

Klinik und Poliklinik für Hals-Nasen-Ohrenheilkunde
Klinik der Ludwig-Maximilians-Universität München
Direktor: Prof. Dr. med. Martin Canis

**Epithelial-Mesenchymal Transition Phenotypes Impact on Metastases
Formation in Metastatic Breast Cancer**

Dissertation
zum Erwerb des Doktorgrades der Humanbiologie
an der Medizinischen Fakultät der
Ludwig-Maximilians-Universität zu München

vorgelegt von

Xiao Liu

aus China

2020

Mit Genehmigung der Medizinischen Fakultät der
Universität München

Berichterstatter: Prof. Dr. *rer. nat.* Olivier Gires
Mitberichterstatter: Prof. Dr. med. Martin Angele
Priv. Doz. Dr. Med. Ulrich Andergassen
Prof. Dr. med. Claus Belka

Dekan: Prof. Dr. med. dent. Reinhard HICKEL

Tag der mündlichen Prüfung: 09.06.2020



LUDWIG-
MAXIMILIANS-
UNIVERSITÄT
MÜNCHEN

Dean's Office
Medical Faculty



Affidavit

I, Xiao Liu, hereby declare that the submitted thesis entitled

Epithelial-Mesenchymal Transition Phenotypes Impact on Metastases Formation in Metastatic Breast Cancer

is my own work. I have only used the sources indicated and have not made unauthorised use of services of a third party. Where the work of others has been quoted or reproduced, the source is always given.

I further declare that the submitted thesis or parts thereof have not been presented as part of an examination degree to any other university.

Munich, 12.08.2019

Place, date

Xiao Liu

Signature doctoral candidate

CONTENTS

1. INTRODUCTION.....	1
1.1 Breast cancer	1
1.2 Tumor metastasis.....	3
1.3 Epithelial to mesenchymal transition.....	4
1.4 EMT and cancer organotropism during metastasis formation	5
1.5 Epithelial cell adhesion molecule (EpCAM)	6
1.6 CTC enrichment	9
1.7 Study design	10
2. MATERIALS AND METHODS.....	12
2.1 Materials	12
2.1.1 Chemicals	12
2.1.2 List of experimental kits.....	13
2.1.3 List of antibodies.....	13
2.1.4 List of qRT-PCR primers	14
2.1.5 List of cell lines.....	16
2.1.6 List of equipment	19
2.1.7 List of consumables.....	20
2.1.8 List of software applications.....	21
2.2 Methods.....	22
2.2.1 Cell culture.....	22
2.2.2 Flow cytometry	22
2.2.3 Proliferation assay	23
2.2.4 Cell metabolism	23
2.2.5 Single cell deposition for single cell-derived clones.....	23
2.2.6 Transwell invasion assay	24
2.2.7 Adhesion assay.....	24
2.2.8 Endothelial adhesion assay	25

2.2.9 Chemoresistance assay	25
2.2.10 2D colony formation assay	25
2.2.11 3D soft agar colony formation assay	26
2.2.12 mRNA isolation	26
2.2.13 Reverse transcription polymerase chain reaction (RT-PCR).....	27
2.2.14 Quantitative Real-Time PCR (qRT-PCR)	28
2.2.15 Animal experiments	29
2.2.16.1 Clinical cohort.....	33
2.3 Statistical analysis	35
3. RESULTS.....	36
3.1 The 4T1 MBC mouse model.....	36
3.2 EMT in CTC1 cells is accompanied by increased migration, but reduced proliferation and tumor formation ability.....	40
3.3 EMT heterogeneity in DTC1-derived CTC lines.....	49
3.4 Mesenchymal-type CTCs are more resistant to chemotherapy than epithelial-type CTCs.....	54
3.5 E/m-type CTCs possess highest lung metastasis formation ability <i>in vivo</i>	55
3.6 EMT in CTC lines is not a mere reflection of cell heterogeneity in 4T1 and DTC1 cells	59
3.7 EpCAM expression is negatively correlated with EMT.....	61
3.8 EpCAM expression is higher in metastatic sites compared with primary tumors	61
3.8 The EpCAM-positivity rates in DTCs is related to metastasis and predicts survival in patients with MBC	64
4. DISCUSSION.....	71
4.1 The 4T1 metastatic mouse model.....	71
4.2 EMT phenotypes of DTC-derived CTC lines	73
4.3 E/m-type CTCs have the highest capacity to form lung metastasis	75
4.4 The influence of EMT on metastasis	77
4.5 EpCAM expression during tumor metastasis.....	79
4.6 Characterization of CTCs and DTCs isolated by SE-iFISH.....	80
5. CONCLUSIONS.....	83

5.1. Summary	83
5.2. Zusammenfassung	84
6. REFERENCES	85
7. LIST OF FIGURES	100
8. TABLE LIST	102
9. ABBREVIATIONS.....	103
10. PUBLICATIONS.....	106
11. ACKNOWLEDGEMENTS.....	107

1. INTRODUCTION

1.1 Breast cancer

Cancer is a life-threatening disease worldwide. It was reported that in 2019, 1,762,450 patients will be newly diagnosed with carcinoma and 606,880 tumor-related deaths are estimated to be registered in the United States (Siegel et al., 2019). Breast cancer is the leading cancer type in women (Siegel et al., 2019). Owing to early detection techniques and improved therapy methods, the mortality of breast cancer patients has decreasing by 40% over the past 30 years (DeSantis et al., 2017). Once breast cancer is suspected, mammography, magnetic resonance imaging (MRI), and tissue biopsy can be used to further diagnose (Breast Cancer Treatment (PDQ®): Health Professional Version (2002)). The American Cancer Society (ACS) highly recommends that women who have reached 45 years of age should start regular mammography screening. Women aged ≥ 55 years should have annual screening for breast cancer (Oeffinger et al., 2015; Smith et al., 2019). Screening mammography could contribute to a 19% decrease of cancer mortality, for women over 60 years breast cancer mortality was even reduced by 32% through mammography (Pace and Keating, 2014).

Final diagnosis of breast cancer is always dependent on further evidence of pathology departments. When breast cancer is assured, staging breast cancer is the next step that is undertaken to help physicians determining the appropriate treatment that should be given to the patients. The TNM staging system is comprised of the physical extent of the tumor and nodal and distant metastases, where T stands for the primary tumor (range from T_0 to T_4), N is the status of tumor invasion of nearby lymph nodes (range from N_0 to N_3), and M means distant metastasis (*i.e.* lung, liver, kidney, bone; either M_0 or M_1) (Sobin LH, 2009).

According to molecular and phenotypic characteristics, breast cancer can be further classified into 3 group:

- The Luminal-type, where tumors express hormone receptors (*i.e.* estrogen (ER) or progesterone receptor (PR)). Patients with the luminal subtype can receive endocrine therapy.
- The HER2-type, where tumors over-express the human epidermal growth factor receptor 2

(HER2). ER and PR receptor can be positive or negative. This subtype can receive HER2-targeted therapy.

- The basal-type, or Triple Negative (TN), which comprises tumors that are negative for all three major receptor types, *i.e.* HER2, estrogen and progesterone receptors (Sobin LH, 2009).

The management of breast cancer includes surgery, medication, and radiation (Goetz et al., 2019). Surgery is the essential management for breast cancer that comprises the standard surgical interventions mastectomy (excision of the entire breast), quadrantectomy (remove a one-quarter part of the breast), and lumpectomy (cut out the primary tumor) (McDonald et al., 2016). Medication includes hormone blocking therapy, chemotherapy, and application of monoclonal antibodies. For patients who have a high predicted probability of tumor recurrence, an adjuvant chemotherapy, which is chemotherapy after definitive surgery, should generally be considered (McDonald et al., 2016). Patients with high expression of HER2 in their primary tumors are eligible for an adjuvant treatment with HER2-specific monoclonal antibodies (*i.e.* trastuzumab). Similarly, for patients with ER or PR positive breast tumors, related hormone blocking treatment is recommended (McDonald et al., 2016; Warriar et al., 2016). Patients with an advanced tumor stage (*i.e.* tumor size larger than 5 cm or fixed to the chest wall) would benefit from neoadjuvant therapy (medical intervention before surgery) (McDonald et al., 2016). Neoadjuvant treatment can help to release tumor burden, so that patients can become eligible for surgical excision of the primary tumor (Leal et al., 2015).

The patients' outcome depends on several factors including age, tumor stage, and tumor subtype. For patients at tumor stages I to III, which means they suffer from primary breast cancer without or with invasion of regional lymph nodes, but no distant metastases, 5-year overall survival rates are generally good (above 70%) (Kang and Pantel, 2013; Lim and Hortobagyi, 2016). In contrast, patients with distant metastasis (*i.e.* lung, liver, bone), which is referred to as tumor stage IV, have a 5-year survival of less than 25% (Kang and Pantel, 2013; Lim and Hortobagyi, 2016). Therefore, improving the management of patients with metastases would have a great influence on breast cancer patients' survival (Rashid and Takabe, 2014).

1.2 Tumor metastasis

Despite the expansion of new technologies and therapeutic methods, the development of metastases remains the major threat with significant impact on the patients' overall survival until today (Liu et al., 2017). The "seed and soil" theory describes the process of metastases formation and is widely cited and accepted. The initial step during the formation of distant metastasis resides in the release of single tumor cell or cell clusters from the primary tumor sites, which eventually intravasate into the blood stream; when carcinoma cells are in blood stream, they are called circulating tumor cells (CTCs) (Azevedo et al., 2015). Most of the CTCs will be under immune surveillance and will be recognized and eradicated by cells of the immune system. Only a small proportion of carcinoma cells can survive in the blood stream, succeed in extravasating, and finally disseminate to distant sites. Tumor cells that have settled in distant sites are then called disseminated tumor cells (DTCs) (Massague and Obenauf, 2016). DTCs are generally regarded as the major source of metastatic outgrowth and tumor relapse (Chaffer and Weinberg, 2011; Joosse et al., 2015; Lambert et al., 2017).

One study showed that even small tumors could release millions of tumor cells in the circulation, yet a lot of cancer patients never relapse, or develop metastases after a long time of latency (Nagrath et al., 2007). This suggests that metastasis formation is, at the single cell level, a rather inefficient process. In other words, metastasis initiating cells (MIC) originate from CTCs and DTCs, but only represent a proportion of the entire CTC/DTC population that is released from primary tumors. In the seed and soil hypothesis, low efficiency of metastatic colonization is explained by the fact that the distant "soil" might not be compatible with a foreign "seed", referring to differences between local and distant microenvironments. In fact, the most favorable soil may be the primary tumor site itself (Kim et al., 2009). In 2009, Kim and colleagues first described the preferential re-seeding of CTCs back to the primary tumor site or existing metastasis, over homing to tumor-free sites, and termed this process tumor "self-seeding" (Kim et al., 2009). The self-seeding program might enhance tumor malignancy via a renewed seeding of the most aggressive and therapy resistant cell clones in primary site or in metastases (Kim et al., 2009; Obenauf et al., 2015)

CTCs are termed DTCs after their successfully homing to distant sites, for instance, in the bone

marrow, lungs, spleen, *etc.* When and where metastases develop, may differ substantially depending on the heterogeneity of primary and systemic tumor cells, tumor types, treatment, and individual patients (Chambers et al., 2002; Dasgupta et al., 2017). Therefore, detecting and characterizing the actual functions and metastatic capacity of CTCs and DTCs could broaden our knowledge of metastasis formation, and eventually improve therapeutic regimens.

1.3 Epithelial to mesenchymal transition

Epithelial to mesenchymal transition (EMT) is known as a cellular program, through which epithelial cells can transdifferentiate to adopt a mesenchymal phenotype (Hay, 1995). During embryonic development, EMT allows cells to relocate within the embryo and, ultimately, to form the three germ layers, endo-, meso-, and ectoderm, which will further differentiate to form the mature organism (Kim et al., 2017). In cancer progression, EMT is considered to associate with various tumor functions, containing the malignant progression, the generation of cancer stem cell, tumor cell migration, intravasation and extravasation from blood vessels, metastasis formation, and resistance to treatment (Huang et al., 2013; Pastushenko and Blanpain, 2019; Pastushenko et al., 2018; Thiery and Lim, 2013; Ye et al., 2017). Importantly, through EMT cancer cells are equipped with increased migratory and invasive abilities, which could be beneficial for the initial steps of metastasis (Lambert et al., 2017). Recent studies demonstrated that, instead of adopting extreme epithelial or mesenchymal phenotypes, carcinoma cells rather undergo a more fluid transition, where cells can adopt a hybrid and more intermediate E/M phenotype (Brabletz et al., 2018b; Campbell, 2018).

The function of EMT in the metastatic process is, however, still in debate. Zheng *et al.* showed that pancreatic cancer cells could form lung metastasis even when the two EMT-inducing transcription factors Twist1 and Snai1 were deleted (Zheng et al., 2015). Similarly, Fischer *et al.* have shown that the overexpression of micro-RNA mir-200, which inhibits EMT, had no influence on lung metastasis formation (Fischer et al., 2015). Particularly these two publications have sparked a vivid discussion and debate on the role of EMT in tumor metastases (Brabletz et al., 2018b; Derynck and Weinberg, 2019; Ye et al., 2017).

At the molecular level, EMT is under the control of multiple pathways, including transforming growth

factor β (TGF- β), the Wnt pathway and β -catenin, the Notch signaling pathway, and tyrosine kinases receptors (*e.g.* EGFR) (De Francesco et al., 2018; Gonzalez and Medici, 2014; Moustakas and Heldin, 2016; Wu et al., 2016; Zhang et al., 2016). Eventually, EMT is induced by a set of core EMT transcription factors (EMT-TFs) including Snail, Slug, Zeb1 and 2, and Twist1 (Kalluri and Weinberg, 2009; Stemmler et al., 2019; Tam and Weinberg, 2013). Multiple genes of the above-mentioned pathways and of the core EMT-TFs directly participate in the regulation of EMT. Epithelial proteins like E-cadherin, Cytokeratins, Claudins, Rab25, and Occludin, play important roles in maintaining epithelial features of cells, like promoting cell-cell contact formation and cell proliferation. EMT-TFs can influence the expression of epithelial (*i.e.* E-cadherin, Cytokeratins) and mesenchymal (*i.e.* Vimentin) proteins, thus, regulating the EMT statuses of cells (Brabletz et al., 2018b; Gonzalez and Medici, 2014). As a result, a down-regulation of E-cadherin can lead to a significant loss of the adhesive ability of cells and, in turn, can reduce cell-cell contacts. Recently, it was reported that small, non-coding RNAs had an effect on the EMT program. For example, miR-9 can inhibit the expression of E-cadherin and promote EMT, while the miR-200 family can interact with ZEB1/2 and suppress EMT (Yan *et al.*, 2013).

1.4 EMT and cancer organotropism during metastasis formation

Different cancer types have differing preferential metastatic organs, *i.e.* patients with MBC frequently have bone and lung metastases, colon cancer preferentially metastasize in the liver, whereas patients with lung cancer frequently develop brain metastasis. This is also known as organotropism during metastasis formation (Gao et al., 2019; Nan et al., 2019; Siegel et al., 2019). Recent studies showed that the development of organotropism during the metastatic process depended on the tumor cells' intrinsic properties, the distinct organ microenvironment, and the interaction between carcinoma cells and distant organs (Akhtar et al., 2019; Chen et al., 2018; Nan et al., 2019). For instance, patients with breast cancer can develop metastasis in multiple sites (*i.e.* bone, liver, and lung), which reflects the fact that cells in individual primary tumors have distinct abilities to metastasize in different organs. Furthermore, some organs are more difficult for tumor cells to access and to settle in. For example, the blood-brain barrier, which is made up of tightly connected endothelia, astrocytes, and pericytes, can

hinder the invasion of tumor cells from numerous cancers (Dong, 2018).

Cell-cell adhesive interactions can also influence organotropism during metastasis formation (Chen et al., 2018). During the EMT procedure, cancer cells will frequently decrease the expression of E-cadherin, and in parallel enhance N-cadherin expression (Gloushankova et al., 2017), which is referred to as the “cadherin switch”. E-cadherin plays a key role in tight cell-cell junctions in epithelial cells, while N-cadherin directly mediates homotypic and heterotypic cell-cell adhesion, and is predominantly expressed in neural, endothelial, and stromal cells (Chen et al., 2018; Gloushankova et al., 2017; Mrozik et al., 2018). Recent studies demonstrate that tumor cells can acquire partial or hybrid EMT phenotypes, which means epithelial and mesenchymal markers, including different cadherins, can co-exist in individual cells (Jordan et al., 2011; Pastushenko and Blanpain, 2019; Tan et al., 2014). CTCs/DTCs are considered as the source of metastatic seeds. Theoretically, CTCs/DTCs with a hybrid EMT phenotype may adapt more easily to the distant microenvironment by, for example, transiently changing their expression of cadherins (E- or N- cadherin). Such changes could allow them to tighten or release contacts to cells of various microenvironments, and thus to better cope with changing conditions along the metastatic cascade. In the last decade, many efforts have been made to capture and characterize CTCs and, to lesser extent, DTCs of various tumor types. However, the biological characterization and capture of CTCs and DTCs are still technically challenging (Gabriel et al., 2016). In most CTC isolation techniques, CTCs are captured through the enrichment of rare epithelial cells from blood, by using epithelial markers, *e.g.* epithelial cell adhesion molecule (EpCAM). Despite being of great value for further prognosis of patients, these techniques might preferentially enrich epithelial cells, thus introducing a bias in the actual nature of CTCs that are enriched. Therefore, EpCAM-dependent CTC enrichment technologies do not allow to study the full spectrum of systemic tumor cells. Hence, the actual contribution of EMT to metastasis formation remains only partially addressed.

1.5 Epithelial cell adhesion molecule (EpCAM)

EpCAM was first described by Herlyn *et al.* in 1979 and it was recognized as a humoral antigen that is highly expressed in human colon cancer (Herlyn et al., 1979). EpCAM is a membrane protein with

an apparent molecular weight of 30- to 40-kDa, and which was found expressed in most epithelial tissues and epithelial tumors (Imrich et al., 2012; Munz et al., 2009). EpCAM consists of a large extracellular domain (EpEX) and a small intracellular domain (EpICD) connected by a single transmembrane domain (Maetzel et al., 2009). Since its first description in 1979, numerous cellular functions have been assigned to EpCAM, ranging from the name-giving cell adhesion property, regulation of proliferation, maintenance of morphological and epithelial integrity, and regulation of cell differentiation, including the maintenance of the pluripotency of progenitor and embryonic stem cells (Gonzalez et al., 2009; Huang et al., 2011; Lu et al., 2010; Sarrach et al., 2018; Wang et al., 2018b; Wu et al., 2013). EpCAM signaling, which is required for its mitogenic and pluripotency-maintaining capacities, is initiated by regulated intramembrane proteolysis (RIP) (Maetzel et al., 2009; Munz et al., 2009). This cleavage is initiated by shedding of the extracellular domain EpEX by ADAM proteases, which generates a membrane-tethered C-terminal fragment (CTF) of EpCAM. EpCAM-CTF, as a substrate of the gamma-secretase complex, can be cleaved to release the EpICD part into the cytoplasm. EpICD can further form a complex together with FHL2, β -catenin, and Lef1, which can translocate in the cell's nucleus to further regulate gene expression (*i.e.* C-myc, Cyclin D1, amongst others) (Chaves-Perez et al., 2013; Huang et al., 2011; Maetzel et al., 2009; Munz et al., 2009).

More recently, a spatiotemporal regulation of EpCAM expression has been determined during the three-dimensional differentiation of pluripotent embryonic stem cells (ESC). EpCAM was shown to be entirely lost in mesodermally differentiating cells at an early time point of gastrulation, while it was retained in endodermal cells (Sarrach et al., 2018). Further analysis disclosed that both, EpCAM-negative and -positive cells are required to ensure proper differentiation of mesodermal cells, for example in order to terminally differentiate into contracting cardiomyocytes (Sarrach et al., 2018). Hence, EpCAM is a transmembrane protein with multiple functions in stem and epithelial cells.

