation programs such as EMT [24]. Cancer cells that have undergone an EMT frequently exhibit features of stemness [24, 141, 142]. Breast cancer cells, after having undergone an EMT, were found to express high levels of CD44 and low levels of CD24 ($CD44^+/CD24^-$), a combination which is associated with both normal mammary gland and breast CSCs [143-146]. However, stemness is not a unique feature of the mesenchymal phenotype and CSCs represent a heterogeneous population of epithelial and mesenchymal phenotypes. [147-149]. Interestingly, one study reported how epithelial CSCs marked by aldehyde dehydrogenase (ALDH) could acquire a $CD44^+/CD24^-$ mesenchymal CSC phenotype and vice versa, indicating that stem-cells could possibly transition between epithelial and mesenchymal states while retaining their stemness [148]. Furthermore, several studies suggest that cells in a fully mesenchymal state are less stem cell-like and MET might actually confer them with the stemness required for metastatic outgrowth [43, 150]. These studies suggest that stemness could be associated with epithelial-mesenchymal plasticity or certain intermediate EMT states rather than with the mesenchymal phenotype itself.

1.2.3 EMT and Metastasis

Once tumor cells disseminate and seed distant metastasis, cancer turns into a systemic, mostly terminal disease. More than 90% of cancer-related deaths are due to metastasis [151]. While malignant transformation and tumor progression depend on the sequential accumulation of mutations, metastasis frequently relies on epigenetic changes and has been closely linked to EMT [4]. Metastasis is a complex, multistep process that includes migration and invasion of cancer cells into the surrounding tissue, intravasation, survival in the circulation, extravasation and metastatic outgrowth. These sequential processes are collectively termed "the invasion-metastasis cascade" (Figure 4) and EMT/MET may contribute to certain aspects of every single step [151].

a) Migration and Invasion: The first step of tumor cell dissemination is the detachment from the primary tumor and invasion into the surrounding tissue. This usually involves the degradation of the basal lamina and requires epithelial cancer cells to acquire migratory and invasive capabilities. In general, cancer cells can migrate and invade as single cells or as a collective group [14, 152, 153].

A full EMT generates motile and invasive mesenchymal single cells. **Mesenchymal cell migration** and invasion is characterized by integrin mediated focal adhesion and degradation of the ECM by proteases secreted by the cell. If multiple cells undergo a full EMT, a leader cell may initially form a track through which other cells may follow. Such “multicellular streaming” can typically be observed in invasive lobular breast cancer in which context pathologists refer to it as “Indian Files” [14, 154]. Besides the mesenchymal type of migration, single cells may adopt an **amoeboid migration** mode by undergoing a process termed mesenchymal-to-amoeboid transition (MAT). This type of migration is characterized by low adhesion and high contractility. Cells adopt a spherical shape lacking actin stress-fibers and focal adhesions and squeeze themselves through gaps in the ECM either as single cells or multicellular streams. Amoeboid migration tends to occur in the absence of proteolytic ECM degradation and is faster than mesenchymal migration. It mostly occurs in cancers of haemopoietic or neuroectodermal origin. Like EMT, also MAT is a reversible process allowing single cells to switch between modes of migration [14].

Collective cell migration and invasion is characterized by coordinated movement of cohesive groups of cells with intact cell-cell junctions. Depending on cell number, differentiation and tissue structure, cell collectives may move as small clusters or solid strands [155]. In many cases, one or several mesenchymal leader cells at the tip of the strands direct the movement and mediate degradation of the ECM. Although direct evidence is limited, a partial EMT program could allow cells to become motile while retaining cell-adhesion, possibly driving collective cell migration [108, 155-159]. Whether an EMT is required at some point for collective cell migration is still not clear. Cheung and colleagues recently identified a basal epithelial program in leader cells of collectively migrating cells and circulating tumor cell (CTC) clusters [160, 161]. These cells express E-cadherin and the basal-epithelial marker cytokeratin 14 but none of the conventional mesenchymal cell-markers. Similarly, Wicki et al. described how podoplanin expression induces collective cell migration in the absence of EMT [162]. It is however possible that a “hidden” partial-EMT program contributed to migration in

these cases. Finally, activated stromal fibroblasts may remodel the ECM to pave the way for collectively migrating cancer cells in the absence of a leader cell that has undergone an EMT [163-165].

Whether cancer cells disseminate as single cells or clusters, initially depends on microenvironmental factors, particularly tumor-stroma interactions, ECM composition and matrix stiffness. Stromal cells at the invasive front secrete a variety of growth factors and cytokines. These may act as chemoattractant, may induce EMT and promote migration [166]. Importantly, cells may switch between different modes of collective and single cell migration in response to microenvironmental conditions and upon changes in cell intrinsic signaling affecting e.g. cell-adhesion or the cytoskeleton [153, 167].

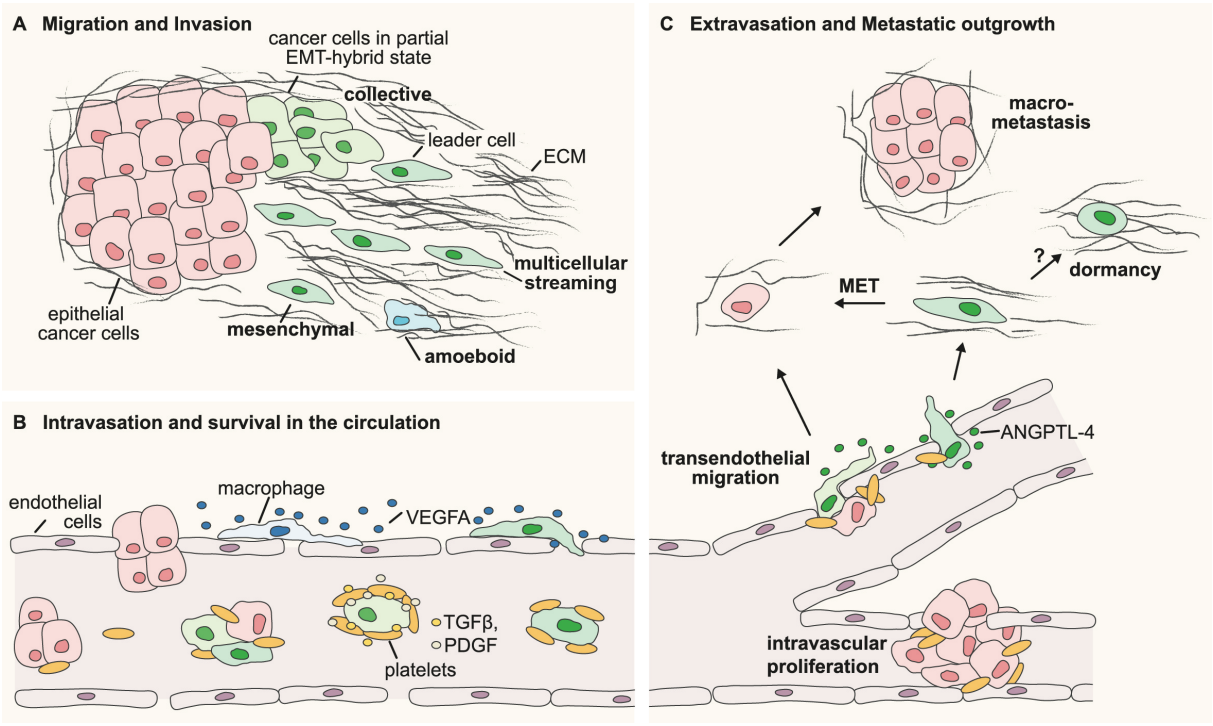


Figure 4: The invasion metastasis cascade: A) Cancer cells migrate and invade as cell collectives or single cells. Cells in a partial EMT hybrid state may migrate as cohesive groups often led by a mesenchymal-type cell (cancer cell or fibroblast). Mesenchymal cancer cells mostly migrate as single cells and may switch to amoeboid migration mode upon undergoing a mesenchymal-to-amoeboid transition. Single cells may also invade as multicellular streams one cell following the other. **B)** Mesenchymal cancer cells may actively intravasate. VEGFA secreted by mesenchymal cancer cells or macrophages facilitates intravasation by promoting vascular leakiness. Due to the abnormally leaky neovasculature, cancer cells may also be passively shed into the circulation. Cancer cells may travel as single CTCs or CTC clusters and frequently express epithelial and mesenchymal markers. CTCs are rapidly coated with platelets which may protect from shear-stress and immune attack as well as from anoikis by promoting an EMT. Platelets also facilitate cancer cell extravasation. **C)** Cancer cells that have survived in the circulation may possibly extravasate similar to leukocytes. Mesenchymal cancer cells may secrete ANGPTL-4 which promotes vessel leakiness and facilitates extravasation. Finally, cancer cells may also get trapped in the capillaries and may proliferate inside the lumen. Successful metastatic outgrowth most likely requires an epithelial

phenotype meaning mesenchymal cancer cells are required to undergo a MET, while tumor dormancy might possibly be associated with a mesenchymal phenotype.

b) Intravasation and survival in the circulation: Cancer cells may actively invade the blood or lymphatic vasculature and disseminate as individual cells or multi-cellular clusters. Whether lymphatic spread contributes to distant metastasis or is rather a surrogate marker for metastatic spread is still under debate [151]. Efficient intravasation is facilitated by secretion of VEGFA by mesenchymal tumor cells and macrophages, which promotes vascular permeability and angiogenesis [168, 169]. This neovasculature is abnormally permeable. As a result, considerable numbers of cells might be passively shed into the circulation in the absence of invasion [170, 171].

Having entered into the circulation, CTCs need to resist shear stress, attacks from natural killer cells and anoikis – apoptosis induced by lack of cell adhesion. CTCs are rapidly coated with platelets which offer protection from immune attacks and shear-stress. Furthermore, secretion of PDGF and TGFβ by platelets may promote the induction or maintenance of EMT in cancer cells [172] which may in turn protect cells from anoikis [82]. CTCs expressing EMT markers are widely observed in cancer patients [173-176]. Interestingly, particularly the presence of CTCs with hybrid E/M-phenotypes were found to correlate with poor prognosis [177-180]. Besides circulating as single cells, cells very rarely travel as multicellular clusters. These CTC clusters are derived from oligoclonal precursors rather than being the progeny of a single invading cancer cell or the result of intravascular aggregation [160, 181, 182]. They have an up to 50-fold increased metastatic potential compared to single CTCs and greatly contribute to metastasis despite their rarity [181]. A diverse spectrum of epithelial, EMT hybrid and fully mesenchymal phenotypes could be observed in CTC clusters. Interestingly, the mesenchymal cell-signature has been found to be even more pronounced in CTC clusters than single CTCs thereby supporting a role of EMT in collective dissemination [177, 180, 183]. Taken together, EMT programs may promote cell survival in the circulation.

c) extravasation: At some point, CTCs get trapped inside the capillaries or stick to the endothelial wall upon which they may extravasate and invade the tissue parenchyma. A mesenchymal phenotype may be beneficial for transendothelial migration due to mesenchymal cell adhesion which facilitates contact with endothelial

cells and invasion of the surrounding tissue [184]. Furthermore, angiopoietin-like 4 (ANGPTL-4) secreted by mesenchymal cancer cells increases vascular permeability and facilitates extravasation [185]. CTC clusters are cleared from the blood more efficiently, possibly due to their size, although it is not clear whether they actively extravasate [181]. Cells which have traveled in clusters may for example start proliferating in microvessels and then rupture the endothelial wall to access the parenchyma. This process might not require migration at all [151].

Depending on the anatomy of capillaries at the target side, CTCs need to meet different requirements to successfully extravasate. The fenestrated bone marrow sinusoid capillaries are more permeable than the capillaries of the lung while the blood-brain barrier provides an extra challenge. In line with this, breast cancer cells which show a tropism for the brain express high levels of certain genes that facilitate passage through the blood-brain barrier [151, 186]. Although EMT has been linked to metastatic organotropism, the precise mechanisms need to be further investigated [187].

d) Metastatic outgrowth: Invasion of the tissue parenchyma at the metastatic side does not guarantee successful establishment of macrometastasis. It has been estimated that less than 0.02% of extravasated cells eventually form metastasis [188, 189]. A mesenchymal phenotype may promote cell survival and secretion of ECM components which may be beneficial for metastatic niche formation during the early steps of colonization. As described above, mesenchymal or hybrid EMT states are associated with increased stemness, which is a requirement for secondary tumor formation. However, most metastasis can phenotypically not be distinguished from the primary tumor. This suggests that upon arrival at the secondary site, any mesenchymal cancer cells would have to undergo a MET in order to form macrometastasis. Indeed, a plethora of studies have highlighted the requirement of MET for metastatic outgrowth [43, 190-194]. In line with this Del Pozo and colleagues revealed how crosstalk between mesenchymal cancer cells and the microenvironment at the secondary site promotes metastasis formation. Mesenchymal cancer cells initially trigger niche induction by activating resident fibroblasts which in turn induce them to undergo a MET which is required for metastatic outgrowth. This suggests that MET is not a passive reaction due to lack of EMT-inducing stimuli at the metastatic site, but can be actively triggered by the local microenvironment [194].

Metastases do not necessarily form immediately upon extravasation. Disseminated breast cancer cells frequently remain dormant only to grow out with a latency of years to decades [186]. Tumor dormancy is still poorly understood but has been linked to mesenchymal CSCs, which can be induced to undergo a MET to initiate active proliferation [195].

Compared to single disseminated cells, cells which have travelled as clusters do more efficiently form macrometastatic lesions [181]. This might simply be a result of being multiple rather than one cell to start with but could also be due to their epithelial or possibly EMT hybrid phenotype.

1.2.4 How Relevant is EMT for Metastasis Formation?

As described above, EMT/MET programs may contribute to various aspects of the invasion-metastasis cascade. However, there are multiple routes towards metastasis and the overall relevance of EMT remains obscure. Previously, the functional aspects of EMT *in vivo* have been studied by transplantation of cell lines manipulated *in vitro* by overexpression/knockdown of key EMT regulators. Whenever EMT was modulated, metastasis was affected as well. However, these models do neither recapitulate spontaneous tumor formation nor tumor heterogeneity and keep cells in a fixed, artificial state. To overcome these limitations, conditional genetic modifications and *in vivo* lineage tracing models have recently been established [96] (Table 1).

To understand whether EMT is sufficient to drive metastasis, Tsai and colleagues induced EMT by means of doxycycline-inducible Twist1 expression in a transgenic mouse model of skin squamous cell carcinoma. Topical or oral administration of doxycycline allowed to restrict Twist1 expression to the primary tumor site or to maintain its expression at metastatic sites. Interestingly, Twist1 expression induced dissemination but had to be downregulated at the metastatic sites to allow colonization [193]. Likewise, Tran et al. reported that Snail is sufficient to drive breast cancer dissemination while its downregulation is critical for metastatic outgrowth [191]. In an indirect approach, Title et al. promoted EMT by conditional knockout (cKO) of the miR200 loci or mutation of the miR200 binding sites in Zeb1 mRNA. This led to increased metastasis in pancreatic cancers [196]. Taken together, these data highlight

the importance of MET in metastatic outgrowth and suggest that a reversible EMT contributes to metastasis.

However, the relevance of EMT for metastasis formation remained questionable as a spontaneous EMT is relatively rare *in vivo*. To explore whether cancer cells may metastasize in the absence of EMT, the key transcriptional regulators Snail, Zeb and Twist have been targeted. Indeed, cKO of Snail and Twist1 in breast cancer cells drastically reduced metastasis formation, indicating a substantial contribution of EMT to metastasis [191, 197]. Similarly, cKO of Zeb1 in PDAC cells greatly reduced EMT, invasiveness and metastasis [198]. Intriguingly, while Zeb1 was found to be critical, cKO of Snai1 or Twist1 reduced EMT in primary tumors but did not affect metastasis in the very same mouse model of PDAC [199]. These inconsistencies may be explained by tissue-specific functions and overlapping roles of EMT-inducing transcription factors, which could compensate for each other [140, 200]. However, these data indicate a significant contribution of certain EMT-transcription factors – and hence EMT – to metastasis depending on the tumor type.

Within the past years, lineage tracing models have become a popular tool to track cells of interest in an unperturbed system. Using a cancer cell-specific lineage tracer combined with CFP-tagged E-cadherin, Beerling et al. observed 60% of CTCs and disseminated tumor cells being E-cadherin low while all lung metastasis of a size beyond three cells were E-cadherin high [106]. Similar findings were reported in PDAC and prostate cancer, where about 50% of CTCs or single cell metastasis had undergone an EMT [109, 115]. The fate of these mesenchymal cells - whether they undergo MET and contribute to macrometastasis, die or remain dormant - could not be addressed by these models as they only captured the acute state (Figure 3).

In an attempt to track the fate of breast cancer cells that have undergone an EMT, two groups relied on labeling of mesenchymal cells using Fsp-Cre in combination with a tdTomato-to-GFP color-switching reporter. Although GFP positive mesenchymal cancer cells were highly invasive [113] and enriched in CTCs relative to primary tumors, only 1 in 150 lung metastases was derived from a cell that had previously undergone an EMT [112]. A similar fate mapping approach based on Fsp or α-smooth muscle actin (SMA)-Cre in a mouse model of PDAC revealed that cells which had undergone an EMT only formed micrometastasis in the liver with the majority of established metastasis being derived from more epithelial cells [116]. These data indicate low efficiency of metastatic outgrowth of mesenchymal compared to epithelial

cancer cells and suggest that a full EMT might not be essential for metastasis. These studies were however limited by the use of Fsp or αSMA as EMT markers, which might not be universally expressed during EMT. Particularly cells in an early EMT hybrid state were likely to be missed [96, 201, 202].

Indeed, distinct EMT states might contribute to metastasis to a different extent. Using a Vimentin-GFP construct to visualize mesenchymal cells in prostate cancer, Ruscetti et al. observed rare populations of both EpCAM positive, GFP positive EMT hybrid cells and EpCAM negative, GFP positive mesenchymal cancer cells. Upon intravenous injection, which allows to directly assess metastatic outgrowth, epithelial and EMT hybrid cells could readily establish metastasis while mesenchymal cells could not [115]. Similarly, Pastushenko et al. identified six distinct populations of cells associated with different EMT hybrid states *in vivo*. Although cells with a late stage EMT phenotype were the most invasive, cells residing in the two earliest EMT hybrid states accounted for the majority of CTCs. Upon intravenous injection, all populations gave rise to epithelial lung metastasis. But consistent with previous observations, the two early EMT hybrid phenotypes showed the most efficient metastatic outgrowth [87]. These data suggest that different EMT states vary in their tendency to metastasize, with the early EMT states possibly having the highest intrinsic abilities to form macrometastasis. These differences could possibly be due to the distinct mechanisms of dissemination employed by the cells. In line with this notion, Aiello and colleagues recently observed that in certain well-differentiated PDAC tumors, cancer cells which lack membranous E-cadherin do not show evidence of a transcriptionally mediated EMT but acquire a partially mesenchymal phenotype by relocalization of E-cadherin. Cells in such a partial EMT state predominantly disseminated as clusters. In contrast, highly invasive PDAC tumors generated “fully” mesenchymal cancer cells by a classical EMT program which disseminated as single cells [108]. These findings indicate that the predominant mechanism of dissemination might depend on tumor type. Along similar lines, data presented by Reichert et al. suggest that mesenchymal PDAC cancer cells might be able to seed lung but not liver metastasis when being prevented from undergoing a MET [203]. This highlights the role of the microenvironment at the metastatic site as a gatekeeper for metastatic outgrowth and connects EMT with metastatic tropism.

Cancer	Approach	Contribution of EMT to metastasis	
BC [191]	Inducible overexpression & cKO (Snai1)	yes	Essential for metastasis but needs to be transient.
SCC [193]	Inducible overexpression (Twist)	yes	Sufficient for metastasis but needs to be transient.
PNET, PDAC [196]	cKO (miR200 & miR200 binding site of Zeb1)	yes	Contributes to metastasis
BC [197]	cKO (Twist1)	yes	Essential for metastasis
PDAC [198]	cKO (Zeb1)	yes	Essential for metastasis
PDAC [199]	cKO (Twist1 /Snai1)	no	Dispensable for metastasis
PDAC [203]	cKO (p120 catenin) + cancer cell lineage tracer	maybe	Epithelial phenotype is essential for liver but not lung metastasis formation.
PDAC [109]	Cancer cell lineage tracer	maybe	50% of single cell metastasis are mesenchymal. Macrometastases are epithelial.
PDAC [108]	Cancer cell lineage tracer	maybe	50% of partial EMT cells migrate as clusters, complete EMT cells migrate as single cells
BC [106]	Cancer cell lineage tracer + acute EMT-marker	maybe	60% of single cell metastasis are mesenchymal. Macrometastases are epithelial.
SCC, [87] BC	Cancer cell lineage tracer	maybe	Early EMT states have the highest intrinsic capabilities to metastasize.
PC [115]	Acute mesenchymal cell tracer	maybe	Partially but not fully mesenchymal cells may establish macrometastases
BC [112]	Mesenchymal lineage tracer	no	No evidence for contribution to metastasis
PDAC [116]	Mesenchymal lineage tracer	no	No evidence for contribution to metastasis

Table 1: Evidence for and against a contribution of EMT to metastasis in mouse models. BC: breast cancer, SCC: squamous cell carcinoma, PNET: pancreatic neuroendocrine tumor, PDAC: pancreatic ductal adenocarcinoma, PC: prostate cancer, cKO: conditional knockout.

Taken together, the contribution of EMT to metastasis is likely context-dependent. Cells in a mesenchymal state are possibly more invasive and more likely to disseminate than epithelial cells. However, survival in the systemic circulation and particularly the colonization step are a major bottleneck for metastasis formation. Experimental evidence suggests that epithelial and partial EMT hybrid cells – especially when travelling as clusters – establish macrometastasis more efficiently than mesenchymal cancer cells, and might have the overall highest intrinsic abilities to form metastases. Consistent with this it has recently been shown that the loss of E-cadherin promotes local dissemination but at the same time reduces proliferation, survival, colonization of distant organs and metastatic outgrowth [204]. At this point, one should consider a spectrum of cells along the E/M axis which take preferred routes on the way to metastasis. Microenvironmental factors at distinct steps of the invasion-

metastasis cascade might select for certain EMT states and modes of dissemination. Furthermore, the phenotype of cancer cells may change at times. The final contribution of a certain program of dissemination towards metastatic burden – which may include a partial or full EMT– would therefore depend on microenvironmental factors and could also change during tumor progression. Activation of an EMT program in a minority of cells might also promote metastasis of neighboring epithelial cells by facilitating dissemination at the primary tumor site e.g. by leading collective invasion strands or promoting ECM degradation. It is possible that few mesenchymal cancer cells escort their epithelial neighbors to the metastatic site where they are being outcompeted. The ultimate relevance of EMT in a given tumor may only be addressed by ablation of cells as soon they undergo an EMT. In any case, the contribution of transiently activated partial or full EMT programs towards metastasis warrants further investigation.

1.2.5 EMT and Therapy Resistance

Surgical removal of the tumor is the most important measure to combat cancer. If detected before cancer cells have spread, surgery is often curative. However, in most cases adjuvant therapy is indicated which commonly means chemo-, radio-, targeted- or immunotherapy. Sadly, even after an initial response, many tumors develop resistance to therapy which is associated with significant mortality.

EMT signatures may predict treatment response and relapsed tumors are frequently dedifferentiated indicating that EMT contributes to intrinsic and possibly acquired resistance [24, 137, 180, 205-207]. A relatively recent study provided direct evidence for increased chemo-resistance of mesenchymal cancer cells by *in vivo* lineage tracing [112]. EMT may contribute to resistance in multiple ways. Slow proliferation and stemness associated with EMT renders cancer cells more resilient. Furthermore, EMT activates the expression of drug-efflux pumps, particularly members of the ATP-binding cassette (ABC) transporters, which protect from chemotherapy and cause multidrug resistance [208, 209]. EMT transcription factors Snail, Twist and Zeb protect cells from apoptosis and DNA-damage induced cell-death, which can contribute to chemo- and radioresistance [72, 210-212]. Consistent with this,

cKO of Snail and Twist significantly increased chemo sensitivity in a mouse model of PDAC [199].

Besides these general survival mechanisms which may account for multidrug resistance, EMT may contribute to targeted drug and immunotherapy resistance by more specific means. Snail-mediated expression of the tyrosine kinase receptor Axl promotes resistance to EGFR inhibitors in EGFR mutant non-small-cell lung carcinoma due to compensatory signals provided by Axl [213, 214]. It is possible that similar mechanisms mediate resistance to other targeted treatments. Similarly, EMT directly induces expression of PD-L1 and is possibly involved in the regulation of other checkpoint ligands [215-217]. Due to the redundancy of immune checkpoints, EMT may cause resistance to currently available checkpoint inhibitors which target only one ligand e.g. PD-L1 or CTLA4 [218]. Furthermore, tumors derived from mesenchymal cell lines were found to express lower levels of MHC-I, thereby being less immunogenic than tumors derived from epithelial cell lines [219]. Importantly, not all immune cells have anti-tumor functions. Mesenchymal cancer cells promote an immunosuppressive microenvironment by secretion of immunoregulatory cytokines and chemokines which activate immunosuppressive cells such as regulatory T cells (T_{reg} cells), alternatively activated M2 macrophages and myeloid-derived suppressor cells (MDSCs). Thereby, low numbers of mesenchymal cancer cells may protect themselves and their epithelial neighbors [4, 218, 220].

Taken together, EMT promotes therapy resistance in a variety of ways. Survival of pre-existing mesenchymal cancer cells may be accountable for intrinsic drug resistance in some cases. As it constitutes an epigenetically regulated survival mechanism, it is plausible that EMT particularly contributes to initial phases of resistance development by allowing mutations to arise which ultimately cause rapid tumor relapse. To which extent EMT is induced by different therapies *in vivo* remains to be determined. New technologies such as EMT lineage tracing approaches combined with intra-vital imaging or *ex vivo* live-imaging to observe EMT happening in real time might be used to shed light on this question.

1.2.6 Partial EMT and EMT Hybrid States *in vivo*

Elizabeth Hay first observed an “epithelial mesenchymal transformation” several decades ago [221]. Ever since, the understanding of EMT and MET has evolved from the concept of a mere transformation to being viewed as a complex epigenetic program of cell plasticity. As described above, increasing amount of evidence suggests that cancer cells may reside stably in intermediate EMT hybrid states. However, little is known about the functional characteristics of distinct EMT states and how they relate to certain aspects of malignancy *in vivo*.

Some of the most paramount questions right now are to which extent cancer cells undergo a partial or full EMT *in vivo* and how distinct EMT states compare to one another in terms of proliferation, invasiveness, stemness, plasticity, metastasis and drug-resistance.

With the identification and functional characterization of six EMT states *in vivo*, Pastushenko and colleagues were the first to thoroughly address these questions [87]. Compared to the fully epithelial state, proliferation was markedly reduced in all transition states, including the very early ones (Figure 5). *In vitro* invasion assay revealed a gradual increase of invasiveness along the E/M axis. Upon orthotopic transplantation, all EMT hybrid populations showed similar and about five times greater tumor initiating frequency compared to epithelial cells, indicating that stemness could possibly be a feature of all EMT hybrid and mesenchymal states. Yet, several other studies suggest an increase in stemness at the intermediate hybrid states but a concomitant decrease when cells are pushed into a highly mesenchymal state [43, 150, 191, 193, 222, 223]. Cell plasticity allows to adapt to microenvironmental changes and is critical for tumor progression. Pastushenko et al. associated the two most intermediate hybrid EMT states with the highest degree of plasticity as they could give rise to all other EMT hybrid states with similar frequencies upon transplantation. Interestingly, cells that had once acquired a partial EMT hybrid phenotype would transition between EMT hybrid states but rarely acquire a fully epithelial phenotype when orthotopically transplanted. However, all EMT hybrid and mesenchymal populations were able to revert back to a fully epithelial phenotype in the lung indicating that epithelial-mesenchymal plasticity depends on microenvironmental factors.

As discussed above, the contribution of distinct EMT states to metastasis depends on many factors and remains to be explored. Data obtained by Pastushenko and Ruscetti et al. suggest that cells residing in early EMT states might have the overall highest inclination to metastasize [87, 115]. However, the overall contribution of cells residing in a certain EMT state to total metastatic burden not only depends on the cell intrinsic capability to metastasize but also microenvironmental factors determining the frequency to which cells may acquire this phenotype. Moreover, an association of certain EMT states – particularly the mesenchymal state – and metastatic dormancy should be addressed. Finally, virtually nothing is known about the functional contribution of distinct EMT states towards therapy resistance. It is possible that the intermediate EMT states are associated with refractory disease due to their increased stemness, consistent with the hybrid state being observed in patients presenting with drug resistance [24, 137, 180, 207, 222] However, whether certain EMT states are associated with a particularly high degree of resistance and whether certain drugs actively induce cells to undergo a partial or full EMT remains to be evaluated.

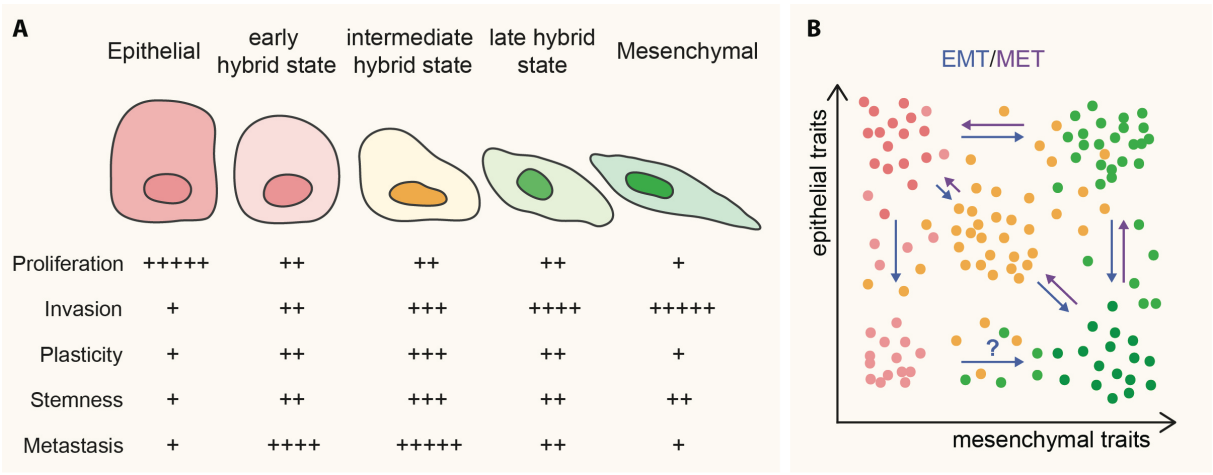


Figure 5: The emerging characteristics of distinct EMT states. A) Distinct EMT states exhibit differences in functional characteristics. (+ very low, +++++ very high intrinsic capacity) Inspired by Pastushenko et al. [87, 91] **B)** EMT is not one stereotypical process and EMT hybrid states are a heterogeneous population. Cancer cells might follow different trajectories within a 2-dimensional EMT-grid. EMT and MET might not necessarily follow the same trajectories [224]. *In vivo* characterization of EMT states and mathematical modeling might eventually shed light on the complex roadmap connecting epithelial and mesenchymal cell states. Inspired by Jolly et al. [224].