In order to fulfil these functions, EpCAM can interact with other cell adhesion molecules and influence adhesive structures between cells and cell-matrix (Wu et al., 2013). For instance, the transmembrane domain can interact with claudin-7, and the reduction of EpCAM expression leads to a decrease of

Claudin-7 protein expression, too (Ladwein et al., 2005). EpCAM can enhance the proliferation of carcinoma cells and accordingly a down-regulation of EpCAM expression can cause the decrease of proliferation in FaDu cells (Maetzel et al., 2009; Munz et al., 2004). Previously, a study demonstrated that EpCAM can crosstalk with the Phosphatase and Tensin Homolog (PTEN)/Protein Kinase B (AKT)/Mechanistic target of Rapamycin Kinase (mTOR) pathway and regulates EMT, metastatic formation, and stemness in nasopharyngeal cancer through interaction with (Wang et al., 2018b).

Apart from epithelial cells, embryonic stem cells and precursor cells also express EpCAM (Balzar et al., 1999; Gires and Stoecklein, 2014; Gonzalez et al., 2009; Lu et al., 2010; Sarrach et al., 2018). On the base of a strong and frequent expression of EpCAM in epithelial cancer cells and because it represents a marker for epithelial cells which is lacking on blood cells, EpCAM is the most widely used membrane-associated anchor molecule to detect CTCs from blood cells (Banko et al., 2019; Gires and Stoecklein, 2014; Went et al., 2004). However, in analogy to its spatiotemporal expression in ESC and during embryonic differentiation, the first prove of a dynamic expression of EpCAM in human CTCs originated from a xenograft mouse model, where cells that had undergone EMT had lost the expression of EpCAM (Gorges et al., 2012). The dynamic regulation of EpCAM expression in tumor, especially in systemic cancer cells, has been addressed in a review paper by Wang *et al.* in 2017, who compiled the knowledge on EpCAM repression during EMT. Here, the general consensus is that EMT is accompanied by a loss of epithelial markers, including EpCAM. Beyond such concomitant loss, a causal role of EpCAM in the regulation of EMT is currently addressed more thoroughly. Hsu *et al.* described an EGFR-dependent cleavage of EpCAM in an endometrial cell line to generate an intracellular domain EpICD that actively induces EMT at the transcriptional level (Hsu et al., 2016). Such cleavage of EpCAM is based on regulated intramembrane proteolysis, as described by Maetzel *et al.* However, the described EGFR-dependent cleavage of EpCAM was neither observed in several head and neck, breast, colon, and prostate cancer cell lines, nor in the published RL95-2 endometrial cell line. In sharp contrast, the soluble ectodomain of EpCAM was described as a novel EGFR ligand that counteracts EGF-mediated EMT in head and neck squamous cell carcinomas (Pan et al., 2018).

Despite opposing results on the functions of EpCAM in EMT, it is ultimately of importance to note

that the reported loss of EpCAM during EMT in combination with its use as a marker to capture CTCs from the blood could hamper the study of CTCs that have undergone EMT and lost EpCAM expression (Wang et al., 2017a).

1.6 CTC enrichment

The presence of CTCs can serve as “liquid biopsy” to obtain instant information about systemic cancer in patients (Wang et al., 2017a). The CELLSEARCH® system is the first US Food and Drug Administration (FDA)-certified EpCAM-based CTC detection system (Gabriel et al., 2016; Liljefors et al., 2005). CELLSEARCH® is an immunofluorescence-based method in which CTCs are enriched with EpCAM-specific antibodies from the peripheral blood and are further defined as cells with nuclei (the nuclear marker, 4',6-diamidino-2-phenylindole (DAPI) is positive), that express epithelial markers (*i.e.* Cytokeratin, EpCAM) and lack the leukocyte marker CD45. In 2004, Allard *et al.* first demonstrated that CELLSEARCH® could be used to evaluate CTCs numbers in cancer patients, and a cutoff value of ≥ 2 CTCs was determined to distinguish healthy or non-malignant disease from cancer (Allard et al., 2004). From then on, plenty of studies have been conducted to investigate the prognostic ability of CELLSEARCH-detected CTCs in various tumor types and most of them showed promising results. Generally, comparably low cutoff values such as 1-5 cells per 7.5 mL of peripheral blood correlated with reduced overall survival (Cohen et al., 2008; Cristofanilli et al., 2004; de Bono et al., 2008; Gradilone et al., 2011; Janni et al., 2016b; Wang et al., 2016).

Recently, more attention was paid to the significance of EpCAM-negative CTCs, which are considered as CTCs with a more mesenchymal phenotype. It was reported that mesenchymal CTCs were related to poor response to treatment and tumor progression in patients with MBC (Bednarz-Knoll et al., 2012; Fischer et al., 2015; Yu et al., 2013). However, EpCAM-dependent CTC isolation methods are not suitable for the isolation of these mesenchymal CTCs that have undergone EMT. Accordingly, the development of non-EpCAM-based CTC capture techniques would benefit the isolation and detection of cells that underwent EMT. Currently, EpCAM-independent CTC isolation methods include microfiltration, immunomagnetic methods, and microchip technologies (Chikaishi et al., 2017; Gabriel et al., 2016). On the basis of a usually larger size of carcinoma cells than erythrocytes and leukocytes,

microfiltration enrichment methods capture CTC through enriching large size cells (Gabriel et al., 2016; Khetani et al., 2018). Immunomagnetic methods are based on the interaction of antigen-antibody and use antibody-conjugated magnetic beads to interact with specific antigens on the cell surface (Gabriel et al., 2016). Microchip technology uses a microfluidic device comprising antibody-functionalized (*i.e.* EGFR, prostate-specific antigen, HER-2) micro-posts to interact with the CTC surface and thus enables non-invasive isolation (Gabriel et al., 2016; Stott et al., 2010). The subtraction enrichment and immunostaining-fluorescence in situ hybridization (SE-iFISH) is an antigen-independent CTC enrichment strategy (Lin, 2015). By detecting aneuploid CTCs with abnormal numbers of Centromere 8, SE-iFISH can capture CTCs regardless of EpCAM expression (Ge et al., 2015; Lin, 2015; Wang et al., 2018a). These EpCAM-independent CTC capture techniques provide new opportunities to detect both EpCAM positive and negative CTCs.

1.7 Study design

CTCs and DTCs are accepted as main sources of MICs that give rise to life-threatening metastases in the course of cancer progression. As both, primary and systemic tumor cells, can undergo EMT to variable degrees, the question emerged as to which degree of EMT-associated changes in carcinoma cells promotes the formation of distant metastases (Aiello et al., 2017; Brabletz et al., 2018a; Ye et al., 2017). One major aim of our study was therefore to investigate EMT phenotypes and their influence on the metastasis formation capacity of systemic cancer cells in the 4T1 breast cancer mouse model. A further aim was to assess whether EpCAM expression in CTCs and DTCs could be used as a surrogate marker to define the EMT status of clinical samples of systemic tumor cells and to predict the outcome of patients with MBC.

To do so, 4T1 murine breast cancer cells were subcutaneously transplanted in the flank of BABL/c mice. Following the formation of primary tumors and the subsequent development of metastases, primary tumors, blood, bones, and organs (lung, liver, kidney, spleen) were collected and 4T1-derived cells were selected *ex vivo*. EMT phenotypes of systemic, 4T1-derived CTCs/DTCs were defined *in vitro* and their influence on proliferation, adhesion, migration, invasion, and tumor and metastases formation ability were analyzed *in vitro* and, following subcutaneous and intravenous re-

transplantation, *in vivo*. In a clinical setting, the EpCAM-independent CTCs/DTCs isolation technique SE-iFISH was used to analyze the proportion of EpCAM-positive cells in blood and bone marrow of patients with MBC (n = 34). In order to assess whether the expression of EpCAM, as a surrogate marker for the EMT status of carcinoma cells, could predict the clinical outcome of patients, EpCAM expression proportion were correlated to the metastatic statuses and clinical outcome of patients with MBC.

2. MATERIALS AND METHODS

2.1 Materials

2.1.1 Chemicals

Chemicals	Company
6-Thioguanine	Sigma, Saint Louis, USA
Agarose	Biosciences, Heidelberg, Germany
Ammonium chloride-based lysis reagent	BD Pharm Lyse™, BD Biosciences, Heidelberg, Germany
Calcein-AM	PromoKine/PromoCell GmbH, Heidelberg
Cisplatin	Santa Cruz Biotechnolog, Heidelberg, Germany
Collagenase	Sigma, Steinheim, Germany
Crystal Violet	Sigma, Saint Louis, USA
4',6-diamidino-2-phenylindole (DAPI)	Sigma-Aldrich GmbH, Taufkirchen
Deoxyribonuclease (DNase)	Sigma, Steinheim, Germany
Doxorubicin	Sigma, Steinheim, Germany
Dulbecco's Modified Eagle Medium (DMEM 4,5 g/L glucose/with L-glutamine)	Biochrom AG, Berlin, Germany
Fetal bovine serum (FBS)	Biochrom AG, Berlin, Germany
Ethanol	Abcam, Cambridge, UK
Gelatin	Sigma, Saint Louis, USA
Matrigel	Becton Dickinson, Heidelberg, Germany
3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) solution	Sigma, Saint Louis, USA
Penicillin-Streptomycin	Biochrom AG, Berlin, Germany
Phosphate buffer saline (PBS)	Apotheke Klinikum Großhadern, München Germany
Propidiumiodid (PI)	Sigma-Aldrich GmbH, Taufkirchen
Tumor necrosis factor alpha (TNF- α)	Thermo Fisher Scientific, Bleiswijk, Netherlands

Chemicals	Company
Trypsin	Biochrom AG, Berlin, Germany
Triton-X 100	Sigma-Aldrich GmbH, Taufkirchen

2.1.2 List of experimental kits

Products	Company
Light Cycler 480 SYBR Green I Master, Roche	Mannheim, Germany
QiaShredder	Qiagen, Hilden, Germany
QuantiTect Reverse Transcription Kit	Qiagen, Hilden, Germany
RNeasy Kit	Qiagen, Hilden, Germany

2.1.3 List of antibodies

Antibodies	Species	Company
Anti-mouse EpCAM	Rat IgG2a, κ cloneG8.8	Becton Dickinson, Heidelberg, Germany
Anti-mouse CD45	Rat IgG2b, κ 30-F11	BD Pharmingen, Heidelberg, Germany
Anti-mouse E-Cadherin	Rabbit IgG, clone24E10	Cell Signaling Technology, Danvers, USA
Anti-mouse Vimentin	Rabbit Monoclonal IgG, EPR3776	Abcam, Cambridge, USA
Anti-mouse Cytokeratin	Rabbit polyclonal	Invitrogen, Waltham, USA
Fluorescein rabbit-anti-rat	IgG (H&L)	BD bioscience, Heidelberg, Germany

2.1.4 List of qRT-PCR primers

Primers	Sequences
Ddr1	FW: 5'-TCC ATA GAC CAG AGG GAT C-3'
	BW: 5'-CAG GGC ATA GCG GCA CTT GG-3'
E-cadherin	FW: 5'-CAG GTC TCC TCA TGG CTT TGC-3'
	BW: 5'-CTT CCG AAA AGA AGG CTG TCC-3'
Epcam	FW: 5'-CAG TGT ACT TCC TAT GGT ACA CAG AAT ACT-3'
	BW: 5'-CTA GGC ATT AAG CTC TCT GTG GAT CTC ACC-3'
ErbB-2	FW: 5'-TCC CCA GGG AGT ATG TGA GG-3'
	BW: 5'- GAG GCG GGA CAC ATA TGG AG-3'
ErbB-3	FW: 5'-GCC CAA TCC TAA CCA GTG CT-3'
	BW: 5'-AGC CTG TAA TCT CCC GGA CT-3'
Grhl1	FW: 5'- GCT GAG ACA CTG GAA GTA CTG-3'
	BW: 5'-CGT GAA GGA AAT GGC GTT ATA AG-3'
Gusp	FW: 5'-CAA CCT CTG GTG GCC TTA CC-3'
	BW: 5'-GGG TGT AGT AGT CAG TCA CA -3'
Krt19	FW: 5'-CTA CCT TGC TCG GAT TGA GGA G-3'
	BW: 5'- AGT CTC GCT GGT AGC TCA GAT G-3'
N-cadherin	FW: 5'-AGG GTG GAC GTC ATT GTA GC-3'
	BW: 5'-CTG TTG GGG TCT GTC-3'
Rab25	FW: 5'-TGA GCC AAG ATG GGG AAT CG-3'
	BW: 5'-GGA GAA CTC AAC CCC GAT GG-3'
Slug	FW: 5'-TCC CAT TAG TGA CGA AGA-3'
	BW: 5'-CCC AGG CTC ACA TAT TCC-3'
Snail	5'-GCG GAA GAT CTT CAA CTG CAA ATA TTG TAA C-3'
	5'-GCA GTG GGA GCA GGA GAA TGG CTT CTC AC-3'

Primers	Sequences
Twist	FW: 5'-CGG GTC ATG GCT AAC GTG-3'
	BW: 5'-CAG CTT GCC ATC TTG GAG TC-3'
Vimentin	FW: 5'-CGG AAA GTG GAA TCC TTG CA-3'
	BW: 5'-CAC ATC GAT CTG GAC ATG CTG T-3'
Zeb1	FW: 5'-CCA TAC GAA TGC CCG AAC T-3'
	BW: 5'-ACA ACG GCT TGC ACC ACA-3'
Zeb2	FW: 5'-CCG TTG GAC CTG TCA TTA CC-3'
	BW: 5'-GAC GAT GAA GAA ACA CTG TTG TG-3'

2.1.5 List of cell lines

Cell lines	Description
4T1	Mouse breast cancer cell
NIH3T3	Mouse Fibroblasts
WEHI-231	Mouse B cell lymphoma
bEnd.3	Mouse endothelialpolyoma middle T antigen transformed
CTC1	4T1-derived CTC line
DTC1	4T1-derived DTC line
CTC8-1	DTC1-derived CTC line (from mouse number DTC1-8)
CTC8-5	DTC1-derived CTC line (from mouse number DTC1-8)
CTC8-6	DTC1-derived CTC line (from mouse number DTC1-8)
CTC8-12	DTC1-derived CTC line (from mouse number DTC1-8)
CTC6-11	DTC1-derived CTC line (from mouse number DTC1-6)
CTC6-6	DTC1-derived CTC line (from mouse number DTC1-6)
CTC7-1	DTC1-derived CTC line (from mouse number DTC1-7)
CTC10-1	DTC1-derived CTC line (from mouse number DTC1-10)
CTC10-2	DTC1-derived CTC line (from mouse number DTC1-10)
CTC8-2	DTC1-derived CTC line (from mouse number DTC1-8)
CTC8-3	DTC1-derived CTC line (from mouse number DTC1-8)
CTC8-8	DTC1-derived CTC line (from mouse number DTC1-8)
CTC8-11	DTC1-derived CTC line (from mouse number DTC1-8)
CTC8-13	DTC1-derived CTC line (from mouse number DTC1-8)
CTC8-14	DTC1-derived CTC line (from mouse number DTC1-8)
CTC8-15	DTC1-derived CTC line (from mouse number DTC1-8)
CTC6-4	DTC1-derived CTC line (from mouse number DTC1-6)
CTC6-5	DTC1-derived CTC line (from mouse number DTC1-6)

Cell lines	Description
CTC6-6	DTC1-derived CTC line (from mouse number DTC1-6)
CTC6-9	DTC1-derived CTC line (from mouse number DTC1-6)
CTC6-10	DTC1-derived CTC line (from mouse number DTC1-6)
CTC6-12	DTC1-derived CTC line (from mouse number DTC1-6)
CTC6-13	DTC1-derived CTC line (from mouse number DTC1-6)
CTC6-14	DTC1-derived CTC line (from mouse number DTC1-6)
CTC6-15	DTC1-derived CTC line (from mouse number DTC1-6)
DTC1-7-1	DTC1-derived DTC line (from mouse number DTC1-7)
DTC1-7-3	DTC1-derived DTC line (from mouse number DTC1-7)
DTC1-7-4	DTC1-derived DTC line (from mouse number DTC1-7)
DTC1-7-5	DTC1-derived DTC line (from mouse number DTC1-7)
DTC1-7-6	DTC1-derived DTC line (from mouse number DTC1-7)
DTC1-8-1	DTC1-derived DTC line (from mouse number DTC1-8)
DTC1-8-4	DTC1-derived DTC line (from mouse number DTC1-8)
DTC1-8-5	DTC1-derived DTC line (from mouse number DTC1-8)
DTC1-8-8	DTC1-derived DTC line (from mouse number DTC1-8)
DTC1-8-9	DTC1-derived DTC line (from mouse number DTC1-8)
4T1-PT1	4T1-derived primary tumor cell line (from mouse number 4T1-1)
4T1-PT2	4T1-derived primary tumor cell line (from mouse number 4T1-2)
4T1-PT3	4T1-derived primary tumor cell line (from mouse number 4T1-3)
4T1-PT4	4T1-derived primary tumor cell line (from mouse number 4T1-4)
4T1-PT5	4T1-derived primary tumor cell line (from mouse number 4T1-5)
4T1-PT6	4T1-derived primary tumor cell line (from mouse number 4T1-6)
4T1-PT7	4T1-derived primary tumor cell line (from mouse number 4T1-7)
4T1-PT8	4T1-derived primary tumor cell line (from mouse number 4T1-8)

Cell lines	Description
T1LM	4T1-derived lung metastasis cell line (from mouse number 4T1-1)
T2LM	4T1-derived lung metastasis cell line (from mouse number 4T1-2)
T3LM	4T1-derived lung metastasis cell line (from mouse number 4T1-3)
T4LM	4T1-derived lung metastasis cell line (from mouse number 4T1-4)
T2LNM	4T1-derived lymph node metastasis cell line (from mouse number 4T1-2)
CTC1-PT1	CTC1-derived primary tumor cell line (from mouse number CTC1-1)
CTC1-PT2	CTC1-derived primary tumor cell line (from mouse number CTC1-2)
CTC1-PT3	CTC1-derived primary tumor cell line (from mouse number CTC1-3)
CTC1-PT4	CTC1-derived primary tumor cell line (from mouse number CTC1-4)
CTC1-7 LM	CTC1-derived lung metastasis cell line (from mouse number CTC1-7)
CTC1-9 LM	CTC1-derived lung metastasis cell line (from mouse number CTC1-9)
CTC1-15 LM	CTC1-derived lung metastasis cell line (from mouse number CTC1-15)
CTC1-10 SM	CTC1-derived spleen metastasis cell line (from mouse number CTC1-10)
D1PT	DTC1-derived primary tumor cell line (from mouse number DTC1-1)
D2PT	DTC1-derived primary tumor cell line (from mouse number DTC1-2)
D3PT	DTC1-derived primary tumor cell line (from mouse number DTC1-3)
D7KM	DTC1-derived kidney metastasis cell line (from mouse number DTC1-7)
D10LM	DTC1-derived lung metastasis cell line (from mouse number DTC1-10)
D10SM	DTC1-derived spleen metastasis cell line (from mouse number DTC1-10)
D8LM	DTC1-derived lung metastasis cell line (from mouse number DTC1-8)
D8KM	DTC1-derived kidney metastasis cell line (from mouse number DTC1-8)
D9LM	DTC1-derived lung metastasis cell line (from mouse number DTC1-9)

2.1.6 List of equipment

Equipment	Company
Autoclave Systec 95	Systec GmbH, Wetzlar, Germany
Centrifuge Mikro 22R	Hettich Lab Technology, Tuttlingen, Germany
Centrifuge Rotanta 46 R	Hettich Lab Technology, Tuttlingen, Germany
ChemiDoc XRS+ imaging system	BioRad, Hercules, USA
Cell culture incubator HeraCell 240	Heraeus, Hanau, Germany
Freezer (-20°C, -80°C)	Liebherr, Ochsenhausen, Germany
Laminar flow cabinet	Heraeus Holding GmbH, Hanau, Germany
Light Cycler 480 System	Roche, Mannheim, Germany
Microliter pipettes	Gilson Inc., Middleton, USA
Microplate reader	Molecular Devices, Sunnyvale, CA, USA
Microwave	Sharp Electronics GmbH, Hamburg, Germany
Nanophotometer	Implen GmbH, München, Germany
Phase contrast microscope “Axiovert 25”	Carl Zeiss AG, Jena, Germany
Power supply E835	Consort bvba, Turnhout, Belgium
Thermomixer Comfort	Eppendorf, Hamburg, Germany
Water bath Exotherm U3e1	Julabo, Seelbach, Germany
Flow cytometer “FACS-Calibur”	BD Biosciences, Heidelberg, Germany
Vortex Mixer	IKA Works Inc., Wilmington, USA

2.1.7 List of consumables

Products	Company
1.5 mL Tube (nuclease-free)	Costar, New York, USA
Micro Tube (1.5 mL/2 mL)	Eppendorf AG, Hamburg, Germany
Pipette Tips (10 μ L, 20 μ L, 100 μ L, 1000 μ L)	Starlab, Hamburg, Germany
Safe seal tips professional (10 μ L, 20 μ L, 100 μ L, 1000 μ L)	Biozym Scientific GmbH, Hessisch Oldendorf, Germany
12-well plate (flat bottom)	Nunc, Wiesbaden, Germany
15 mL/50 mL Tube	Becton Dickinson, Heidelberg, Germany
24-well plate (flat bottom)	Nunc, Wiesbaden, Germany
40 μ m, 100 μ m sterile filters	Millipore, Wiesbaden, Germany
6-well plate (flat bottom)	Nunc, Wiesbaden, Germany
96-well plate (flat bottom)	Nunc, Wiesbaden, Germany
96-well plate (round bottom)	Nunc, Wiesbaden, Germany
Cell culture flasks and dishes	Nunc, Wiesbaden, Germany
Corning® Costar® stripettes	Sigma-Aldrich GmbH, Taufkirchen, Germany
Cryo tubes	Becton Dickinson, Heidelberg, Germany
Cryomold Tissue-Tek®, Biopsy (10x10x5mm)	Sakura Finetek, Staufen, Germany
FACS-tubes	Becton Dickinson, Heidelberg, Germany
Gauge needle Microlance™ 3	Millipore, Schwalbach, Germany
Gloves	Sempermed, Vienna, Austria
Microlance 3/23G, 3/24G	Becton Dickinson, Heidelberg, Germany
Neubauer chamber	Sondheim, Rhön, Germany
“Super Frost” slides	Nunc, Wiesbaden, Germany
Parafilm	American National Can, Menasha, USA
Quadriperm	Sarstedt, Nümbrecht, Germany

Products	Company
Reagent reservoirs	Costar, New York, USA
Scalpels	Feather/ PFM, Cologne, Germany
Syringes	Braun, Melsungen, Germany
Transwell chambers (8 μ m)	Corning, Berdorf, US

2.1.8 List of software applications

Software	Company
Adobe illustrator CC 2018	Adobe company, USA
EndNote	Thomson ResearchSoft, Stanford, USA
BD Cell Quest Pro Version 5.2.1	Becton Dickinson, Heidelberg, Germany
GraphPad Prism 5	GraphPad Software, Inc., San Diego, CA, USA
Image J	Wayne Rasband, Bethesda, USA
LightCycler® 480 SW 1.5	Roche, Mannheim, Germany
Microsoft Office (Word, PowerPoint, Excel)	Microsoft, Redmond, WA, USA
Photoshop CS6	Adobe Systems Inc., San Jose, USA
SPSS	SPSSInc., Chicago, USA

2.2 Methods

2.2.1 Cell culture

2.2.1.2 Cell culture medium

Standard culture medium was prepared as Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10 % fetal bovine serum (FBS) and 1 % penicillin/streptomycin. Selection medium is based on standard culture medium supplemented with 60 μ M 6-thioguanine (6-TG). 4T1 cells were cultured in standard cell culture medium without 6-TG. 4T1-derived cell lines from primary tumors, CTCs, DTCs, and metastases were initially selected in 6-TG-containing medium and thereafter maintained in standard culture medium.

2.2.1.2 Cell passaging

Cells were washed twice by using each 10 mL of PBS buffer after removal of the cell culture medium. Then, 3 mL trypsin solution was added to let cells detach from the culture flask. After a 3 to 20 min incubation in trypsin solution, trypsin was neutralized by mixing with 3 times the volume of culture medium containing FBS. The cell suspension was split as required and varying in ratio from 1:6 to 1:10.

Several DTC1-derived CTC lines grew in a semi-adherent way. Therefore, cell culture medium from these cell lines was collected and centrifuged for collecting cell pellets of poorly adherent cells. Thereafter, poorly adherent cells were suspended in 3 mL of trypsin solution and were put back to flask. Then, poorly adherent cells together with adherent cells were split as mentioned above.

All cell lines were cultured in an incubator at a 5 % CO₂ atmosphere at 37 °C.

2.2.2 Flow cytometry

In the present study, flow cytometry was applied for testing the expression of transmembrane proteins (EpCAM and CD45) at the cell surface.

General procedure: Cells were collected and washed two times in FACS buffer (PBS containing 3% FBS). Cells were suspended in 50 μ L FACS buffer, next mixed with 1 μ L EpCAM- or CD45-specific antibody and incubated for 15 min. After centrifuging and washing twice in 500 μ L FACS buffer, cells were suspended in 50 μ L FACS buffer and incubated with 1 μ L fluorescein isothiocyanate (FITC)-

conjugated secondary antibody for 15 min. Then, cells were washed twice before resuspending in 500 μ L FACS buffer with propidium iodide (PI, 1 mg/mL). Finally, the cell suspension was analyzed in a FACS Calibur cytometer to measure the expression of EpCAM or CD45. The control group was performed using a secondary antibody only.

2.2.3 Proliferation assay

Cells were harvested from growing cultures at a confluency of 50 - 80 %. Cell numbers were measured by a trypan blue exclusion assay. To do so, an aliquot of 10 μ L of cell suspension was taken from each cell sample and mixed with 10 μ L trypan blue solution (0.4%), from which 10 μ L were counted in a Neubauer chamber. Finally, 5,000 cells were seeded per well in a 6-well plate format in $n = 10$ repeats for each experimental cell line. Cell numbers were counted from day 1 to day 5 in duplicates for each day. For semi-adherent cell lines, the used media were collected and centrifuged to include non-adherent cells.

2.2.4 Cell metabolism

To assess the metabolic activity of cells, 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay was used. MTT is a yellow tetrazole, which can be metabolized to purple formazan in living cells (Stockert et al., 2018). Initially, the trypan blue assay was used to obtain absolute cell numbers, and 1,000 cells were plated in triplicates in $3 \times$ wells of 96-well plates per cell line. On day 1, day 3, and day 5, MTT solution (5 mg/mL) was mixed with cells to reach a concentration of 0.5 mg/mL, was incubated for 4 h at 37 °C, and thereafter 200 μ L MTT solvent (0.1 N HCl in isopropanol) were added to dissolve formazan crystals. The dissolved purple formazan was measured in a microplate reader at an optical density (OD) of 570 nm wavelength and 690 nm (as a reference) (Mosmann, 1983).

2.2.5 Single cell deposition for single cell-derived clones

Cell numbers of 4T1, CTC1, and DTC1 were measured by trypan blue exclusion assay. Thereafter, cell suspensions were diluted to 150 cells in 30 mL of culture medium. Cell suspensions were seeded in 96-well plates in a volume of 100 μ L/well, corresponding to 0.5 cell/well. After 7 - 10 days, cell

colonies were observed under a microscope. Colonies growing to high confluence were transferred independently to 6-well plates and further expanded in culture flask.

2.2.6 Transwell invasion assay

The transwell invasion assay was conducted with Matrigel-coated transwell inserts. To prepare the gel layer, 200 μL of 0.9 mg/mL Matrigel in DMEM was coated on 24-well plate inserts with 8 μm pores. Cell numbers were assessed and 1×10^5 cells were seeded in 200 μL serum-free DMEM into the upper chamber. The lower chamber of the invasion chamber was filled with 800 μL standard cell culture medium containing 10% FBS. Thereafter, cells were incubated at 37 °C for 16 h to allow migration and invasion. Following this incubation, cells on the surface of the upper chamber were flushed and washed with PBS. Next, migrated cells on the insert membrane were fixed and stained with 1 % crystal violet/ 70 % methanol solution for 20 min. Afterwards, insert membranes were washed twice in PBS and air dried. Pictures of the stained inserts were taken before membranes were transferred to 96-well plates. Thereafter, 200 μL acetic acid were added to dissolve crystal violet-stained cells. The OD at 590 nm was assessed in a microplate reader.

2.2.7 Adhesion assay

Calcein-AM is a fluorescent dye, which can be used for the short-term labeling of living cells. The fluorescence intensity of Calcein-AM is proportional to the amount of living cells. To prepare the bottom layer, 96-well plates with flat bottom were coated with 50 μL of Gelatin (0.2 % in PBS), Matrigel (100 $\mu\text{L}/\text{mL}$ in PBS), or were left uncoated. Cells were harvested and 5×10^5 cells were incubated with Calcein-AM (10 μM) for 30 min at 37 °C. Thereafter, cells were washed three times with PBS and suspended in 1 mL culture medium. Next, 2.5×10^4 labeled cells in 50 μL culture medium were seeded in triplicates in 96-wells with Gelatin, Matrigel, or no coating, and incubated in the dark for 2 h at 37 °C for cell-matrix adhesion. Next, media were removed and 96-well plates were carefully washed twice with PBS. Adherent cells were lysed by adding 100 μL Triton X-100 (2 % in distilled water). In the input control wells, cells were directly lysed by adding 1 μL Triton X-100 solution without removing media. The Calcein-AM fluorescence was measured in a Victor Wallac instrument.

2.2.8 Endothelial adhesion assay

To prepare the endothelial cell monolayer, 1×10^5 murine endothelial bEnd.3 cells in 100 μL of culture medium were plated in 96-well plates and cultured for 24 - 48 h to form an endothelial cell layer. Before the addition of tumor cells, endothelial cells were activated by adding TNF- α (10 ng/mL) for 5 h. After the removal of the old medium containing TNF- α , 2.5×10^4 Calcein-AM-labeled cells in 50 μL medium were plated per well and incubated for 2h to allow cells to adhere. Further measurements of adherent cells were performed as described in the adhesion assay protocol (Paragraph 2.2.7).

2.2.9 Chemoresistance assay

In order to test for chemoresistance, cells were cultured in medium with Doxorubicin or Cisplatin, and MTT assay was conducted to measure the cell viability after defined time periods. Briefly, 5,000 cells/well were seeded 8 times in triplicates in 96-well plates and incubated for 24 h. Next, cells were cultured in 100 μL medium containing 120, 60, 30, 15, 7.5, 3.75, or 1.875 μM Doxorubicin or Cisplatin for 48 h, and the control group was kept untreated. After treatment with chemotherapeutic agents, MTT solution was added to reach a final concentration of 0.5 mg/mL and incubated with cells for 4 h at 37 $^{\circ}\text{C}$. Then, the MTT solvent was added. The OD at 570 nm wavelength and 690 nm (as a reference) was measured in a microplate reader. Cell viability curves were calculated at different treatment concentrations. The IC_{50} (the drug concentration that can induce 50 % death of treated cells) value was calculated by using GraphPad Prism 5. Cell viability was calculated according to the following equation:

$$\text{Cell viability} = \frac{\text{mean(OD of treated cells - OD of blank)}}{\text{mean(OD of control - OD of blank)}}$$

2.2.10 2D colony formation assay

Cell numbers were assessed by trypan blue assay and diluted to lower concentration ($\leq 1 \times 10^4$ cells/mL). Thereafter, 50, 100, and 200 cells were suspended in 10 mL culture medium and plated in culture dishes (diameter of 10 cm). After incubation for 11 days, media were removed, and dishes were washed twice with PBS. Colonies were stained with 1 % crystal violet/70 % methanol solution for 20 min. Then, plates were washed with PBS twice and air-dried. Numbers of colonies were counted manually,

where clusters containing ≥ 30 cells were defined as a cell colony. The sizes of colonies were assessed by image J software. Plating efficiency was calculated by the following equation:

Plating efficiency = no. of colonies formed / no. of cells seeded x 100%.

2.2.11 3D soft agar colony formation assay

Firstly, 5 % and 3.5 % low melting point (LMP) agarose (weight (g)/volume (mL)) solution were prepared by dissolving 5 g or 3.5 g LMP agarose in 100 mL PBS, followed by autoclaving for 2 h. Before the experiment, the agarose solution was liquefied by heating and kept in a 40 °C water bath. The agarose bottom layer was generated by mixing 1 mL of 5 % agarose with 9 mL culture medium to reach a final concentration of 0.5 % agarose/medium solution. The 0.5 % agarose/medium solutions were quickly plated into culture dishes and solidified at room temperature for 30 min. Afterwards, 1×10^4 cells were resuspended in 9 mL of culture medium and quickly mixed well with 1 mL 3.5 % LMP agarose, then plated into 0.5 % agarose/medium solution coated plates. After 7 days, colony numbers were observed and counted under a microscope. Cell clusters containing more than 30 cells were considered as a colony. The sizes of the colonies were calculated by Image J software.

2.2.12 mRNA isolation

The RNeasy Kit containing QiaShredder columns was applied to isolate mRNA from cells according to the manufacturer's protocol (<http://www.bea.ki.se/documents/EN-RNeasy%20handbook.pdf>). Briefly, cells were harvested and lysed in RLT buffer (supplied in the RNeasy Kit). To homogenize the lysate, the cell lysate was pipetted to a QIAshredder spin column, put in a 2 mL tube (supplied in the RNeasy Kit), and centrifuged for 2 min at high speed ($> 8,000g$). Then, 1 volume of 70 % ethanol was added to the homogenized lysate and were mixed thoroughly. Afterwards, 700 μ L of the sample were added to an RNeasy spin column and were put in a 2 mL tube. After centrifugation for 15 sec at $\geq 8,000$ g, the flow-through was discarded. Next, 700 μ L of RW1 solution were added to the RNeasy spin column and centrifuge for 15 sec at $\geq 8,000$ g. The flow-through was discarded. To wash the spin column membrane, 500 μ L of RPE buffer were added and centrifuged for 15 sec at $\geq 8,000$ g. This washing step was repeated twice. Then, the RNeasy spin column was put in a new 1.5 mL RNase-free tube. To elute RNA, 30 - 50 μ L RNase-free water was added into the spin column membrane. After

incubating for 2 min, the RNeasy spin column was centrifuged for 1 min at $\geq 8,000$ g. The flow-through contained the extracted mRNA and was stored at -20°C for further use.

2.2.13 Reverse transcription polymerase chain reaction (RT-PCR)

Extracted mRNA was first reverse transcribed to cDNA, and cDNA was applied for qRT-PCR measurements to test gene expression levels. The QuantiTect® Reverse Transcription kit was used for reverse transcription. The reaction process was conducted following the manufacturer's protocol (QuantiTect® Reverse Transcription Handbook - Qiagen). First, the concentration of mRNA sample was measured a GeneQuantPro spectrophotometer and 1 μg mRNA was used to prepare the Reaction Mix 1 (table below). The Reaction Mix 1 was incubated for 2 min at 42°C , then immediately put on ice.

Reaction Mix 1 (Genomic DNA elimination reaction components):

Component	Volume/reaction	Final concentration
gDNA Wipeout buffer, 7x	2 μL	1x
Template RNA	1 μg	
RNase-free water	Variable	
Total volume	14 μL	-

Next, the Reaction Mix 1 was used to prepare the Reaction Mix 2 according to the table below, and was incubated for 20 - 30 min at 42°C .

Reaction Mix 2 (Reverse-transcription reaction components):

Component	Volume/reaction	Final concentration
Quantiscript reverse transcriptase	1 μL	
Quantiscript RT Buffer, 5x	4 μL	1x
RT Primer Mix	1 μL	

Template RNA (Reaction Mix 1)	14 μ L	
Total volume	20 μ L	–

To inactivate the Quantiscript reverse transcriptase, the Reaction Mix 2 was incubated for 3 min at 95°C, then immediately put on ice. Afterwards, 20 μ L cDNA samples were diluted 1:10 with 180 μ L ddH₂O and short-term stored at -20 °C.

2.2.14 Quantitative Real-Time PCR (qRT-PCR)

cDNA samples were applied for qRT-PCR to study the expression of genes of interest. SYBR Green PCR Master Mix kit was used to quantify gene expression.

Standard master-mix (per reaction) was prepared as below:

Component	Volume/reaction
cDNA template	1 μ L
Primer mix	2 μ L
SYBR Green master-mix (2x)	5 μ L
ddH ₂ O	2 μ L
Total	10 μ L

Primer mix: a mix of forward and backward primers for each gene of interest (each primer 1 μ L corresponding to 5 μ M).

Gene expression was tested in the Light Cycler 480 System (Roche, Mannheim, Germany). The reaction setup was as below

Procedure	Time	Temperature
Initial denaturation	10 min,	95°C
Denaturation	30 sec	95°C
Annealing and elongation	60 sec	72°C
45 repeats		back to denaturation step
Cooling/Storage	∞	4°C

Gene expression level of glucuronidase-beta (Gusb) was used as a reference. Gene expression levels were calculated according to the equation $2^{-\Delta\Delta CT}$;

$$\Delta CT = CT_{\text{gene of interest}} - CT_{\text{endogenous control}}$$

$$\Delta\Delta CT = \Delta CT_{\text{gene of interest}} - \Delta CT_{\text{reference}}$$

ΔCT values were used for statistical comparison.

2.2.15 Animal experiments

2.2.15.1 Ethical statements

Mouse experiments were approved by the local Ethical Committee-Regierung von Oberbayern, Munich, Germany (Az 55.2.1.54-2532-90/12 and 177/15) at the laboratory of PD Dr. Sebastian Kobold, Division of Clinical Pharmacology, Ludwig-Maximilians-University, Munich, Germany.

2.2.15.2 Tumorigenicity assay

Experimental animals (female BALB/c mice) were generally between 6 - 8 weeks old. Cells numbers for transplantations were calculated by trypan blue assay, then 1.25×10^6 cells were suspended in 1 mL of PBS. Cell suspensions were stored a short time on ice. Then, 100 μL PBS/cell suspension (1.25×10^5 cells) were subcutaneously injected into the flank of the mice. The sizes of the tumors were assessed every 2 - 3 days. Mice were sacrificed after 27 days. To do so, mice were anesthetized by inhaling 0.4 % isoflurane, then blood was collected from the orbital sinus, and femurs and tibiae were collected. Primary tumors were weighted. Primary tumors and organs, including lung, spleen, kidney, and liver, were collected in order to cryopreserve samples for IHC staining and to establish *ex vivo* sublines.

In the second round of mouse experiment, 1.25×10^6 4T1 and DTC1 cells, and 5×10^6 , 1×10^7 , and 2×10^7 CTC1 cells were resuspended in 1 mL PBS. Then, 100 μL of cells in PBS were subcutaneously transplanted into the flank of BALB/c mice. Tumor size was measured every 1 - 2 days. After 15 days, mice were sacrificed, primary tumors were weighted. Lungs were harvested and a short time kept in culture medium on ice to further perform a metastatic colony formation assay.