It is tempting to picture the EMT spectrum as a sequence of cell states along a linear path through which cells transition from one state to the next. However, as cells residing in an EMT state are a very heterogeneous population, this an unlikely scenario [87, 104, 224]. By definition, a full EMT encompasses the loss of epithelial and the acquisition of mesenchymal traits. The detection of cells with co-expression of both epithelial and mesenchymal markers suggests that the gain of mesenchymal markers may happen concomitantly with the loss of epithelial traits. However, cells could also lose all epithelial traits before eventually acquiring mesenchymal ones [224, 225]. In line with this, cancer cells having lost epithelial cell adhesion in the absence of major EMT transcription factor activity or even without any transcriptional reprogramming have been observed [104, 108, 203]. Whether these cells could eventually undergo a full EMT and therefore represent truly early EMT states is not clear. Considering all of the above, it has been proposed to view EMT at least on a 2-dimentional landscape with mesenchymal and epithelial traits on the x and y axis and multiple paths which cells might take on their journey from a fully epithelial to a fully mesenchymal state [85, 224] (Figure 5). Importantly EMT and MET may or may not follow the same routes. Hence, we need to better understand the characteristics of distinct EMT states, how they relate to each other and which stimuli promote stable residence in a given state or trigger the individual transitions.

1.2.7 Therapeutic Targeting of EMT

Due to the plethora of malignant traits associated with it, major effort is being put into devising strategies to pharmaceutically target EMT. Approaches to prevent or possibly revert EMT mainly involve the inhibition of signaling pathways critical to the process, such as TGFβ and Wnt. Novel antibodies and small molecule inhibitors targeting these pathways are currently being developed and entering clinical trials with promising results [24, 226]. Other potential targets include kinases, such as the HGF-receptor c-MET, AXL and Src [24]. Intravital imaging of lineage-traced EMT cells revealed that the multikinase inhibitor cabozatinib, whose targets include c-MET and AXL, significantly repressed EMT [113, 227]. Other promising approaches include selective killing of mesenchymal cancer cells. Withaferin A disrupts the mesenchymal vimentin intermediate filaments and abrogated invasion and metastasis of breast cancer cells

[24, 228, 229]. Similarly, a high-throughput screen identified salinomycin as a potent agent to kill mesenchymal cancer cells [230]. Finally, great effort has been made to develop therapies to induce differentiation thereby reversing an EMT. Activation of PKA signaling by cholera toxin and forskolin treatment has been found to induce a MET with concomitant reduction of invasiveness and increased chemo-sensitivity of transformed mammary epithelial cells [231]. It will be interesting to see whether this can be translated to the clinics. A prime example of differentiation therapy that is of major clinical importance, is all-trans retinoic acid (ATRA) treatment of patients with acute promyelocytic leukemia. ATRA induces terminal differentiation of leukaemic promyelocytes and translates into a high rate of remission [232]. However, in contrast to ATRA treatment which ultimately leads to cell death, induction of MET in carcinoma cells might initially shift them into an aggressive partial-EMT state which could promote metastatic outgrowth. Recently a very elegant approach by Ishay-Ronen et al. exploited the plasticity associated with the partial EMT phenotype in order to trans-differentiate highly aggressive cells into harmless, post-mitotic adipocytes which resulted in greatly reduced invasion and metastasis [233]. Nevertheless, better understanding of the functional and molecular characteristics of partial and full EMT states is required to devise strategies to target these problematic populations. In future, novel therapies targeting EMT could be combined with established treatments to overcome therapeutic resistance and metastasis to ultimately improve patient survival.

2 Aims of this Study

EMT is a gradual process of dedifferentiation which conveys carcinoma cells with increased migratory and invasive capabilities and is associated with tumor progression and metastasis. Rather than being a binary switch, an EMT covers a spectrum of intermediate “partial” EMT hybrid states. Recent studies indicate that distinct EMT states differ in their proliferative, invasive and metastatic potential and could contribute to distinct aspects of malignancy. However, due to the transient and reversible nature of the process, the extent to which cancer cells undergo a spontaneous EMT *in vivo* and the contribution of EMT towards metastasis remain obscure.

With my PhD work, I have aimed at assessing the role of EMT in breast cancer progression and metastasis to the lung using a lineage tracing approach. I have established two novel lineage tracing mouse models for studying the fate of cancer cells that have undergone an EMT to understand:

- to which extent cancer cells undergo a spontaneous partial or full EMT *in vivo*

and

- whether partial or full EMT might contribute to lung metastasis.

3 Results

3.1 Lineage Tracing Reveals the Contribution of Partial Epithelial-to-Mesenchymal Transition to Breast Cancer Metastasis

Fabiana Lüönd[1], Nami Sugiyama[1], Carolina Hager[1], Ruben Bill[1], Thomas Bürglin[1], Telma Lopes[1], Gerhard M. Christofori[1]

[1] Department of Biomedicine, University of Basel, Switzerland

in preparation

3.1.1 Abstract

Epithelial-to-mesenchymal transition (EMT) is a multistep process of transdifferentiation which conveys carcinoma cells with enhanced malignant characteristics. Due to the transient and reversible nature of the process, the extent to which cancer cells may undergo a spontaneous EMT *in vivo* and the contribution of EMT towards metastasis remain obscure. We have established two lineage tracing systems to label cancer cells undergoing a partial or full EMT. Within epithelial tumors few cancer cells transition between epithelial and EMT hybrid states but rarely undergo a full EMT. While cells that have undergone a partial EMT are found enriched in metastasis, cells which have completed a full EMT did not colonize the lung. Our data indicate that although EMT might not be a prerequisite, a partial EMT may contribute to metastasis.

3.1.2 Introduction

Despite major advances in understanding the mechanisms of malignant dissemination, metastasis remains the leading cause of breast cancer-related death. The epithelial-to-mesenchymal transition (EMT) is a developmental program of transdifferentiation which may contribute to the systemic spread of cancer [94]. During an EMT, epithelial cancer cells lose their cell-cell adhesions and acquire a migratory and invasive, mesenchymal phenotype. Mesenchymal cancer cells have an increased ability to travel to distant sites *in vivo*, and induction of a transient EMT is sufficient for metastasis formation [106, 109, 115, 191, 193]. However, whether EMT contributes to or is required for metastasis is still under debate [4, 96, 97, 112, 199, 201, 202].

Recent studies indicate that rather than being a binary switch, EMT proceeds through intermediate hybrid states associated with distinct functional characteristics [85-88]. Cells residing in early EMT hybrid states were observed to metastasize particularly efficiently despite being less invasive than “fully” mesenchymal cells [87, 115]. However, to which extent cancer cells may transition between epithelial, partial-EMT and fully mesenchymal cell states *in vivo* is unknown and particularly the contribution of “partial” and “full” EMT towards metastasis warrants investigation.

The identification of cancer cells that have undergone an EMT is technically challenging, as it is both a transient and a reversible process. EMT lineage tracing systems based on recombination of lox-stop-lox reporter alleles by Cre-driven from a mesenchymal specific promoter have been generated to irreversibly label cancer cells undergoing an EMT [105, 112, 113, 116]. This allowed to trace cancer cells which have previously undergone an EMT including cells that have reverted back to an epithelial phenotype by undergoing a mesenchymal-to-epithelial transition (MET). However, mesenchymal lineage labelling is not cancer cell-specific which means that fully mesenchymal cancer cells could hardly be distinguished from stromal cells.

Here, using a combination of Flp and tamoxifen-inducible Cre recombinases, we have developed two cancer cell-specific EMT lineage tracing mouse models for studying the fate of cells that have undergone a "partial" or "full" EMT. Our data indicate that cancer cells mostly transition between partial EMT hybrid states and rarely undergo a full EMT. Furthermore, while a partial EMT can contribute to lung metastasis, a full EMT may not be required.

3.1.3 Results

3.1.3.1 Visualization of EMT by dual recombinase-mediated lineage labeling

To irreversibly label breast cancer cells undergoing an EMT, we established a cancer cell-specific fluorescent color switching system in the MMTV-PyMT mouse model of metastatic breast cancer. Combination of the Flp and Cre recombinase responsive "*RC::FrePe*" reporter [234, 235] with *MMTV-PyMT* [236] , *MMTV-Flpo* [237] and one of two EMT-specific tamoxifen inducible Cre alleles, *tenascin C (Tnc)*- or *N-cadherin (Cdh2)-CreER*^{T2}, leads to Flp-mediated mCherry expression in all breast cancer cells and a Cre-mediated irreversible switch from mCherry to GFP expression in cancer cells undergoing an EMT (Figure 1a). Cells keep expressing GFP, even if they revert back to an epithelial phenotype, thus allowing the visualization of EMT and MET. The *Tnc* and *Cdh2* promoters were chosen to drive Cre expression as these genes are newly expressed already in early stages or only in late stages of a TGFβ-induced EMT *in vitro*, respectively (Supplementary Figure 1a, b and GSE112797 [86], GSE1145722 [238]). This temporal difference allowed us to trace EMT events including partial EMT

(*Tnc* model) or bona-fide "full" EMT as marked by a classical cadherin switch (*Cdh2* model) [10].

Specificity of Cre activity in the newly generated *Tnc*- and *Cdh2-CreER^{T2}* lines was validated by crossing these strains to a loxP-stop-loxP-tdTomato reporter line (Supplementary Figure 1c). As *Tnc* is highly expressed by cancer associated stromal cells [239], we administrated tamoxifen to 11 week old tumor-bearing *Tnc-CreER^{T2}*/*lsl-tdTomato*/*MMTV-PyMT* mice. As expected, the majority of stromal cells was recombined (Supplementary Figure 1d). Despite recombination of a few cells in tumors of *Cdh2-CreER^{T2}*/*lsl-tdTomato*/*MMTV-PyMT* mice, no N-cadherin protein expression could be detected. However, high recombination efficiency was observed in N-cadherin-positive liver and heart tissue [240], with little leakiness in vehicle-treated mice (Supplementary Figure 1e, f).

As we had confirmed specificity and robust functionality of CreER^{T2} in *Tnc*- and *Cdh2-CreER^{T2}* mouse lines, we sought to visualize cancer cells residing in an EMT state *in vivo*. To this end, we treated *Tnc-CreER^{T2}/MMTV-Flpo/RC::FrePe/MMTV-PyMT* (hereafter Tnc-CreER^{T2}) and *Cdh2-CreER^{T2}/MMTV-Flpo/RC::FrePe/MMTV-PyMT* (hereafter Cdh2-CreER^{T2}) female mice with tamoxifen on three consecutive days before sacrifice (Figure 1b). In Tnc-CreER^{T2} mice a low number of GFP$^+$ cells could be observed. These cells were mostly localized in peritumoral regions. In Cdh2-CreER^{T2} mice only very few single GFP$^+$ cells were found dispersed in the tumor, indicating that cancer cells rarely undergo a full EMT (Figure 1c). GFP$^+$ cells in both early and late-stage tracing models had mostly an elongated, mesenchymal morphology and expressed no or only low levels of E-cadherin, indicating successful tracing of mesenchymal cancer cells. In both models, only a low number of E-cadherin$^+$ GFP$^+$ cells could be found. These cells may have undergone a MET or they may represent a rare subpopulation of EMT hybrid cells co-expressing E-cadherin and tenascin C or N-cadherin, respectively (Figure 1d, e). Interestingly, despite the predominantly E-cadherin$^-$ phenotype of GFP$^+$ cells in both models, the majority of GFP$^+$ cells did not express Vimentin or α-smooth muscle actin (α-SMA). This indicates that the majority of cells might either lose E-cadherin expression relatively early during an EMT *in vivo*, before the classical mesenchymal cell markers are being expressed, or that an EMT does not necessarily involve the upregulation of these commonly used

EMT markers in this mouse model [202] (Supplementary Figure 2). Taken together, our novel EMT lineage tracing model allows the visualization of cancer cells in a partial or full EMT state in an unperturbed system *in vivo*.

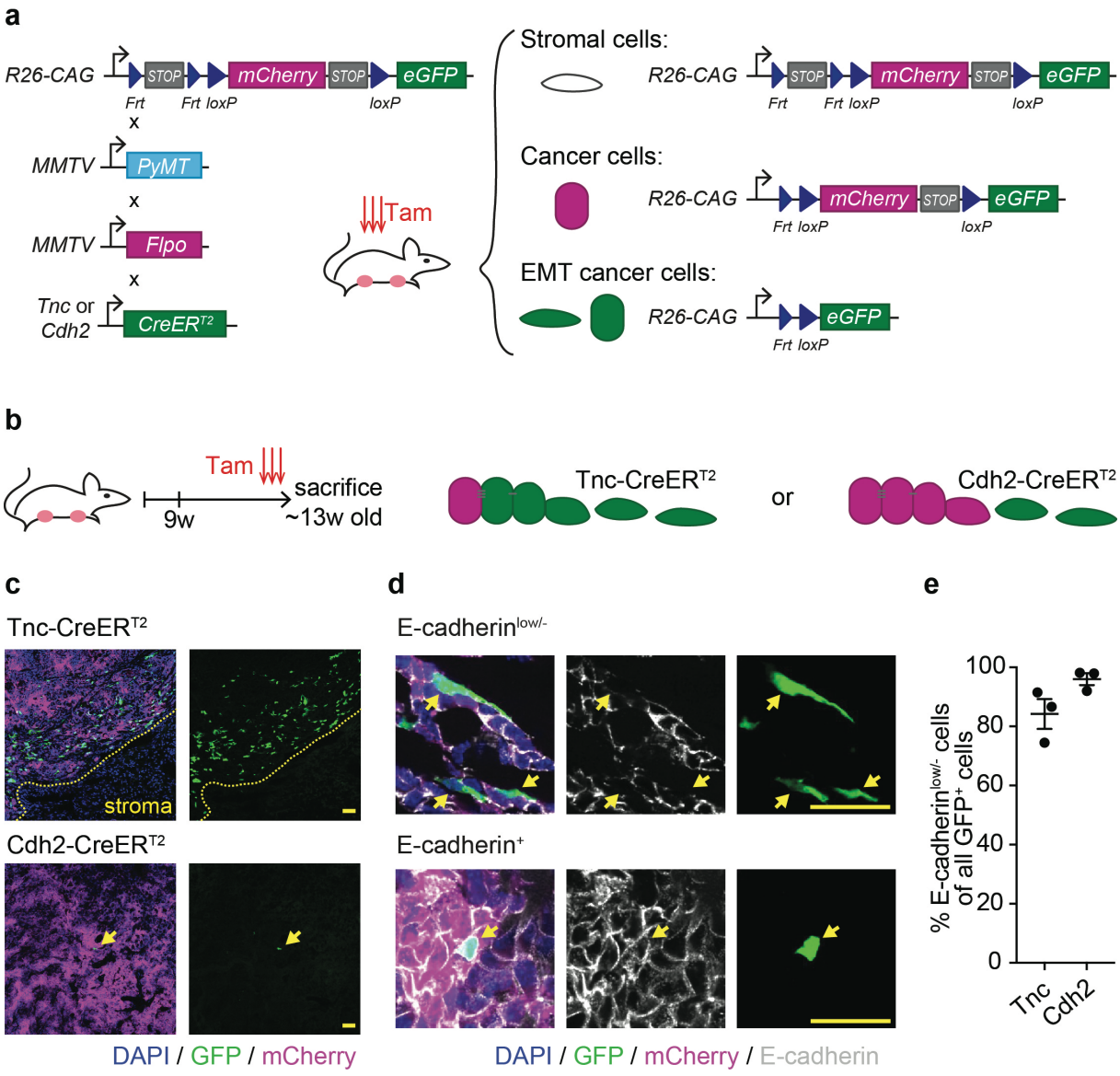


Figure 1: Visualization of partial EMT and mesenchymal cancer cells by lineage tracing. a Strategy for EMT lineage tracing. Combination of *RC::FrePe*, *MMTV-PyMT*, *MMTV-Flpo* and either *Tnc-CreER^{T2}* or *Cdh2-CreER^{T2}* alleles leads to mCherry expression in cancer cells and a CreER^{T2} mediated irreversible switch to GFP expression in cancer cells undergoing an EMT. **b** Strategy for visualization of mesenchymal cancer cells. Tnc-CreER^{T2} or Cdh2-CreER^{T2} mice were treated with tamoxifen (Tam) before sacrifice to induce GFP expression in mesenchymal cancer cells including early EMT states (Tnc-CreER^{T2}) or only late EMT states (Cdh2-CreER^{T2}). **c** Representative images of mCherry^{+} and GFP^{+} cells in tumors of Tnc-CreER^{T2} and Cdh2-CreER^{T2} mice. Scale bars: 50µm. **d** Representative images of GFP^{+}E-cadherin$^{low/-}$ and GFP^{+}E-cadherin^{+} cells. Scale bars: 25µm. Arrows indicate GFP^{+} cells. **e** Quantification of GFP^{+}E-cadherin$^{low/-}$ cells in tumors of Tnc-CreER^{T2} (Tnc) and Cdh2-CreER^{T2} (Cdh2) mice. n = 3 mice, lines represent mean +/- SEM.

3.1.3.2 Cancer cells mostly transition between EMT hybrid states and rarely undergo a full EMT

To trace EMT events occurring during tumor progression, we treated Tnc-CreER^{T2} and Cdh2-CreER^{T2} mice with tamoxifen starting at 9 weeks of age and continuing until sacrifice (Figure 2a). GFP expression was neither observed in tumors of vehicle-treated nor Flpo-negative mice (Supplementary Figure 3). In tumor sections of Tnc-CreER^{T2} mice, GFP^{+} cells were frequently found locally enriched and sometimes forming differentiated colonies. In contrast, mostly single GFP^{+} cells were observed in the Cdh2-CreER^{T2} model (Figure 2b and Supplementary Figure 4a). Consistent with our previous observations, the total numbers of GFP^{+} cells were significantly smaller in Cdh2-CreER^{T2} compared to Tnc-CreER^{T2} mice, indicating that cancer cells rarely undergo a full EMT (Figure 2c). To assess the phenotype of GFP^{+} cells, we examined their cellular morphology and expression levels of E-cadherin. Interestingly, about 90% of GFP^{+} cells in the Tnc-CreER^{T2} model had an E-cadherin^{+} epithelial phenotype, while around 70% of GFP^{+} cells in Cdh2-CreER^{T2} mice had an E-cadherin$^{low/-}$ mesenchymal phenotype, suggesting that “partial” EMT hybrid cells but not “fully” mesenchymal cells frequently undergo a MET in the primary tumor (Figure 2d, e). Interestingly, epithelial, E-cadherin^{+}GFP^{+} cells were mostly found in well-differentiated rather than dedifferentiated tumor areas. In contrast, E-cadherin$^{low/-}$GFP^{+} partial EMT cells were mostly localized to the stromal interface and less frequently found within the mCherry^{+} epithelial tumor mass, indicating that the stroma could be an important source of EMT-inducing stimuli (Figure 2d). Furthermore, rare populations of E-cadherin^{-}GFP^{+} fully mesenchymal cells could be found within the tumor stroma (Figure 2d, e). Of note, across all phenotypic states, greater numbers of GFP^{+} cells were observed in Tnc-CreER^{T2} versus Cdh2-CreER^{T2} mice, with the differences being particularly significant for E-cadherin^{+} epithelial cells (Figure 2f). Phospho-histone H3 (pH3)-positive proliferating GFP^{+} cells could be found in well-differentiated epithelial tumor mass, indicating that cell-proliferation contributes to the formation of epithelial GFP^{+} colonies and the much greater numbers of E-cadherin^{+}GFP^{+} cells in Tnc-CreER^{T2} compared to Cdh2-CreER^{T2} mice (Supplementary Figure 4b). Interestingly, although 30% of GFP^{+} cells in the Cdh2-CreER^{T2} mice express E-cadherin, epithelial GFP^{+} colonies were hardly found in these tumors.

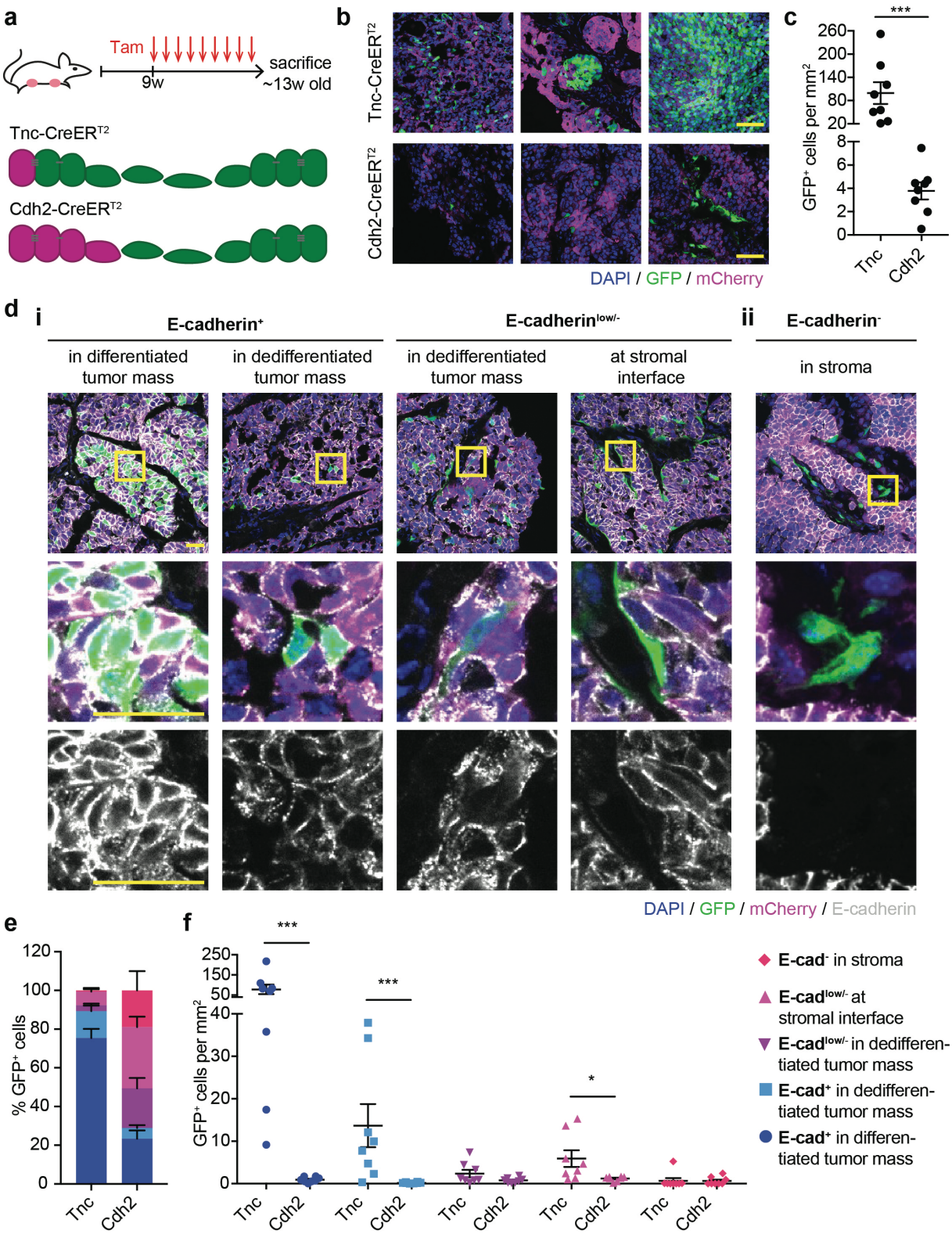


Figure 2: Cancer cells mostly transition between epithelial and EMT hybrid states and rarely undergo a full EMT. a Strategy for tracing of cells that have undergone an EMT during ~4 weeks. **b** mCherry$^+$ and GFP$^+$ cells in tumor sections of Tnc-CreER^{T2} (Tnc) and Cdh2-CreER^{T2} (Cdh2) mice. Representative images of areas with single GFP$^+$ cells or GFP$^+$ colonies are shown. Scale bars: 50µm. **c** Quantification of GFP$^+$ cells in tumors of Tnc and Cdh2 mice. n=8 mice, lines represent mean +/- SEM. ***p < 0.001, Mann-Whitney *U* test. **d** Representative images of E-cadherin$^+$ epithelial cells as well as E-cadherin$^{low/-}$ and E-cadherin$^-$ mesenchymal GFP$^+$ cells and their respective localizations in a tumor of a Tnc mouse **(i)** or a Cdh2 mouse **(ii)**. Boxes indicate the area included as

magnified insets. Scale bars: 25µm. **e, f** Distribution of the phenotypes of GFP$^+$ cells **(e)** and quantification of total number of GFP$^+$ cell types **(f)** observed in tumors of Tnc and Cdh2 mice. n=8 mice, lines represent mean +/- SEM. */***p <0.05/0.001, Mann-Whitney *U* test.

As *Tnc* is upregulated rather early during an EMT, cancer cells might upregulate *Tnc* expression before E-cadherin is lost, thereby making it difficult to distinguish GFP$^+$ epithelial cells from GFP$^+$ EMT hybrid cells based solely on E-cadherin expression. In order to confirm that GFP$^+$E-cadherin$^+$ cells in Tnc-CreER^{T2} mice may undergo a MET, we analyzed *Tnc* expression by RNA-FISH. While elongated GFP$^+$ cells at the stromal interface expressed high levels of *Tnc*, GFP$^+$ cells with an epithelial morphology were mostly *Tnc* negative, indicating that these cells had undergone a MET (Supplementary Figure 4c).

To further validate our observations, we analyzed the expression levels of the epithelial cell marker EpCAM by flow cytometry (Supplementary Figure 5a). Only 0.04% and 0.2% of cancer cells were GFP$^+$ in Cdh2- and Tnc-CreER^{T2} mice, respectively. Thus, consistent with our previous observations, a significantly smaller number of GFP$^+$ cells were found in tumors of Cdh2-CreER^{T2} compared to Tnc-CreER^{T2} mice (Supplementary Figure 5b). More than 80% of GFP$^+$ cells in Cdh2-CreER^{T2} mice had an EpCAM$^{low/-}$ mesenchymal phenotype, while the majority of GFP$^+$ cells in Tnc-CreER^{T2} mice as well as mCherry$^+$ cells in both mouse models were EpCAM$^+$ (Supplementary Figure 5c). Moreover, about 5% of GFP$^+$ cells in Tnc-CreER^{T2} mice expressed high levels of EpCAM, while this population was not observed in six out of seven Cdh2-CreER^{T2} mice, further demonstrating that “fully” mesenchymal cells may rarely revert back to an epithelial phenotype (Supplementary Figure 5a, c). Importantly, only about 1% of total EpCAM$^-$ fluorescent cells in both mouse models were GFP$^+$, indicating that the vast majority of mesenchymal cancer cells did not switch color in either of the models (Supplementary Figure 5d).

3.1.3.3 Partial EMT may occur during the early stages of tumor development

In addition to the contribution of EMT to the advanced stages of tumor progression, it has been shown that EMT may occur already early during tumor development [107]. Consistent with this, particularly in the Tnc-CreER^{T2} model, GFP$^+$ cells could

occasionally be found in early-stage lesions, which suggests a possible contribution of a partial EMT towards tumor development in a subset of lesions (Figure 3a). To further examine whether both partial and full EMT may occur during early stages of disease, we treated mice with tamoxifen for three consecutive days at nine weeks of age followed by tracing for 4-5 weeks. In both models, we observed a much lower number of GFP$^+$ cells compared to mice treated with tamoxifen only before sacrifice. These data suggest that although EMT may occur early during tumor progression, it is predominantly associated with advanced tumor stages. Consistent with previous results, the majority of GFP$^+$ cells in Tnc-CreER^{T2} mice were epithelial and sometimes forming colonies, while GFP$^+$ cells in Cdh2-CreER^{T2} mice had mostly retained a mesenchymal phenotype (Figure 3b). These data suggest that cells which have undergone a partial EMT could contribute to primary tumor growth when reverting back to an epithelial phenotype.

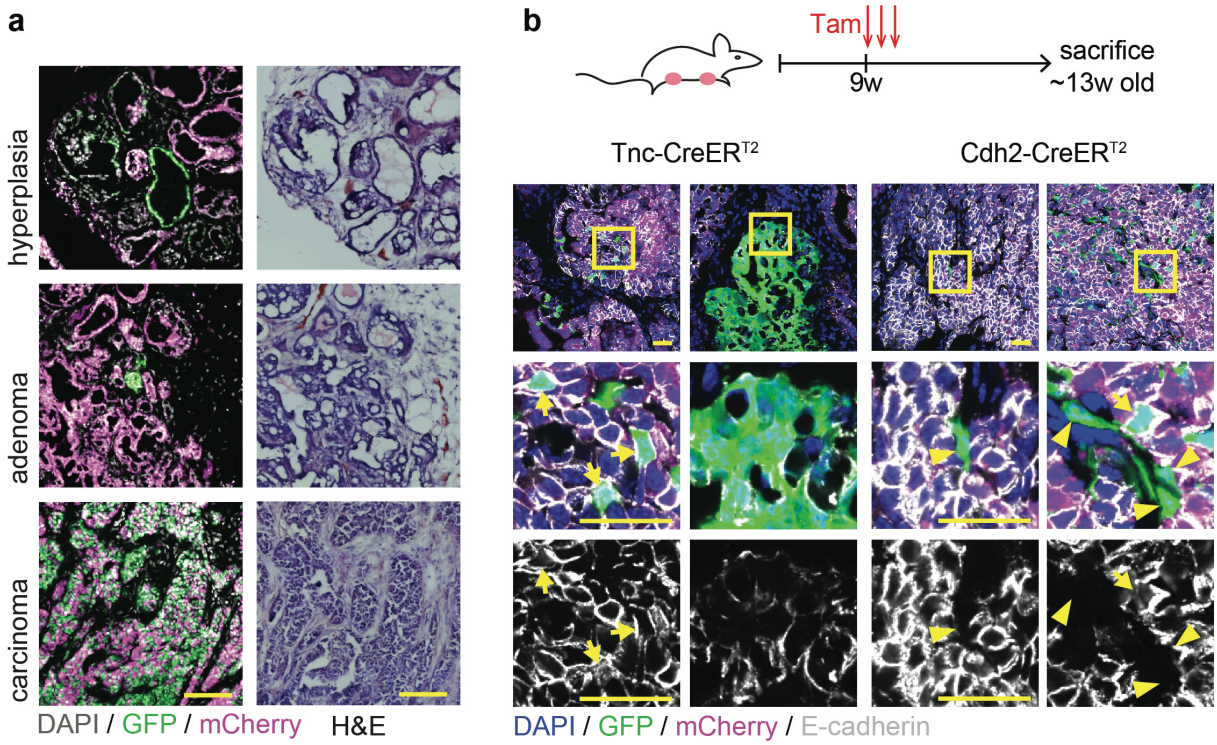


Figure 3: EMT may occur during early stages of tumor development. a Fluorescence analysis of GFP$^+$ and mCherry$^+$ cells and H&E staining on serial tumor sections of Tnc-CreER^{T2} mice. Representative images show GFP$^+$ cells in different tumor stages. Scale bars: 100µm. **b** Representative images of E-cadherin staining on tumor sections from Tnc-CreER^{T2} (n=6) and Cdh2-CreER^{T2} (n=2) mice treated with tamoxifen at 9 weeks of age followed by four weeks of lineage tracing. Arrows and arrowheads indicate E-cadherin$^+$GFP$^+$ and E-cadherin$^-$GFP$^+$ cells, respectively. Boxes indicate the area included as magnified insets. Scale bars: 25µm.