2.2.15.3 *Ex vivo* establishment of CTCs lines

To establish stable CTC lines *ex vivo*, blood was collected from BALB/c mice at the end of the tumorigenicity experiments. Blood volumes varied between 0.8 mL and 1 mL per mouse. Blood was collected in heparin covered sterile tubes and stored for short time on ice. Blood samples were first eliminated RBCs by adding ammonium chloride-based lysis reagent. After 2 min incubation time, cells were centrifuged and washed twice with PBS. Next, cell pellets (containing white blood cells and potential CTCs) were resuspended in 10 mL selection medium containing 60 μ M 6-TG, and were plated in 2-fold serial dilutions in 96-well plates with a starting volume of 100 μ L. Cell colonies were observed under a microscope and were expanded individually to 12-well plates and later to 6-well plates. Finally, selected cells were established as *ex vivo* CTC lines.

2.2.15.4 *Ex vivo* establishment of DTCs lines

To establish DTC lines *ex vivo*, mice were sacrificed, and femurs and tibiae were collected. After removing skin and muscles, the joints of hips and knees were cut out from femurs and tibiae, and bone marrows were flushed out with PBS by using a syringe. The flushed-out cells were centrifuged and supplemented with 10 mL of ammonium chloride-based lysing reagent to deplete the cell suspension of RBCs. After washing with PBS, cell pellets were resuspended in 10 mL selection medium containing 60 μ M of 6-TG, and cell lines were generated as described in the previous paragraph for CTCs.

2.2.15.5 *Ex vivo* establishment of primary tumor and metastasis cell lines

To establish tumors and metastases cell lines *ex vivo*, primary tumors and organs (lung, spleen, kidney and liver) were collected and short-time stored in tubes with culture medium on ice. Primary tumors and organs were cut into small pieces in 10 cm² dishes, then minced through a 100 μ m filter. After washing twice with PBS and centrifuging, cell pellets containing tissue cells and tumor cells were suspended in 10 mL selection medium. Each tumor or organ cell suspension was plated independently in a 10 cm² culture dish. Growing cells were observed under a microscope. When selected cells grew to confluence, they were transferred to flasks for further maintenance.

For all established cell lines *in vitro*, passage numbers less than 5 were used for re-injection *in vivo*, and passage numbers below 10 were used for performing functional assays.

2.2.15.6 Intravenous (*i.v*) injection and lung metastasis formation

To study the capability of tumor cells to form lung metastasis, cells were intravenously injected in BALB/c mice (age-matched 6 - 8 weeks) through the tail vein. Brief procedure: cells numbers were assessed by trypan blue assay, then 5×10^5 cells were suspended in 1 mL of PBS. 100 μ L of cells in PBS (5×10^4 cells) were injected into the tail vein. Mice were observed every 1 - 2 days and sacrificed at day 19. Superficial metastatic nodules in the lungs were assessed by visual inspection and lungs were harvested to further perform a metastasis colony formation assay. Alternatively, 5×10^4 cells in 100 μ L of PBS were injected into the tail vein of BALB/c mice. After injection, mice were checked daily for signs of endpoints to the experiment in each group. The defined endpoints were: (1) substantial weight loss (more than 5% in more than two mice out of five), (2) weakness, *i.e.* tiredness and unresponsiveness (more than 2 mice out of five), (3) ≤ 2 remaining experimental groups. Mice weights were measured every 1 - 2 days to calculate weight-curves. For each group, mice were anesthetized before being sacrificed, and autopsy was performed to study the metastatic status of each mouse. Metastatic nodules in lungs were assessed by visual inspection and lungs were collected for metastasis colony formation assay.

2.2.15.7 Metastasis colony formation assay

To quantify lung metastases, the metastasis colony formation assay was conducted. Entire lungs were collected after autopsy, and were cut into small pieces and incubated in RPMI medium containing collagenase (5 mg/mL) and DNase (1 mg/mL) for 30 min. Next, lung fragments were minced through a 100 μ m filter. Cells remaining in the filter were rinsed with 5 mL of PBS. Tissue fragments were further minced through a 40 μ m filter and rinsed with 5 mL of PBS. Cell suspensions were centrifuged and cell pellets were incubated with ammonium chloride-based lysing reagent for 2 min to deplete the cell suspension of RBCs. After centrifugation and washing, cells were suspended in 10 mL of selection medium (60 μ M 6-TG), and 1 mL of cells was taken to dilute to 1:100 in selection medium.

Sequentially, 3 mL of each concentration were seeded in 6-well plates in triplicates. After 10 days for subcutaneous and after 4 days for *i.v* transplantation of tumor cells, metastatic colonies were stained with 1% crystal violet/70 % methanol solution for 20 min. Then, plates were washed with PBS twice and air dried. Numbers of colonies were counted manually, where clusters containing more than 20 cells were considered as a colony.

2.2.15.8 Immunocytochemistry and immunohistochemistry staining

Sample preparation: to perform immunohistochemistry staining, tumors or organs were collected and preserved in cryomolds. Then, tissues were covered with Tissue Tek gel avoiding the formation of air bubbles, and carefully frozen in liquid nitrogen. Frozen tissue samples were cut to serial sections of 3 - 5 μm thickness and mounted on glass slides. For immunocytochemistry, cells were seeded on glass slices, covered with culture medium, and incubated overnight to allow cells to adhere. Immunostaining was done by our technician Gisela Kranz and was conducted according to the avidin-biotin-peroxidase method (Vectastain, Vector laboratories, Burlingame, CA, USA) (<https://docs.abcam.com/pdf/protocols/ihc-immunostaining.pdf>). Briefly, the sample slides were gently washed twice for 5 min in Tris-Buffered Saline (TBS) plus 0.025% Triton X-100 to reduce surface tension. Then, samples were blocked in 10% normal serum with 1% Bovine Serum Albumin (BSA) in TBS for 2 h. Slides were air-dried for several seconds and tissue paper was used to wipe away excess of TBS buffer. Primary antibodies (EpCAM, E-Cadherin, Vimentin, Cytokeratin) were diluted in TBS with 1% BSA (concentrations depend on manufacturer's recommendations, normally, 0.5 - 10 $\mu\text{g}/\text{mL}$) and incubated with sample slides overnight at 4°C. Afterward, slides were rinsed twice for 5 min in TBS with 0.025% Triton. The secondary antibody was diluted in TBS with 1% BSA (concentrations depend on manufacturer's recommendations, normally, 0.5 - 5 $\mu\text{g}/\text{mL}$) and was incubated with sample slides for 1 h at room temperature. After staining with chromogen for 10 min, slides were rinsed in tap water for 5 min. After dehydration, clear and mount process, slides could be observed under the microscope.

Immunohistochemistry intensity scores (IHC scores) were calculated as the product of protein expression intensity (0 to 3; 0, negative; 1, weak; 2, moderate; 3, strong) and the percentage of

expressing cells per area (score 0 = 0 ~ 5 %, 1 = 5 ~ 25 %, 2 = 25 ~ 50 %, 3 = 50 ~ 75%, 4 = 75 ~ 100 %).

2.2.15.9 EMT scoring

EMT score was applied to compare the level of mesenchymal transition of cells. It was defined as the product of the percentage of mesenchymal-, spindle-like cells (0 - 100%) and the level of cell-cell contact (disseminated cells represent 0 ~ 25% - 1, 25 ~ 50% - 2, 50 ~ 75% - 3, 75 ~ 100% - 4). Thereby, an EMT score of 0 represents cells in a highly epithelial state, whereas a score of 400 represents strongly mesenchymal cells.

2.2.16 Clinical study

The clinical study was performed in cooperation with the laboratory of Prof. Hongxia Wang, Department of Oncology, Shanghai General Hospital, Shanghai, China. The clinical study was approved by the Ethics Committee of Shanghai General Hospital (ethics #2018KY153). The clinical study was conducted obeying the Declaration of Helsinki Principles.

2.2.16.1 Clinical cohort

Breast cancer patients, who were at tumor stage III - IV and received inpatient treatment, were included in the clinical study with informed written consent. From September 2015 to April 2017, 34 metastatic breast cancer patients were enrolled at Shanghai General Hospital. Patients' information was anonymized and de-identified prior to analysis. Patients were followed from September 2015 to January 2018.

2.2.16.2 Subtraction enrichment of CTC and DTC

Enrichment of CTC and DTC was conducted following the manufacturers' instructions (Cytelligen, San Diego, CA, USA) (Lin et al., 2017) at Shanghai General Hospital by our cooperation partner Dr. Junjiang Li. Blood and bone marrow samples were collected before patients received any treatment. To avoid contamination of epithelial cells, the first 2 mL of blood or bone marrow sample were discarded. For each patient, 6 mL of blood and 3 mL of bone marrow samples were collected at the same time. Blood and bone marrow samples were first centrifuged at 600 g for 5 min. Cell pellets were depleted RBC by going through a non-hematopoietic cell separation matrix. Then, the remaining cells were incubated with anti-WBC immunomagnetic beads. After centrifugation at 400 g

for 5 min, supernatants were collected. The supernatants containing rare cells were centrifuged at 500 g for 2 min. Cell pellets were carefully collected and fixed in CTC slides (Cytelligen) for further staining.

2.2.16.3 Identification of CTC and DTC by SE-iFISH

The experiment was conducted at Shanghai General Hospital by our cooperation partner Dr. Junjiang Li and Dr. Peter Lin from Cytelligen (San Diego, USA). Centromere Probe (CEP8), anti-CD45, anti-EpCAM antibodies were used for cell staining. Cell nuclei were stained with DAPI. CTCs and DTCs were defined as cells with a DAPI-positive nucleus, a lack of CD45 expression, heteroploidy at chromosome 8, with or without EpCAM expression, or diploid CEP8 signals with EpCAM expression. Cell clusters were defined as visualizing more than two cells in cell-cell contact.

2.3 Statistical analysis

Statistical analysis was performed by computer software, including Prism5, Microsoft Excel, and SPSS. The Kolmogorov–Smirnov test was used to define whether data sets obeyed normal distribution. If data sets had a normal distribution, unpaired or paired t-test was used for comparing between unpaired or paired two groups, one-way ANOVA and post hoc Bonferroni t-test for multiple comparisons. If data sets did not obey a normal distribution, Mann–Whitney U test was used to compare the difference between two independent groups, the Wilcoxon signed-rank test was used to compare two related samples, and Kruskal-Wallis test with post hoc Dunn’s test was applied for comparing multiple groups. The correlation between EMT score and EpCAM expression was analyzed by Spearman’s rank test. Receiver operating characteristic (ROC) curves assessed the capability of the EpCAM-positivity rate of cells to predict patients’ outcomes. Patients’ overall survival was analyzed by the Kaplan–Meier curve. Results were shown as mean values \pm standard deviation (SD.) from at least three independent experiments. P-values < 0.05 were considered significant.

Statistics shown in figures within this thesis are: ns: not significant, * < 0.05 , ** < 0.01 , *** < 0.001 .

3. RESULTS

In this study, the aim was to analyze the contribution(s) of an EMT to the metastatic potential of MBC cells. To do so, a syngeneic mouse model of MBC was used to recapitulate all stages of the cancer progression, from the formation of primary tumors to the growth of distant metastases in the lungs.

3.1 The 4T1 MBC mouse model

4T1 cells are transplantable murine breast cancer cells that were first generated by Fred Miller and colleagues (Miller et al., 1981; Miller et al., 1987). After syngeneic transplantation into immunocompetent BALB/c mice, 4T1 cells form primary tumors and induce distant metastases at multiple sites (*i.e.* lung, liver, kidney, bones). 4T1 cells have been selected for their resistance to 6-TG, a guanine analogue that inhibits the syntheses of DNA and RNA, and thus, eventually causes cell death. This inherent property of 4T1 cells to proliferate in the presence of 6-TG allows for the *ex vivo* isolation of 4T1-derived tumor cells by selecting cells in medium supplemented with 6-TG. In order to establish *ex vivo* cultures of CTC and DTC lines, 4T1 cells were subcutaneously transplanted to the flank of BALB/c mice (age-matched between 6 - 8 weeks old). After 3 - 4 weeks, with the generation of the primary tumors, mice were sacrificed, and blood, bones, and organs were collected for selections of CTC, DTC, and metastasis-derived cell lines (**Figure 1**).

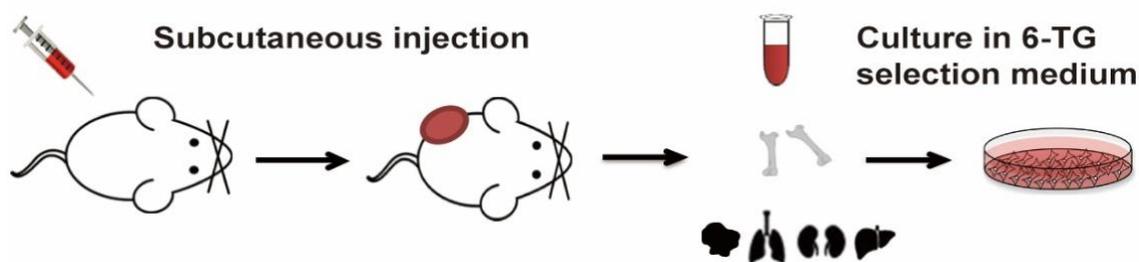


Figure 1: Scheme of the 4T1 mouse model: 4T1 cells were subcutaneously transplanted in the flank of BALB/c mice.

After 3 - 4 weeks, mice were sacrificed, blood, bone, primary tumors, and organs were harvested for the isolation of 4T1-derived sublines from each localization.

In the first round of animal experiments, $n = 5$ BALB/c mice were transplanted with 4T1 cells. All mice developed primary tumors, and, hence, blood and bones were used for the isolation of CTCs and DTCs. As a result, one CTC line and one DTC cell line were successfully generated as *ex vivo* cultures from blood and bone marrow, namely CTC1 and DTC1, respectively (**Figure 2**). The culture procedure

was described in detail in the Methods section 2.2.15. CTC1 and DTC1 cells grew in an adherent manner and an obvious difference of EMT phenotypes was observed between 4T1, CTC1, and DTC1 cells. Parental 4T1 cells show a classical epithelial phenotype with a cobblestone-like morphology and tight cell-cell junctions (Epithelial, E-type), whereas CTC1 cells displayed a typical mesenchymal phenotype, in which cells are spindle-shaped and have lost cell-cell contact (Mesenchymal, M-type). DTC1 cells showed a partial EMT compared to 4T1 cells, and hence the phenotype of DTC1 cells was defined as an E/m-type, reflecting that cells have gone through a restricted EMT and still maintained a major epithelial type.

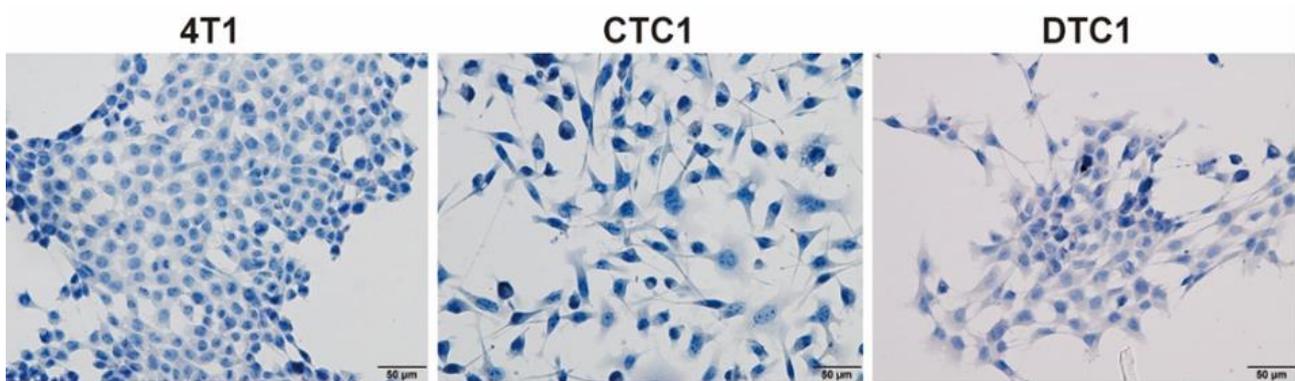


Figure 2: EMT phenotypes of 4T1, CTC1, and DTC1 cells *in vitro*. Hematoxylin and eosin staining of 4T1, CTC1, and DTC1 cells. Shown are representative photographs of the morphology of each cell line grown under standard cell culture conditions. Scale bars are indicated in each picture.

To assess a potential contamination of the generated cell lines with white blood cells, the leukocyte marker CD45 was measured by flow cytometry in 4T1 cells (as a negative control), CTC1, and DTC1 cells (**Figure 3**). The result confirmed that CD45 was not expressed in any of these three cell lines.

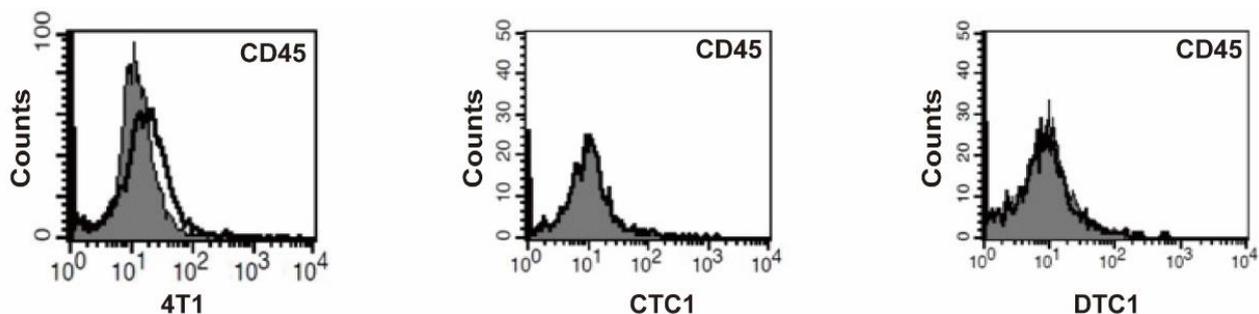


Figure 3: CD45 expression in 4T1, CTC1, and DTC1 cells was tested by flow cytometry. Shown are representative histograms of CD45 expression in each cell line from $n = 3$ independent experiments. CD45 antibodies (black) and isotype controls (grey).

Morphological changes in CTC1 and DTC1 cells could be caused by culture condition, *e.g.* long-term 6-TG treatment might have an influence on 4T1 cell morphology. To test this eventuality, 4T1 cells were cultured in medium containing 6-TG for 28 days and the morphology of the cells was recorded every 7 days as micrographs. The result demonstrated that 6-TG *per se* did not induce an EMT in 4T1 cells (**Figure 4**).

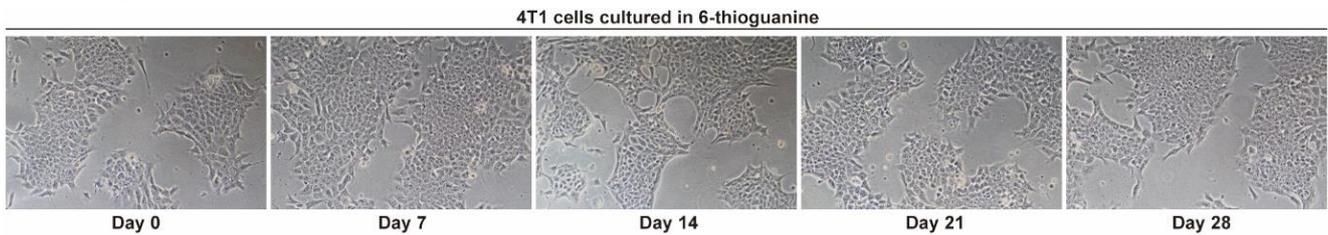


Figure 4: 6-TG treatment of 4T1 cells. 4T1 cells were cultured in medium containing 6-TG for 28 days and cell morphology was assessed at the indicated time points. Representative pictures taken at each indicated time point are shown.