3.1.3.4 Cells which have undergone a partial EMT contribute to metastasis

To assess the contribution of a spontaneous EMT towards metastasis *in vivo*, we next analyzed the lung metastases in our transgenic mouse models. In Cdh2-CreER^{T2} mice, metastases were exclusively mCherry^{+} with no evidence of cells having undergone a full EMT as marked by N-cadherin expression. In Tnc-CreER^{T2} mice approximately 68% of metastases were mCherry^{+}. Only 1% of metastases were GFP^{+}, however, 31% of metastasis consisted of both GFP^{+} and mCherry^{+} cells (Figure 4a). These mixed-color metastases contained either few, mostly E-cadherin$^{low/-}$ mesenchymal GFP^{+} cells, or presented with a mosaic-like pattern of E-cadherin^{+}GFP^{+} and mCherry^{+} cells (Figure 4b). Notably, GFP^{+} cells were found highly enriched in lung metastases compared to the primary tumor (Figure 4c). Interestingly, GFP^{+}/mCherry^{+} mixed metastases were significantly larger than pure mCherry^{+} metastases (Figure 4d). This could be due to few cells undergoing an EMT in established metastases or due to cells disseminating as mixed epithelial/mesenchymal cell clusters which could colonize the lungs more efficiently [181]. However, in this transgenic mouse model, the possibility of cancer cells undergoing an EMT and MET in primary tumors and metastasizing as epithelial GFP^{+} cells irrespective of their previous EMT state cannot be excluded.

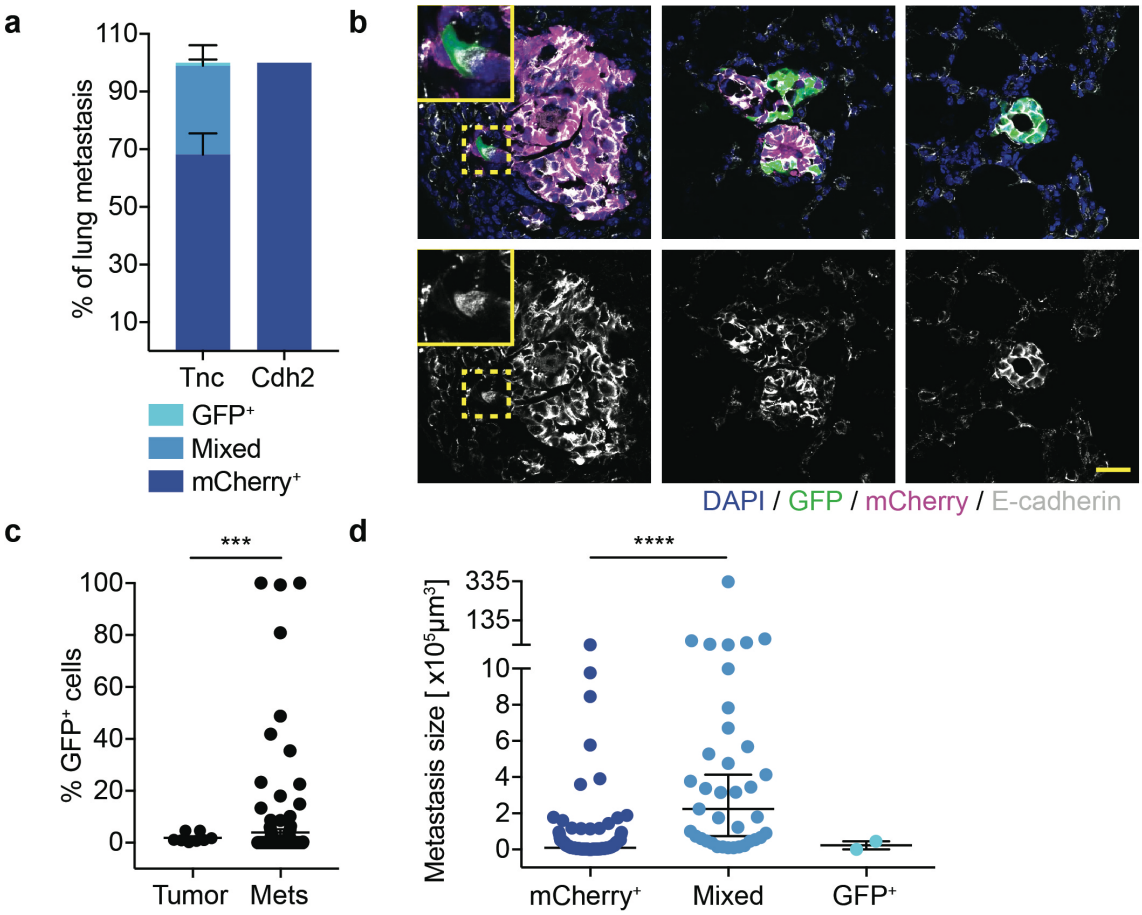


Figure 4: GFP⁺ cells are enriched in lung metastases of partial-EMT lineage tracing mice. a Quantification of GFP⁺, mixed and mCherry⁺ lung metastasis in Tnc-CreER^{T2} (Tnc) and Cdh2-CreER^{T2} (Cdh2) mice. n=8 mice, mean + SEM are shown. **b** Representative images of GFP⁺ cells in lung metastases stained for E-cadherin with magnified insets for E-cadherin⁻ GFP⁺ cell. Scale bar: 25µm. **c** Quantification of GFP⁺ cells in primary tumors and lung metastases (Mets) of Tnc mice. n=8 mice, dots represent mean values of 2 tumors per mouse and n=171 individual metastases from the same mice. ***p < 0.001 by Mann-Whitney *U* test. **d** Size distributions of lung metastases. Dots represent n=171 individual metastases, lines represent mean +/- SEM. ****p < 0.0001 by Kruskal-Wallis test.

To further address the contribution of a partial EMT towards lung metastasis, we established a transplantation model in which primary tumors were resected to narrow the time window of metastatic dissemination and allow the outgrowth of macrometastases. Tumor pieces of untreated Tnc-CreER^{T2} mice were orthotopically transplanted into immunodeficient NSG mice. Once a tumor size of approximately 200mm^3 was reached, one cohort of mice received tamoxifen for three consecutive days and primary tumors were surgically removed the next week. A second cohort of mice received tamoxifen only one week after tumor removal (Figure 5a). Approximately one month after tumor removal mice had developed overt metastases. Primary tumors resected post tamoxifen treatment contained few GFP⁺ cells with phenotypes similar

to the ones observed in transgenic mice (Figure 5b). Comparable to transgenic mice, the majority of established lung metastases were mCherry$^+$ in mice treated with tamoxifen prior to or after resection of the primary tumor. However, consistent with our previous results, GFP$^+$ cells were significantly enriched in lung metastases compared to primary tumors (Figure 5c-e). Of note, GFP$^+$ cells were found in the lung even when the tumor had been removed prior to tamoxifen administration. These GFP$^+$ cells must consequently have resided in a mesenchymal cell state in the lung, which indicates that they had either undergone a partial EMT at the primary tumor site and arrived in the lung as mesenchymal cells or that they had undergone a partial EMT as an adaptive process at the metastatic site (Figure 5d, e). Taken together, our data indicate that a partial EMT could contribute to the formation of a considerable number of metastases.

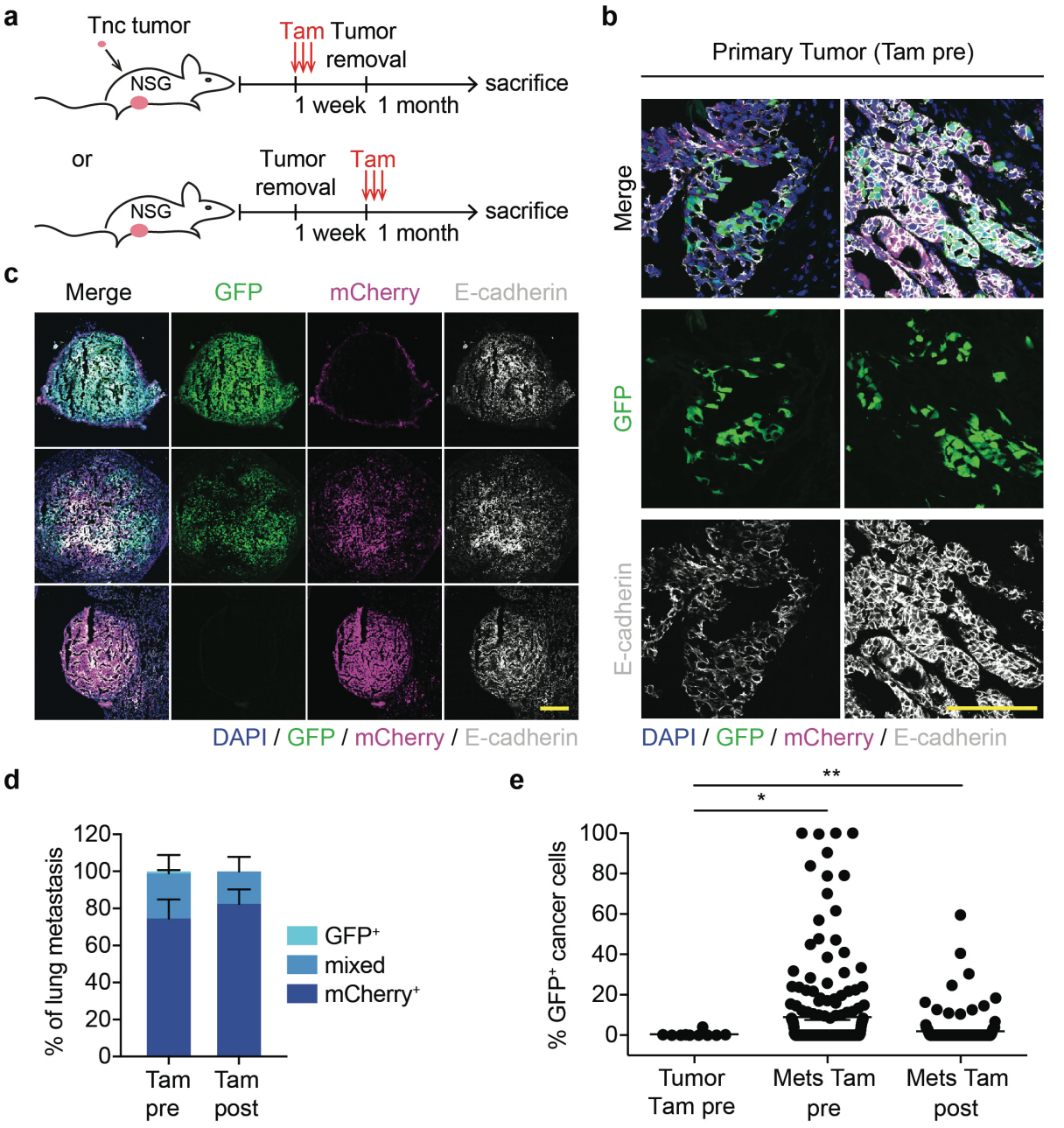


Figure 5: EMT tracing in transplantation model. a Schema of experimental setup. Tumor pieces of Tnc-CreER^{T2} transgenic mice were transplanted into NSG mice that were subsequently treated with tamoxifen (Tam) for three consecutive days either one week prior to (Tam pre) or one week after (Tam post) surgical tumor removal. **b** Representative images of GFP$^+$ cells in tumors of Tam pre mice stained for E-cadherin. Scale bar: 100µm. **c** Representative images GFP$^+$, mixed and mCherry$^+$ metastasis in Tam pre mice stained for E-cadherin. Scale bar: 200µm. **d** Quantification of GFP$^+$, mixed and mCherry$^+$ lung metastasis. n=10 (Tam pre) or n=6 (Tam post) mice, mean + SEM are shown. **e** Quantification of the percentage of GFP$^+$ cells in primary tumors and lung metastases (Mets). Dots represent mean values of n=10 primary tumors and n=201 (Tam pre) or n=159 (Tam post) individual metastases, lines represent mean +/- SEM. */**p < 0.05/0.01, Kruskal-Wallis test.

3.1.3.5 Mesenchymal cancer cells are intrinsically more invasive than epithelial cancer cells

To further study the metastatic capabilities of partial EMT hybrid and fully mesenchymal cells, we derived epithelial cell lines from Tnc-CreER^{T2} and Cdh2-CreER^{T2} mice (Figure 6a, b). When stimulated with TGFβ to induce an EMT, the majority of Tnc-CreER^{T2} cells expressed GFP within few days of TGFβ treatment (Figure 6c, d) concomitantly with detectable expression of tenascin C, yet sometimes before losing E-cadherin expression (Supplementary Figure 6a, b). After 10 to 20 days of TGFβ treatment, 85% of cells were GFP^{+}, consistent with *Tnc* being a robust early-stage marker of a TGFβ-induced EMT *in vitro* (Figure 6d). Even in the absence of exogenous TGFβ, a low number of Tnc-CreER^{T2} cells spontaneously underwent an EMT. This could be prevented by treatment with a TGFβ-receptor inhibitor (TGFβRI), indicating that endogenous TGFβ accounted for this spontaneous EMT (Figure 6d, Supplementary Figure 6c, d). In contrast, even upon long-term TGFβ treatment and complete loss of E-cadherin expression, only 30% of Cdh2-CreER^{T2} cells expressed GFP and N-cadherin, which demonstrated the rarity to which cancer cells underwent a full EMT (Figure 6c, d, Supplementary Figure 6e, f).

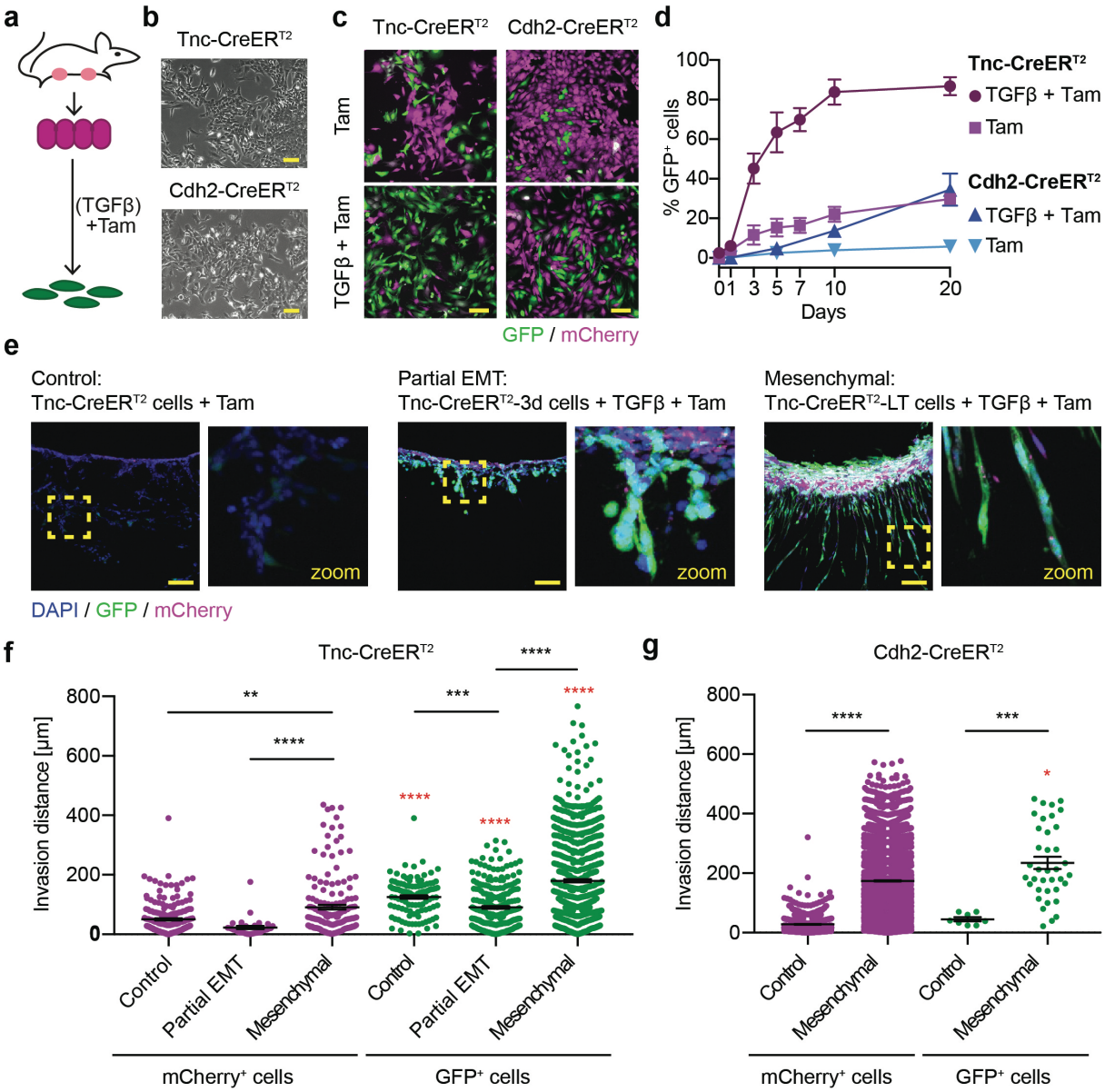


Figure 6: Mesenchymal cells are intrinsically more invasive than epithelial cancer cells. a Cell lines isolated from transgenic Tnc-CreER^{T2} or Cdh2-CreER^{T2} mice switch color upon undergoing an EMT *in vitro*. **b** Representative phase contrast images of established epithelial Tnc-CreER^{T2} and Cdh2-CreER^{T2} cell lines. Scale bars: 100µm. **c** Representative images of Tnc-CreER^{T2} and Cdh2-CreER^{T2} cell lines treated with TGFβ and/or tamoxifen (Tam) for 20 days. Scale bars: 100µm **d** Kinetics of color switch in Tnc-CreER^{T2} and Cdh2-CreER^{T2} cell lines treated with TGFβ and/or Tam over a time course of 20 days. **e** Representative images of collagen I invasion assay of Tnc-CreER^{T2} cells. Cells were kept under TGFβ and/or Tam during the assay. Boxes indicate the area included as magnified insets. Scale bars: 100µm. **f, g** Quantification of invasion distances of mCherry$^+$ or GFP$^+$ Tnc-CreER^{T2} cells (**f**) and Cdh2-CreER^{T2} cells (**g**). Dots represent individual cells that have invaded into the collagen I matrix, lines represent mean +/- SEM. *Comparison between corresponding GFP$^+$ and mCherry$^+$ cells. */**/***/**** p < 0.05/0.01/0.001/0.0001, Kruskal-Wallis test.

In order to assess the intrinsic invasive capabilities of epithelial cells as well as partial EMT hybrid and mesenchymal cancer cell, we performed a collagen I invasion assay

in vitro. Cells were incubated with TGFβ and tamoxifen for three days to induce a partial EMT state (Tnc-CreER^{T2}-3d cells) or were treated with TGFβ for >20 days to obtain mesenchymal sublines (Tnc-CreER^{T2}-LT and Cdh2-CreER^{T2}-LT (long-term) cells). These sublines and corresponding parental cell lines were seeded on top of a 3D cross-linked collagen matrix, which typifies the ECM of a collagen-rich stroma [241, 242], and were incubated with TGFβ and/or tamoxifen for 7 days. Incubation with TGFβ increased the number of GFP$^+$ cells in Tnc-CreER^{T2}-3d cells (incubation for total 10 days) as well as Tnc-CreER^{T2}-LT and Cdh2-CreER^{T2}-LT cells (incubation >30 days) compared to tamoxifen-treated parental cells, consistent with our previous observations in cultured cells (Supplementary Figure 7a, d). Interestingly, Tnc-CreER^{T2}-3d cells invaded as cell collectives, while fully mesenchymal Tnc-CreER^{T2}-LT and Cdh2-CreER^{T2}-LT cells invaded as single cells following one another in multicellular streams (Figure 6e and Supplementary Figure 7c). Even though only a not significantly higher percentage of GFP$^+$ compared to mCherry$^+$ cells had invaded into the collagen, collectively invading cells were predominantly led by GFP$^+$ cells, suggesting a possible role for cells in a partial EMT state as leader cells in collective invasion, as previously reported [108, 156-158] (Figure 6e and Supplementary Figure 7b, e). Moreover, across all samples GFP$^+$ cells had invaded significantly further into the collagen compared to the respective pool of mCherry$^+$ cells, indicating that cells in a partial EMT or fully mesenchymal state were per se more invasive than mCherry$^+$ epithelial cells (Figure 6f, g).

3.1.3.6 Partial EMT and fully mesenchymal cells efficiently seed metastases, while epithelial cells have a higher potential for metastatic outgrowth

Metastatic colonization, which includes metastatic seeding and outgrowth, is a bottleneck for metastasis formation, and a MET is crucial for these processes [43, 188, 190-194]. To assess the colonization ability of cells spontaneously undergoing an EMT, cultured Tnc-CreER^{T2} cells were injected into the tail vein of NSG mice and tamoxifen was administrated to the mice to label cancer cells residing in an EMT state *in vivo* (Figure 7a). Six weeks post injection few GFP$^+$ and GFP$^+$/mCherry$^+$ mixed metastases were found in the lungs, consistent with the low number of cells undergoing a spontaneous EMT *in vitro* (Figure 7b, c). Interestingly, GFP$^+$ metastases were epithelial, and significantly smaller than mCherry$^+$ metastases (Figure 7b, d). However,

there were no differences in cell proliferation between GFP⁺ and mCherry⁺ cells in established metastatic colonies (Supplementary Figure 8). These data suggest that GFP⁺ cells may acquire proliferative abilities similar to mCherry⁺ cells upon undergoing a MET at the metastatic site.

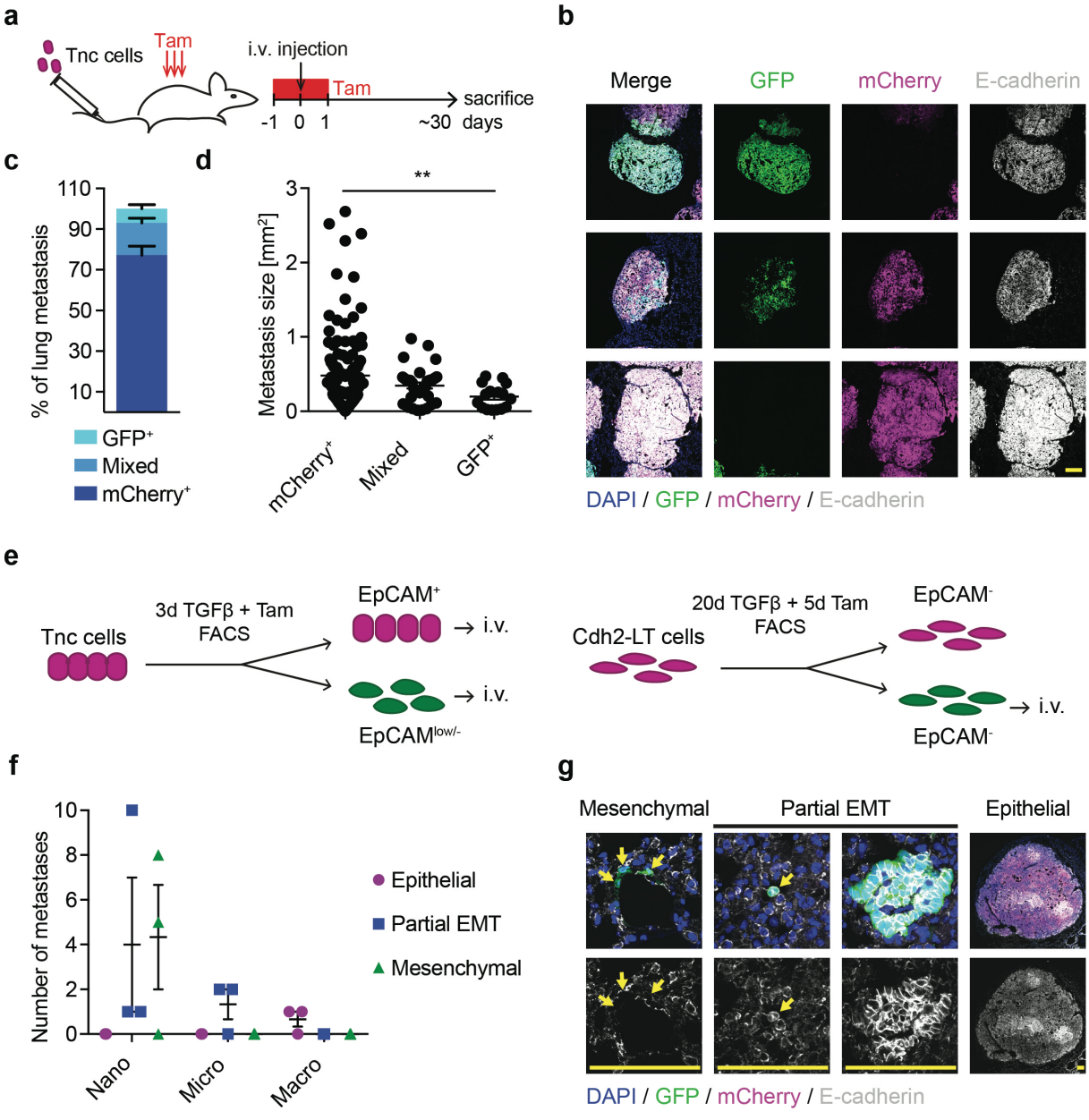


Figure 7: Partial EMT and fully mesenchymal cells efficiently seed metastases, while epithelial cells have a higher potential for metastatic outgrowth. a Strategy for assessing metastatic outgrowth of epithelial cells and cells undergoing a spontaneous EMT. **b** Representative images of GFP⁺, mixed and mCherry⁺ lung metastases stained for E-cadherin, Scale bar: 200µm. **c** Quantification of GFP⁺, mixed and mCherry⁺ metastases, mean + SEM of n = 3 mice are shown **d** Quantification of metastasis size. Dots represent individual metastasis of n = 3 mice, lines represent mean +/- SEM. **p < 0.01, Kruskal-Wallis test. **e** Strategy for isolation of epithelial cells as well as partial-EMT and fully mesenchymal cells for i.v. injection. **f,** Quantification of number and type of lung metastases

per mouse obtained by i.v. injection of indicated cell populations. n = 3 mice, lines represent mean +/- SEM. **g** Representative images of lung metastases showing size and E-cadherin expression. Scale bars: 100µm.

To further assessed the seeding and colonization abilities of cells in an epithelial as well as a partial EMT and a fully mesenchymal cell state, Tnc-CreER^{T2} and Cdh2-CreER^{T2} cells were treated with TGFβ *in vitro* and distinct epithelial and mesenchymal cell populations were isolated by FACS for subsequent tail vein injection. Tnc-CreER^{T2} cells were treated with TGFβ and tamoxifen for 3 days and sorted for mCherry^{+}EpCAM^{+} epithelial cells and GFP^{+}EpCAM$^{low/-}$ partial EMT cells. GFP^{+}EpCAM^{-} fully mesenchymal cells were isolated from Cdh2-CreER^{T2}-LT cells that had been treated with TGFβ for >20 days and tamoxifen for 5 days (Figure 7e, Supplementary Figure 9). Two months post transplantation, mesenchymal cancer cells had only formed E-cadherin^{-} mesenchymal nanometastases (1-7 cells), while cells in a partial EMT hybrid state had formed E-cadherin^{+} nanometastases and few micrometastases (>40 cells). In contrast, mCherry^{+} epithelial cells only formed macrometastases (visible by eye), but the number of metastastatic colonies was much lower compared to partial EMT- and fully mesenchymal cells (Figure 7f, g). Thus, both GFP^{+} partial EMT and fully mesenchymal cells could efficiently seed metastases in the lung. While cells in a partial EMT hybrid state frequently acquired an epithelial phenotype, which possibly allowed them to establish micrometastasis, fully mesenchymal cells did not undergo a MET, consistent with what we have observed in primary tumors of our transgenic mice. Epithelial cells, in comparison, had a greater ability for metastatic outgrowth, but not for seeding.

3.1.4 Discussion

EMT has emerged as a key process involved in tumor progression and metastasis based on experimental manipulation of the process *in vitro* and *in vivo*. However, only a minority of cancer cells undergo an EMT in an unperturbed system *in vivo* and the overall relevance of a spontaneously occurring EMT for metastasis remains obscure [4, 24, 91, 96]. Moreover, EMT is a complex process which covers a heterogeneous spectrum of transition states associated with distinct functional characteristics and metastatic potential [85-88, 115]. However, the spatiotemporal dynamics associated with EMT and MET, particularly the extent to which cells undergo a partial or full EMT, and the contribution of partial and full EMT to metastasis are poorly understood.

Here, we have developed two cancer cell-specific lineage tracing systems which allow to identify cells that have undergone an EMT *in vivo*. Using two distinct mesenchymal-specific $CreER^{T2}$ lines to label cells which have undergone a partial or full EMT, we show that cancer cells rarely undergo a full EMT, but upon undergoing a partial EMT, frequently redifferentiate and acquire an epithelial phenotype. Cells which have undergone a full EMT, in contrast, mostly retain a mesenchymal phenotype. Although similar lineage tracing systems have been described before [105-107, 110-112, 115, 116], this is the first breast cancer cell-specific EMT tracing model. As a consequence, we have been able to specifically visualize cancer cells that have undergone an EMT/MET *in situ*, while clearly distinguishing cancer cells from stromal cells. This has allowed us to study the spatiotemporal localization of cells which have undergone an EMT during primary tumor growth and metastasis formation.

Within primary tumors, cells in a partial EMT or fully mesenchymal state are predominantly localized to the stromal interface or found within the stroma, respectively. Cells that have undergone a MET, in contrast, are mostly part of the well differentiated, epithelial tumor mass and sometimes forming GFP^+ colonies indicative of a proliferative phenotype. Of note, in the partial EMT tracing model, GFP^+ cells have been found locally enriched. This distinct localization of cells in different EMT states indicates an important contribution of the tumor microenvironment to epithelial/mesenchymal cell fate decisions. Moreover, these observations, together with the results from the invasion assay, suggest that cancer cells undergoing a partial EMT at the stromal interface may contribute to local dissemination, possibly by acting

as leader cells in collective invasion. Upon invasion, these cells could eventually undergo a MET and initiate proliferation, thereby contributing to the local spread of cancer cells.

EMT covers a heterogeneous population of intermediate hybrid states. Based solely on E-cadherin expression, it is difficult to distinguish cells that have reacquired a fully epithelial phenotype from cells residing in a certain hybrid state co-expressing epithelial and mesenchymal markers. As Tenascin C is a particularly early EMT marker, it is possible that E-cadherin$^+$GFP$^+$ cells in this model represent a very early EMT hybrid state. However, RNA-FISH has confirmed the absence of *Tnc* expression in GFP$^+$ cells with an epithelial morphology. Furthermore, the emergence of GFP$^+$ colonies, which are associated with cell proliferation and has not been observed in mice treated with tamoxifen acutely, indicates that these cells frequently undergo a full MET. As it is, however, not possible to distinguish GFP$^+$ cells which have themselves undergone an EMT/MET from their progeny, the extent to which cells have undergone a MET may be overestimated. Although there exists a rare population of E-cadherin$^+$ cells that have undergone a full EMT marked by N-cadherin expression, these are mostly single cells which do not form colonies. A full EMT might hence be associated with a lower rate of proliferation, thus a more quiescent cell phenotype. Moreover, it is even possible that these single E-cadherin$^+$GFP$^+$ cells represent rare EMT hybrid populations which may explain why these cells are also observed in mice treated with tamoxifen acutely before sacrifice. In any case, the fate of cells that have undergone a partial or full EMT seems different. Moreover, the high frequency to which cells in a partial EMT state undergo a MET reflects a high degree of cellular plasticity. Whether cells that have undergone a full EMT are intrinsically less plastic remains to be determined. It is possible that the microenvironmental niche that promotes the progression of cells towards a fully mesenchymal cell state concomitantly stabilizes the mesenchymal phenotype.

Recent studies have challenged the functional importance of EMT in metastasis formation. Specifically, no evidence for a contribution of EMT towards metastases has been found using lineage tracing approaches with late-stage mesenchymal markers, such as Fsp1, vimentin, or α-SMA [112, 116]. Consistent with these reports, our data also indicate that a full EMT is not a prerequisite for metastases formation. However, our results show that cells that have undergone a partial EMT are highly enriched in

lung metastases compared to primary tumors. Moreover, the mosaic like composition of mCherry$^+$ and GFP$^+$ cells of a subset of metastases indicates the possibility of collective invasion and dissemination of GFP$^+$ and mCherry$^+$ cancer cells as mixed circulating tumor cell (CTC) clusters. Consistent with this, the invasion patterns observed in the *in vitro* invasion assay imply a role for cells in a partial EMT state as leader cells in collectively invading cell cohorts. Even though CTCs are too rare in our mouse models to directly validate this hypothesis *in vivo*, previous studies have associated CTC clusters with a mixed epithelial/mesenchymal phenotype [177, 180, 183], and a partial EMT has been implicated in collective dissemination [108, 156]. Taken together, our data indicate that a partial EMT may contribute to lung metastasis, possibly by taking the lead in collective invasion.

Metastasis is a complex multistep process and particularly the colonization step is a bottleneck for metastasis formation. In order to successfully establish metastatic nodules, mesenchymal cells need to undergo a MET [43, 188, 190-194]. Previous studies suggested a reduced ability of mesenchymal cancer cells to establish metastases compared to cells in a partial EMT state and epithelial cells [87, 115]. Similarly, genetic loss of E-cadherin increases invasion but reduces overall metastatic burden [204]. Consistent with this, we have found that upon i.v. injection, cells in a partial EMT or fully mesenchymal state are associated with a smaller metastases size compared to epithelial cells. Interestingly, while cells in a partial EMT state may undergo a MET and eventually establish micrometastases, mesenchymal cancer cells survive in a mesenchymal state but fail to grow out. This suggests that special stimuli may be required to induce a MET in fully mesenchymal cells, which might not be readily provided by the microenvironment at the metastatic site. A fully mesenchymal cell state may therefore be associated with the clinically relevant phenomenon of metastatic dormancy and late recurrences in breast cancer patients after therapy [243, 244]. Interestingly, partial EMT and fully mesenchymal cancer cells seed metastases more efficiently than epithelial cells, which suggests that these cell states are associated with increased cell survival and a higher ability to extravasate. The colonizing ability of partial EMT cells which disseminate as cell collectives remains to be determined. Of note, we have not detected any cells that have undergone a full EMT in lungs of Cdh2-CreER^{T2} transgenic mice. It is possible that the specific microenvironment required to induce a fully mesenchymal phenotype at the primary tumor site does not induce invasion, even though these cells could be invasive per se.

Due to biological and technical reasons, our lineage tracing system has certain limitations. First, *Tnc* expression may not be exclusively associated with an EMT, as it has also been found upregulated in a basal-epithelial program implicated in collective dissemination [160, 161]. However, GFP^+ cells in Tnc-$CreER^{T2}$ mice treated with tamoxifen only before sacrifice were predominantly E-$cadherin^{low/-}$ and it is possible that the previously described basal-epithelial program is associated with a certain partial-EMT phenotype. More importantly, flow cytometry analysis revealed that the vast majority of $EpCAM^{low/-}$ mesenchymal cells were $mCherry^+$. Since cancer cells may rarely undergo a full EMT, this is to be expected in the Cdh2-$CreER^{T2}$ model. However, this indicates either inefficient Cre-recombinase-induced tracing in the Tnc-$CreER^{T2}$ model or that the majority of cells are undergoing a partial EMT in the absence of *Tnc* expression. EMT can be triggered by a variety of microenvironmental stimuli which may be associated with different EMT programs and distinct EMT states *in vivo* [24], thereby giving rise to a heterogeneous cell population [87]. While tenascin C is consistently upregulated by TGFβ treatment *in vitro*, it might not be a universal EMT marker *in vivo*. Taking these impediments into consideration, a much greater number of metastases may originate from cells that have undergone a partial EMT.

Finally, although increasing evidence suggests that metastases can be derived from cells that have not undergone an EMT, the mechanisms by which these cells manage to escape from the primary tumor remain to be determined. Importantly, cancer cells that have undergone a partial EMT are frequently localized at the invasive front. They may potentially facilitate the dissemination of $mCherry^+$ epithelial cancer cells without metastasizing themselves or they may be outcompeted at the metastatic site by the fast proliferating epithelial cells.

In conclusion, our study provides novel insights into the dynamics of EMT and MET *in vivo* and shows that cancer cells mostly transition between partial EMT states and rarely undergo a full EMT, at least in the MMTV-PyMT transgenic mouse model of metastatic breast cancer. Furthermore, our data suggest that cells which have undergone a spontaneous partial EMT may contribute to metastasis, while cells which have undergone a full EMT mostly retain a mesenchymal phenotype and are associated with a more quiescent cell state.

3.1.5 Materials and Methods

Animals

Mouse colonies were maintained at the animal facility of the Department of Biomedicine, University of Basel, Switzerland. All experiments were carried out in accordance with the guidelines of the Swiss Federal Veterinary Office and the Cantonal Veterinary Office of Basel-Stadt (Licenses 1878, 1907, 1908). MMTV-PyMT (FVB/N) mice [236] were a gift of N. Hynes (FMI, Basel, Switzerland). RC::FrePe mice (FVB/N; B6; 129S6 - Gt(ROSA)26Sor$^{tm8(CAG-mCherry,-EGFP)Dym}$/J) were a kind gift of S. Dymecki (Department of Genetics, Harvard Medical School, Boston, Massachusetts) [234, 235]. The B6;129S6-Gt(ROSA)26Sor$^{tm9(CAG-tdTomato)Hze}$/J strain [245] was kindly provided by V. Taylor (Department of Biomedicine, University of Basel, Switzerland). MMTV-Flpo mice (FVB/N-Tg(MMTV-Flpo)9Gcr) were described before [237]. Tnc-CreER^{T2} (FVB/N; B6; 129S6-Tnc$^{tm1(cre/ERT2)Gcr}$) and Cdh2-CreER^{T2} (FVB/N; B6; 129S6-Cdh2$^{tm1(cre/ERT2)Gcr}$) knock-in mice were generated by Ingenious Targeting Laboratory (Ronkonkoma, New York) using ES cell electroporation and homologous recombination. Vectors encoding for *Cre-ER^{T2}-RPA* (rabbit beta-globin gene polyadenylation signal) were targeted to the *Tnc* and *Cdh2* loci, respectively. Knock-in mice of B6/129S6 mixed origin were crossed into the FVB/N background for a minimum of 3 generations prior to starting experiments. Cre alleles were detected by PCR using primers 5'–CGG TCG ATG CAA CGA GTG ATG AGG–3' (forward) and 5'–CCA GAG ACG GAA ATC CAT CGC TCG – 3' (reverse). Cre mice were bred to the MMTV-Flpo strain to obtain double transgenic mice. These were further bred to the MMTV-PyMT and RC::FrePe strains to obtain *Tnc-CreER^{T2}/MMTV-Flpo/RC::FrePe/MMTV-PyMT* or *Cdh2-CreER^{T2}/MMTV-Flpo/RC::FrePe/MMTV-PyMT* females. All alleles were kept strictly heterozygous. NOD/Scid;common γ receptor$^{-/-}$ (NSG) mice were used for transplantation experiments.

Tamoxifen treatment

To induce Cre recombinase activity, mice were treated with 2mg of Tamoxifen (Sigma) dissolved in 100µl of 10% ethanol in sunflower seed oil (Sigma) by intraperitoneal injection. Transgenic mice received Tamoxifen either on three consecutive days just before sacrifice, at 11 or 9 weeks of age or were treated twice per week starting at 9 weeks of age and continuing until sacrifice. Mice were sacrificed before a tumor volume

of 1500mm^3 was reached, usually around 13-14 weeks of age. Tamoxifen treatment of NSG mice is described below.

Tumor piece transplantation and primary tumor removal

Tumors from a *Tnc-CreER^{T2}/MMTV-Flpo/RC::FrePe/MMTV-PyMT* female were cut in approximately 1mm^3 pieces and transplanted into the 4th mammary gland of NSG mice. Tumors were propagated for several rounds in NSG mice before one propagated tumor was selected for the experiment. For transplantation of tumor pieces and primary tumor removal, mice were anaesthetized with isoflurane and received meloxicam (5 mg/kg body weight per day) for two or four days for pain relief. The incision/wound was closed with metal clips (7mm, Alzet Wound Closure System) which were removed after 10 days. Tumor size was measured by caliper twice per week. Tumors were surgically removed at a size of (~500mm^3). Mice had either received tamoxifen for three consecutive days one week prior to surgery at a tumor size of ~200mm^3 or received tamoxifen one-week after surgery. Mice were sacrificed approximately four weeks after tumor removal to assess lung metastasis.

Tail vein injections

For i.v. injections of spontaneous EMT cells, 500'000 Tnc-CreER^{T2} cells suspended in 100µl PBS were injected into the tail-vein of NSG mice. Mice were treated with tamoxifen for three consecutive days starting one day prior to i.v. injection and were sacrificed 6 weeks after. For epithelial and mesenchymal populations isolated by FACS, 30'000 cells were injected into NSG mice. Lungs were collected two months after transplantation.

Flow cytometry analysis and FACS

Dissected tumors were minced and digested with 1mg/ml collagenase D (Roche, 11088858001) and 0.4X Trypsin/EDTA (Sigma, T4174) diluted in DMEM/F12 (Sigma) containing 2.5% heat-inactivated FBS, 10ug/ml gentamycin (Sigma, G1397), and antibiotic-antimycotic (ThermoFisher, 15240062) for 1 hour at 37°C while shaking. Digested tumors were incubated for 5 min with 0.1 mg/ml DNaseI (Roche, 11284932001) at room temperature and filtered through a 70µm cell strainer (BD 352350). Cells were resuspended in 2% FBS in PBS containing 2mM EDTA (FACS buffer) followed by blocking with CD16/32 (1:100, Biolegend, 101302) for 5 min at room

temperature. Lineage-negative cells were depleted by incubation with biotinylated antibodies against CD31 (1:100, Biolegend, 102404), CD45 (1:100, Biolegend, 103104) and Ter119 (1:100, Biolegend, 116204) for 15 min on ice, followed by immunomagnetic pulldown using streptavidin-conjugated dynabeads (ThermoFisher, 11047). Cancer cells were washed with FACS buffer and stained with allophycocyanin (APC)-conjugated anti-EpCAM antibody (1:100, BioLegend, 118213) and BV650-conjugated streptavidin (1:100, BioLegend, 405232) for 20min in the dark on ice. Cells were washed with FACS buffer, filtered through a 40µm mesh and stained with Draq7 (1µM, Biostatus) to label dead cells before loading on a FACSAria II (BD Biosciences). For FACS of cultured cells, Tnc-CreER^{T2} and Cdh2-CreER^{T2} cells were trypsinized, resuspended in FACS buffer and incubated for 20min with APC-conjugated anti-EpCAM antibody (1:100, BioLegend, 118213). Cells were washed with FACS buffer and filtered (40µm pore size). Dead cells were labeled with 4′,6-Diamidin-2-phenylindol (DAPI 2µg/ml) or Draq7 (1µM, Biostatus). The samples were subjected to sorting on FACSAria II. Cells were kept on ice throughout the whole procedure. Data was processed and analyzed with FlowJo^{TM} software (Becton, Dickinson and Company).

Tissue processing and immunofluorescent staining of frozen sections

Tissues were fixed in 4% paraformaldehyde (PFA) for two hours followed by overnight incubation with 20% sucrose at 4 °C for cryoprotection. Tissues were embedded in Tissue-Tek O.C.T compound (Sakura), snap-frozen and stored at −80 °C. 10µm thick sections were cut and dried for 30 min. Before staining, sections were rehydrated in PBS followed by permeabilization with 0.2% Triton X-100 in PBS for 20 min and blocking with 5% normal goat serum (NGS, Sigma, G6767) in PBS for one hour. Sections were incubated with primary antibody overnight at 4 °C followed by incubation with secondary antibody for 1 h at room temperature. All antibodies were diluted in 5% NGS in PBS. The following primary antibodies were used: Rat-anti-E-cadherin (clone ECCD-2, Invitrogen, 13-1900, 1:400), chicken-anti-Vimentin (Novus Biologicals, NB300-223, 1:100), rabbit-anti-α-SMA (Abcam, ab 5694, 1:100), rabbit-phospho-histone H3 (Merck, Millipore, 06-570, 1:200). Rat-anti-Tenascin C (clone MTn-12, [246]) was kindly provided by G. Orend (INSERM, Strasbourg). Secondary antibodies directed against the species of the primary antibody were coupled to Alexa Fluor 647 (Invitrogen, 1:400) or 633 (goat-anti-chicken, Invitrogen, 1:400). Nuclei were stained with DAPI (Sigma, 1µg/ml) for 15 min at room temperature. tdTomato, GFP and

mCherry were detected by endogenous fluorescence. The sections were mounted with Dako fluorescence mounting medium (Agilent). Confocal micrographs of immunofluorescence staining were acquired using a spinning disk confocal microscope (Visitron) with a 20x, 0.75 NA or 60x, 1.2 NA objective or a Leica confocal SP5 microscope with a 20x 0.8 NA objective (tenascin C staining).

For N-cadherin staining, PBS was replaced by HBS-Ca^{2+} (10 mM Hepes pH7.4, 150mM NaCl, 1mM $CaCl_2$). Upon rehydration, sections were rinsed with double distilled H_2O (ddH_2O) and incubated for 20 min in 10 mM citrate buffer pH6 at 80°C for antigen retrieval. Tissues were cooled down for 10 min at room temperature and rinsed with ddH_2O prior to permeabilization. Sections were blocked with M.O.M kit (Vector laboratories, BMK-2202) according to the manufacturer's instructions and incubation with mouse-anti-N-cadherin antibody (MAB 3B9, ThermoFisher, 33-3900, 1:100) and rabbit-anti-tdTomato (anti-DsRed, Clontech, 632496, 1:200) overnight at 4°C followed by incubation with anti-rabbit Alexa Fluor 568 and anti-mouse Alexa Fluor 633 secondary antibodies. Images were acquired using a Leica confocal SP5 microscope with 63x, 1.4 NA objective.

Hematoxylin and Eosin (H&E) staining of frozen sections

Frozen sections cut at 10µm thickness were rehydrated for 30min in ddH_20 and stained with hematoxylin (Thermo Scientific) for 15 s. Sections were rinsed with tap water, briefly dipped in 0,5% HCl in water, rinsed with tap water and incubated in 0.1% lithium carbonate for 10 s for blueing. Sections were then washed with ddH_2O, dehydrated through a series of 30%, 70% and 80% ethanol and stained with eosin (Sigma Aldrich) for 30s. Sections were further dehydrated with 90 and 100% ethanol prior to 1h incubation with Xylene and mounting with Cytoseal XYL (Thermo Scientific). Images were acquired using a Zeiss Axio Imager Scanning Microscope with a 40x, 0.95 NA objective.

RNA-FISH on frozen tissue sections

Custom Stellaris® FISH Probes were designed against mouse *Tnc* (exon 2 and 3) by utilizing the Stellaris® RNA FISH Probe Designer (Biosearch Technologies, Inc., Petaluma, CA) available online at www.biosearchtech.com/stellarisdesigner (Version 4.2). The probes were labeled with Quasar 670 Dye. RNA-FISH was performed following the manufacturer's instructions available online at

www.biosearchtech.com/stellarisprotocols. Nuclei were visualized with DAPI (1:5000). Images were acquired using a spinning disc confocal microscope (Visitron) with a 60x, 1.2 NA objective.

Image processing and analysis

All obtained micrographs were processed using the ImageJ software (ImageJ, Wayne Rasband, National Institutes of Health, USA). The individual image quantifications were performed as described below.

Quantification of E-cadherin, Vimentin and α-SMA expression of GFP$^+$ cells in primary tumors of mice treated with Tamoxifen before sacrifice.

A minimum of 40 random imaging fields (60x) containing GFP$^+$ cells were analyzed per mouse. n=3 mice were analyzed.

Quantification of E-cadherin expression and classification of GFP$^+$ cells in primary tumors of mice treated continuously with tamoxifen.

Each one tissue sections of 2-3 tumors per mouse and a total of n=8 mice per line were analyzed. Images (60x) of all GFP$^+$ cells per section were classified and counted. To determine absolute numbers of GFP$^+$ cells per mm^2, the total tumor area was approximated as the mCherry$^+$ area. To this end, the entire tumor cross-section was imaged using a Zeiss Axio Imager Scanning Microscope with 10x 0.45 NA objective and processed using Zeiss Zen 2 software (blue edition). The mCherry$^+$ area was determined by global thresholding based on Otsu algorithm (manually corrected) using ImageJ.

Quantification of GFP$^+$ cells in metastasis and corresponding primary tumors

The lungs of transgenic mice were serial sectioned at 10µm thickness. All frozen sections were screened for lung metastasis by examining fluorescence of GFP and mCherry using a Leica DMI 4000 microscope. The individual metastases were imaged at 20x magnification. To approximate the percentage of GFP$^+$ cells per individual metastasis, GFP$^+$ and mCherry$^+$ areas were determined by global thresholding based on Otsu algorithm (manually corrected) using ImageJ. Metastasis size was calculated as cumulative area of the individual metastasis on all serial sections multiplied by section thickness. Each one tumor cross-section of two primary tumors per mouse

were stained with DAPI and imaged using a Zeiss Axio Imager Scanning Microscope with 10x 0.45 NA objective for quantification of GFP^+ and $mCherry^+$ area as described above. n=8 mice were analyzed per line. For NSG mice transplanted with tumor pieces, cross-sections of each one section of three lung lobes and three sections of primary tumors with 250µm interval were analyzed as described above. n=10 and n=6 mice were analyzed for tamoxifen treatment pre- and post-tumor removal, respectively. For NSG mice i.v. injected with 500'000 Tnc-$CreER^{T2}$ mice, one section per lung was analyzed per mouse. n=3 mice were analyzed. For NSG mice injected with 30'000 cells isolated by FACS, 300 serial sections per mouse were screened for metastases. n=3 mice were analyzed per cell population.

Cell culture

All cells were cultured in DMEM high glucose (Sigma, D5671) supplemented with 10% fetal bovine serum (FBS, Sigma), 2mM glutamine (Sigma, G7513) 100 U penicillin and 0.1 mg/ml streptomycin (Sigma, P4333). Py2T cells have been derived from a mammary tumor of a female MMTV-PyMT mouse [247]. For isolation of cell lines from *Tnc-$CreER^{T2}$/MMTV-Flpo/RC::FrePe/MMTV-PyMT* and *Cdh2-$CreER^{T2}$/MMTV-Flpo/RC::FrePe/MMTV-PyMT* mice, tumors were minced and enzymatically digested as described above. Primary cells were cultured in normal culture medium supplemented with 10% horse serum (Bioconcept Amimed). After epithelial clones became dominant, horse-serum was removed and $mCherry^+EpCAM^+$ epithelial cells were purified by FACS. The following reagents were used for *in vitro* experiments: recombinant human TGFβ (2 ng/ml, R&D Systems, 240-B), TGFβ receptor kinase inhibitor SB431542 (10µM, Sigma, S4317) and Tamoxifen (4-Hydroxytamoxifen, 100nM, Sigma, T176-10MG). All experiments were carried out in biological triplicates.

Immunofluorescent staining of cells

Cells grown on uncovered glass coverslips (#1, round, 10mm) were washed with PBS (or Hank's Balanced Salt Solution (HBSS, Gibco) for N-cadherin staining) and fixed with 4% PFA for 15 min. Cells were permeabilized with 0.5% NP40 in PBS for 5 min and blocked with 3% bovine serum albumin in PBS for one hour at room temperature. Antibodies were diluted in blocking solution and cells were incubated with primary antibody overnight at 4 °C followed by 1 hour of incubation with secondary antibody at room temperature. The following antibodies were used: Rat-anti-E-cadherin (clone

ECCD-2, Invitrogen, 13-1900, 1:400), rat-anti-Tenascin C (clone MTn12, GeneTex, 1:400), mouse-anti-N-cadherin (BD Transduction Laboratories, 610921, 1:400), rabbit-anti-mCherry (anti-DsRed, Clonetech, 632496, 1:200). Secondary antibodies directed against the species of the primary antibody were coupled to Alexa Fluor 647 or 568 (Invitrogen, 1:400). GFP was detected by endogenous fluorescence and nuclei were stained with DAPI for 10 min at room temperature. The coverslips were mounted with Dako fluorescence mounting medium (Agilent). Images were acquired using a Leica DMI 4000 microscope with a 20x, 0.4 NA or 40x, 1.3 NA objective and processed with ImageJ. An ImageJ macro was developed to count GFP^+ cells in 20x images.

Invasion assay

For 3D collagen invasion assay, Rat tail collagen I (Cultrex 3D Culture Matrix Rat Collagen I, R&D Systems) was neutralized with NaOH and diluted to a final concentration of 2.2 mg/ml in RPMI. 150µl of collagen solution was cast into Falcon cell culture inserts (8.0µm pore size, 24-well, Corning) placed in a 24-well plate (Corning). Gels were allowed to polymerize for one hour at 37°C. Cells ($6x10^5$ per insert) were suspended in 300µl culture medium containing 1% FBS and were seeded on top of the gel. 800µl normal culture medium (containing 10% FBS) supplemented with recombinant human bFGF (25ng/ml, Sigma) as a chemoattractant was added to the bottom chamber. TGFβ and/or tamoxifen were added to the top and bottom chambers. Cells were cultured for 7 days and media were replenished every two to three day. Gels were fixed with 4% PFA for one hour at 4°C, dehydrated in 20% sucrose overnight at 4°C and embedded in Tissue-Tek O.C.T. compound. Serial cryosections were cut at 50µm thickness. Dried sections were incubated with ice-cold acetone:methanol (1:1) for 30 seconds, washed with PBS and nuclei were stained with DAPI (5µg/ml) for 2 hours at room temperature. Sections were mounted with ibidi mounting medium (Ibidi) and sealed with Cytoseal XYL. Images were acquired using a spinning disk confocal (Visitron) using a 10x, 0.45 NA or 20x, 0.75 NA objective. Stacks of 81 sections with 0.5µm step-size were acquired. Images were processed with ImageJ and a combination of all serial optical sections with maximum projection was generated for image display. The invasion assay was performed three times and individual replicates were carried out in technical duplicates. An ImageJ macro was developed to quantify the numbers of invading cells and invasion distances of individual cells. For quantification of percent $mCherry^+$ and GFP^+ invading cells,

images of three times three serial sections of different locations within the collagen gel were analyzed per technical duplicate (total 18 images per biological replicate). For quantification of invasion distances of mCherry$^+$ and GFP$^+$ Tnc-CreER^{T2} cells, 1 image was analyzed per biological replicate. Due to very low number of GFP$^+$ Cdh2-CreER^{T2} cells, invasion distances of all GFP$^+$ cells (18 images per biological replicate) were measured and 3 images per replicate were analyzed for invasion distance of mCherry$^+$ Cdh2-CreER^{T2} cells.

RNA Isolation and qPCR

Total RNA was isolated using TRI reagent (Sigma) according to the manufacturer's instruction. cDNA was synthesized using ImProm-II Reverse Transcriptase (Promega). qPCR was performed on a StepOnePlus machine (Applied Biosystems) using Power-Up SYBR green (Applied Biosystems). Target gene expression levels were normalized to Rpl19. Fold changes were calculated using the comparative Ct method ($\Delta\Delta$Ct). The following primers were used:

Rpl19 5'-CTCGTTGCCGGAAAAACA-3'
5'-TCATCCAGGTCACCTTCTCA-3'
Cdh1 5'-CGACCCTGCCTCTGAATCC-3'
5'-TACACGCTGGGAAACATGAGC-3'
Tnc 5'- GGGCTATAGAACACCGATGC-3'
5'- CATTTAAGTTTCCAATTTCAGGTTC-3'
Cdh2 5'- CTGCCATGACTTTCTACGGAGA-3'
5'- CAATGACGTCCACCCTGTTCT-3'.

Western blot analysis

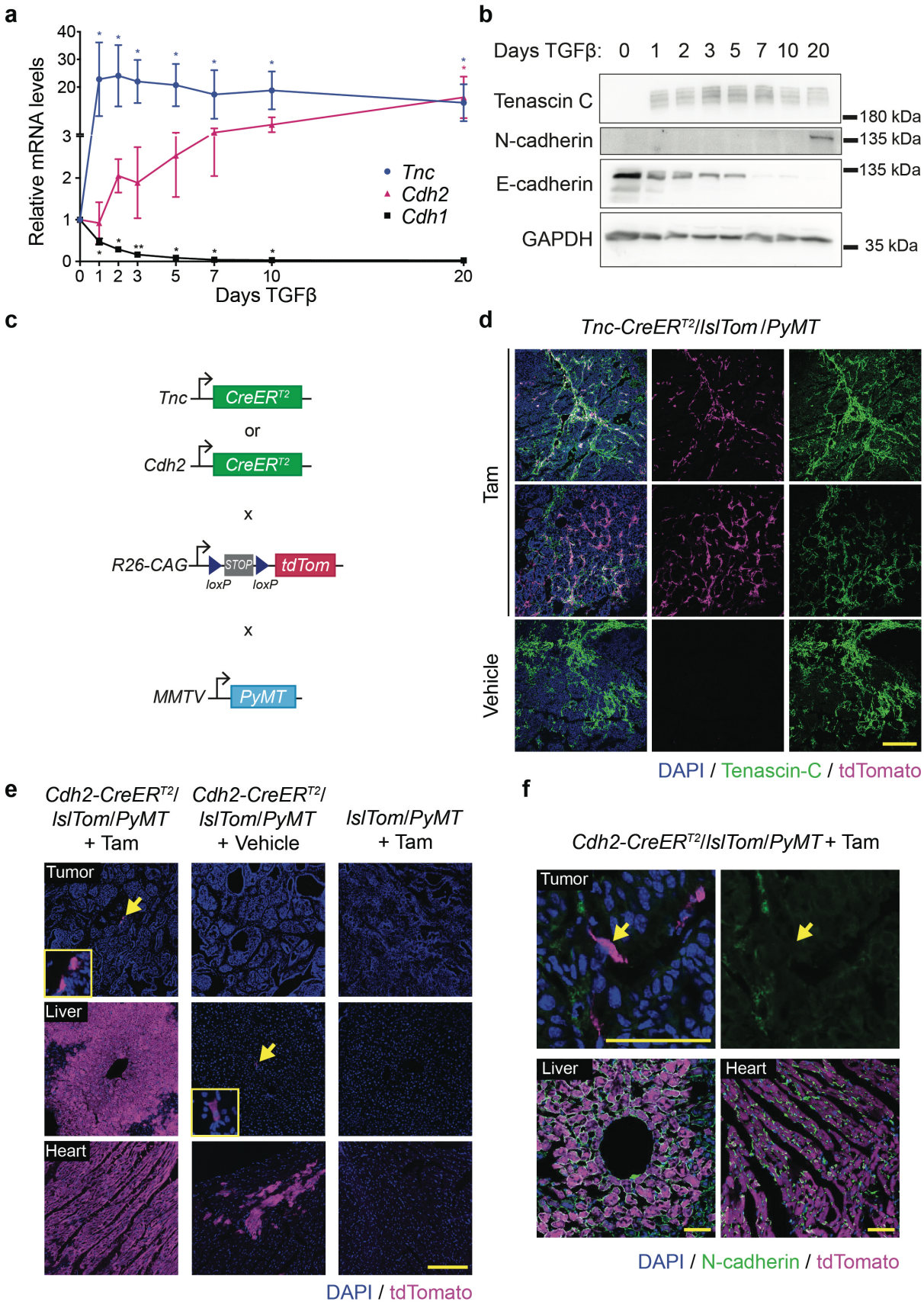
Cells were lysed by boiling in lysis buffer (300mM Tris-HCl pH6.8, 6% SDS, 25% glycerol) for 5 min at 95°C. Protein concentration was determined by Bradford assay (Bio-Rad). Equal amount of protein was resolved by SDS-PAGE (10% gel) and transferred to nitrocellulose membranes (Amersham Protran, Sigma-Aldrich) by wet-transfer. Membranes were blocked in 5% milk in 0.05% Tween-20 in PBS for one hour at room temperature followed by incubation with primary antibodies overnight at 4°C and incubation with HRP conjugated secondary antibodies for one hour at room temperature. Antibodies were diluted in blocking solution. Protein was detected by

chemiluminescence using a Fusion F67 chemiluminescence reader (Vilber Lourmat, France). The following antibodies were used: mouse-anti-E-cadherin (BD Transduction Laboratories, 610182, 1:1500), rat-anti-Tenascin C (clone MTn12, GeneTex, 1:500), mouse-anti-N-cadherin (BD Transduction Laboratories, 610921, 1:2500), rabbit-anti-GAPDH (Abcam, ab9485, 1:2500). Secondary antibodies directed against the species of the primary antibody were coupled to horseradish peroxidase (Jackson ImmunoResearch, 1:10'000).