EMT is usually accompanied by switching the expression of EMT-related genes. Therefore, immunocytochemistry staining was conducted to study the expression of epithelial and mesenchymal markers. **Figure 5** shows that 4T1, CTC1, and DTC1 cells expressed substantial amounts of epithelial marker Cytokeratin and mesenchymal marker Vimentin. In addition, the epithelial markers EpCAM and E-cadherin were highly expressed in 4T1 and DTC1 cells, but were entirely lacking in CTC1 cells.

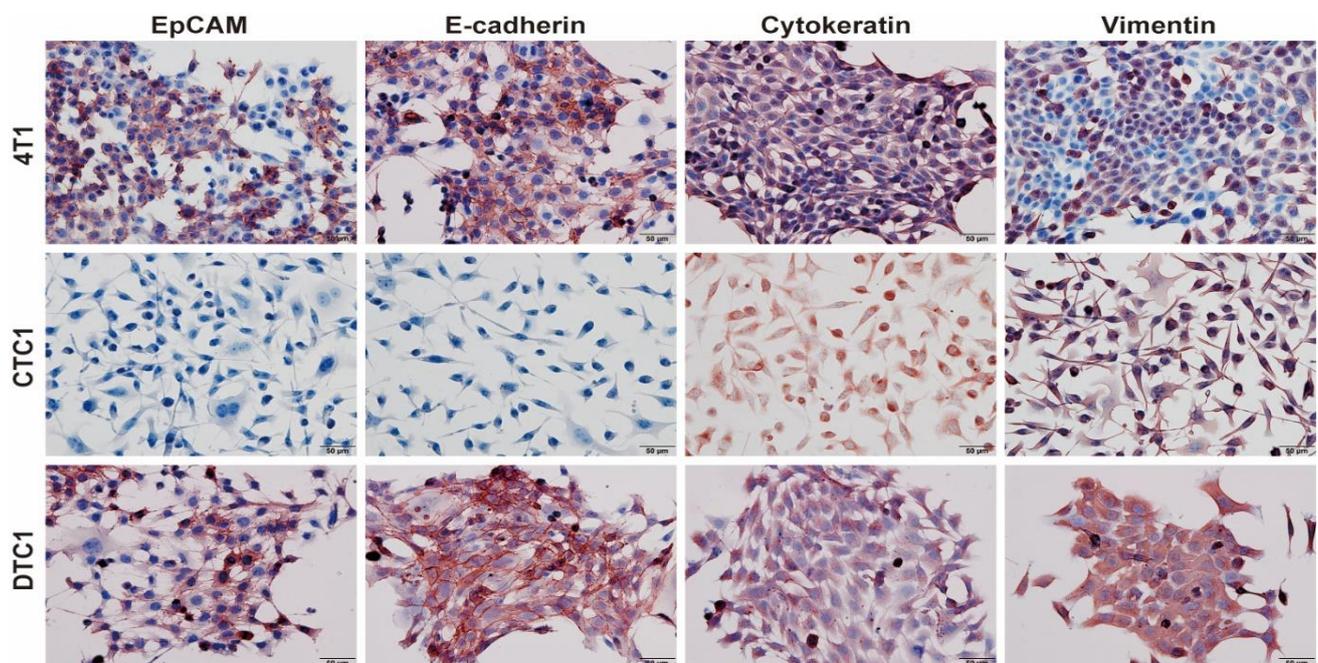


Figure 5: Immunocytochemistry staining of EpCAM, E-cadherin, Cytokeratin, and Vimentin in 4T1, CTC1, and DTC1 cell lines. Representative pictures from $n = 3$ independent experiments are shown with scale bars.

EpCAM is an epithelial marker that is widely applied for the isolation and detection of CTCs. We tested the cell surface expression of EpCAM among 4T1, CTC1, and DTC1 cells by flow cytometry. Parental 4T1 cells expressed EpCAM to a high level, while EpCAM was entirely negative in CTC1 cells, and DTC1 contained a mix of EpCAM^{high} cells, EpCAM^{intermediate} cells, and a small proportion of EpCAM^{negative} cells (**Figure 6**). Interestingly, two distinct populations of EpCAM-positive cells with differing expression levels could be visualized as high and intermediate in DTC1 cells. This result confirmed the findings of the immunocytochemistry staining of EpCAM.

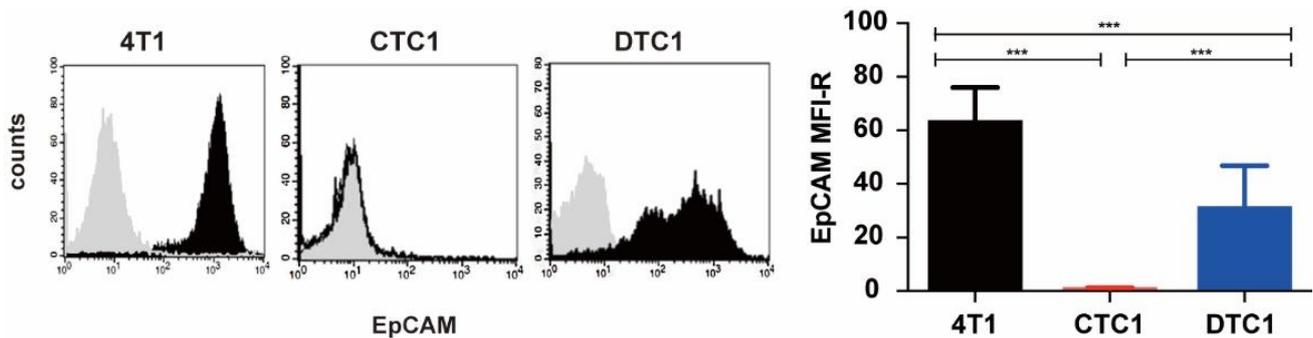


Figure 6: Flow cytometry analysis of EpCAM expression in 4T1, CTC, and DTC1 cells. The left panels show representative histograms of EpCAM expression (EpCAM in black, controls in grey). Quantification of EpCAM expression on 4T1, CTC1, and DTC1 cell lines are shown as means \pm SD from $n \geq 5$ independent experiments in the right panel.

Next, mRNA levels of epithelial markers EpCAM, E-cadherin, Rab25, Ddr1, Grhl2, and Krt19, and of mesenchymal markers N-cadherin, Vimentin, Slug, Zeb1/2, Erbb2/3, Snail, and Twist were measured in 4T1, CTC1, and DTC1 cells by quantitative RT-PCR. As shown in **Figure 7**, a significant reduction of the expression of epithelial markers EpCAM, E-cadherin, Rab25, and Grhl2 and an increase of the expression of mesenchymal markers N-cadherin, Vimentin, Slug, ZEB1, ZEB2 was measured in CTC1 cells. Furthermore, no significant difference was found for the expression of Ddr1, ErbB2, ErbB3 between 4T1, CTC1, and DTC1 cells (**Figure 7**). Krt19 was unexpectedly highly expressed, and EMT-TF Snail and Twist were expressed to lower levels in CTC1 and DTC1 cells compared to 4T1 cells (**Figure 7**).

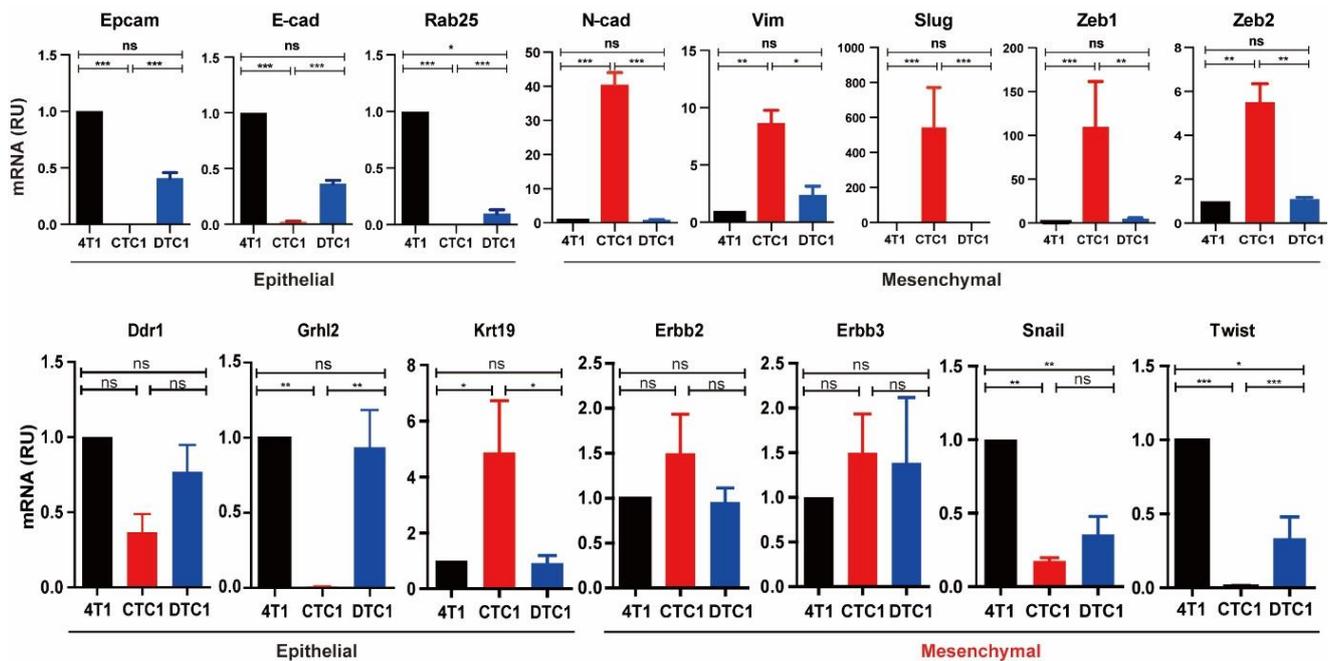


Figure 7: mRNA expression levels of epithelial and EMT markers in 4T1, CTC1, and DTC1 cells. mRNA expression levels of the indicated genes were assessed by qRT-PCR in 4T1, CTC1, and DTC1 cells. Gene expression of Gusp was applied as a reference to normalize all samples. Bar charts show results with means \pm SD from $n = 3$ independent experiments. Normalized gene expression of 4T1 cells was set to 1 as reference.

3.2 EMT in CTC1 cells is accompanied by increased migration, but reduced proliferation and tumor formation ability

To understand the influence of EMT on the behaviour and functional capacities of tumor cells, *in vitro* functional assays and *in vivo* tumorigenic studies were conducted. First, proliferation and cell metabolism were analyzed by cell counting and MTT assay, respectively. Results from the MTT assay demonstrated that 4T1 cells represented the highest cell metabolism, followed by DTC1 and CTC1 cells (**Figure 8A**). Cell numbers were assessed by cell counting over a total time period of 5 days. Results of cell counting were similar to the cell metabolism assay, where 4T1 had the highest cell numbers and CTC1 had the lowest cell counts, while DTC1 cells showed an intermediate cell count (**Figure 8B**).

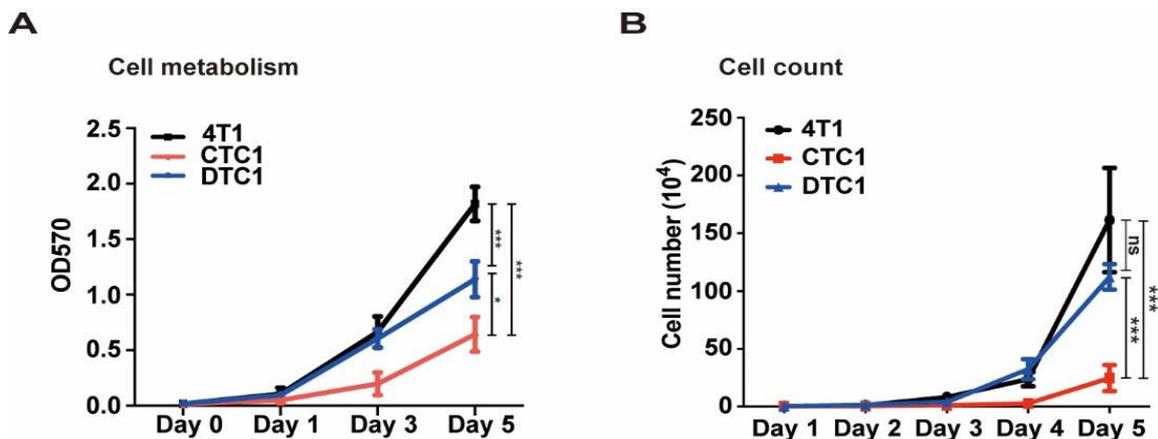


Figure 8: Comparison of cell metabolism and proliferation between 4T1, CTC1, and DTC1 cells. (A) Cell metabolism of 4T1, CTC1, DTC1 cells was measured by MTT assay (initial seeding number, 1,000 cells). (B) The proliferation of 4T1, CTC1, DTC1 was tested by cell counting (initial seeding number, 5,000 cells). Line charts show means \pm SD from $n \geq 3$ independent experiments.

The two-dimensional (2D) colony formation assay essentially assesses every single cell for its capacity to undergo “unlimited” division. The three-dimensional (3D) colony formation assay or soft agar colony formation assay tests the cells’ ability to grow in an anchorage-independent condition. In 2D colony formation assays, 4T1, CTC1, and DTC1 cells had a similar capacity in anchorage-dependent “unlimited” cell growth (**Figure 9**). When observing the cell colonies under the microscope, the morphology of colonies showed substantial differences among these cell lines. CTC1 cells formed the biggest cell colonies with loose cell-cell contact, while 4T1 cells generated the smallest colonies with tight cell-cell connections. DTC1 cells showed intermediate colony size and relatively loose cell-cell connection compared to 4T1 cells. To compare the size of cell colonies, pictures were taken and the size of colonies was assessed by the Image J software. The results confirmed that the average colony size of CTC1 cells was $11.07 \pm 2.68 \text{ mm}^2$ and it was the highest among these cell lines, followed by DTC1 cells ($7.17 \pm 3.16 \text{ mm}^2$) and 4T1 cells ($5.45 \pm 2.09 \text{ mm}^2$) (**Figure 9**, upper right panel). The colony size of CTC1 cells was significantly larger than the parental 4T1 cells, while no significant difference was observed compared to DTC1 cells (**Figure 9**, upper right panel).

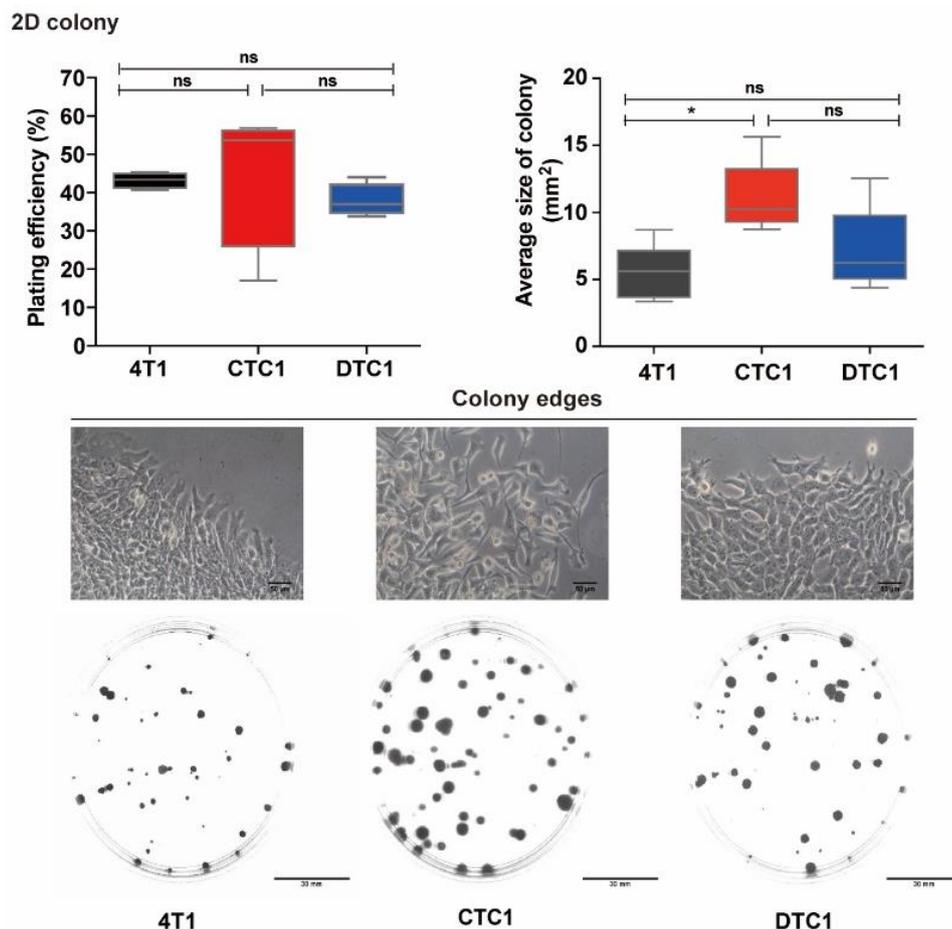


Figure 9: 2D colony formation capacity of 4T1, CTC1, and DTC1 cells. Upper left panel: Shown is the plating efficiency of 4T1, CTC1, and DTC1 cells from $n = 4$ independent experiments in box-plot whiskers graphs with means \pm SD. Upper right panel: Shown is the colony size as calculated using the Image J software. Box-plot whiskers graphs show means \pm SD from $n = 4$ independent experiments. Lower panels: Shown are representative images of 2D colony edges and crystal violet-stained colonies from 4T1, CTC1, and DTC1 cells (initial seeding number of 200 cells) from $n = 4$ independent experiments.

In the 3D colony formation assay, CTC1 cells showed highest ability in anchorage-independent cell growth (**Figure 10A**). To compare the size of 3D-colonies, 12 colonies per cell line were randomly chosen and the size of area for each colony was measured by the Image J software. The result reflected that 4T1 cells formed the smallest 3D colonies with clearly defined edges. CTC1 cells generated the largest colonies with loose edges, while DTC1 cells generated intermediately sized colonies and their colony edges had both sharply and loosely defined areas (**Figure 10B**).

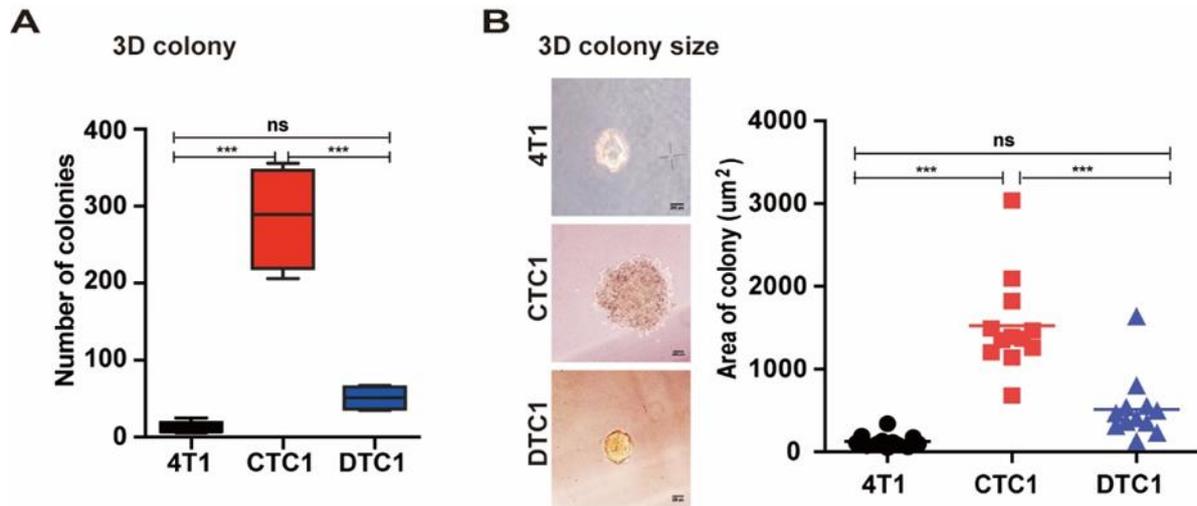


Figure 10: 3D colony formation capacity of 4T1, CTC1, and DTC1 cells. (A) Box-plot whiskers graphs show numbers of colonies from 4T1, CTC1, and DTC1 cells as means \pm SD from $n = 4$ independent experiment. (B) Shown are representative images of colonies from 4T1, CTC1, and DTC1 cells (left), and quantification of 3D colony size in dot plots (right) from $n = 12$ randomly selected colonies.

Adhesion is an important trait of tumor cells in order to intravasation and extravasation from blood vessels. To test the adhesion of 4T1, CTC1, and DTC1 cells, an adhesion assay was performed. Murine endothelial cells (bEnd.3), Matrigel, and Gelatin were used as adhesion matrix. The result demonstrated that in comparison with 4T1 and DTC1 cells, CTC1 cells were characterized by significantly reduced adhesion to Matrigel and Gelatin (**Figure 11**). In addition, DTC1 cells showed the highest adhesion to endothelial cells, 1.38-fold higher than 4T1 and 2.01-fold higher than CTC1 cells (**Figure 11**).

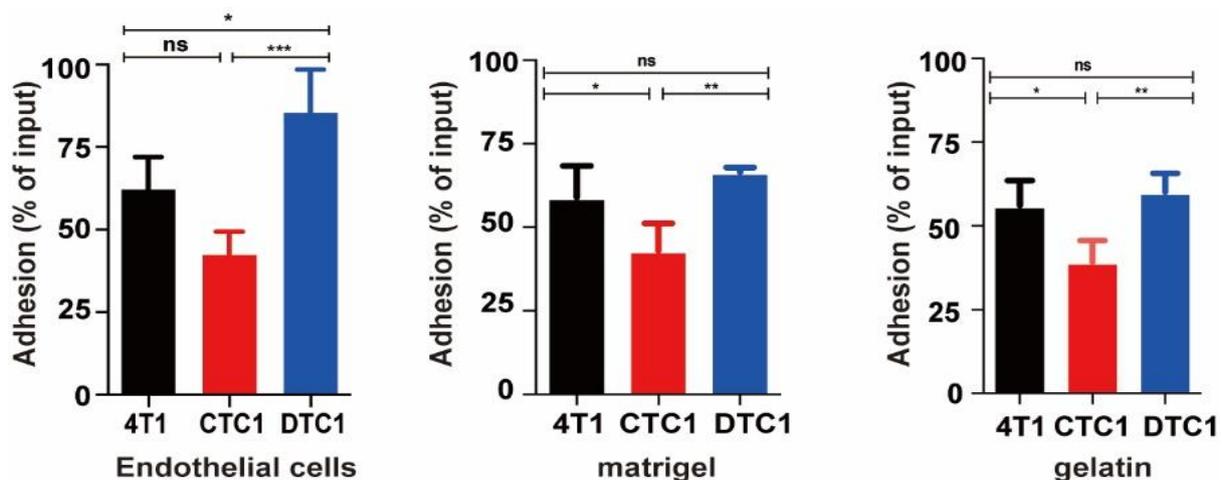


Figure 11: Adhesion of 4T1, CTC1, and DTC1 cells to bEnd.3 endothelial cells, Matrigel, and Gelatin. Shown are means \pm SD adhesion rates of 4T1, CTC1, and DTC1 cells from $n \geq 3$ independent experiments.

Invasion is an important characteristic of tumor cells that is critical for the process of metastasis formation. A Boyden chamber invasion assay with Matrigel-coated inlays was performed to measure the migration ability of 4T1, CTC1, and DTC1 cells. The result of the invasion assay showed that CTC1 and DTC1 cells possessed similarly high invasive capacity on average, while 4T1 cells were characterized by the lowest invasion into Matrigel (**Figure 12**).

Invasion

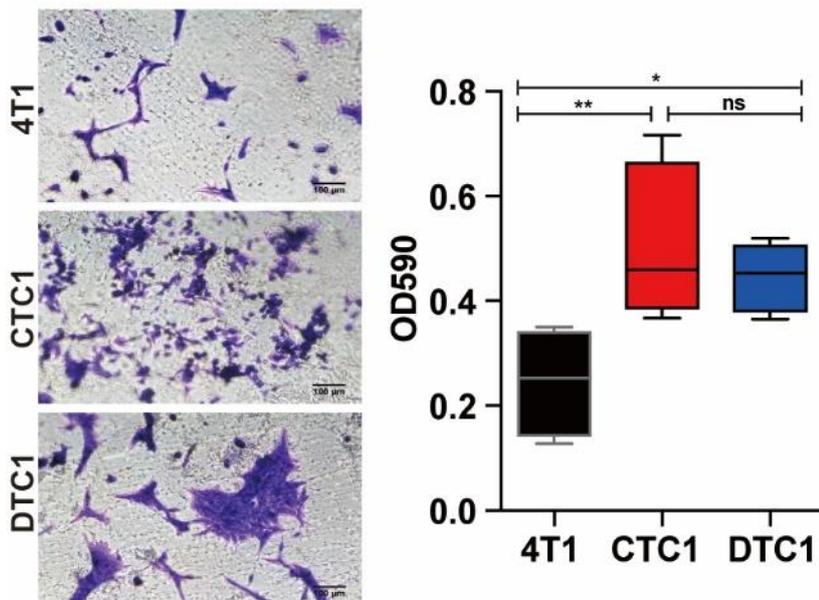


Figure 12: Invasion capacity of 4T1, CTC1, and DTC1 cells. Left panel shows representative images of invasive 4T1, CTC1, and DTC1 cells stained with methylene blue. Right panel: Box-plot whiskers graphs show the quantification of invaded cells as OD₅₉₀ from $n \geq 3$ independent experiments.

As a short summary of the *in vitro* functional assays, one can deduce that EMT in CTC1 cells was reflected by a reduction of adhesion and proliferation, and by an increase in anchorage-independent invasion and growth. DTC1 cells with a partial EMT phenotype had generally improved capacities, with retained proliferation, increased adhesion, especially to endothelial cells, enhanced migration, invasion, and anchorage-independent cell growth.

With the aim to compare the tumorigenic ability of 4T1, CTC1, and DTC1 cells and to obtain additional *ex vivo* CTC or DTC lines to study their EMT phenotypes, identical cell numbers were subcutaneously transplanted into the flank of BALB/c mice. After three weeks, mice were sacrificed and tumor weights were quantified. In this experiment, primary tumors, blood, bones, and organs were collected for the *ex vivo* selection of the primary tumor, CTC, DTC, and metastasis cell lines (See Methods 2.3).

All 4T1- and DTC1-transplanted mice generated primary tumors (13/13 and 8/8, respectively). However, only 7 of 17 CTC1-injected mice generated primary tumors (41.2% frequency) (**Figure 13A**). With respect to tumor weights, the DTC1-injected group of mice displayed the highest tumor

weights and also showed the biggest tumor sizes, which were determined with a Caliper measurement at the indicated time points *in vivo* (**Figure 13A and B**). The average tumor weight and size of the CTC1-injected group of mice were significantly reduced compared to the 4T1- and DTC1-injected groups of mice (**Figure 13A and B**).

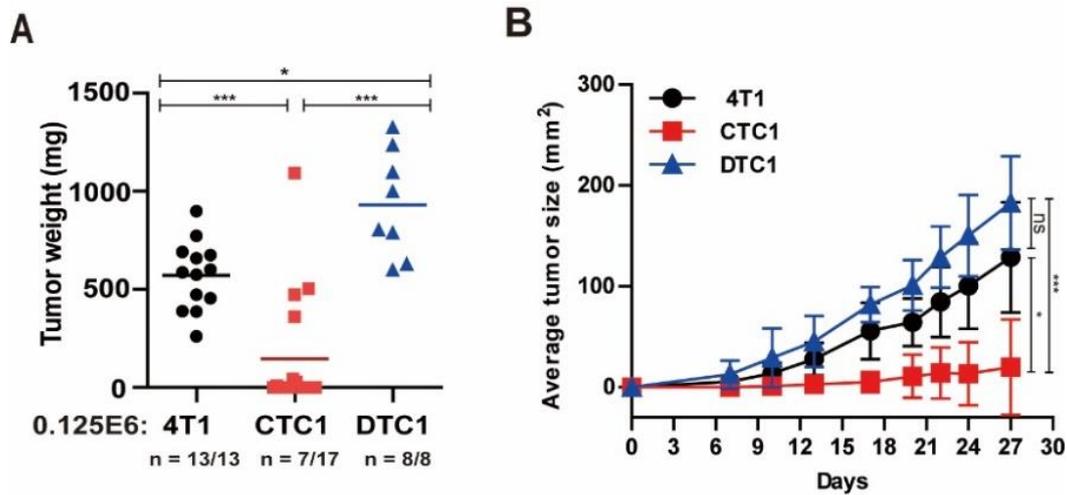


Figure 13: Tumorigenic ability of 4T1, CTC1, and DTC1 cells *in vivo*. (A) 4T1, CTC1, and DTC1 cell (1.25×10^5 cells) were subcutaneously transplanted in BALB/c mice. Shown are tumor weights with mean (horizontal line) for each group at the end of the experiment, including numbers of transplanted mice and numbers of tumor-bearing mice. (B) Line chart shows the tumor growth curves for 4T1-, CTC1-, and DTC1-injected groups of mice. Shown are means \pm SD.

IHC staining of epithelial marker EpCAM and mesenchymal marker Vimentin were performed in 4T1, CTC1, and DTC1 primary tumors. The result demonstrated that CTC1-derived primary tumors remained EpCAM-negative, similar to their parental CTC1 cells, whereas 4T1- and DTC1-derived primary tumors expressed EpCAM to a high level (**Figure 14**). 4T1-, CTC1-, DTC1-derived primary tumors had similar expression levels of Vimentin (**Figure 14**). This result provided proof that CTC1 cells could generate primary tumors without re-expressing the epithelial marker EpCAM, however with a reduced penetrance of 42%. Furthermore, the CTC1-derived primary tumors- and metastases-derived *ex vivo* cell lines retained a mesenchymal phenotype similar to their parental CTC1 cells.

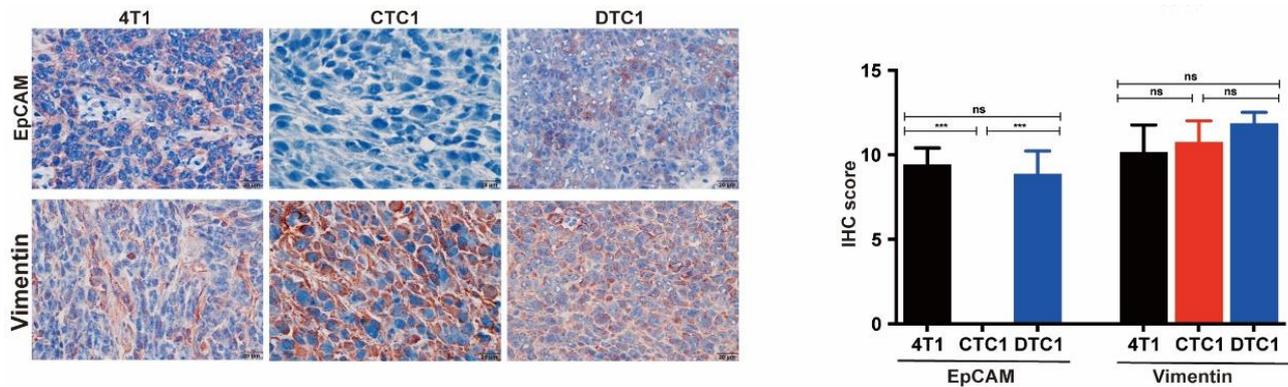


Figure 14: Immunohistochemistry (IHC) staining of EpCAM and Vimentin in primary tumors of 4T1, CTC1, and DTC1 cells-injected groups of mice. Left panels show representative images of IHC staining of EpCAM and Vimentin of 4T1, CTC1, and DTC1 primary tumors. IHC scores were quantified (see Methods 2.3.8) and are shown as bar charts of means ± SD in the right panel.

Incidences of *ex vivo* cultures retrieved from primary tumors, organs, blood, and bone marrows are shown in **Figure 15A and B**. Briefly, CTC1 and DTC1 were the only cell lines that were established from blood and bone marrow from 4T1-injected mice, and CTC1-transplanted mice group failed to establish any CTC or DTC cell line. In contrast, we could establish 26 CTC and 10 DTC lines in 4 of 8 and in 2 out of 8 DTC1-injected mice, respectively. In general, lungs displayed the highest frequency for the retrieval of metastatic cell lines as compared to liver, kidney, and spleen (**Figure 15**).

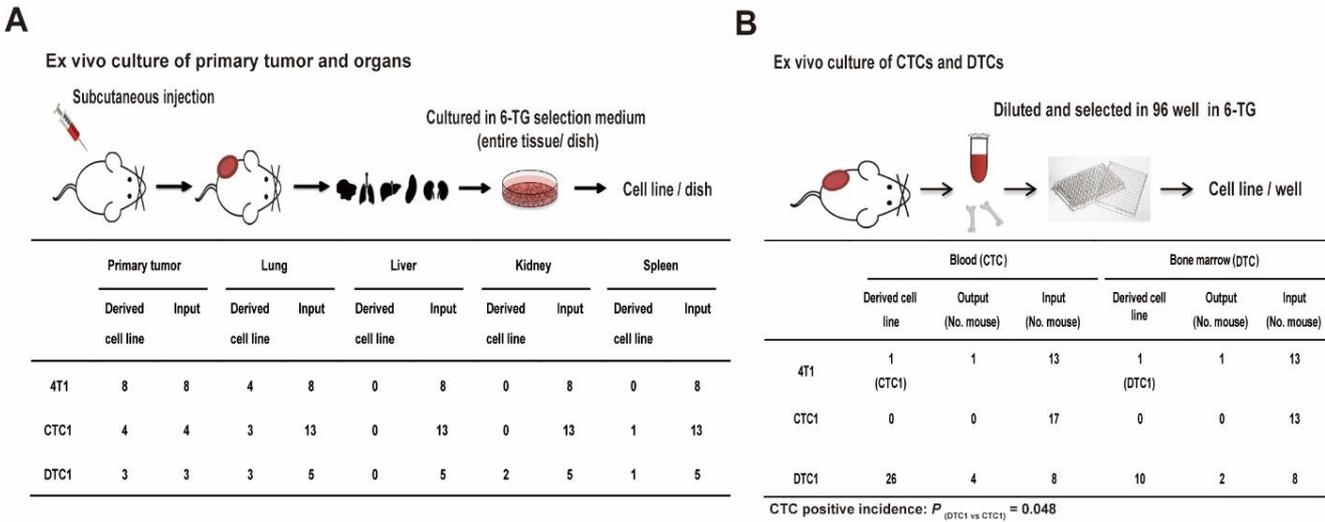


Figure 15: Frequencies of *ex vivo* cultures of 4T1, CTC1, and DTC1 cells following syngeneic transplantation. (A) *Ex vivo* establishment of cell lines from primary tumors and metastatic lesions. Schematic representation of *ex vivo* retrieval of cell lines from primary tumors and metastatic sites (lung, spleen, liver, and kidney). The table shows the frequencies of successfully established cell lines from transplanted mice (input). (B) *Ex vivo* establishment of CTCs and DTCs. Schematic representation of *ex vivo* retrieval of CTC and DTC lines from blood and bone marrow. The table shows the frequencies of successfully established cell lines from transplanted mice (input).

Transplantation of identical cell numbers (1.25×10^5) of 4T1, CTC1, and DTC1 cells resulted in CTC1-injected groups with significantly reduced tumorigenic potential compared to 4T1- and DTC1-injected groups of mice. *In vitro*, CTC1 showed comparably reduced proliferation, which could contribute to the formation of smaller tumors *in vivo*. In order to further explore the tumorigenic and metastatic potential of CTC1 cells, we performed a second round of transplantation experiments with increased numbers of injected of CTC1 cells. To do so, 1.25×10^5 (identical numbers for 4T1 and DTC1 cells), 5×10^5 , 1×10^6 , and 2×10^6 CTC1 cells were subcutaneously transplanted into the flank of BALB/c mice (See Methods 2.3). After 15 days, mice were sacrificed, tumor weights were measured, and lungs were collected for metastasis colony formation assay. With increased numbers of injected cells, 1×10^6 and 2×10^6 CTC1 groups reached similar tumor weights as 4T1- and DTC1-injected groups of mice (1.25×10^5 injected cells) (**Figure 16**). However, the group of DTC1-injected mice with an injection of 1.25×10^5 cells had significant bigger tumors than the group of CTC1-injected mice with an injection of 5×10^5 cells (**Figure 16**).

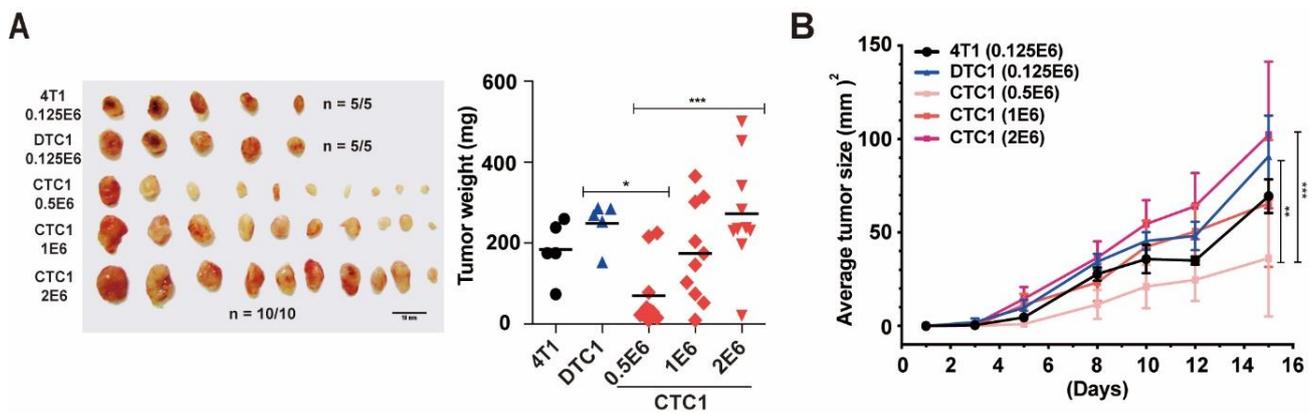


Figure 16: Tumorigenic ability of 4T1, DTC1, and CTC1 cells with increased cell numbers in the CTC1-injected group of mice. (A) 4T1 (1.25×10^5 cells; $n = 5$ mice), CTC1 (5×10^5 cells, 1×10^6 , 2×10^6 cells, $n = 10$ mice per group) and DTC1 cells (1.25×10^5 cells; $n = 5$ mice) were transplanted subcutaneously in BALB/c mice. Shown is a picture of all primary tumors. Dot plot on the right shows tumor weights with means (lines) of each group. (B) The line chart shows tumor growth curves with means \pm SD, which were assessed *in vivo* with a Caliper measurement.