Statistical Analysis

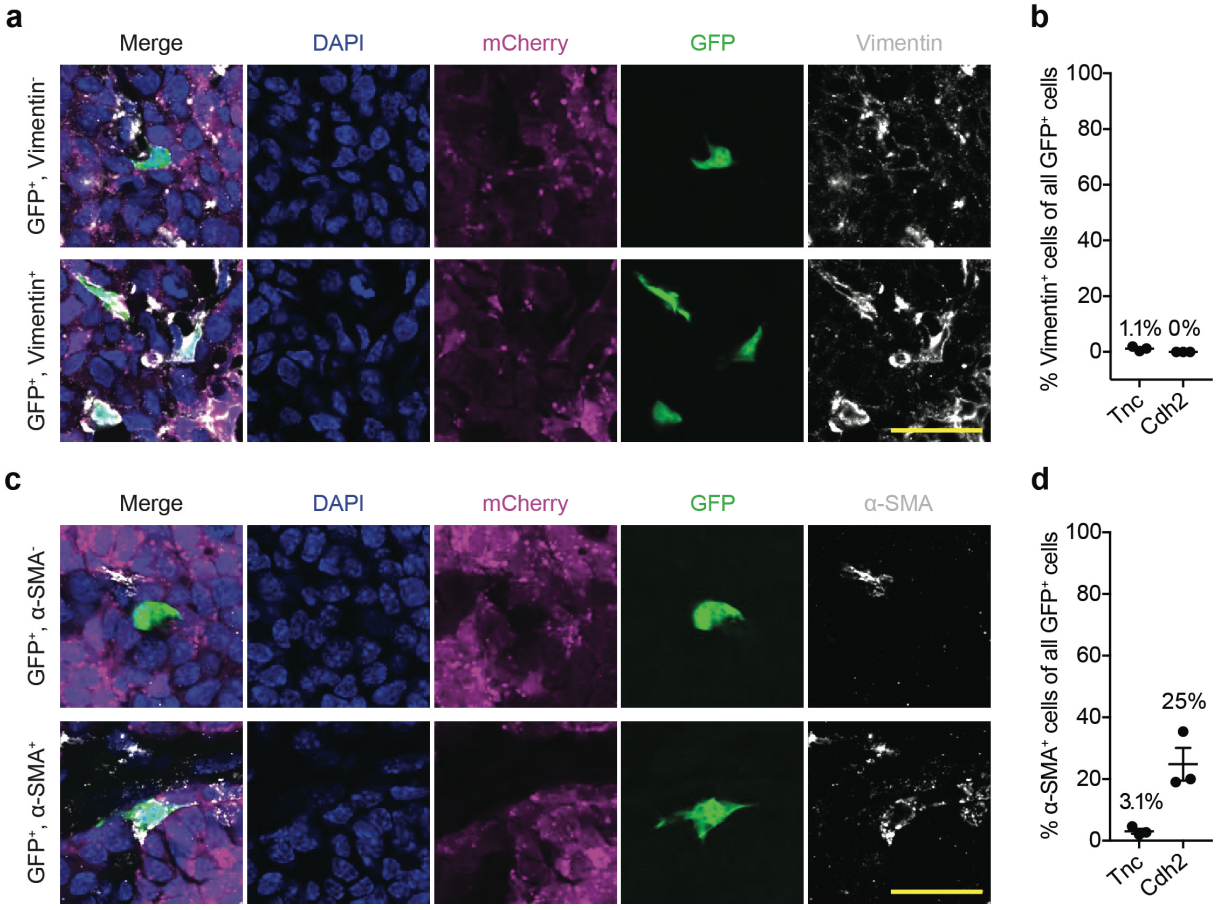
Statistical analysis was performed using GraphPad Prism (GraphPad Software Inc.). Graphs display single measurement points or mean values per animal as indicated in figure legends. Error-bars represent SEM. For mouse experiments, differences between two groups were analyzed by two-tailed non-parametric t-tests (Mann-Whitney *U* test). Multiple groups were analyzed by Kruskal-Wallis test. (Ratio-)paired t-tests were performed for *in vitro* experiments.

3.1.6 Supplementary Figures

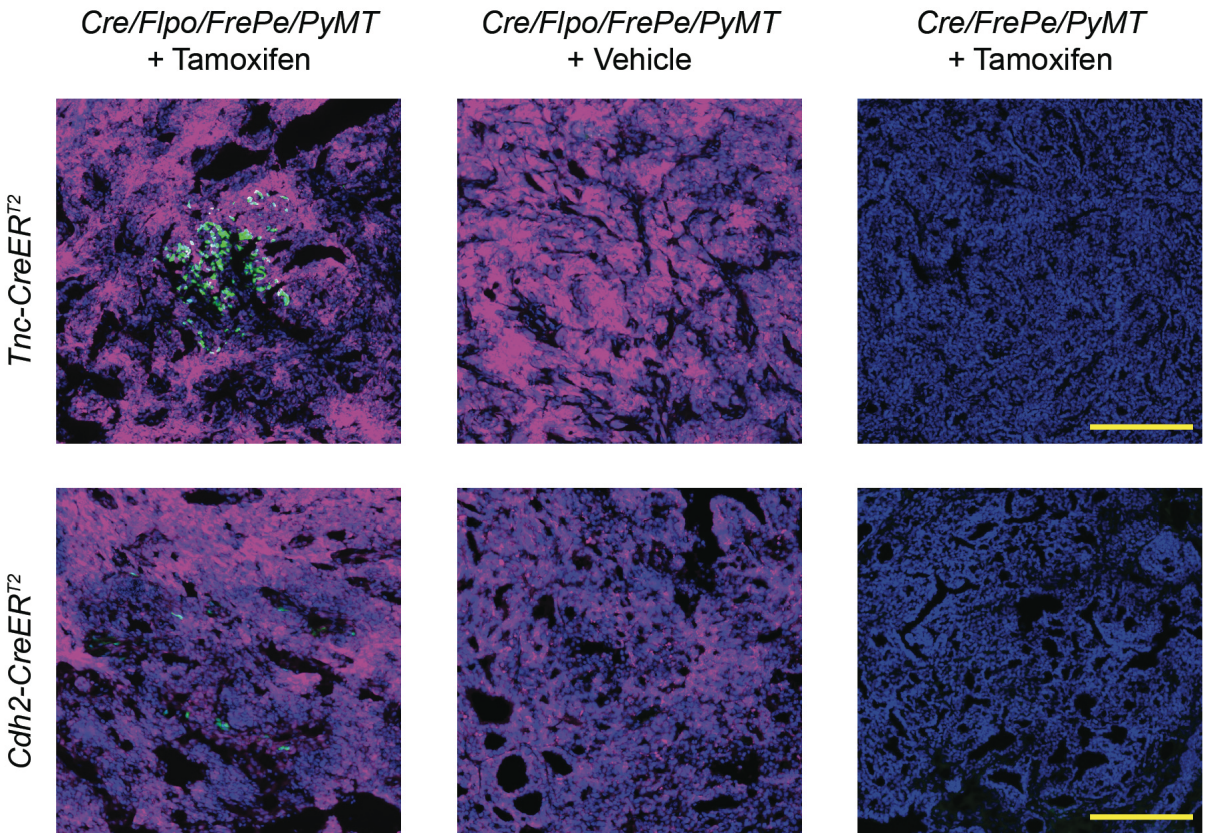


Supplementary Figure 1: Validation of novel Tnc-CreER^{T2} and Cdh2-CreER^{T2} mouse lines for EMT lineage tracing. a, b qPCR and immunoblot analysis of Tenascin C, N-cadherin and E-cadherin mRNA and protein

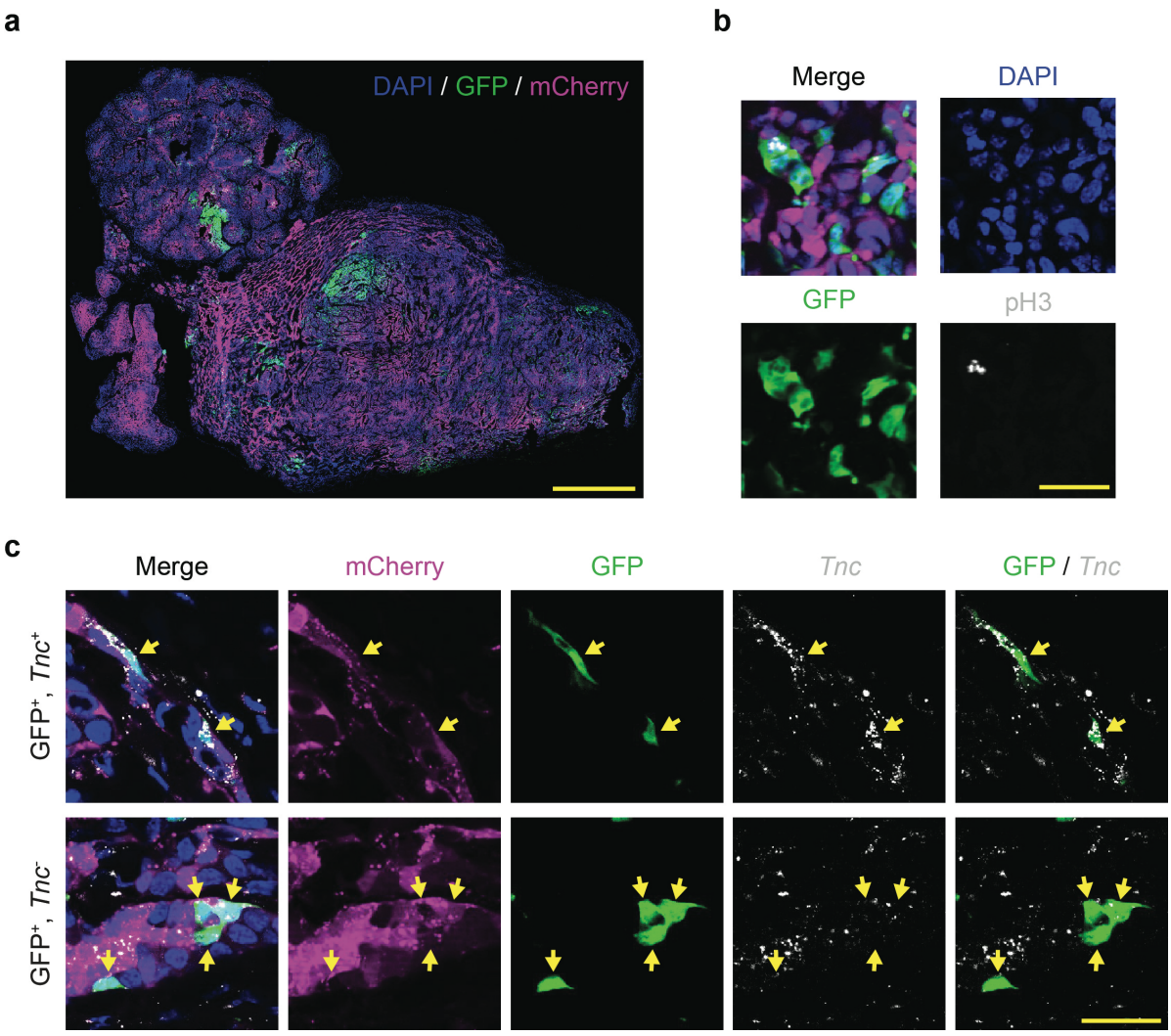
expression levels during a TGFβ-induced EMT in Py2T cells. GAPDH was used as a loading control. */**p < 0.05/0.01, ratio-paired t-test. **c** Strategy for validation of Cre activity of novel CreER^{T2} mouse lines. **d** Tenascin C staining on tumors of *Tnc-CreER^{T2}*/*lslTom*/*PyMT* mice treated with tamoxifen (Tam) or vehicle at 11 weeks of age. Scale bar: 200µm. **e** tdTomato expression in tissues of *Cdh2-CreER^{T2}*/*lslTom*/*PyMT* and *lslTom*/*PyMT* mice treated with Tam or vehicle at 11 weeks of age. Arrows indicate tdTomato-positive cells shown in magnified insets. Scale bar: 200µm. **f** N-cadherin staining on tissue sections of Cdh2-*CreER^{T2}*/*lslTom*/*PyMT* mice treated with Tam. Arrows indicate tdTomato-positive cell. Scale bar: 50µm.



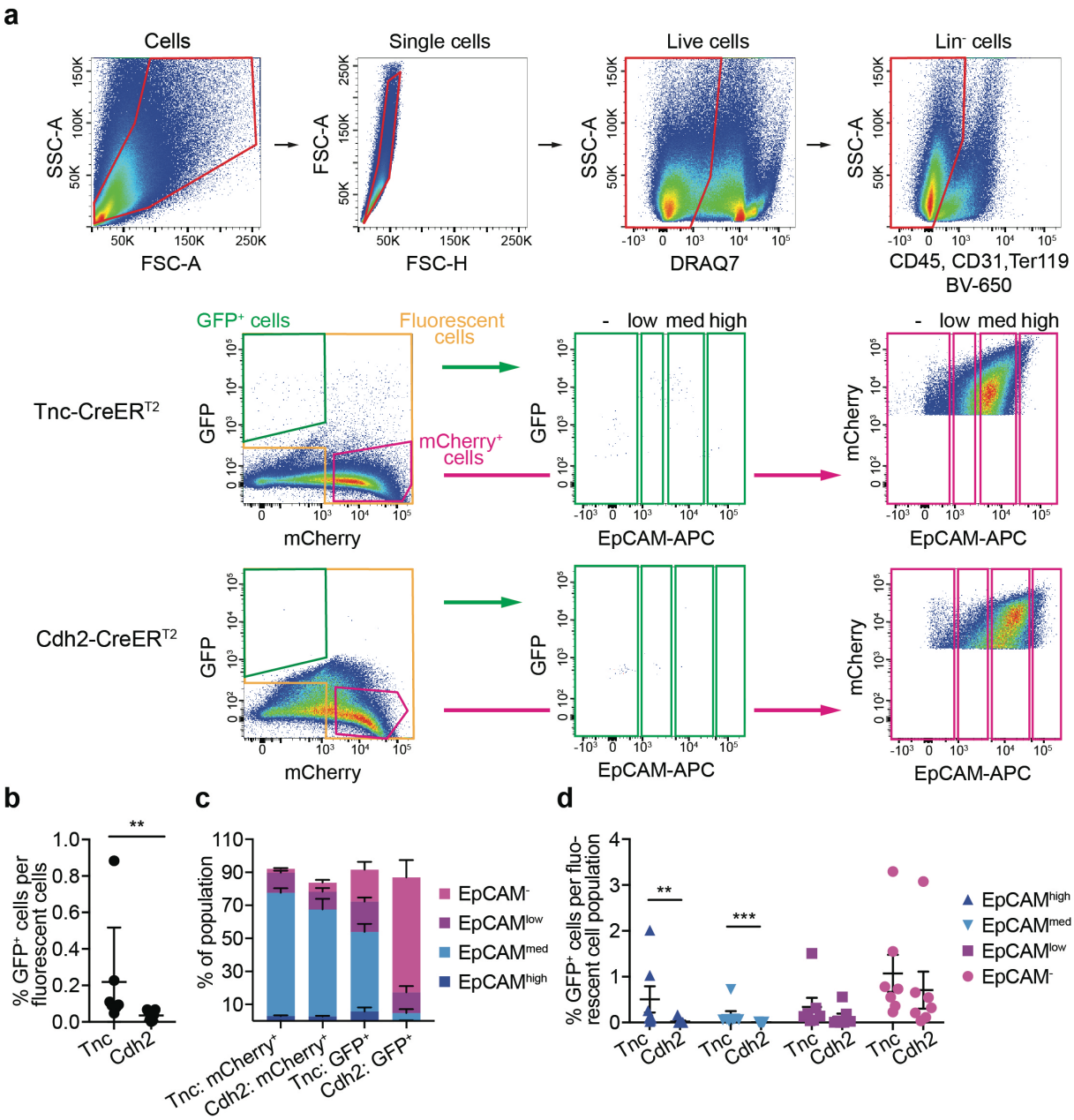
Supplementary Figure 2: Characterization of cells in partial and full EMT states. a, c Representative images of Vimentin$^-$ or Vimentin$^+$ GFP$^+$ cells **(a)** and α-SMA$^-$ or α-SMA$^+$ GFP$^+$ cells **(c)** in tumors of mice treated with tamoxifen before sacrifice. Scale bars: 25µm. **b, d** Quantification of GFP$^+$ Vimentin$^+$ cells **(c)** or GFP$^+$ α-SMA$^+$ cells **(d)** in Tnc-CreER^{T2} (Tnc) and Cdh2-CreER^{T2} (Cdh2) mice. n = 3 mice, lines represent mean +/- SEM. Mean is indicated on top of column.



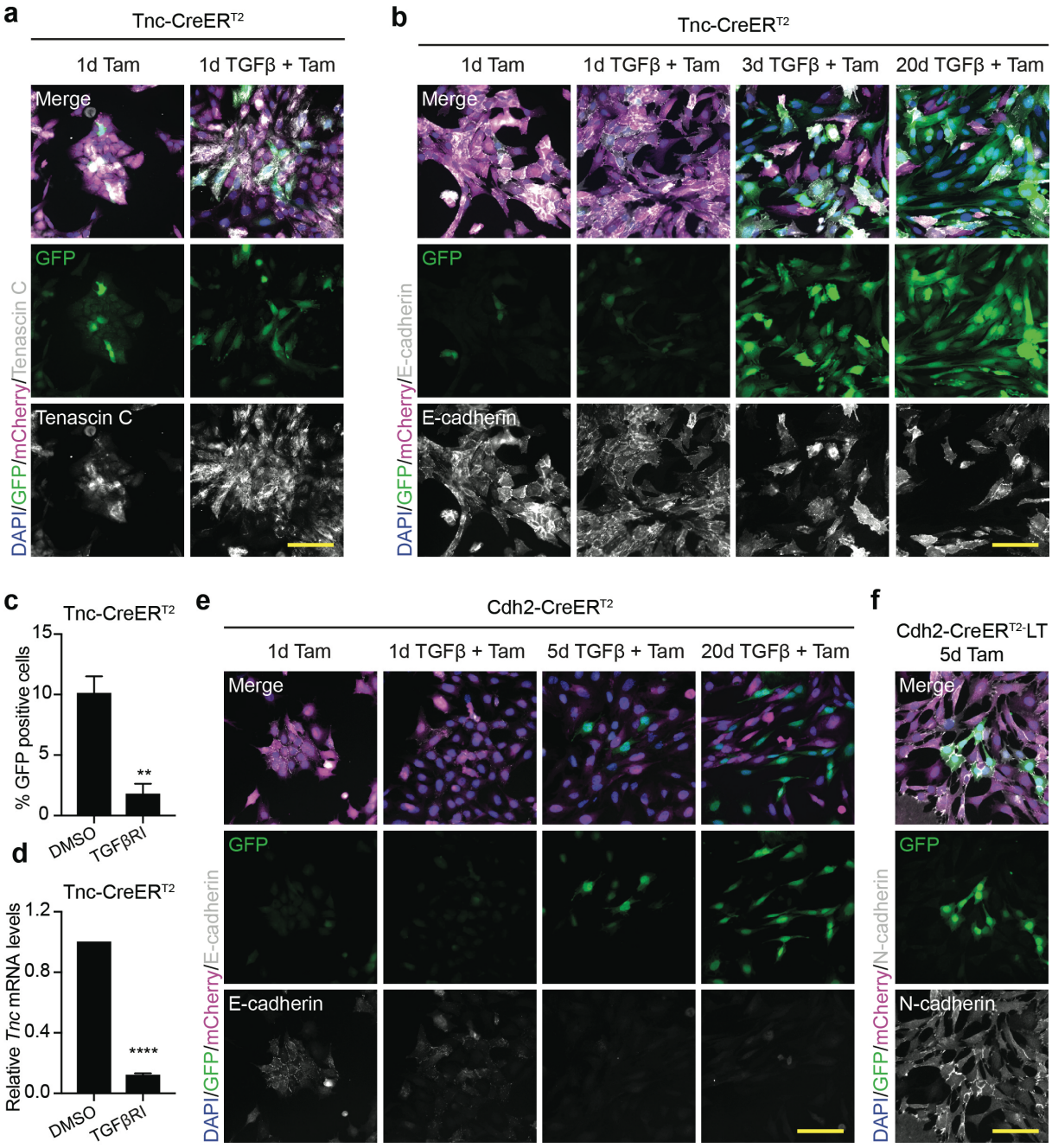
Supplementary Figure 3: No leakiness of GFP or mCherry expression in mice treated with vehicle or negative for *Flpo*. GFP and mCherry expression in tumors of Tnc-CreER^{T2} and Cdh2-CreER^{T2} mice. Mice were treated with tamoxifen or vehicle twice per week starting at nine weeks of age and continuing until sacrifice. Scale bars: 200µm



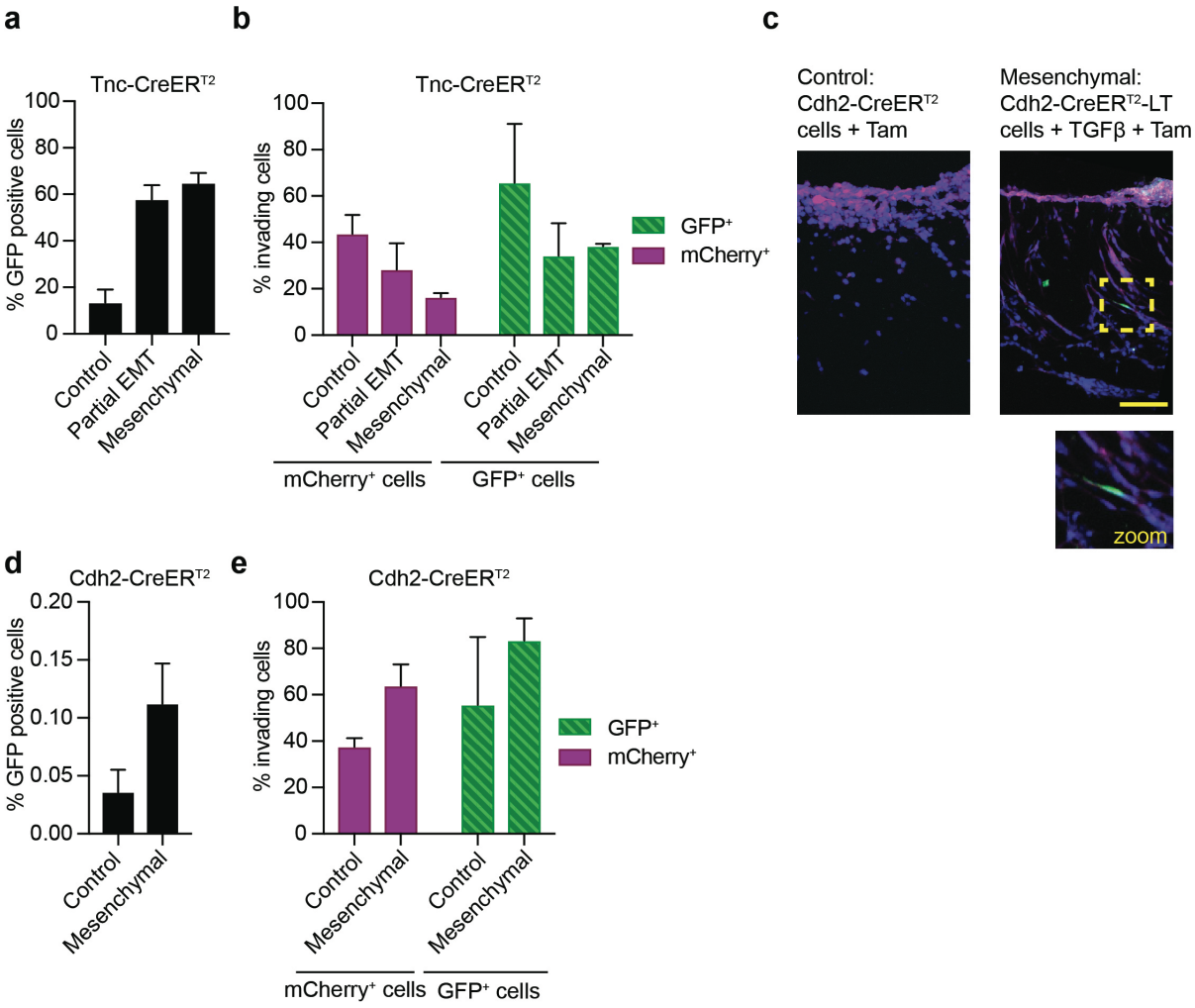
Supplementary Figure 4: Characterization of cancer cells that have undergone an EMT. a Representative image of tumor section from a Tnc-CreER^{T2} mouse. Scale bar: 1mm. **b** Representative image of pH3^{+} GFP^{+} cell in a Tnc-CreER^{T2} tumor. Scale bar: 25µm. **c** RNA-FISH for *Tnc* mRNA on tumor sections of Tnc-CreER^{T2} mice. Representative images of GFP^{+} *Tnc*$^{+}$ and GFP^{+} *Tnc*$^{-}$ cells are shown. Arrows indicate GFP^{+} cells. Scale bar: 25µm.



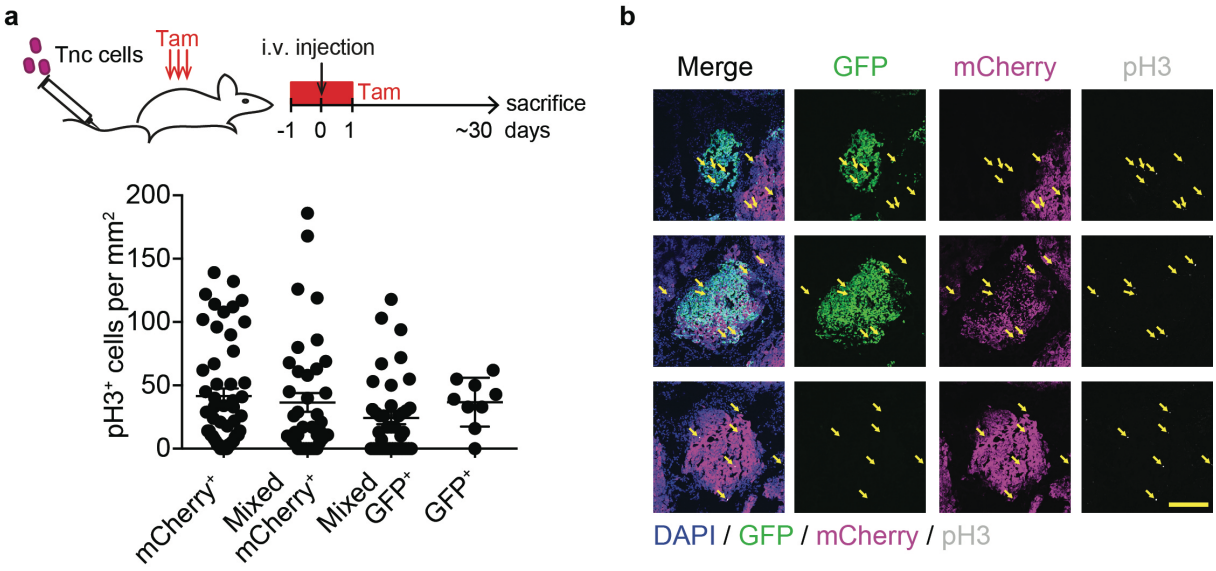
Supplementary Figure 5: Flow cytometry analysis of primary tumor cells. a Gating strategy to identify EpCAM negative, low, medium and high populations of GFP$^+$ and mCherry$^+$ tumor cells. **b** Percentage of GFP$^+$ cells per all fluorescent cells (orange gate) in Tnc-CreER^{T2} (Tnc) and Cdh2-CreER^{T2} (Cdh2) mice. n = 7 mice, lines represent mean +/- SEM. **p < 0.01, Mann-Whitney *U* test. **c** Percentage of EpCAM negative, low, medium and high subpopulations of mCherry$^+$ and GFP$^+$ tumor cell populations in Tnc-CreER^{T2} and Cdh2-CreER^{T2} mice. n=7 mice, mean + SEM are shown **d** Percentage of GFP$^+$ subpopulations in total fluorescent cell populations. n = 7 mice, lines represent mean +/- SEM. **p < 0.01, Mann-Whitney *U* test.



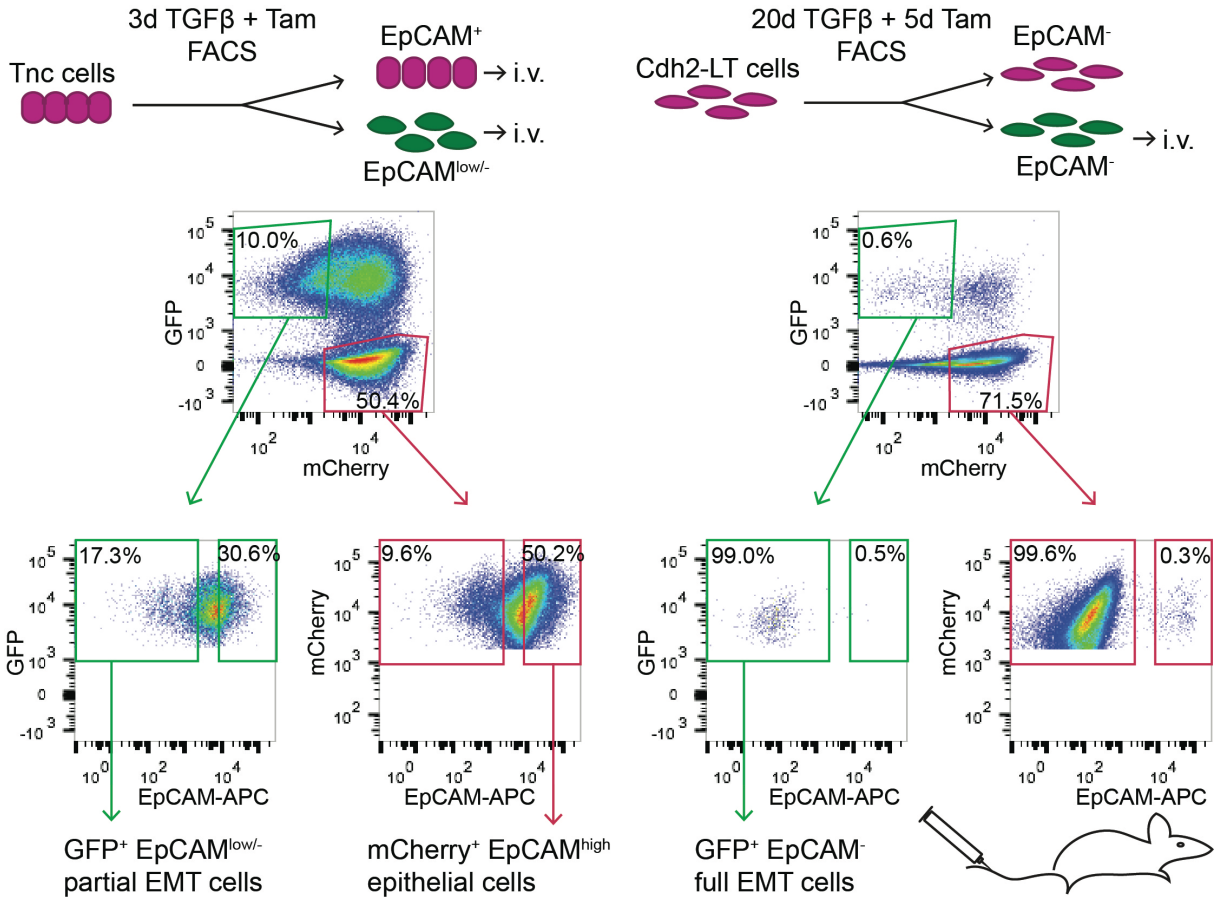
Supplementary Figure 6: Characterization of color-switching cell lines. a Tenascin C staining of Tnc-CreER^{T2} cells treated with tamoxifen (Tam) and/or TGFβ for one day. Scale bar: 100µm. **b** Representative images of E-cadherin staining of Tnc-CreER^{T2} cells treated with TGFβ and Tam over a time course of 20 days. Scale bar 100µm. **c, d** Quantification of percent of GFP^{+} cells **(c)** and qPCR analysis of *Tnc* mRNA levels **(d)** in Tnc-CreER^{T2} cell line after treatment with a TGFβRI and Tam for 5 days. **p < 0.01, paired t-test, ****p < 0.0001, ratio-paired t-test. **e** Representative images of E-cadherin staining of Cdh2-CreER^{T2} cells treated with TGFβ and Tam over a time course of 20 days. **f** Representative images of N-cadherin staining of mesenchymal Cdh2-CreER^{T2}-LT cells treated with Tam for 5 days. Scale bar: 100µm.