Metastasis formation assay was conducted to compare the metastatic ability of 4T1-, DTC1-, and CTC1-injected groups of mice. Four of 5 mice had generated lung metastases (80 %), both in the 4T1- and in the DTC1-injected group of mice. In the CTC1-injected group of mice, 5×10^5 and 1×10^6 cells generated lung metastases in 3 of 10 mice (30 % frequency), and mice after injection of 2×10^6 cells had lung metastases in 8 of 10 mice (80 % frequency) (**Figure 17A**). The average number of metastatic colonies was highest in DTC1-injected group, followed by 2×10^6 CTC1-injected group, 4T1-injected group, 1×10^6 CTC1-injected group, and the 5×10^5 CTC1-injected group showed the lowest number of the metastatic colonies (**Figure 17B**). The metastatic index per cell was defined as numbers of lung metastatic colonies divided by the numbers of injected cells. The metastatic index was used to compare the metastatic ability across all mice groups. As shown in **Figure 17A**, the metastatic index was highest in DTC1 mice group and lowest in 5×10^5 CTC1 group. In addition, metastatic indexes of CTC1-injected groups were reduced compared to parental 4T1 cells (**Figure 17A**). To sum up, CTC1 cells have significantly decreased tumorigenic and metastatic capability compared to 4T1 and DTC1 cells.

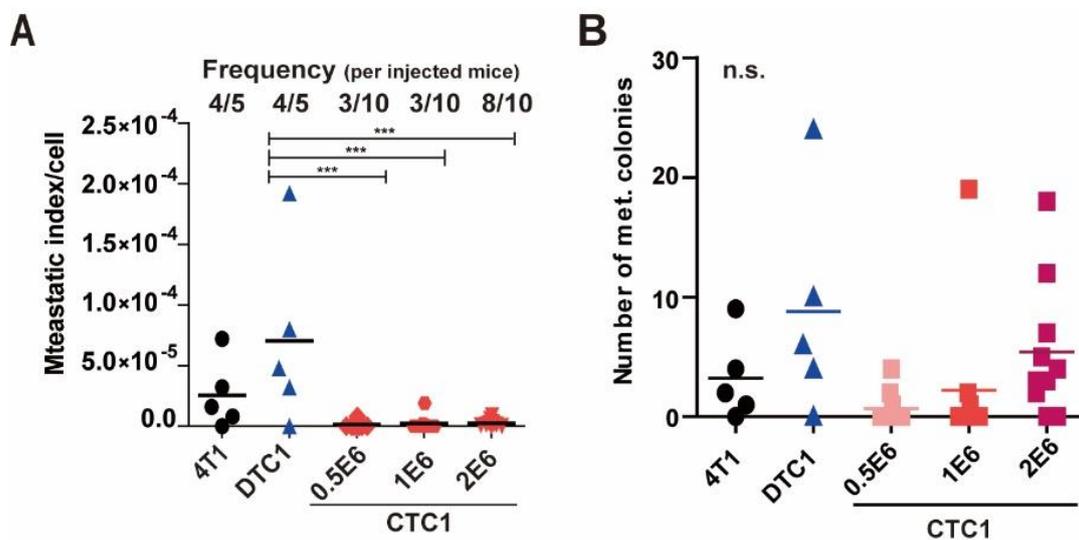


Figure 17: Metastatic potential of 4T1, DTC1, and CTC1 cells. (A) Dot plot graph shows metastatic index/cell with mean values (lines) and frequencies of lung metastasis per mouse. Metastatic index per cell was defined as numbers of lung metastatic colony divided by initially injected cell numbers. (B) Dot plot graph shows numbers of superficial metastatic colonies in each group.

3.3 EMT heterogeneity in DTC1-derived CTC lines

In the present study, blood and bone marrows were collected for the isolation of CTC and DTC lines. In DTC1-injected mice, 26 CTC lines were isolated from 4 of a total 8 mice that had been transplanted. To confirm that the cell lines did not comprise white blood cells, the leukocyte marker CD45 was tested in all these cell lines by flow cytometry. Murine B cell lymphoma WEHI-231 cells represented a positive control for CD45 staining. The results demonstrated that all 26 CTC lines were CD45-negative and thus, distinguished them from white blood cells (**Figure 18**).

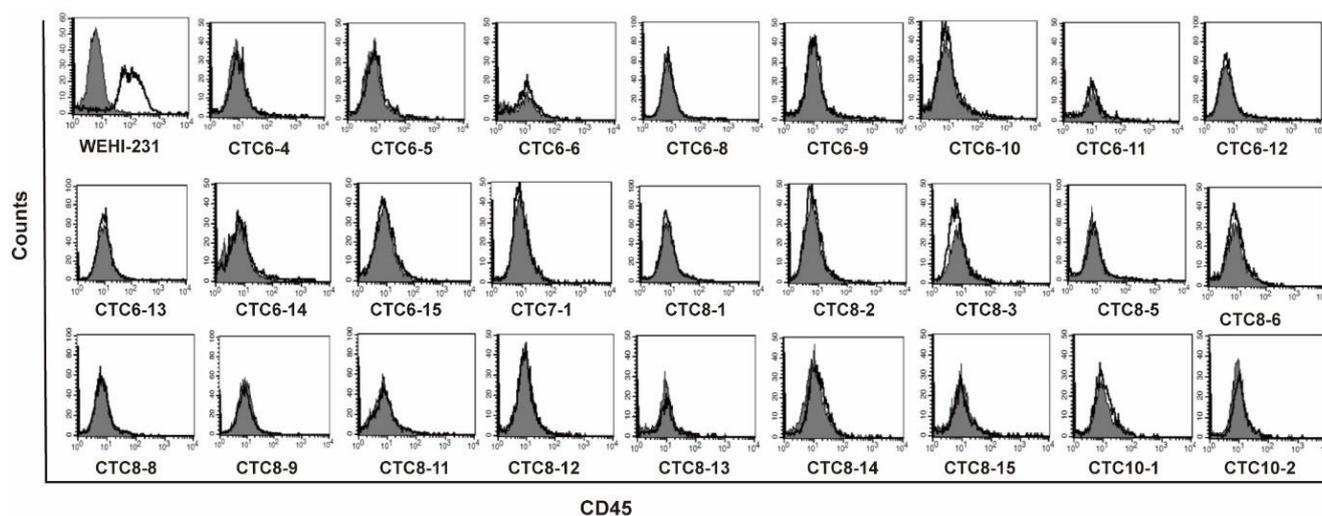


Figure 18: CD45 expression on DTC1-derived CTC lines. CD45 expression was measured in DTC1-derived CTC lines by flow cytometry with specific antibodies. CD45 staining is displayed in black lines, control staining is displayed in grey filled curves. Shown are representative histograms from $n = 3$ independent experiments.

When observed under the microscope, these 26 CTC lines differed considerably with respect to their EMT phenotypes, which enabled us to further study the impact of EMT statuses on cellular functions. In order to classify the degree of EMT in these CTC lines, an EMT scoring system was introduced. This EMT scoring system combined the percentage of mesenchymal, spindle-shaped cells (0 - 100%) and the level of cell-cell contact (1 - 4; see Methods 2.2.15.9), to achieve an EMT score ranging from 0 (epithelial, 4T1) to 400 (mesenchymal, CTC1). **Figure 19B** shows representative pictures of *ex vivo* cultured CTC lines with differing EMT scores to exemplify the phenotypic transition from epithelial to mesenchymal. The quantification of EMT scores is shown in **Figure 19C**, and the result reflected that EMT scores across all CTC lines were evenly distributed. EMT phenotypes of CTC lines varied

in different mice, but different EMT scores were also present in one individual mouse, *i.e.* in mice number 6 and 8 (**Figure 19C**). This demonstrated that CTCs with heterogeneous EMT phenotypes could co-exist in the blood of individual mice.

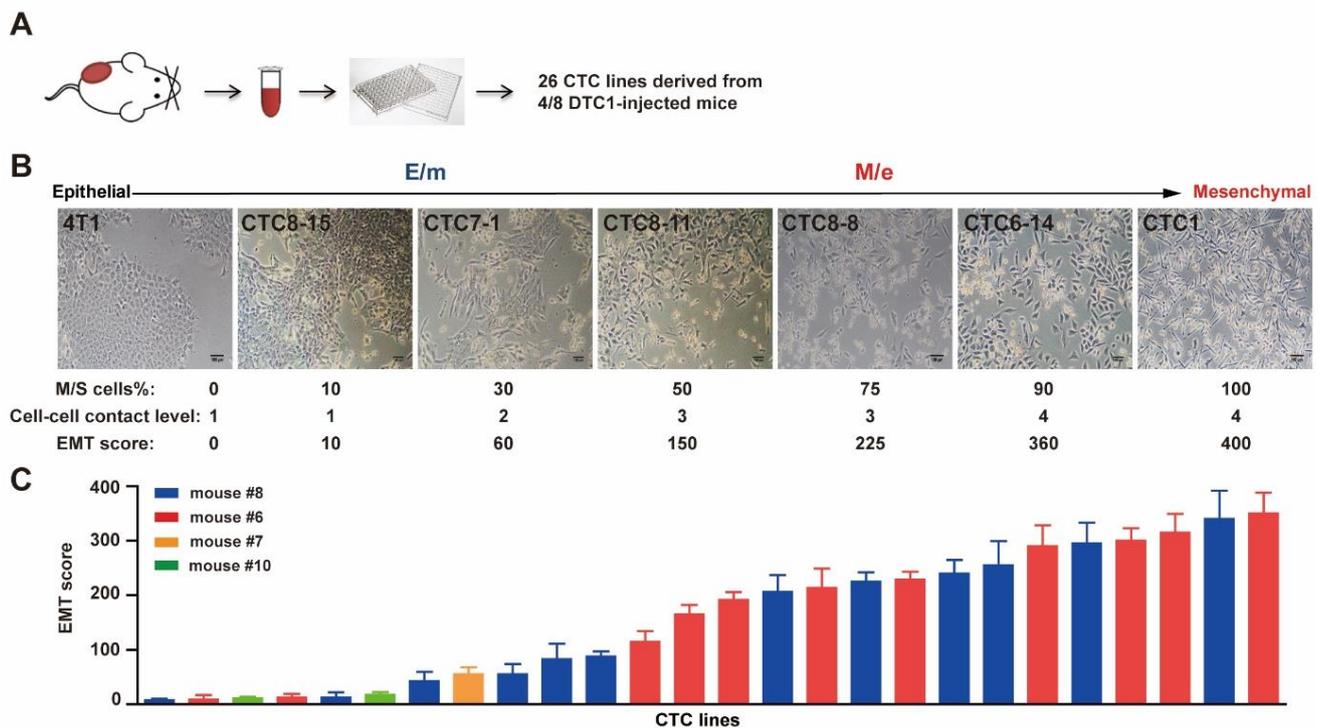


Figure 19: EMT scores of DTC1-derived CTC lines. (A) Scheme of the experimental setup for the isolation of CTC lines. (B) Representative pictures of 4T1 and CTC1 cells, and DTC1-derived CTC lines representing various levels of EMT. The individual EMT scores are indicated in the bottom right of each picture. (C) Shown are mean EMT scores \pm SD of 4T1-, CTC1-, and DTC1-derived CTC lines from $n = 4$ independent measurements. Mice of origin are color-coded as indicated in the bar graph.

Based on the EMT scoring system, EMT phenotypes of CTC lines were divided into E-type (Epithelial, parental 4T1 cells), E/m-type (defined as primarily epithelial with a moderate transition to mesenchymal traits), M/e-type (defined as primarily mesenchymal with a moderate transition to epithelial traits), and M-type (mesenchymal, CTC1 cells). We selected three E/m- (CTC6-6, CTC6-11, CTC8-12) and three M/e- (CTC8-6, CTC8-5, CTC8-1) type DTC1-derived CTC lines to study whether EMT traits were associated with different cellular functions.

First, immunocytochemistry staining was performed to test the protein expression of epithelial markers EpCAM and E-cadherin, and mesenchymal marker Vimentin. In M/e-type CTCs, a substantial reduction of the expression of EpCAM and E-cadherin was observed, however EpCAM and E-cadherin were still expressed in a small proportion of cells. In contrast, E/m-type CTCs showed high

levels of EpCAM and E-cadherin (**Figure 20**). Vimentin expression was highly positive in all selected cell lines (**Figure 20**).

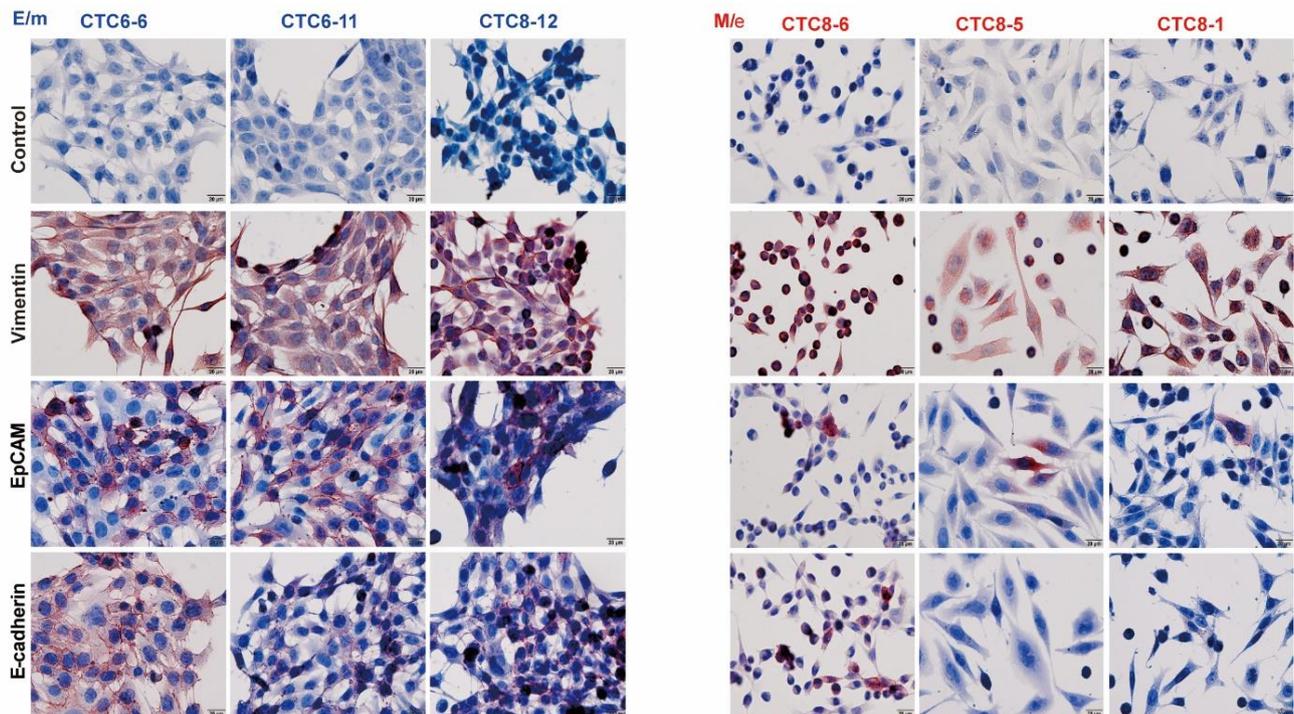


Figure 20: Immunocytochemistry staining of control, Vimentin, EpCAM, and E-cadherin in E/m type (CTC6-6, CTC6-11, CTC8-12) and M/e type (CTC8-6, CTC8-5, CTC8-1) CTCs. Shown are representative pictures from $n = 3$ independent experiment. Antigen staining is displayed in brown, cytoplasm and nuclei are counterstained with hematoxylin and eosin in blue.

To compare the expression of classical EMT genes between E/m-type (CTC6-6, CTC6-11, CTC8-12) and M/e-type (CTC8-6, CTC8-5, CTC8-1) CTCs, quantitative RT-qPCR was conducted. The results reflected that the expression of the epithelial markers EpCAM, E-cadherin, and Rab25 was higher in E/m-type CTCs compared to M/e-type cells (**Figure 21**). mRNA levels of Ddr1, Grhl2, and Krt19 were similar between E/m-type and M/e-type CTCs. EMT-related genes Vimentin, Slug, and Zeb2 were significantly higher in the M/e-type CTCs. No statistical difference was found in expression of N-cad, Zeb1, ErbB2, ErbB3, Snail, and Twist (**Figure 21**).

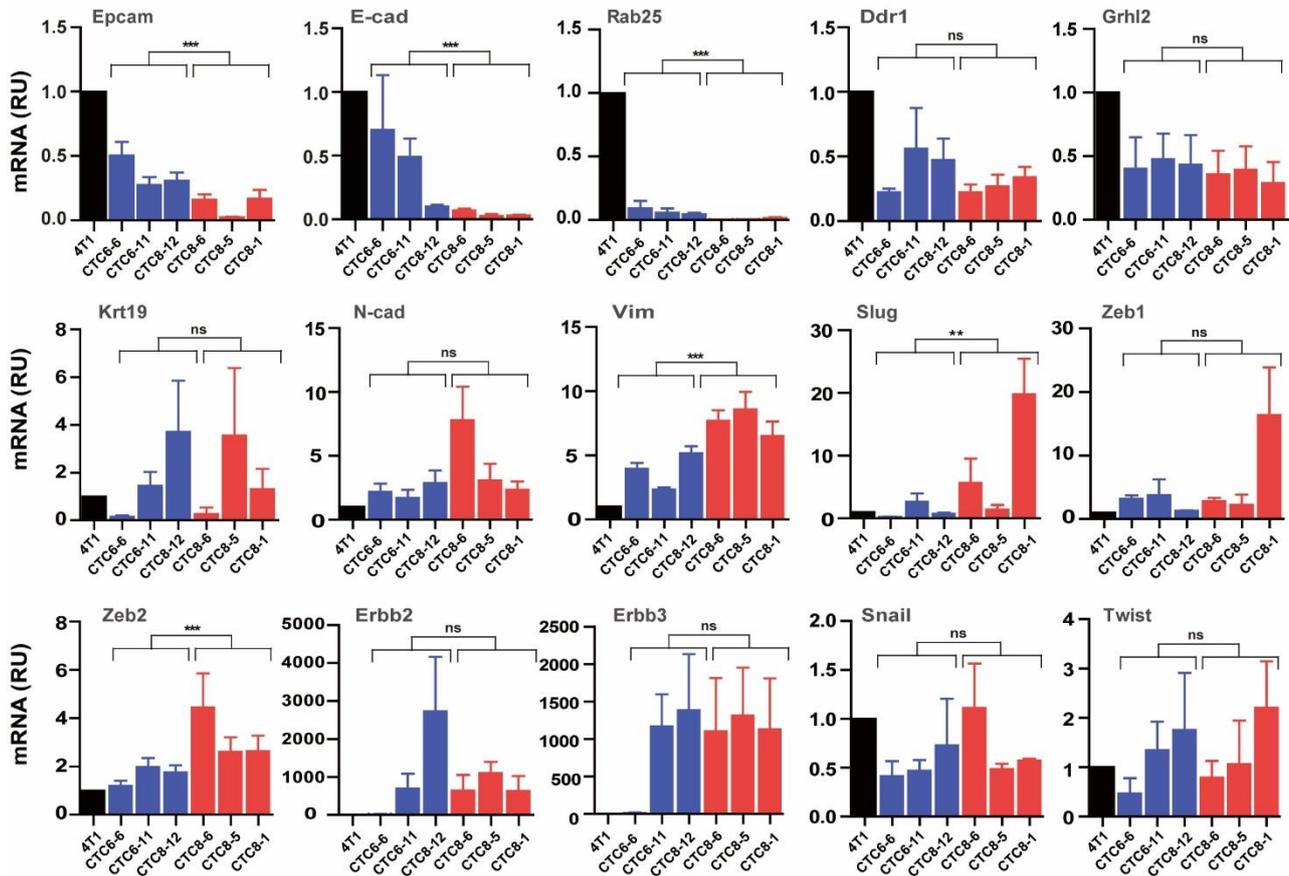


Figure 21: Expression of EMT-related genes in E/m- and M/e-type CTCs. mRNA transcript levels of epithelial markers Epcam, E-cad, Rab25, Ddr1, Grhl2, and Krt19, and mesenchymal marker N-cad, Vimentin, Slug, Zeb1/2, Erbb2/3, Snail, and Twist in E/m-type (CTC6-6, CTC6-11, CTC8-12) and M/e-type (CTC8-6, CTC8-5, CTC8-1) CTCs derived from the blood of DTC1-transplanted mice were assessed by quantitative RT-PCR. Gusp mRNA expression served as house-keeping gene to normalize all samples. 4T1 cells were set to 1 as a reference. Bar charts show means \pm SD from $n = 3$ independent experiments performed in triplicates.