Supplementary Figure 7: Invasion assay of Tnc-CreER^{T2} and Cdh2-CreER^{T2} color switching cell lines. a, d Percent of GFP$^+$ cells per total Tnc-CreER^{T2} cells **(a)** and Cdh2-CreER^{T2} cells **(d)**. **b, e** Percent of invading mCherry$^+$ or GFP$^+$ cells relative to the total number of mCherry$^+$ or GFP$^+$ Tnc-CreER^{T2} cells **(b)** and Cdh2-CreER^{T2} cells **(e)**. Control: Tnc-CreER^{T2} or Cdh2-CreER^{T2} parental cells treated with Tam during the assay. Partial EMT: Tnc-CreER^{T2}-3d cells kept under TGFβ and Tam during the assay. Mesenchymal: Tnc-CreER^{T2}-LT or Cdh2-CreER^{T2}-LT cells kept under TGFβ and Tam during the assay. Mean and SEM of n=3 experiments are shown. **c** Representative images of collagen I invasion assay of Cdh2-CreER^{T2} cells. Box indicates the area included as magnified inset. Scale bar: 100µm.



Supplementary Figure 8: GFP$^+$ and mCherry$^+$ cells in established metastasis proliferate at a similar rate. a Quantification of pH3$^+$ cells in metastases obtained by tail vein injection of Tnc-CreER^{T2} cells. mCherry$^+$ and GFP$^+$ areas of mixed-color metastases were analyzed separately. Dots represent individual metastases of n=3 mice, lines represent mean +/- SEM. **b** Representative images of pH3 staining. Arrows indicate pH3$^+$ cells. Scale bar: 200µm.



Supplementary Figure 9: Gating strategy for sorting of epithelial, partial-EMT hybrid and fully-mesenchymal cells from cell lines by FACS. Tnc-CreER^{T2} (Tnc) cells treated with TGFβ and tamoxifen (Tam) for 3 days *in vitro* were sorted for mCherry^{+}EpCAM^{+} epithelial and GFP^{+}EpCAM$^{low/-}$ partial EMT hybrid cells. GFP^{+}EpCAM^{-} fully mesenchymal cells were sorted from Cdh2-CreER^{T2} (Cdh2)-LT cells that had been treated with TGFβ for >20 days and tamoxifen for 5 days.

3.1.7 Author Contributions

F.L. designed and performed the experiments, analyzed the data, prepared the Figures and wrote the manuscript. N.S. designed experiments, performed the flow cytometry analysis, analyzed the data and edited the manuscript. C.H. performed some of the tissue culture experiments. R.B. conceived the idea and designed some of the experiments. T.B. and T.L. assisted with FACS and flow cytometry. G.C. conceived the idea, supervised the project, designed the experiments, analyzed the data and edited the manuscript.

3.2 A Transgenic MMTV-Flippase Mouse Line for Molecular Engineering in Mammary Gland and Breast Cancer Mouse Models

Fabiana Lüönd[1], Ruben Bill[1], Andrea Vettiger[1,2], Heide Oller[3], Pawel Pelczar[3], Gerhard M. Christofori[1]

1 Department of Biomedicine, University of Basel

2 Focal Area Infection Biology, Biozentrum University of Basel

3 Center for Transgenic Models, University of Basel

Published in Journal of Mammary Gland Biology and Neoplasia,
March 2019, Volume 24, Issue 1, pp 39–45

3.2.1 Abstract

Genetically engineered mouse models have become an indispensable tool for breast cancer research. Combination of multiple site-specific recombination systems such as Cre/loxP and Flippase (Flp)/Frt allows for engineering of sophisticated, multilayered conditional mouse models. Here, we report the generation and characterization of a novel transgenic mouse line expressing a mouse codon-optimized Flp under the control of the mouse mammary tumor virus (MMTV) promoter. These mice show robust Flp-mediated recombination in luminal mammary gland and breast cancer cells but no Flp activity in non-mammary tissues, with the exception of limited activity in salivary glands. These mice provide a unique tool for studying mammary gland biology and carcinogenesis in mice.

3.2.2 Introduction

Despite major advances in understanding the mechanisms of malignant carcinogenesis, breast cancer is still a leading cause of cancer-related death in women worldwide. In the past decades, genetically engineered mouse models have contributed significantly towards understanding breast cancer development and progression. These include amongst others the widely used MMTV-driven transgenic mouse models, such as MMTV-PyMT or MMTV-Neu, as well as conditional models based on the knockout of tumor-suppressor genes, such as *Brca1*, *p53,* and *Cdh1* [248]. Recently, genetic lineage tracing approaches using conditional fluorescent reporters for the visualization of cell-populations of interest have contributed towards our understanding of the origin of breast cancer [118, 119, 249], breast cancer stem cells [250], metastatic dissemination [106, 160] and chemoresistance [112].

Site-specific recombinase systems such as Cre/*loxP* and Flp/*Frt* are powerful tools for the genetic manipulation of somatic cells *in vivo.* Briefly, these systems consist of a DNA recombinase and its unique short DNA target sequences. Cre recombinase of bacteriophage P1 recognizes and acts on 34 base pair *loxP* sites, while Flp recombinase of Saccharomyces cerevisiae specifically recombines 48 base pair *Frt* sites. Consequently, DNA sequences flanked by two *loxP* or *Frt* sites will be deleted upon Cre and Flp-mediated recombination, respectively, if the recognition sites are

oriented in the same direction. In case the recognition sites are oriented in opposite directions, the flanked DNA sequence will be inverted. Although in murine models Cre has been initially found more efficient in catalyzing recombination than Flp, temperature-optimized version of Flp (Flpe) and mouse codon-optimized variants of Flpe (Flpo) are now almost as efficient as Cre [251-253].

To date, conditional breast cancer mouse models and genetic lineage tracing approaches most commonly rely on the widely used Cre/*loxP* system. In these models, Cre recombinase mediates excision of DNA to e.g. delete tumor suppressor genes or to remove a STOP cassette that prevents the expression of a fluorescent reporter or an activated oncogene. The choice of the promoter driving Cre expression is essential, as it determines the cell-lineage to be targeted in the mouse.

More recently, different site-specific recombination systems such as Cre/*loxP* and/or Flp/*Frt* or Dre/*Rox* have been combined in order to generate increasingly sophisticated model systems which combine the activation or inactivation of specific genes in a temporal and spatial manner [254-258]. Such dual-recombinase mediated approaches potentially allow to simultaneously target different cell-lineages, such as tumor and stromal cells or to modulate gene expression in the same cell-lineage in a sequential manner. The latter allows for the time-dependent introduction of mutations in tumor cells, for example by using inducible recombinases such as $CreER^{T2}$, or for the specific targeting of distinct tumor cell sub-populations in an individual tumor. Thus, the generation of novel Flp-driver lines and *Frt*-flanked alleles is required in order to establish “next-generation” dual recombinase-mediated mouse models for studying the tumor microenvironment, tumor progression and tumor heterogeneity in breast cancer. Here, we have generated a novel mouse line expressing a mouse codon-optimized Flp (Flpo) under the control of the MMTV promoter for the specific genetic manipulation of mammary gland epithelial cells and, in combination with adequate oncogene and/or tumor suppressor gene manipulations, in breast cancer cells.

3.2.3 Results and Discussion

In order to express the Flp recombinase in mammary gland epithelial cells, we placed the mouse codon-optimized *Flp* (*Flpo*) under the control of the MMTV promoter.

Transgenic mice were generated by pronuclear microinjection. 17 founder animals were obtained, of which ten showed germline transmission.

To analyze Flpo activity in normal mammary gland epithelial cells, the founder lines were crossed with the RC::FrePe reporter strain [234, 235] (Fig. 1a). Briefly, this dual recombinase-responsive reporter allele encodes for mCherry and GFP whose expression can be controlled by Flp and Cre activity in a sequential manner. Since transcription of mCherry is prevented by a *Frt*-flanked STOP site, Flp-mediated recombination leads to mCherry expression. Upon fluorescence microscopy analysis of mCherry expression in mammary gland epithelial cells, two founder lines (9 and 11) with high Flpo activity were identified (Fig. 1b & 1c). Further analysis revealed that in animals of line 9 Flpo activity was restricted to cytokeratin 8/18 (CK-8/18)-positive luminal mammary gland cells, while Flpo was found active in both luminal and cytokeratin 14 (CK-14)-positive myoepithelial cells in mice of line 11 (Fig. 1b). Besides the mammary gland epithelium, also the pulmonary- and most notably the salivary gland epithelium are sites of MMTV promoter activity, which poses a risk of unwanted off-target effects. Interestingly, we did not detect any recombination in the lung and only very limited recombination in the salivary gland of line 9. However, we observed considerable Flpo activity in the lung and very strong activity in the salivary gland of line 11 (Fig. 1c). Consistent with this, Flp expression levels were overall higher in animals of line 11 with particularly strong expression in salivary glands. Although we did not observe Flpo-mediated recombination in pulmonary epithelium of line 9, Flpo expression was detected by quantitative RT-PCR and, thus, low Flpo activity in the lung cannot be fully excluded (Fig. 1d).

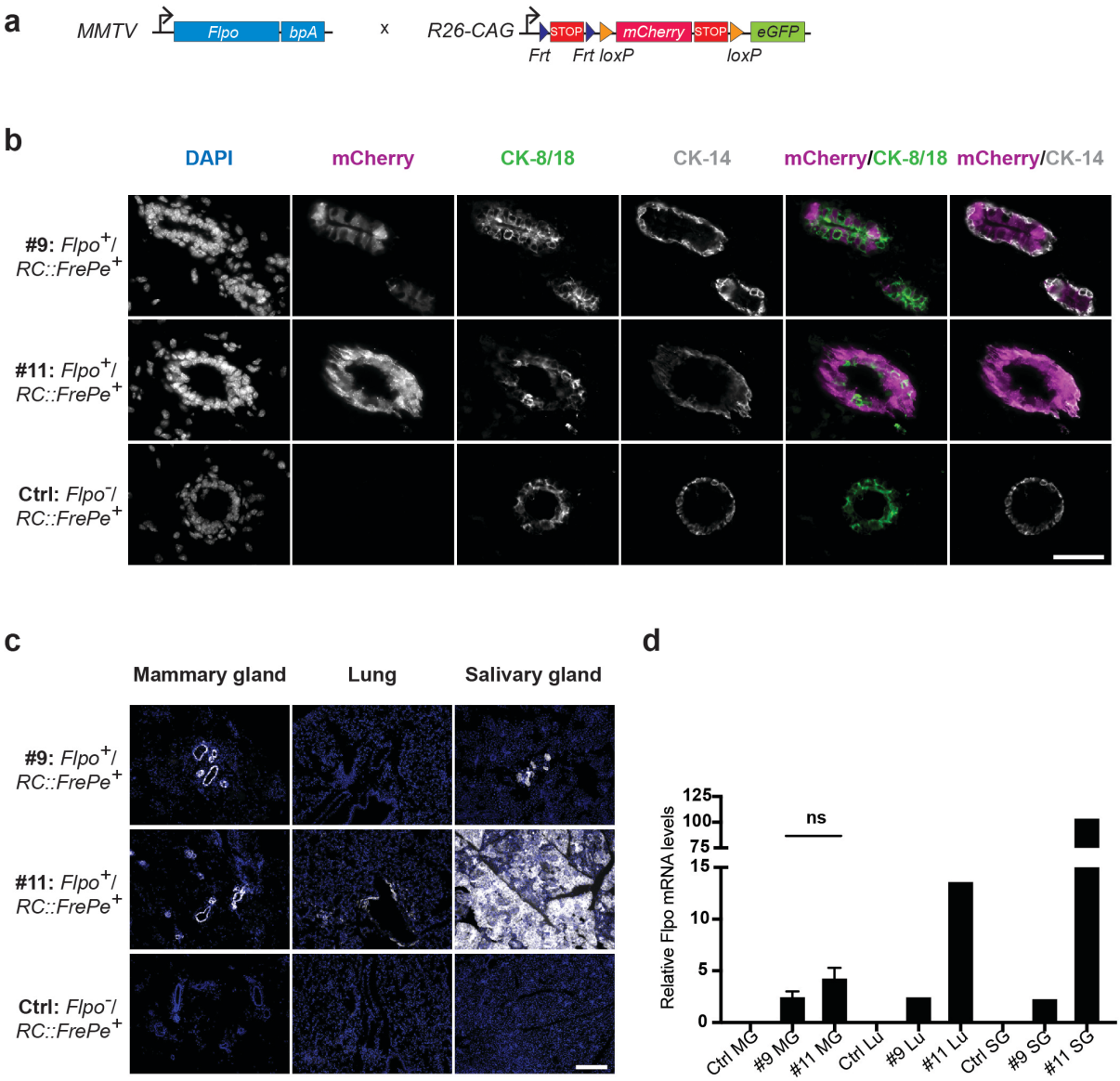


Figure 1: Flpo activity in the mammary gland, lung and salivary gland. a Strategy for genetic lineage tracing of Flpo activity. Combination of MMTV-Flpo (left) and RC::FrePe (right) leads to stable mCherry expression in Flpo-expressing cells. **b** Co-staining of mammary glands with the luminal epithelial cell marker CK-8/18 (green) and basal epithelial cell marker CK-14 (grey). mCherry: magenta. Scale bar, 50µm. **c** Analysis of mCherry expression in mammary glands, lungs and salivary glands of adult female mice of MMTV-Flpo lines 9 and 11. mCherry: grey, DAPI: blue. Scale bar, 200µm. **d** Analysis of Flpo expression by quantitative RT-PCR in mammary glands (MG), lungs (Lu) and salivary glands (SG) of adult female mice of MMTV-Flp lines 9 and 11. Mean and SEM of n=5 mice are shown for mammary glands. Salivary gland and lung tissue of n=5 mice were pooled prior to RNA extraction. Flpo expression levels were normalized to Gapdh. P-value, 0.166; unpaired, two-tailed t-test.

The MMTV-PyMT transgenic mouse model is one of the most widely used models of metastatic breast cancer [236]. To generate mice developing tumors of the mammary glands, we crossed MMTV-Flpo/RC::FrePe double-transgenic mice with the MMTV-PyMT strain (Fig. 2a). Primary tumors and lung metastasis of both Flpo founder

lines showed strong mCherry expression (Fig. 2b). Immunofluorescence staining for the stromal cell marker PDGFRβ and for the epithelial cell marker E-cadherin confirmed that Flpo activity of both mouse lines is specific to cancer cells. No leakiness of Flpo expression was observed in stromal cells (Fig. 2c).

Thus, both our MMTV-Flpo transgenic founder lines show high recombination efficiency in mammary gland epithelial cells and cancer cells in the MMTV-PyMT mouse model. However, Flpo activity of line 9 is spatially more restricted compared to line 11 and hence at lower risk for off-target effects. We therefore continued employing line 9 for further experimentation.

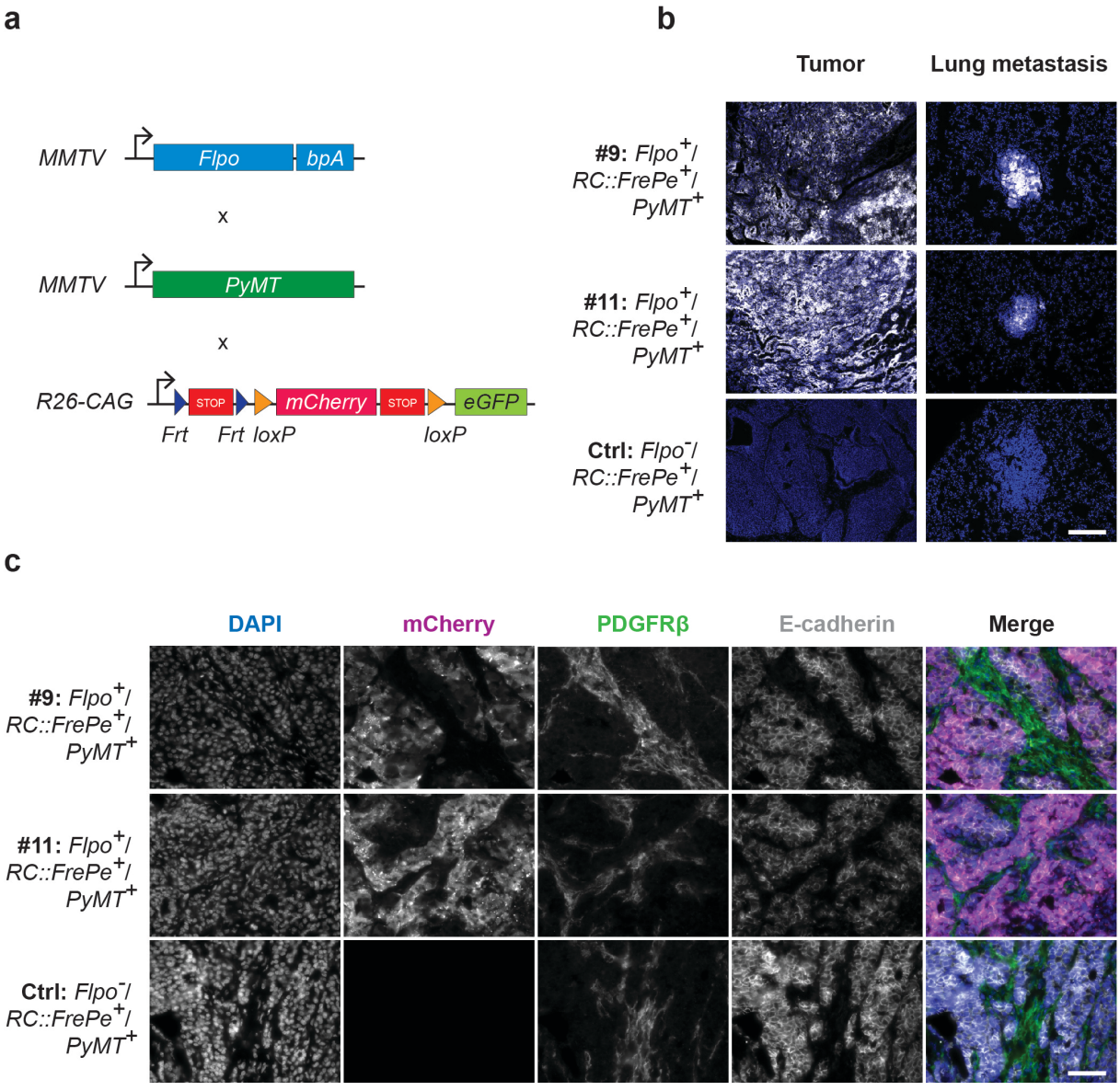


Figure 2: Flpo activity in breast cancer cells of MMTV-PyMT mice. a Strategy for genetic lineage tracing of Flpo activity in breast cancer cells of MMTV-PyMT transgenic mice. Combination of MMTV-Flpo (top), MMTV-PyMT

(middle) and RC::FrePe (bottom) leads to stable mCherry expression in Flp-expressing tumor cells. **b** Analysis of mCherry expression in primary tumors and lung metastasis of MMTV-Flp mouse lines 9 and 11. mCherry: grey, DAPI: blue. Scale bar, 200µm. **c** Co-staining of primary tumors for the stromal cell marker PDGFRβ (green) and the epithelial cell marker E-cadherin (grey). mCherry: magenta, DAPI: blue. Scale bar, 50µm.

When breeding both male and female transgenic mice in strictly heterozygous matings we observed stable transgene transmission according to Mendelian inheritance. Follow-up analysis after at least four additional generations of breeding confirmed robust Flpo activity throughout the mammary glands (Fig. 3a). Overall, approximately 64% of luminal mammary gland epithelial cells showed mCherry expression in later generation MMTV-Flpo/RC::FrePe mice. There was no significant difference in recombination efficiencies at second and sixth filial generations (65.20 ± 1.22% at F_2; 64.04 ± 4.38% at F_6). Similarly, approximately 67.45 ± 4.71%, of cancer cells in PyMT-driven breast tumors showed mCherry expression (Fig. 3b & 3c).

In the various MMTV-driven transgenic strains generated in the past years, differences in the onset of transgene expression have been observed. We hence performed time course analysis to investigate the onset of Flpo expression in the mammary glands of MMTV-Flpo line 9. Flpo expression levels of four-week-old females were comparable to expression in eight-week-old females and low levels of Flpo expression could be detected already in two-week-old mice (Fig. 3d). Consistent with this, we observed mCherry expression in mammary glands of two and four-week-old mice indicating Flpo is active in mammary glands of mice as young as two weeks old (Fig. 3e & 3f).

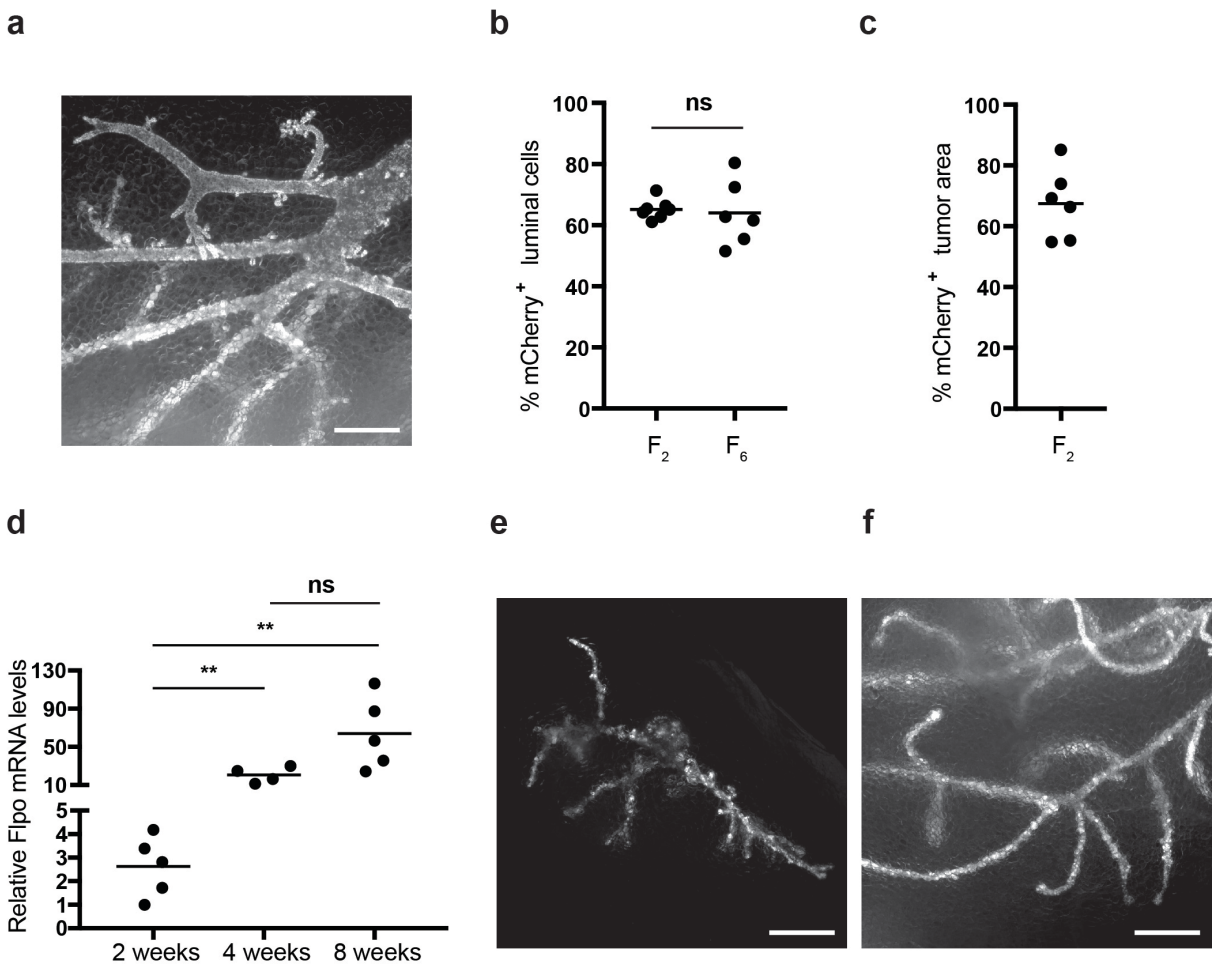


Figure 3: Efficiency and onset of Flpo activity in the MMTV-Flpo mouse line 9. a Mammary glands of MMTV-Flpo$^+$/RC::FrePe$^+$ adult (8 week old) female mice (whole mounted tissue). mCherry: grey. Scale bar, 500µm **b** Flpo recombination efficiency in mammary glands of 8-week-old mice at filial generations F2 and F6. Means of n=7 (F2) and n=6 (F6) individual mice are shown. Five imaging fields (63x magnification) were analyzed per mouse. P-value, 0.7902; unpaired, two-tailed t-test. **c** Flpo recombination efficiency in breast tumors of MMTV-PyMT$^+$/MMTV-Flpo$^+$/RC::FrePe$^+$ triple-transgenic mice. The mean of n=6 individual mice are shown. Two tumor sections were analyzed per mouse. **d** Analysis of Flpo expression by quantitative RT-PCR in mammary glands of 2, 4 and 8-week old female mice. Means of n=5 (2 weeks), n=4 (4 weeks) and n=5 (8 weeks) mice are shown. Flp expression levels were normalized to Gapdh. P-values: 0.0018 (2 vs 4 weeks), 0.0068 (2 vs 8 weeks), 0.0620 (4 vs 8 weeks). **e, f** Mammary glands of MMTV-Flpo$^+$/RC::FrePe$^+$ 2 week (**e**) and 4 week (**f**) old female mice (whole mounted tissue). mCherry: grey. Scale bars, 500µm.

In summary, we have generated and characterized a novel mouse line expressing Flpo under the control of the MMTV promoter. This line shows robust Flpo activity in luminal mammary gland epithelial cells. Since Flpo is already expressed in the developing mammary glands of prepubertal mice, the mouse line may also be useful for research focusing on pubertal mammary gland development. Furthermore, the mouse line may be crossed to strains harboring *Frt*-flanked tumor suppressor

genes, such as *p53*$^{Frt/Frt}$ or an activated version of oncogenes [259] for the generation of novel conditional breast cancer mouse model. Importantly, our new mouse line offers new opportunities for genetic lineage tracing in breast cancer, particularly in the widely used MMTV-driven mouse models. For example, combination of MMTV-Flpo, RC::FrePe and a Cre driver line of choice may be readily used to visualize cancer cell subpopulations of interest, for example in MMTV-PyMT-driven tumors. With novel *Frt*-flanked conditional alleles and increasingly complex site-specific recombinase-responsive reporter alleles becoming available, our mouse line has the potential to become a versatile tool for genetic manipulation and lineage tracing in the murine mammary gland and in mammary gland cancer.

3.2.4 Materials and Methods

Animals

Mouse colonies were kept at the animal facility of the Department of Biomedicine, University of Basel, Switzerland. All animal experiments were carried out in accordance with the guidelines of the Swiss Federal Veterinary Office (SFVO) and the Cantonal Veterinary Office of Basel-Stadt (Licenses 1878, 1023G1). MMTV-PyMT (FVB/N) mice were a kind gift of N. Hynes (FMI, Basel, Switzerland). RC::FrePe mice (B6;129S6-Gt(ROSA)26Sortm8(CAG-mCherry,-EGFP)Dym/J) were a kind gift of S. Dymecki (Department of Genetics, Harvard Medical School, Boston, Massachusetts). The mice were crossed six to seven generations into the FVB/N background and then used for the experiments. All experiments were performed on female mice. For analysis of the mammary gland, mice were sacrificed at 8 weeks, 4 weeks or 2 weeks of age. Tumor bearing mice were sacrificed before a tumor volume of 1500mm^3 was reached, usually at 12-13 weeks of age.

Construction of the Flp expression vector

The mouse codon-optimized Flp (*Flpo*) gene was placed under the control of the *MMTV LTR* promoter. The pPGKFlpobpA plasmid was a kind gift from Philippe Soriano (Addgene plasmid # 13793). The vector containing the *MMTV-LTR* and *BGHpA* was a kind gift from Kari Alitalo. The *Flpo* fragment was PCR amplified from the pPGKFlpobpA plasmid and inserted into the vector as an XhoI fragment.

Generation and genotyping of transgenic mice

The *MMTV-Flpo-BGHpA* was released from the construct as a HindIII-NdeI fragment and purified using the QIAEX II Gel Extraction Kit (Qiagen) according to the manufacturer's instructions. The fragment was then microinjected into the pronuclei of FVB/N zygotes and the embryos were transferred into pseudopregnant CD1 females at the Center for Transgenic Models (University of Basel, Switzerland). Transgenic founders were identified by PCR using primers 5'-CCA AGG TGC TGG TGC GGC AGT TCG TG-3' and 3`-GAT CTC CCA GAT GCT CTC GCC CTC GGA C-5 resulting in a 436bp PCR product in transgenic mice. The established MMTV-Flpo line (FVB/N-Tg(MMTV-Flpo9Gcr) was kept heterozygous and in a pure FVB/N background. Transgenic progeny was identified by PCR as described above or alternatively using primer pair 5'-TGA GCT TCG ACA TCG TGA AC-3' and 3'-TCA GCA TCT TCT TGC TGT GG-3' generating a 231bp product.

Immunofluorescent staining

Tissues were fixed in 4% PFA for two hours and incubated in 20% sucrose overnight at 4°C. Tissues were then embedded in Tissue-Tek O.C.T compound (Sakura), snap-frozen and stored at -80°C. 7µm thick tissue sections were cut, dried for 30min, rehydrated in PBS, permeabilized with 0.2% Triton X-100 in PBS for 20min and blocked with 5% normal goat serum (NGS; Sigma-Aldrich; G6767) in PBS for one hour. Next, sections were incubated with primary antibody overnight at 4°C followed by 1h incubation with secondary antibody at room temperature. All antibodies were diluted in 5% NGS in PBS. The following primary antibodies were used: guinea pig-anti-CK8/18 (Fitzgerald Industries; 20R-CP004; 1:50), rabbit-anti-CK14 (Thermo Scientific; RB-9020; 1:50), rat-anti-E-cadherin (clone ECCD-2; Invitrogen; 13-1900; 1:400), rabbit-anti-PDGFRβ (clone Y92; abcam; ab32570; 1:100). Secondary antibodies directed against the species of the primary antibody were coupled to Alexa 488 or Alexa 647 (Invitrogen; 1:400). Nuclei were stained with 4′,6-Diamidin-2-phenylindol (DAPI; Sigma-Aldrich; 1:5000) for 10min at room temperature. The sections were mounted with Dako fluorescence mounting medium (Agilent). Images were acquired with a Leica DMI 4000 microscope and processed with ImageJ.

RNA isolation and quantitative RT-PCR

Total RNA was isolated using RNAeasy Mini Kit (Qiagen). On column DNA digestion using RNAse-Free DNase Set (Qiagen) was performed. Prior to isolation, tissue of thoracic and abdominal mammary glands was pooled for every individual mouse. Salivary gland and lung tissue of five mice was pooled. cDNA was synthesized using ImProm-II Reverse Transcriptase (Promega). Quantitative RT-PCR was performed on a StepOnePlus machine (Applied Biosystems) using Power-Up SYBR green (Applied Biosystems). Flpo expression levels were normalized to Gapdh. *Flpo*-specific primers 5'-ccgagaagatcctgaacagc-3' (forward) and 5'-ggtacagggtcttggtcttgg-3' (reverse) and *Gapdh*-specific primers 5'- agcttgtcatcaacgggaag-3' and 5'-tttgatgttagtggggtctcg-3' were used.

Mammary gland whole mounts

Isolated mammary glands were mounted on a glass-slide and imaged with a Leica DMI 4000 microscope.

Quantification of recombination efficiency

For quantification of Flpo-mediated recombination efficiency in the mammary gland, the total numbers of cytokeratin 8/18 positive (luminal) cells and the number of mCherry expressing cytokeratin 8/18 positive cells were counted. Five imaging fields each of mammary glands (63x magnification) from a total of 7 (F_2) and 6 (F_6) mice were analyzed. For quantification of the levels of recombined cells in PyMT tumors, sections of two different tumors per mouse were analyzed. The whole tumor section was imaged with a Zeiss Axio Imager Scanning Microscope (10x magnification). The mCherry-positive area and total tumor areas (DAPI positive area) were determined by global thresholding (Otsu algorithm) using ImageJ. Six mice (F_2) were analyzed.

Statistical analysis

Statistical analysis was performed using GraphPad Prism (GraphPad Software Inc.). Graphs display single measurement points and means or mean + SEM. Unpaired, two-tailed t-tests were performed.

3.2.5 Author Contributions

F.L. designed and performed the experiments, analysed the data and wrote the manuscript. R.B. conceived the idea and designed some of the experiments and analysed the data. A.V. performed experiments related to the construction of the Flp expression vector. H.O. and P.P. performed the pronuclei microinjections. G.C. supervised the project, designed the experiments, analysed the data and wrote the manuscript.

4 General Discussion and Future Directions

In cancer, EMT has emerged as a complex epigenetic program of cell plasticity that allows cells to transition between a heterogeneous population of intermediate hybrid states. As EMT/MET are associated with a variety of malignant traits, better understanding of the dynamics between these cell transitions, the characteristics associated with distinct EMT transition states and the functional consequences thereof would be critical to develop novel therapeutic strategies. For now, our novel lineage tracing systems allow to study the contribution of partial and full EMT towards distinct aspects of cancer progression in an unperturbed system *in vivo*.

Of note, our data indicate that a partial EMT may contribute to breast cancer metastasis to the lung. However, breast cancer frequently metastasizes to the bone, brain and liver. As the microenvironment drastically differs between these organs, it is possible that distinct EMT states have an advantage at other metastatic sites. It would thus be of particular interest to study the contribution of partial and full EMT to metastasis formation in different organs. Along similar lines, we observed that only a small number of cells undergo an EMT in the MMTV-PyMT mouse model. This model, which corresponds best with the luminal B subtype, is one of the most commonly used transgenic models to study breast cancer metastasis. However, breast cancer is a very heterogeneous disease, and particularly the basal and claudin-low rather than the luminal subtypes have been associated with EMT. It is likely that an EMT is of greater relevance for metastasis in these subtypes. It would hence be interesting to combine our color switching system with other breast cancer mouse models. Myc-driven models, including MMTV-Myc or WAP-Myc [260, 261], in which a high number of cancer cells have been shown to undergo an EMT [105], and the highly metastatic p53 knockout model [262] (as $p53^{frt/frt}$ mice are available already [259]) would be of particular interest.

A pivotal question concerns the contribution of EMT to therapy resistance, particularly with regard to chemotherapeutic agents. In patients, EMT signatures may predict treatment response and relapsed tumors are frequently more mesenchymal [24, 137, 180, 205-207]. Our data indicate that cells which have undergone a full EMT may stably reside in a mesenchymal low-proliferative state, which thus could potentially be less susceptible to chemotherapy. It will therefore be of particular interest to determine the contribution of full compared to partial EMT towards cancer

recurrences at both the primary tumor and metastatic site. Chemotherapy treatments to address this question are currently ongoing.

It has recently become clear that EMT is not a binary phenotype-switch but covers a heterogeneous population of EMT hybrid states [87, 104, 224, 225]. Particularly the molecular characteristics of distinct EMT states and their functional implications in tumor progression require further investigation. Our ongoing efforts regarding RNA-sequencing of different mCherry$^+$ epithelial and GFP$^+$ EMT subpopulations from our lineage tracing systems will presumably shed some light on the heterogeneity and molecular characteristics of cells in different EMT states. Furthermore, these investigations could lead to the discovery of novel biomarkers to detect specific EMT states in patient samples, which might have prognostic value. One might speculate that some patients with specific EMT signatures could benefit from novel therapies targeting EMT. For example, our lab has previously demonstrated that the plasticity associated with EMT can be exploited to transdifferentiate aggressive cancer cells into harmless adipocytes [233]. A better understanding of the plasticity associated with distinct EMT transition states could be beneficial to translate such innovative approaches into the clinics.

As discussed above, cancer cells that have undergone an EMT might exert important functions within the tumor microenvironment and might e.g. facilitate dissemination of epithelial cells without metastasizing themselves. The functional relevance of a spontaneous EMT for tumor progression and metastasis *in vivo* may therefore only be addressed by selective ablation of these cells. To this end, we have recently established organoids derived from our breast cancer EMT lineage tracing models that are stably expressing a loxP-stop-loxP-diphteria toxin A construct, allowing for selective killing of cells that are undergoing an EMT upon tamoxifen administration. Transplantation of these organoids into mice and ablation of cells undergoing a partial or full EMT *in vivo* might provide valuable information about the functional implication of EMT in metastasis and might even provide direct evidence for the contribution of a partial EMT to metastasis.

Finally, in combination with Cre alleles distinct from ours, other cell-fate transitions of interest or rare breast cancer cell subpopulations might be studied. With novel floxed and frt-flanked alleles becoming available, our MMTV-Flp mice could become a useful tool to establish complex transgenic breast cancer models. This may e.g. include the sequential introduction of mutations to rebuild breast cancer

progression in the mouse or to fluorescently label different cell-types, for example epithelial cancer cells together with cancer-associated fibroblasts or immune cells to study specific cell-cell interactions by intravital imaging.

In conclusion, our EMT lineage tracing models provide novel opportunities for studying spontaneous partial and full EMT in breast cancer *in vivo*. A better understanding of partial and full EMT programs and their implications in breast cancer progression may ultimately lead to the development of novel therapeutic strategies to combat breast cancer.

5 References

1. Kalluri, R. and R.A. Weinberg, *The basics of epithelial-mesenchymal transition.* J Clin Invest, 2009. **119**(6): p. 1420-8.
2. Thiery, J.P., et al., *Epithelial-mesenchymal transitions in development and disease.* Cell, 2009. **139**(5): p. 871-90.
3. Lopez-Novoa, J.M. and M.A. Nieto, *Inflammation and EMT: an alliance towards organ fibrosis and cancer progression.* EMBO Mol Med, 2009. **1**(6-7): p. 303-14.
4. Dongre, A. and R.A. Weinberg, *New insights into the mechanisms of epithelial-mesenchymal transition and implications for cancer.* Nat Rev Mol Cell Biol, 2019. **20**(2): p. 69-84.
5. Lamouille, S., J. Xu, and R. Derynck, *Molecular mechanisms of epithelial-mesenchymal transition.* Nat Rev Mol Cell Biol, 2014. **15**(3): p. 178-96.
6. Xu, J., S. Lamouille, and R. Derynck, *TGF-beta-induced epithelial to mesenchymal transition.* Cell Res, 2009. **19**(2): p. 156-72.
7. Cavallaro, U. and G. Christofori, *Cell adhesion and signalling by cadherins and Ig-CAMs in cancer.* Nat Rev Cancer, 2004. **4**(2): p. 118-32.
8. Yilmaz, M. and G. Christofori, *EMT, the cytoskeleton, and cancer cell invasion.* Cancer Metastasis Rev, 2009. **28**(1-2): p. 15-33.
9. Hazan, R.B., et al., *Cadherin switch in tumor progression.* Ann N Y Acad Sci, 2004. **1014**: p. 155-63.
10. Wheelock, M.J., et al., *Cadherin switching.* J Cell Sci, 2008. **121**(Pt 6): p. 727-35.
11. Savagner, P., K.M. Yamada, and J.P. Thiery, *The zinc-finger protein slug causes desmosome dissociation, an initial and necessary step for growth factor-induced epithelial-mesenchymal transition.* J Cell Biol, 1997. **137**(6): p. 1403-19.
12. Huang, R.Y., P. Guilford, and J.P. Thiery, *Early events in cell adhesion and polarity during epithelial-mesenchymal transition.* J Cell Sci, 2012. **125**(Pt 19): p. 4417-22.
13. St Johnston, D. and J. Ahringer, *Cell polarity in eggs and epithelia: parallels and diversity.* Cell, 2010. **141**(5): p. 757-74.
14. Friedl, P. and S. Alexander, *Cancer invasion and the microenvironment: plasticity and reciprocity.* Cell, 2011. **147**(5): p. 992-1009.
15. Ridley, A.J., et al., *Cell migration: integrating signals from front to back.* Science, 2003. **302**(5651): p. 1704-9.
16. Hall, A., *Rho GTPases and the control of cell behaviour.* Biochem Soc Trans, 2005. **33**(Pt 5): p. 891-5.
17. Yilmaz, M. and G. Christofori, *Mechanisms of motility in metastasizing cells.* Mol Cancer Res, 2010. **8**(5): p. 629-42.
18. Mendez, M.G., S. Kojima, and R.D. Goldman, *Vimentin induces changes in cell shape, motility, and adhesion during the epithelial to mesenchymal transition.* FASEB J, 2010. **24**(6): p. 1838-51.
19. Frantz, C., K.M. Stewart, and V.M. Weaver, *The extracellular matrix at a glance.* J Cell Sci, 2010. **123**(Pt 24): p. 4195-200.
20. Yang, X., et al., *Regulation of beta 4-integrin expression by epigenetic modifications in the mammary gland and during the epithelial-to-mesenchymal transition.* J Cell Sci, 2009. **122**(Pt 14): p. 2473-80.

21. Huveneers, S., et al., *Binding of soluble fibronectin to integrin alpha5 beta1 - link to focal adhesion redistribution and contractile shape.* J Cell Sci, 2008. **121**(Pt 15): p. 2452-62.
22. Maschler, S., et al., *Tumor cell invasiveness correlates with changes in integrin expression and localization.* Oncogene, 2005. **24**(12): p. 2032-41.
23. Turunen, S.P., O. Tatti-Bugaeva, and K. Lehti, *Membrane-type matrix metalloproteases as diverse effectors of cancer progression.* Biochim Biophys Acta Mol Cell Res, 2017. **1864**(11 Pt A): p. 1974-1988.
24. Shibue, T. and R.A. Weinberg, *EMT, CSCs, and drug resistance: the mechanistic link and clinical implications.* Nat Rev Clin Oncol, 2017. **14**(10): p. 611-629.
25. Yang, M.H., et al., *Direct regulation of TWIST by HIF-1alpha promotes metastasis.* Nat Cell Biol, 2008. **10**(3): p. 295-305.
26. Wei, S.C., et al., *Matrix stiffness drives epithelial-mesenchymal transition and tumour metastasis through a TWIST1-G3BP2 mechanotransduction pathway.* Nat Cell Biol, 2015. **17**(5): p. 678-88.
27. Jung, H.Y., et al., *Apical-basal polarity inhibits epithelial-mesenchymal transition and tumour metastasis by PAR-complex-mediated SNAI1 degradation.* Nat Cell Biol, 2019. **21**(3): p. 359-371.
28. Persa, O.D. and C.M. Niessen, *Epithelial polarity limits EMT.* Nat Cell Biol, 2019. **21**(3): p. 299-300.
29. Lehembre, F., et al., *NCAM-induced focal adhesion assembly: a functional switch upon loss of E-cadherin.* EMBO J, 2008. **27**(19): p. 2603-15.
30. Onder, T.T., et al., *Loss of E-cadherin promotes metastasis via multiple downstream transcriptional pathways.* Cancer Res, 2008. **68**(10): p. 3645-54.
31. Perl, A.K., et al., *A causal role for E-cadherin in the transition from adenoma to carcinoma.* Nature, 1998. **392**(6672): p. 190-3.
32. Berx, G., et al., *E-cadherin is inactivated in a majority of invasive human lobular breast cancers by truncation mutations throughout its extracellular domain.* Oncogene, 1996. **13**(9): p. 1919-25.
33. Yang, J., et al., *Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis.* Cell, 2004. **117**(7): p. 927-39.
34. Navarro, P., E. Lozano, and A. Cano, *Expression of E- or P-cadherin is not sufficient to modify the morphology and the tumorigenic behavior of murine spindle carcinoma cells. Possible involvement of plakoglobin.* J Cell Sci, 1993. **105 (Pt 4)**: p. 923-34.
35. Sarrio, D., et al., *Epithelial-mesenchymal transition in breast cancer relates to the basal-like phenotype.* Cancer Res, 2008. **68**(4): p. 989-97.
36. Peinado, H., D. Olmeda, and A. Cano, *Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype?* Nat Rev Cancer, 2007. **7**(6): p. 415-28.
37. Guaita, S., et al., *Snail induction of epithelial to mesenchymal transition in tumor cells is accompanied by MUC1 repression and ZEB1 expression.* J Biol Chem, 2002. **277**(42): p. 39209-16.
38. Moreno-Bueno, G., et al., *Genetic profiling of epithelial cells expressing E-cadherin repressors reveals a distinct role for Snail, Slug, and E47 factors in epithelial-mesenchymal transition.* Cancer Res, 2006. **66**(19): p. 9543-56.
39. Tam, W.L. and R.A. Weinberg, *The epigenetics of epithelial-mesenchymal plasticity in cancer.* Nat Med, 2013. **19**(11): p. 1438-49.

40. Chakrabarti, R., et al., *Elf5 inhibits the epithelial-mesenchymal transition in mammary gland development and breast cancer metastasis by transcriptionally repressing Snail2.* Nat Cell Biol, 2012. **14**(11): p. 1212-22.
41. Cieply, B., et al., *Suppression of the epithelial-mesenchymal transition by Grainyhead-like-2.* Cancer Res, 2012. **72**(9): p. 2440-53.
42. Roca, H., et al., *Transcription factors OVOL1 and OVOL2 induce the mesenchymal to epithelial transition in human cancer.* PLoS One, 2013. **8**(10): p. e76773.
43. Ocana, O.H., et al., *Metastatic colonization requires the repression of the epithelial-mesenchymal transition inducer Prrx1.* Cancer Cell, 2012. **22**(6): p. 709-24.
44. Tiwari, N., et al., *Sox4 is a master regulator of epithelial-mesenchymal transition by controlling Ezh2 expression and epigenetic reprogramming.* Cancer Cell, 2013. **23**(6): p. 768-83.
45. Sakai, D., et al., *Cooperative action of Sox9, Snail2 and PKA signaling in early neural crest development.* Development, 2006. **133**(7): p. 1323-33.
46. Guo, W., et al., *Slug and Sox9 cooperatively determine the mammary stem cell state.* Cell, 2012. **148**(5): p. 1015-28.
47. Tiwari, N., et al., *Klf4 is a transcriptional regulator of genes critical for EMT, including Jnk1 (Mapk8).* PLoS One, 2013. **8**(2): p. e57329.
48. Mani, S.A., et al., *Mesenchyme Forkhead 1 (FOXC2) plays a key role in metastasis and is associated with aggressive basal-like breast cancers.* Proc Natl Acad Sci U S A, 2007. **104**(24): p. 10069-74.
49. Meyer-Schaller, N., et al., *Foxf2 plays a dual role during transforming growth factor beta-induced epithelial to mesenchymal transition by promoting apoptosis yet enabling cell junction dissolution and migration.* Breast Cancer Res, 2018. **20**(1): p. 118.
50. Bakiri, L., et al., *Fra-1/AP-1 induces EMT in mammary epithelial cells by modulating Zeb1/2 and TGFbeta expression.* Cell Death Differ, 2015. **22**(2): p. 336-50.
51. Zhang, H., et al., *TEAD transcription factors mediate the function of TAZ in cell growth and epithelial-mesenchymal transition.* J Biol Chem, 2009. **284**(20): p. 13355-62.
52. Diepenbruck, M., et al., *Tead2 expression levels control the subcellular distribution of Yap and Taz, zyxin expression and epithelial-mesenchymal transition.* J Cell Sci, 2014. **127**(Pt 7): p. 1523-36.
53. Lehmann, W., et al., *ZEB1 turns into a transcriptional activator by interacting with YAP1 in aggressive cancer types.* Nat Commun, 2016. **7**: p. 10498.
54. Thuault, S., et al., *Transforming growth factor-beta employs HMGA2 to elicit epithelial-mesenchymal transition.* J Cell Biol, 2006. **174**(2): p. 175-83.
55. Thuault, S., et al., *HMGA2 and Smads co-regulate SNAIL1 expression during induction of epithelial-to-mesenchymal transition.* J Biol Chem, 2008. **283**(48): p. 33437-46.
56. Kim, K., Z. Lu, and E.D. Hay, *Direct evidence for a role of beta-catenin/LEF-1 signaling pathway in induction of EMT.* Cell Biol Int, 2002. **26**(5): p. 463-76.
57. Yook, J.I., et al., *A Wnt-Axin2-GSK3beta cascade regulates Snail1 activity in breast cancer cells.* Nat Cell Biol, 2006. **8**(12): p. 1398-406.
58. Lamouille, S., et al., *Regulation of epithelial-mesenchymal and mesenchymal-epithelial transitions by microRNAs.* Curr Opin Cell Biol, 2013. **25**(2): p. 200-7.

59. Siemens, H., et al., *miR-34 and SNAIL form a double-negative feedback loop to regulate epithelial-mesenchymal transitions.* Cell Cycle, 2011. **10**(24): p. 4256-71.
60. Gregory, P.A., et al., *The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1.* Nat Cell Biol, 2008. **10**(5): p. 593-601.
61. Korpal, M., et al., *The miR-200 family inhibits epithelial-mesenchymal transition and cancer cell migration by direct targeting of E-cadherin transcriptional repressors ZEB1 and ZEB2.* J Biol Chem, 2008. **283**(22): p. 14910-4.
62. Burk, U., et al., *A reciprocal repression between ZEB1 and members of the miR-200 family promotes EMT and invasion in cancer cells.* EMBO Rep, 2008. **9**(6): p. 582-9.
63. Bracken, C.P., et al., *A double-negative feedback loop between ZEB1-SIP1 and the microRNA-200 family regulates epithelial-mesenchymal transition.* Cancer Res, 2008. **68**(19): p. 7846-54.
64. Bracken, C.P., et al., *Genome-wide identification of miR-200 targets reveals a regulatory network controlling cell invasion.* EMBO J, 2014. **33**(18): p. 2040-56.
65. Xu, Q., et al., *Long non-coding RNA regulation of epithelial-mesenchymal transition in cancer metastasis.* Cell Death Dis, 2016. **7**(6): p. e2254.
66. Warzecha, C.C. and R.P. Carstens, *Complex changes in alternative pre-mRNA splicing play a central role in the epithelial-to-mesenchymal transition (EMT).* Semin Cancer Biol, 2012. **22**(5-6): p. 417-27.
67. Shapiro, I.M., et al., *An EMT-driven alternative splicing program occurs in human breast cancer and modulates cellular phenotype.* PLoS Genet, 2011. **7**(8): p. e1002218.
68. Zhou, B.P., et al., *Dual regulation of Snail by GSK-3beta-mediated phosphorylation in control of epithelial-mesenchymal transition.* Nat Cell Biol, 2004. **6**(10): p. 931-40.
69. Yang, Z., et al., *Pak1 phosphorylation of snail, a master regulator of epithelial-to-mesenchyme transition, modulates snail's subcellular localization and functions.* Cancer Res, 2005. **65**(8): p. 3179-84.
70. Zhang, K., et al., *Lats2 kinase potentiates Snail1 activity by promoting nuclear retention upon phosphorylation.* EMBO J, 2012. **31**(1): p. 29-43.
71. Hong, J., et al., *Phosphorylation of serine 68 of Twist1 by MAPKs stabilizes Twist1 protein and promotes breast cancer cell invasiveness.* Cancer Res, 2011. **71**(11): p. 3980-90.
72. Zhang, P., et al., *ATM-mediated stabilization of ZEB1 promotes DNA damage response and radioresistance through CHK1.* Nat Cell Biol, 2014. **16**(9): p. 864-75.
73. Gumbiner, B.M., *Regulation of cadherin adhesive activity.* J Cell Biol, 2000. **148**(3): p. 399-404.
74. Fujita, Y., et al., *Hakai, a c-Cbl-like protein, ubiquitinates and induces endocytosis of the E-cadherin complex.* Nat Cell Biol, 2002. **4**(3): p. 222-31.
75. Radisky, E.S. and D.C. Radisky, *Matrix metalloproteinase-induced epithelial-mesenchymal transition in breast cancer.* J Mammary Gland Biol Neoplasia, 2010. **15**(2): p. 201-12.
76. Koenig, A., et al., *Collagen type I induces disruption of E-cadherin-mediated cell-cell contacts and promotes proliferation of pancreatic carcinoma cells.* Cancer Res, 2006. **66**(9): p. 4662-71.

77. Nelson, W.J. and R. Nusse, *Convergence of Wnt, beta-catenin, and cadherin pathways.* Science, 2004. **303**(5663): p. 1483-7.
78. Massague, J., *TGFbeta in Cancer.* Cell, 2008. **134**(2): p. 215-30.
79. Batlle, E. and J. Massague, *Transforming Growth Factor-beta Signaling in Immunity and Cancer.* Immunity, 2019. **50**(4): p. 924-940.
80. Zahnow, C.A., et al., *Overexpression of C/EBPbeta-LIP, a naturally occurring, dominant-negative transcription factor, in human breast cancer.* J Natl Cancer Inst, 1997. **89**(24): p. 1887-91.
81. Gomis, R.R., et al., *C/EBPbeta at the core of the TGFbeta cytostatic response and its evasion in metastatic breast cancer cells.* Cancer Cell, 2006. **10**(3): p. 203-14.
82. Tiwari, N., et al., *EMT as the ultimate survival mechanism of cancer cells.* Semin Cancer Biol, 2012. **22**(3): p. 194-207.
83. Brown, K.A., et al., *Induction by transforming growth factor-beta1 of epithelial to mesenchymal transition is a rare event in vitro.* Breast Cancer Res, 2004. **6**(3): p. R215-31.
84. David, C.J., et al., *TGF-beta Tumor Suppression through a Lethal EMT.* Cell, 2016. **164**(5): p. 1015-30.
85. Jolly, M.K., et al., *Stability of the hybrid epithelial/mesenchymal phenotype.* Oncotarget, 2016. **7**(19): p. 27067-84.
86. Meyer-Schaller, N., et al., *A Hierarchical Regulatory Landscape during the Multiple Stages of EMT.* Dev Cell, 2019. **48**(4): p. 539-553 e6.
87. Pastushenko, I., et al., *Identification of the tumour transition states occurring during EMT.* Nature, 2018. **556**(7702): p. 463-468.
88. Zhang, J., et al., *TGF-beta-induced epithelial-to-mesenchymal transition proceeds through stepwise activation of multiple feedback loops.* Sci Signal, 2014. **7**(345): p. ra91.
89. Hong, T., et al., *An Ovol2-Zeb1 Mutual Inhibitory Circuit Governs Bidirectional and Multi-step Transition between Epithelial and Mesenchymal States.* PLoS Comput Biol, 2015. **11**(11): p. e1004569.
90. Huang, R.Y., et al., *An EMT spectrum defines an anoikis-resistant and spheroidogenic intermediate mesenchymal state that is sensitive to e-cadherin restoration by a src-kinase inhibitor, saracatinib (AZD0530).* Cell Death Dis, 2013. **4**: p. e915.
91. Pastushenko, I. and C. Blanpain, *EMT Transition States during Tumor Progression and Metastasis.* Trends Cell Biol, 2019. **29**(3): p. 212-226.
92. Tian, X.J., H. Zhang, and J. Xing, *Coupled reversible and irreversible bistable switches underlying TGFbeta-induced epithelial to mesenchymal transition.* Biophys J, 2013. **105**(4): p. 1079-89.
93. Lu, M., et al., *MicroRNA-based regulation of epithelial-hybrid-mesenchymal fate determination.* Proc Natl Acad Sci U S A, 2013. **110**(45): p. 18144-9.
94. Nieto, M.A., et al., *Emt: 2016.* Cell, 2016. **166**(1): p. 21-45.
95. Bray, F., et al., *Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.* CA Cancer J Clin, 2018. **68**(6): p. 394-424.
96. Aiello, N.M. and Y. Kang, *Context-dependent EMT programs in cancer metastasis.* J Exp Med, 2019. **216**(5): p. 1016-1026.
97. Diepenbruck, M. and G. Christofori, *Epithelial-mesenchymal transition (EMT) and metastasis: yes, no, maybe?* Curr Opin Cell Biol, 2016. **43**: p. 7-13.
98. Blanco, M.J., et al., *Correlation of Snail expression with histological grade and lymph node status in breast carcinomas.* Oncogene, 2002. **21**(20): p. 3241-6.

99. Lee, T.K., et al., *Twist overexpression correlates with hepatocellular carcinoma metastasis through induction of epithelial-mesenchymal transition.* Clin Cancer Res, 2006. **12**(18): p. 5369-76.
100. Soltermann, A., et al., *Prognostic significance of epithelial-mesenchymal and mesenchymal-epithelial transition protein expression in non-small cell lung cancer.* Clin Cancer Res, 2008. **14**(22): p. 7430-7.
101. Gravdal, K., et al., *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, 2007. **13**(23): p. 7003-11.
102. Yagasaki, R., et al., *Clinical significance of E-cadherin and vimentin co-expression in breast cancer.* Int J Oncol, 1996. **9**(4): p. 755-61.
103. Thomas, P.A., et al., *Association between keratin and vimentin expression, malignant phenotype, and survival in postmenopausal breast cancer patients.* Clin Cancer Res, 1999. **5**(10): p. 2698-703.
104. Puram, S.V., et al., *Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer.* Cell, 2017. **171**(7): p. 1611-1624 e24.
105. Trimboli, A.J., et al., *Direct evidence for epithelial-mesenchymal transitions in breast cancer.* Cancer Res, 2008. **68**(3): p. 937-45.
106. Beerling, E., et al., *Plasticity between Epithelial and Mesenchymal States Unlinks EMT from Metastasis-Enhancing Stem Cell Capacity.* Cell Rep, 2016. **14**(10): p. 2281-8.
107. Rhim, A.D., et al., *EMT and dissemination precede pancreatic tumor formation.* Cell, 2012. **148**(1-2): p. 349-61.
108. Aiello, N.M., et al., *EMT Subtype Influences Epithelial Plasticity and Mode of Cell Migration.* Dev Cell, 2018. **45**(6): p. 681-695 e4.
109. Aiello, N.M., et al., *Metastatic progression is associated with dynamic changes in the local microenvironment.* Nat Commun, 2016. **7**: p. 12819.
110. Chanrion, M., et al., *Concomitant Notch activation and p53 deletion trigger epithelial-to-mesenchymal transition and metastasis in mouse gut.* Nat Commun, 2014. **5**: p. 5005.
111. Latil, M., et al., *Cell-Type-Specific Chromatin States Differentially Prime Squamous Cell Carcinoma Tumor-Initiating Cells for Epithelial to Mesenchymal Transition.* Cell Stem Cell, 2017. **20**(2): p. 191-204 e5.
112. Fischer, K.R., et al., *Epithelial-to-mesenchymal transition is not required for lung metastasis but contributes to chemoresistance.* Nature, 2015. **527**(7579): p. 472-6.
113. Zhao, Z., et al., *In Vivo Visualization and Characterization of Epithelial-Mesenchymal Transition in Breast Tumors.* Cancer Res, 2016. **76**(8): p. 2094-2104.
114. Xue, C., et al., *The gatekeeper effect of epithelial-mesenchymal transition regulates the frequency of breast cancer metastasis.* Cancer Res, 2003. **63**(12): p. 3386-94.
115. Ruscetti, M., et al., *Tracking and Functional Characterization of Epithelial-Mesenchymal Transition and Mesenchymal Tumor Cells during Prostate Cancer Metastasis.* Cancer Res, 2015. **75**(13): p. 2749-59.
116. Chen, Y., et al., *Dual reporter genetic mouse models of pancreatic cancer identify an epithelial-to-mesenchymal transition-independent metastasis program.* EMBO Mol Med, 2018. **10**(10).

117. Tan, T.Z., et al., *Epithelial-mesenchymal transition spectrum quantification and its efficacy in deciphering survival and drug responses of cancer patients.* EMBO Mol Med, 2014. **6**(10): p. 1279-93.
118. Koren, S., et al., *PIK3CA(H1047R) induces multipotency and multi-lineage mammary tumours.* Nature, 2015. **525**(7567): p. 114-8.
119. Van Keymeulen, A., et al., *Reactivation of multipotency by oncogenic PIK3CA induces breast tumour heterogeneity.* Nature, 2015. **525**(7567): p. 119-23.
120. Sinn, H.P. and H. Kreipe, *A Brief Overview of the WHO Classification of Breast Tumors, 4th Edition, Focusing on Issues and Updates from the 3rd Edition.* Breast Care (Basel), 2013. **8**(2): p. 149-54.
121. Bertos, N.R. and M. Park, *Breast cancer - one term, many entities?* J Clin Invest, 2011. **121**(10): p. 3789-96.
122. Perou, C.M., et al., *Molecular portraits of human breast tumours.* Nature, 2000. **406**(6797): p. 747-52.
123. Prat, A. and C.M. Perou, *Deconstructing the molecular portraits of breast cancer.* Mol Oncol, 2011. **5**(1): p. 5-23.
124. Prat, A., et al., *Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer.* Breast Cancer Res, 2010. **12**(5): p. R68.
125. Weigelt, B., et al., *Metaplastic breast carcinoma: more than a special type.* Nat Rev Cancer, 2014. **14**(3): p. 147-8.
126. Choi, Y., et al., *Epithelial-mesenchymal transition increases during the progression of in situ to invasive basal-like breast cancer.* Hum Pathol, 2013. **44**(11): p. 2581-9.
127. Taube, J.H., et al., *Core epithelial-to-mesenchymal transition interactome gene-expression signature is associated with claudin-low and metaplastic breast cancer subtypes.* Proc Natl Acad Sci U S A, 2010. **107**(35): p. 15449-54.
128. Chaffer, C.L., et al., *Poised chromatin at the ZEB1 promoter enables breast cancer cell plasticity and enhances tumorigenicity.* Cell, 2013. **154**(1): p. 61-74.
129. Banerji, S., et al., *Sequence analysis of mutations and translocations across breast cancer subtypes.* Nature, 2012. **486**(7403): p. 405-9.
130. Nik-Zainal, S., et al., *Landscape of somatic mutations in 560 breast cancer whole-genome sequences.* Nature, 2016. **534**(7605): p. 47-54.
131. Cancer Genome Atlas, N., *Comprehensive molecular portraits of human breast tumours.* Nature, 2012. **490**(7418): p. 61-70.
132. Sorlie, T., et al., *Repeated observation of breast tumor subtypes in independent gene expression data sets.* Proc Natl Acad Sci U S A, 2003. **100**(14): p. 8418-23.
133. Foulkes, W.D., et al., *Germline BRCA1 mutations and a basal epithelial phenotype in breast cancer.* J Natl Cancer Inst, 2003. **95**(19): p. 1482-5.
134. Powell, E., D. Piwnica-Worms, and H. Piwnica-Worms, *Contribution of p53 to metastasis.* Cancer Discov, 2014. **4**(4): p. 405-14.
135. Sengodan, S.K., et al., *Regulation of epithelial to mesenchymal transition by BRCA1 in breast cancer.* Crit Rev Oncol Hematol, 2018. **123**: p. 74-82.
136. Jiang, Z., et al., *Rb deletion in mouse mammary progenitors induces luminal-B or basal-like/EMT tumor subtypes depending on p53 status.* J Clin Invest, 2010. **120**(9): p. 3296-309.
137. Creighton, C.J., et al., *Residual breast cancers after conventional therapy display mesenchymal as well as tumor-initiating features.* Proc Natl Acad Sci U S A, 2009. **106**(33): p. 13820-5.

138. Brabletz, T., et al., *Variable beta-catenin expression in colorectal cancers indicates tumor progression driven by the tumor environment.* Proc Natl Acad Sci U S A, 2001. **98**(18): p. 10356-61.
139. Christofori, G., *New signals from the invasive front.* Nature, 2006. **441**(7092): p. 444-50.
140. Puisieux, A., T. Brabletz, and J. Caramel, *Oncogenic roles of EMT-inducing transcription factors.* Nat Cell Biol, 2014. **16**(6): p. 488-94.
141. Scheel, C. and R.A. Weinberg, *Cancer stem cells and epithelial-mesenchymal transition: concepts and molecular links.* Semin Cancer Biol, 2012. **22**(5-6): p. 396-403.
142. Lytle, N.K., A.G. Barber, and T. Reya, *Stem cell fate in cancer growth, progression and therapy resistance.* Nat Rev Cancer, 2018. **18**(11): p. 669-680.
143. Mani, S.A., et al., *The epithelial-mesenchymal transition generates cells with properties of stem cells.* Cell, 2008. **133**(4): p. 704-15.
144. Wellner, U., et al., *The EMT-activator ZEB1 promotes tumorigenicity by repressing stemness-inhibiting microRNAs.* Nat Cell Biol, 2009. **11**(12): p. 1487-95.
145. Ye, X., et al., *Distinct EMT programs control normal mammary stem cells and tumour-initiating cells.* Nature, 2015. **525**(7568): p. 256-60.
146. Guen, V.J., et al., *EMT programs promote basal mammary stem cell and tumor-initiating cell stemness by inducing primary ciliogenesis and Hedgehog signaling.* Proc Natl Acad Sci U S A, 2017. **114**(49): p. E10532-E10539.
147. Kim, J., et al., *Tumor initiating but differentiated luminal-like breast cancer cells are highly invasive in the absence of basal-like activity.* Proc Natl Acad Sci U S A, 2012. **109**(16): p. 6124-9.
148. Liu, S., et al., *Breast cancer stem cells transition between epithelial and mesenchymal states reflective of their normal counterparts.* Stem Cell Reports, 2014. **2**(1): p. 78-91.
149. Sikandar, S.S., et al., *Role of epithelial to mesenchymal transition associated genes in mammary gland regeneration and breast tumorigenesis.* Nat Commun, 2017. **8**(1): p. 1669.
150. Celia-Terrassa, T., et al., *Epithelial-mesenchymal transition can suppress major attributes of human epithelial tumor-initiating cells.* J Clin Invest, 2012. **122**(5): p. 1849-68.
151. Lambert, A.W., D.R. Pattabiraman, and R.A. Weinberg, *Emerging Biological Principles of Metastasis.* Cell, 2017. **168**(4): p. 670-691.
152. Yilmaz, M., G. Christofori, and F. Lehembre, *Distinct mechanisms of tumor invasion and metastasis.* Trends Mol Med, 2007. **13**(12): p. 535-41.
153. Odenthal, J., R. Takes, and P. Friedl, *Plasticity of tumor cell invasion: governance by growth factors and cytokines.* Carcinogenesis, 2016. **37**(12): p. 1117-1128.
154. Martinez, V. and J.G. Azzopardi, *Invasive lobular carcinoma of the breast: incidence and variants.* Histopathology, 1979. **3**(6): p. 467-88.
155. Friedl, P., et al., *Classifying collective cancer cell invasion.* Nat Cell Biol, 2012. **14**(8): p. 777-83.
156. Campbell, K., et al., *Collective cell migration and metastases induced by an epithelial-to-mesenchymal transition in Drosophila intestinal tumors.* Nat Commun, 2019. **10**(1): p. 2311.
157. Bocci, F., et al., *Numb prevents a complete epithelial-mesenchymal transition by modulating Notch signalling.* J R Soc Interface, 2017. **14**(136).

158. Dang, T.T., et al., *DeltaNp63alpha Promotes Breast Cancer Cell Motility through the Selective Activation of Components of the Epithelial-to-Mesenchymal Transition Program.* Cancer Res, 2015. **75**(18): p. 3925-35.
159. Grigore, A.D., et al., *Tumor Budding: The Name is EMT. Partial EMT.* J Clin Med, 2016. **5**(5).
160. Cheung, K.J., et al., *Polyclonal breast cancer metastases arise from collective dissemination of keratin 14-expressing tumor cell clusters.* Proc Natl Acad Sci U S A, 2016. **113**(7): p. E854-63.
161. Cheung, K.J., et al., *Collective invasion in breast cancer requires a conserved basal epithelial program.* Cell, 2013. **155**(7): p. 1639-51.
162. Wicki, A., et al., *Tumor invasion in the absence of epithelial-mesenchymal transition: podoplanin-mediated remodeling of the actin cytoskeleton.* Cancer Cell, 2006. **9**(4): p. 261-72.
163. Gaggioli, C., et al., *Fibroblast-led collective invasion of carcinoma cells with differing roles for RhoGTPases in leading and following cells.* Nat Cell Biol, 2007. **9**(12): p. 1392-400.
164. Dang, T.T., A.M. Prechtl, and G.W. Pearson, *Breast cancer subtype-specific interactions with the microenvironment dictate mechanisms of invasion.* Cancer Res, 2011. **71**(21): p. 6857-66.
165. Matise, L.A., et al., *Lack of transforming growth factor-beta signaling promotes collective cancer cell invasion through tumor-stromal crosstalk.* Breast Cancer Res, 2012. **14**(4): p. R98.
166. Joyce, J.A. and J.W. Pollard, *Microenvironmental regulation of metastasis.* Nat Rev Cancer, 2009. **9**(4): p. 239-52.
167. Friedl, P. and K. Wolf, *Plasticity of cell migration: a multiscale tuning model.* J Cell Biol, 2010. **188**(1): p. 11-9.
168. Harney, A.S., et al., *Real-Time Imaging Reveals Local, Transient Vascular Permeability, and Tumor Cell Intravasation Stimulated by TIE2hi Macrophage-Derived VEGFA.* Cancer Discov, 2015. **5**(9): p. 932-43.
169. Fantozzi, A., et al., *VEGF-mediated angiogenesis links EMT-induced cancer stemness to tumor initiation.* Cancer Res, 2014. **74**(5): p. 1566-75.
170. Bockhorn, M., R.K. Jain, and L.L. Munn, *Active versus passive mechanisms in metastasis: do cancer cells crawl into vessels, or are they pushed?* Lancet Oncol, 2007. **8**(5): p. 444-8.
171. Bill, R. and G. Christofori, *The relevance of EMT in breast cancer metastasis: Correlation or causality?* FEBS Lett, 2015. **589**(14): p. 1577-87.
172. Labelle, M., S. Begum, and R.O. Hynes, *Direct signaling between platelets and cancer cells induces an epithelial-mesenchymal-like transition and promotes metastasis.* Cancer Cell, 2011. **20**(5): p. 576-90.
173. Aktas, B., et al., *Stem cell and epithelial-mesenchymal transition markers are frequently overexpressed in circulating tumor cells of metastatic breast cancer patients.* Breast Cancer Res, 2009. **11**(4): p. R46.
174. Kallergi, G., et al., *Epithelial to mesenchymal transition markers expressed in circulating tumour cells of early and metastatic breast cancer patients.* Breast Cancer Res, 2011. **13**(3): p. R59.
175. Francart, M.E., et al., *Epithelial-mesenchymal plasticity and circulating tumor cells: Travel companions to metastases.* Dev Dyn, 2018. **247**(3): p. 432-450.
176. Gradilone, A., et al., *Circulating tumour cells lacking cytokeratin in breast cancer: the importance of being mesenchymal.* J Cell Mol Med, 2011. **15**(5): p. 1066-70.

177. Lecharpentier, A., et al., *Detection of circulating tumour cells with a hybrid (epithelial/mesenchymal) phenotype in patients with metastatic non-small cell lung cancer.* Br J Cancer, 2011. **105**(9): p. 1338-41.
178. Wu, S., et al., *Classification of circulating tumor cells by epithelial-mesenchymal transition markers.* PLoS One, 2015. **10**(4): p. e0123976.
179. Polioudaki, H., et al., *Variable expression levels of keratin and vimentin reveal differential EMT status of circulating tumor cells and correlation with clinical characteristics and outcome of patients with metastatic breast cancer.* BMC Cancer, 2015. **15**: p. 399.
180. Yu, M., et al., *Circulating breast tumor cells exhibit dynamic changes in epithelial and mesenchymal composition.* Science, 2013. **339**(6119): p. 580-4.
181. Aceto, N., et al., *Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis.* Cell, 2014. **158**(5): p. 1110-1122.
182. Aceto, N., et al., *En Route to Metastasis: Circulating Tumor Cell Clusters and Epithelial-to-Mesenchymal Transition.* Trends Cancer, 2015. **1**(1): p. 44-52.
183. Hou, J.M., et al., *Circulating tumor cells as a window on metastasis biology in lung cancer.* Am J Pathol, 2011. **178**(3): p. 989-96.
184. Shenoy, A.K. and J. Lu, *Cancer cells remodel themselves and vasculature to overcome the endothelial barrier.* Cancer Lett, 2016. **380**(2): p. 534-44.
185. Padua, D., et al., *TGFbeta primes breast tumors for lung metastasis seeding through angiopoietin-like 4.* Cell, 2008. **133**(1): p. 66-77.
186. Nguyen, D.X., P.D. Bos, and J. Massague, *Metastasis: from dissemination to organ-specific colonization.* Nat Rev Cancer, 2009. **9**(4): p. 274-84.
187. Nan, X., et al., *Epithelial-Mesenchymal Plasticity in Organotropism Metastasis and Tumor Immune Escape.* J Clin Med, 2019. **8**(5).
188. Cameron, M.D., et al., *Temporal progression of metastasis in lung: cell survival, dormancy, and location dependence of metastatic inefficiency.* Cancer Res, 2000. **60**(9): p. 2541-6.
189. Luzzi, K.J., et al., *Multistep nature of metastatic inefficiency: dormancy of solitary cells after successful extravasation and limited survival of early micrometastases.* Am J Pathol, 1998. **153**(3): p. 865-73.
190. Takano, S., et al., *Prrx1 isoform switching regulates pancreatic cancer invasion and metastatic colonization.* Genes Dev, 2016. **30**(2): p. 233-47.
191. Tran, H.D., et al., *Transient SNAIL1 expression is necessary for metastatic competence in breast cancer.* Cancer Res, 2014. **74**(21): p. 6330-40.
192. Stankic, M., et al., *TGF-beta-Id1 signaling opposes Twist1 and promotes metastatic colonization via a mesenchymal-to-epithelial transition.* Cell Rep, 2013. **5**(5): p. 1228-42.
193. Tsai, J.H., et al., *Spatiotemporal regulation of epithelial-mesenchymal transition is essential for squamous cell carcinoma metastasis.* Cancer Cell, 2012. **22**(6): p. 725-36.
194. Del Pozo Martin, Y., et al., *Mesenchymal Cancer Cell-Stroma Crosstalk Promotes Niche Activation, Epithelial Reversion, and Metastatic Colonization.* Cell Rep, 2015. **13**(11): p. 2456-2469.
195. Weidenfeld, K. and D. Barkan, *EMT and Stemness in Tumor Dormancy and Outgrowth: Are They Intertwined Processes?* Front Oncol, 2018. **8**: p. 381.
196. Title, A.C., et al., *Genetic dissection of the miR-200-Zeb1 axis reveals its importance in tumor differentiation and invasion.* Nat Commun, 2018. **9**(1): p. 4671.

197. Xu, Y., 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**(43): p. 11494-11499.
198. Krebs, A.M., et al., *The EMT-activator Zeb1 is a key factor for cell plasticity and promotes metastasis in pancreatic cancer.* Nat Cell Biol, 2017. **19**(5): p. 518-529.
199. Zheng, X., et al., *Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer.* Nature, 2015. **527**(7579): p. 525-530.
200. Stemmler, M.P., et al., *Non-redundant functions of EMT transcription factors.* Nat Cell Biol, 2019. **21**(1): p. 102-112.
201. Aiello, N.M., et al., *Upholding a role for EMT in pancreatic cancer metastasis.* Nature, 2017. **547**(7661): p. E7-E8.
202. Ye, X., et al., *Upholding a role for EMT in breast cancer metastasis.* Nature, 2017. **547**(7661): p. E1-E3.
203. Reichert, M., et al., *Regulation of Epithelial Plasticity Determines Metastatic Organotropism in Pancreatic Cancer.* Dev Cell, 2018. **45**(6): p. 696-711 e8.
204. Padmanaban, V., et al., *E-cadherin is required for metastasis in multiple models of breast cancer.* Nature, 2019.
205. Farmer, P., et al., *A stroma-related gene signature predicts resistance to neoadjuvant chemotherapy in breast cancer.* Nat Med, 2009. **15**(1): p. 68-74.
206. Singh, A. and J. Settleman, *EMT, cancer stem cells and drug resistance: an emerging axis of evil in the war on cancer.* Oncogene, 2010. **29**(34): p. 4741-51.
207. Khoo, B.L., et al., *Short-term expansion of breast circulating cancer cells predicts response to anti-cancer therapy.* Oncotarget, 2015. **6**(17): p. 15578-93.
208. Gottesman, M.M., T. Fojo, and S.E. Bates, *Multidrug resistance in cancer: role of ATP-dependent transporters.* Nat Rev Cancer, 2002. **2**(1): p. 48-58.
209. Saxena, M., et al., *Transcription factors that mediate epithelial-mesenchymal transition lead to multidrug resistance by upregulating ABC transporters.* Cell Death Dis, 2011. **2**: p. e179.
210. Kurrey, N.K., et al., *Snail and slug mediate radioresistance and chemoresistance by antagonizing p53-mediated apoptosis and acquiring a stem-like phenotype in ovarian cancer cells.* Stem Cells, 2009. **27**(9): p. 2059-68.
211. Lim, S., et al., *SNAI1-mediated epithelial-mesenchymal transition confers chemoresistance and cellular plasticity by regulating genes involved in cell death and stem cell maintenance.* PLoS One, 2013. **8**(6): p. e66558.
212. Li, Q.Q., et al., *Twist1-mediated adriamycin-induced epithelial-mesenchymal transition relates to multidrug resistance and invasive potential in breast cancer cells.* Clin Cancer Res, 2009. **15**(8): p. 2657-65.
213. Zhang, Z., et al., *Activation of the AXL kinase causes resistance to EGFR-targeted therapy in lung cancer.* Nat Genet, 2012. **44**(8): p. 852-60.
214. Byers, L.A., et al., *An epithelial-mesenchymal transition gene signature predicts resistance to EGFR and PI3K inhibitors and identifies Axl as a therapeutic target for overcoming EGFR inhibitor resistance.* Clin Cancer Res, 2013. **19**(1): p. 279-90.
215. Noman, M.Z., et al., *The immune checkpoint ligand PD-L1 is upregulated in EMT-activated human breast cancer cells by a mechanism involving ZEB-1 and miR-200.* Oncoimmunology, 2017. **6**(1): p. e1263412.

216. Chen, L., et al., *Metastasis is regulated via microRNA-200/ZEB1 axis control of tumour cell PD-L1 expression and intratumoral immunosuppression.* Nat Commun, 2014. **5**: p. 5241.
217. Mak, M.P., et al., *A Patient-Derived, Pan-Cancer EMT Signature Identifies Global Molecular Alterations and Immune Target Enrichment Following Epithelial-to-Mesenchymal Transition.* Clin Cancer Res, 2016. **22**(3): p. 609-20.
218. Soundararajan, R., et al., *Targeting the Interplay between Epithelial-to-Mesenchymal-Transition and the Immune System for Effective Immunotherapy.* Cancers (Basel), 2019. **11**(5).
219. Dongre, A., et al., *Epithelial-to-Mesenchymal Transition Contributes to Immunosuppression in Breast Carcinomas.* Cancer Res, 2017. **77**(15): p. 3982-3989.
220. Gonzalez, H., C. Hagerling, and Z. Werb, *Roles of the immune system in cancer: from tumor initiation to metastatic progression.* Genes Dev, 2018. **32**(19-20): p. 1267-1284.
221. Hay, E.D., *Organization and fine structure of epithelium and mesenchyme in the developing chick embryo.* Fleischmajer. R, Billingham. Baltimore, MD, USA: Williams & Wilkins Co., 1968. **Proceedings of the 18th Hahnemann Symposium**(31-55).
222. Bierie, B., et al., *Integrin-beta4 identifies cancer stem cell-enriched populations of partially mesenchymal carcinoma cells.* Proc Natl Acad Sci U S A, 2017. **114**(12): p. E2337-E2346.
223. Kroger, C., et al., *Acquisition of a hybrid E/M state is essential for tumorigenicity of basal breast cancer cells.* Proc Natl Acad Sci U S A, 2019. **116**(15): p. 7353-7362.
224. Jolly, M.K., S.A. Mani, and H. Levine, *Hybrid epithelial/mesenchymal phenotype(s): The 'fittest' for metastasis?* Biochim Biophys Acta Rev Cancer, 2018. **1870**(2): p. 151-157.
225. Pattabiraman, D.R. and R.A. Weinberg, *Targeting the Epithelial-to-Mesenchymal Transition: The Case for Differentiation-Based Therapy.* Cold Spring Harb Symp Quant Biol, 2016. **81**: p. 11-19.
226. Chaffer, C.L., et al., *EMT, cell plasticity and metastasis.* Cancer Metastasis Rev, 2016. **35**(4): p. 645-654.
227. Lacy, S.A., D.R. Miles, and L.T. Nguyen, *Clinical Pharmacokinetics and Pharmacodynamics of Cabozantinib.* Clin Pharmacokinet, 2017. **56**(5): p. 477-491.
228. Thaiparambil, J.T., et al., *Withaferin A inhibits breast cancer invasion and metastasis at sub-cytotoxic doses by inducing vimentin disassembly and serine 56 phosphorylation.* Int J Cancer, 2011. **129**(11): p. 2744-55.
229. Bargagna-Mohan, P., et al., *The tumor inhibitor and antiangiogenic agent withaferin A targets the intermediate filament protein vimentin.* Chem Biol, 2007. **14**(6): p. 623-34.
230. Gupta, P.B., et al., *Identification of selective inhibitors of cancer stem cells by high-throughput screening.* Cell, 2009. **138**(4): p. 645-659.
231. Pattabiraman, D.R., et al., *Activation of PKA leads to mesenchymal-to-epithelial transition and loss of tumor-initiating ability.* Science, 2016. **351**(6277): p. aad3680.
232. de The, H., *Differentiation therapy revisited.* Nat Rev Cancer, 2018. **18**(2): p. 117-127.

233. Ishay-Ronen, D., et al., *Gain Fat-Lose Metastasis: Converting Invasive Breast Cancer Cells into Adipocytes Inhibits Cancer Metastasis.* Cancer Cell, 2019. **35**(1): p. 17-32 e6.
234. Engleka, K.A., et al., *Islet1 derivatives in the heart are of both neural crest and second heart field origin.* Circ Res, 2012. **110**(7): p. 922-6.
235. Bang, S.J., et al., *Projections and interconnections of genetically defined serotonin neurons in mice.* Eur J Neurosci, 2012. **35**(1): p. 85-96.
236. Guy, C.T., R.D. Cardiff, and W.J. Muller, *Induction of mammary tumors by expression of polyomavirus middle T oncogene: a transgenic mouse model for metastatic disease.* Mol Cell Biol, 1992. **12**(3): p. 954-61.
237. Luond, F., et al., *A Transgenic MMTV-Flippase Mouse Line for Molecular Engineering in Mammary Gland and Breast Cancer Mouse Models.* J Mammary Gland Biol Neoplasia, 2019. **24**(1): p. 39-45.
238. Saxena, M., et al., *PyMT-1099, a versatile murine cell model for EMT in breast cancer.* Sci Rep, 2018. **8**(1): p. 12123.
239. Orend, G. and R. Chiquet-Ehrismann, *Tenascin-C induced signaling in cancer.* Cancer Lett, 2006. **244**(2): p. 143-63.
240. Linnemann, D., et al., *Characterization of N-cadherin messenger RNA and polypeptide expression in rat.* Int J Dev Neurosci, 1994. **12**(5): p. 441-50.
241. Egeblad, M., M.G. Rasch, and V.M. Weaver, *Dynamic interplay between the collagen scaffold and tumor evolution.* Curr Opin Cell Biol, 2010. **22**(5): p. 697-706.
242. Sabeh, F., R. Shimizu-Hirota, and S.J. Weiss, *Protease-dependent versus -independent cancer cell invasion programs: three-dimensional amoeboid movement revisited.* J Cell Biol, 2009. **185**(1): p. 11-9.
243. Giancotti, F.G., *Mechanisms governing metastatic dormancy and reactivation.* Cell, 2013. **155**(4): p. 750-64.
244. Goddard, E.T., et al., *Dormant tumour cells, their niches and the influence of immunity.* Nat Cell Biol, 2018. **20**(11): p. 1240-1249.
245. Madisen, L., et al., *A robust and high-throughput Cre reporting and characterization system for the whole mouse brain.* Nat Neurosci, 2010. **13**(1): p. 133-40.
246. Aufderheide, E. and P. Ekblom, *Tenascin during gut development: appearance in the mesenchyme, shift in molecular forms, and dependence on epithelial-mesenchymal interactions.* J Cell Biol, 1988. **107**(6 Pt 1): p. 2341-9.
247. Waldmeier, L., et al., *Py2T murine breast cancer cells, a versatile model of TGFbeta-induced EMT in vitro and in vivo.* PLoS One, 2012. **7**(11): p. e48651.
248. Borowsky, A.D., *Choosing a mouse model: experimental biology in context--the utility and limitations of mouse models of breast cancer.* Cold Spring Harb Perspect Biol, 2011. **3**(9): p. a009670.
249. Blaas, L., et al., *Lgr6 labels a rare population of mammary gland progenitor cells that are able to originate luminal mammary tumours.* Nat Cell Biol, 2016. **18**(12): p. 1346-1356.
250. Zomer, A., et al., *Intravital imaging of cancer stem cell plasticity in mammary tumors.* Stem Cells, 2013. **31**(3): p. 602-6.
251. Branda, C.S. and S.M. Dymecki, *Talking about a revolution: The impact of site-specific recombinases on genetic analyses in mice.* Dev Cell, 2004. **6**(1): p. 7-28.
252. Olorunniji, F.J., S.J. Rosser, and W.M. Stark, *Site-specific recombinases: molecular machines for the Genetic Revolution.* Biochem J, 2016. **473**(6): p. 673-84.

253. Dymecki, S.M., R.S. Ray, and J.C. Kim, *Mapping cell fate and function using recombinase-based intersectional strategies.* Methods Enzymol, 2010. **477**: p. 183-213.
254. Schonhuber, N., et al., *A next-generation dual-recombinase system for time- and host-specific targeting of pancreatic cancer.* Nat Med, 2014. **20**(11): p. 1340-1347.
255. Shai, A., et al., *TP53 Silencing Bypasses Growth Arrest of BRAFV600E-Induced Lung Tumor Cells in a Two-Switch Model of Lung Tumorigenesis.* Cancer Res, 2015. **75**(15): p. 3167-80.
256. Young, N.P., D. Crowley, and T. Jacks, *Uncoupling cancer mutations reveals critical timing of p53 loss in sarcomagenesis.* Cancer Res, 2011. **71**(11): p. 4040-7.
257. Moding, E.J., et al., *Atm deletion with dual recombinase technology preferentially radiosensitizes tumor endothelium.* J Clin Invest, 2014. **124**(8): p. 3325-38.
258. He, L., et al., *Enhancing the precision of genetic lineage tracing using dual recombinases.* Nat Med, 2017. **23**(12): p. 1488-1498.
259. Lee, C.L., et al., *Generation of primary tumors with Flp recombinase in FRT-flanked p53 mice.* Dis Model Mech, 2012. **5**(3): p. 397-402.
260. Andrechek, E.R., et al., *Genetic heterogeneity of Myc-induced mammary tumors reflecting diverse phenotypes including metastatic potential.* Proc Natl Acad Sci U S A, 2009. **106**(38): p. 16387-92.
261. Schoenenberger, C.A., et al., *Targeted c-myc gene expression in mammary glands of transgenic mice induces mammary tumours with constitutive milk protein gene transcription.* EMBO J, 1988. **7**(1): p. 169-75.
262. Fantozzi, A. and G. Christofori, *Mouse models of breast cancer metastasis.* Breast Cancer Res, 2006. **8**(4): p. 212.

6 Contribution to Other Projects

During my Ph.D., I have had the opportunity to contribute to the following internal research projects:

- Meyer-Schaller N, Cardner M, Diepenbruck M, Saxena M, Tiede S, **Lüönd F**, Ivanek R, Beerenwinkel N, Christofori G. *A Hierarchical Regulatory Landscape during the Multiple Stages of EMT.* Dev Cell, 2019.

- Diepenbruck M, Tiede S, Saxena M, Ivanek R, Kalathur RKR, **Lüönd F**, Meyer-Schaller N, Christofori G. *miR-1199-5p and Zeb1 function in a double-negative feedback loop potentially coordinating EMT and tumour metastasis.* Nat Commun, 2017

By contributing to these projects, I got a better understanding of the regulatory mechanisms of EMT and how these could be targeted to suppress EMT and ultimately tumor growth and metastasis *in vivo.*

7 Acknowledgements

To Prof. Dr. Gerhard Christofori I would like to express my sincere gratitude for trusting me with this exciting and challenging project and guiding me along the way. I was incredibly fortunate to join a project that had been in planning for several years, and yet being given the freedom to pursue my own ideas. I greatly appreciated his enthusiasm, constructive criticism and encouraging support.

I would like to extend my gratitude to my thesis committee members Prof. Dr. Verdon Taylor and Prof. Dr. Lukas Sommer for their assistance and critical input during annual meetings.

My heartfelt appreciation goes to Dr. Nami Sugiyama for sharing excitement and frustration and making major progress together. Despite her busy schedule, there was always time for discussions, which influenced my scientific thinking tremendously.

Many thanks go to all colleagues from the Christofori Lab for the pleasant atmosphere, sharing knowledge and exchanging ideas. I am particularly grateful to Dr. Ruben Bill for introducing me to this project and to Carolina Hager for joining our team as a Master student. Her curiosity and motivation made me truly enjoy this teaching experience. I would also like to especially thank Stefanie Tiede and Dr. Marco Morini for their friendship, the nice time spent together in the lab and helping me out during tough times. I owe many thanks to Eleni Panoussis, Isabel Galm and Ursula Schmieder for their helpful “hands on” in managing endless mouse experiments. Furthermore, I am very thankful to PD Dr. Thomas Bürglin for support with FACS and spending hours sorting our cells. I would also like to thank Prof. Dr. Nicola Aceto and his team for sharing knowledge and technology.

I greatly appreciated the support with microscopy and image processing provided by Pascal Lorentz and Dr. Loïc Sauteur and I am very thankful to Angelika Offinger and Lisa Schneider for taking good care of my mice.

Finally, I am profoundly thankful to my family for supporting me in all respects.