To address the influence of EMT on cellular functions, functional assays were performed *in vitro*. First, metabolism and proliferation were tested by MTT assay and cell counting. The results showed significantly higher cell metabolism and a 1.52-fold higher proliferation rate in the E/m-type group in comparison with the M/e-type group (**Figure 22A and B**). Furthermore, the adhesion and invasion capacities of each cell line were measured. Results disclosed that the adhesion property of E/m-type CTCs to endothelial cells, Matrigel, and Gelatin was significantly higher than that of M/e-type CTCs (**Figure 23A**). Furthermore, the invasion capacity of E/m-type cells was significantly, but only slightly increased (**Figure 23B**).

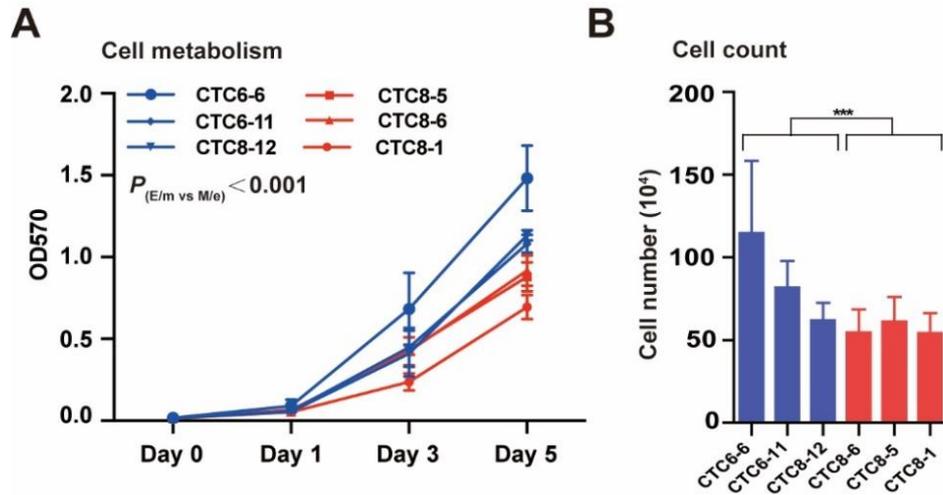


Figure 22: Cell metabolism and proliferation of E/m- and M/e- type DTC1-derived CTC lines. (A) The cell metabolism of E/m-type (CTC6-6, CTC6-11, CTC8-12) and M/e-type (CTC8-6, CTC8-5, CTC8-1) CTCs (initial seeding number 1,000 cells) CTCs was assessed by MTT assay. Line chart shows cell metabolism as OD₅₇₀ as means \pm SD from $n \geq 3$ independent experiments. (B) The proliferation of E/m-type (CTC6-6, CTC6-11, CTC8-12) and M/e-type (CTC8-6, CTC8-5, CTC8-1) CTCs (initial seeding number 5,000 cells) was assessed by cell counting at day 5. Shown are cell numbers as means \pm SD from $n \geq 3$ independent experiments.

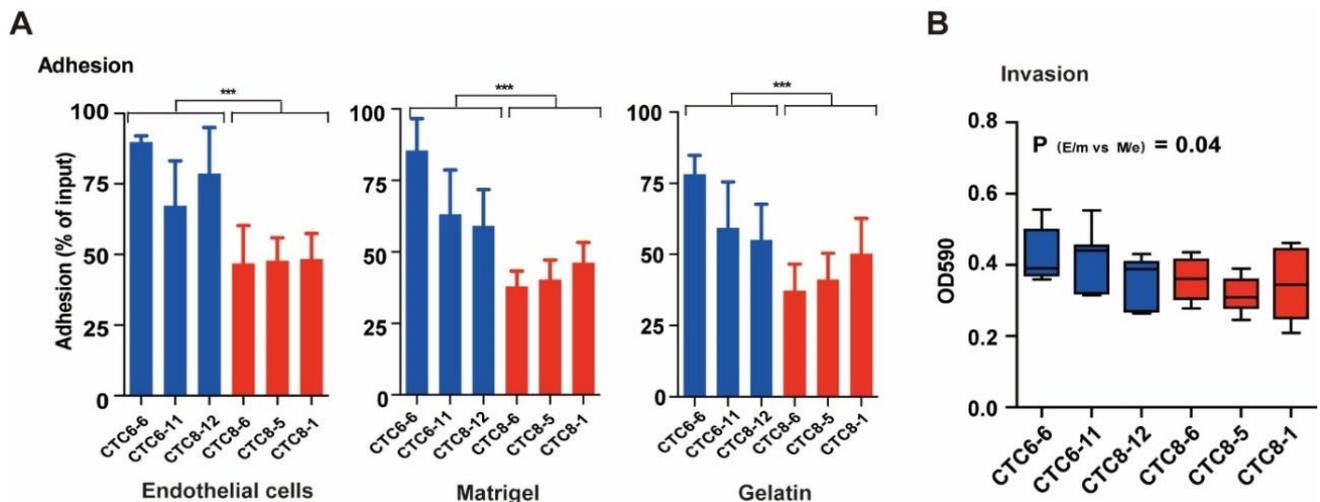


Figure 23: Adhesion and invasion assay of E/m- and M/e- type DTC1-derived CTC lines. (A) Adhesion assay to bEnd.3 endothelial cells, Matrigel and Gelatin was performed with E/m-type (CTC6-6, CTC6-11, CTC8-12) and M/e-type (CTC8-6, CTC8-5, CTC8-1) CTCs. Bar chart shows percentages of adherent cells as means \pm SD from $n \geq 3$ independent experiments. (B) Invasion ability of E/m-type (CTC6-6, CTC6-11, CTC8-12) and M/e-type (CTC8-6, CTC8-5, CTC8-1) CTCs were tested by transwell invasion assay. Box-plot whiskers graph shows OD₅₉₀ as means with ranges from $n \geq 3$ independent experiments.

3.4 Mesenchymal-type CTCs are more resistant to chemotherapy than epithelial-type CTCs

Recent studies reported that EMT in tumor cells was associated with increased chemoresistance (Fischer et al., 2015; Lambert et al., 2017). To test the response to chemotherapeutics in relation to the EMT status of the tested cells, E (4T1), E/m- (CTC6-6, CTC6-11, CTC8-12, DTC1), M/e- (CTC8-6, CTC8-5, CTC8-1), and M-type (CTC1) cells were cultured in medium containing Cisplatin or Doxorubicin at increasing concentrations for 48 h. Metabolic activity, which was an indication of cell viability, was detected by MTT assay. IC_{50} values define the concentration of drug that results in a 50% reduction of cell growth. Mesenchymal-type cells (M-, M/e- type) were characterized by increased chemoresistance in comparison with epithelial-type (E-, E/m- type) cells, both for Cisplatin (IC_{50} mean values: M: $18.81 \pm 3.40 \mu\text{M}$, M/e: $18.12 \pm 7.33 \mu\text{M}$, E: $10.37 \pm 4.15 \mu\text{M}$, E/m: $11.35 \pm 3.16 \mu\text{M}$) and for Doxorubicin (IC_{50} mean values: M: $4.51 \pm 0.30 \mu\text{M}$, M/e: $3.05 \pm 2.04 \mu\text{M}$, E: $0.66 \pm 0.59 \mu\text{M}$, E/m: $0.93 \pm 0.52 \mu\text{M}$) (Figure 24). Hence, M- and M/e- type CTC lines showed increased resistance towards clinical chemotherapeutic drugs, as compared with E- and E/m- type CTCs.

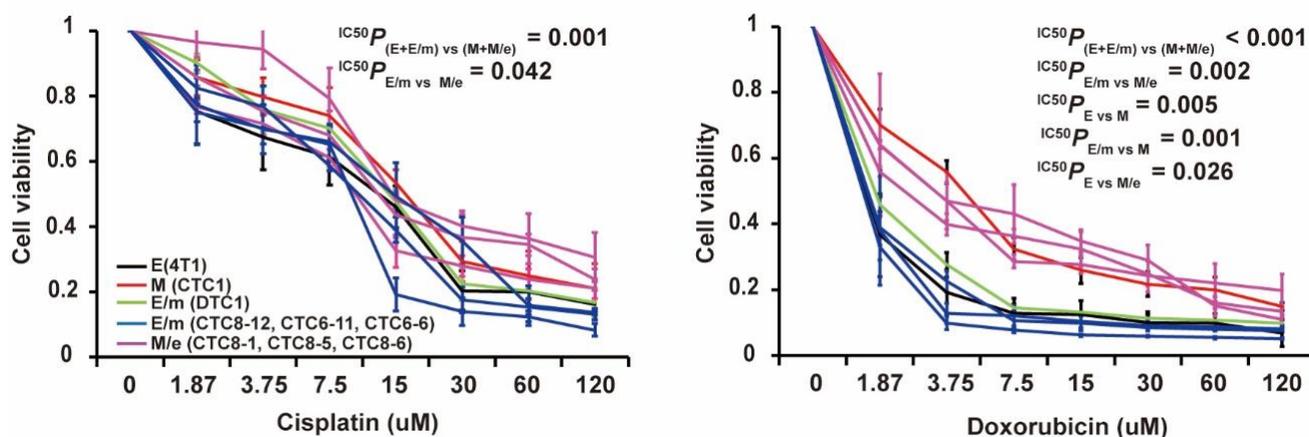


Figure 24: Chemoresistance of E-, E/m-, M/e-, and M- type cells. Chemoresistance towards Cisplatin and Doxorubicin was tested across a concentration range of 1.875 - 120 μM by MTT assay in the indicated cell lines. Line charts show cell viability as means \pm SD from $n \geq 3$ independent experiments. Statistical analysis was done by comparing IC_{50} values.

3.5 E/m-type CTCs possess highest lung metastasis formation ability *in vivo*

In vitro assays showed that E/m-type CTCs had increased proliferation, adhesion, and invasion rates compared to M/e-type CTCs, whereas M/e-type CTCs had increased chemoresistance. CTCs are cells in the bloodstream and are regarded as the main source of metastasis-initiating cells. To compare CTCs with differing EMT phenotypes for their capacity to form lung metastases, we performed intravenous (*i.v.*) injections. To do so, E-type (4T1), E/m-type (CTC6-6, CTC6-11, CTC8-12, and DTC1 as control), M/e-type (CTC8-6, CTC8-5, CTC8-1), and M-type (CTC1) cells were injected in the tail vein of BALB/c mice (identical cell number, 5×10^4 cells) (**Figure 25A**). After 19 days, mice were sacrificed and lung metastasis was assessed by counting superficial metastases and by *ex vivo* metastasis colony formation assay. The metastatic index was calculated as numbers of lung metastatic colonies divided by the numbers of intravenously injected cells. As shown in **Figure 25 B, C, and D**, epithelial-type (E, E/m) cells had a significantly improved metastasis-inducing ability in comparison with mesenchymal cells (M, M/e) cells. Especially, E/m-type CTCs exhibited the best capacity to form lung metastasis and had an average 9.13-fold higher metastatic index than M/e-type cells, a 2.41-fold higher than E-type cells, and a 98.98-fold higher than M-type cells (**Figure 25 B, C, and D**).

Given that M/e- and M-type cells possessed reduced proliferation capacity compared to E- and E/m-type cells, differences in forming lung metastasis might result from longer latencies *in vivo*, and not from inherently decreased metastatic capacities. To address this eventuality, E/m-type CTC6-6 cells (highest metastatic ability; as positive control group), M/e-type cells (CTC8-6, CTC8-5, and CTC8-1), and M-type CTC1 cells were injected intravenously into the tail vein of BALB/c mice (each 5 mice per cell line). The aim of this mouse experiment was to observe the metastatic potential over prolonged time periods. Hence, mice were observed daily for signs for an endpoint (Methods 2.3.6). The defined endpoints were: (1) substantial weight loss (more than 5% in more than two mice out of five), (2) weakness, *i.e.* tiredness, unresponsiveness (more than 2 mice out of five), (3) ≤ 2 remaining experimental groups. After 22 days, 5 of 5 E/m-type CTC6-6-injected mice (100%) and 1 out of 5 CTC8-1-injected mice (20%) displayed substantial weight loss, weakness, and dyspnoea, and had to be sacrificed (**Figure 26A**). Severe lung metastases were defined as ≥ 10 metastases per lung. Upon

autopsy, all mice ended on day 22 had severe lung metastases, and metastatic colony formation assay confirmed these results (**Figure 26 D and E**). On day 25, the remaining 4 mice (80%) in the CTC8-1-injected group and all 5 mice (100%) in the CTC8-5-injected group displayed signs for an endpoint. In these groups, 2 of 9 mice were diagnosed with severe lung metastases (22.2%), 4 of 9 mice (44.4%) had metastases near larger bones, and 3 of 9 mice showed multiple tumor sites (33.3%) (**Figure 26A**). At day 28, all mice in the CTC8-6-injected group presented substantial weight loss and weakness. Mice in the CTC1-injected group did not display weight loss or weakness. These two groups were sacrificed and analyzed at day 28. None of the CTC8-6-injected mice showed severe lung metastases (0%), but all 5 mice had metastases near big bones (100%), and 3 of 5 mice had multiple metastases (60%) (**Figure 26 A and B**). Through autopsy, CTC1-injected mice did not show any metastasis in the lungs, only 2 of 5 mice were bearing small tumors at the injection site in the tail area (**Figure 26A**). Numbers of superficial lung metastases were counted manually, and numbers of *ex vivo* metastatic colonies and metastatic indexes are shown in **Figure 26C** and **Figure 26D and E**. Results confirmed that E/m-type CTC6-6 cells had a significantly higher metastatic index compared to M/e- and M- type cells, even after extended observation time. Substantial weight loss of the injected animals, as a surrogate sign of clinical disease progression, was also first observed in CTC6-6 group (**Figure 26F**).

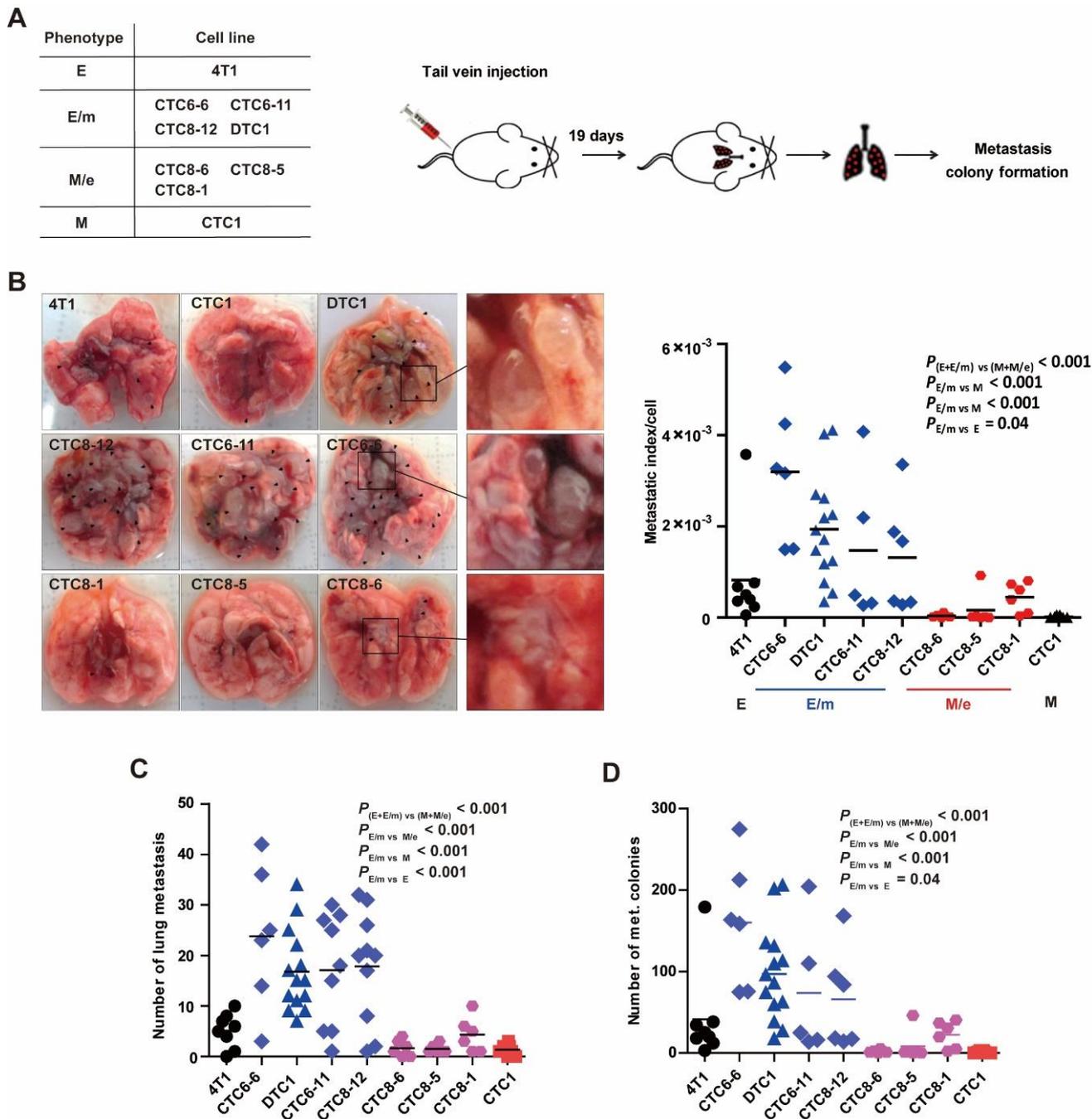


Figure 25: Lung metastasis formation by E-, E/m-, M/e-, M- type cells. (A) Scheme of *i.v.* injection: Epithelial E-type (4T1), E/m-type (CTC6-6, CTC6-11, CTC8-12, DTC1), M/e-type (CTC8-6, CTC8-5, CTC8-1), and mesenchymal M-type (CTC1) cells (5×10^4) were transplanted into BALB/c mice through tail vein intravenous injection. After 19 days, numbers of superficial lung metastasis were counted and lungs were harvested for further metastasis colony formation assay. (B) Metastatic index per cell for each cell line. Left: Shown are representative pictures of lungs following *i.v.* injection. Right: Dot plot graph shows metastatic index with means (lines). (C) Dot plot graph shows numbers of metastases counted in the lungs with means (lines). (D) Dot plot graph shows numbers of colonies counted in metastasis colony formation assay with means (lines). Injected numbers of mice: $n_{4T1} = 8$, $n_{CTC1} = 8$, $n_{DTC1} = 14$, $n_{CTC6-6} = 6$, $n_{CTC6-11} = 5$, $n_{CTC8-12} = 6$, $n_{CTC8-6} = 6$, $n_{CTC8-5} = 6$, $n_{CTC8-1} = 6$.

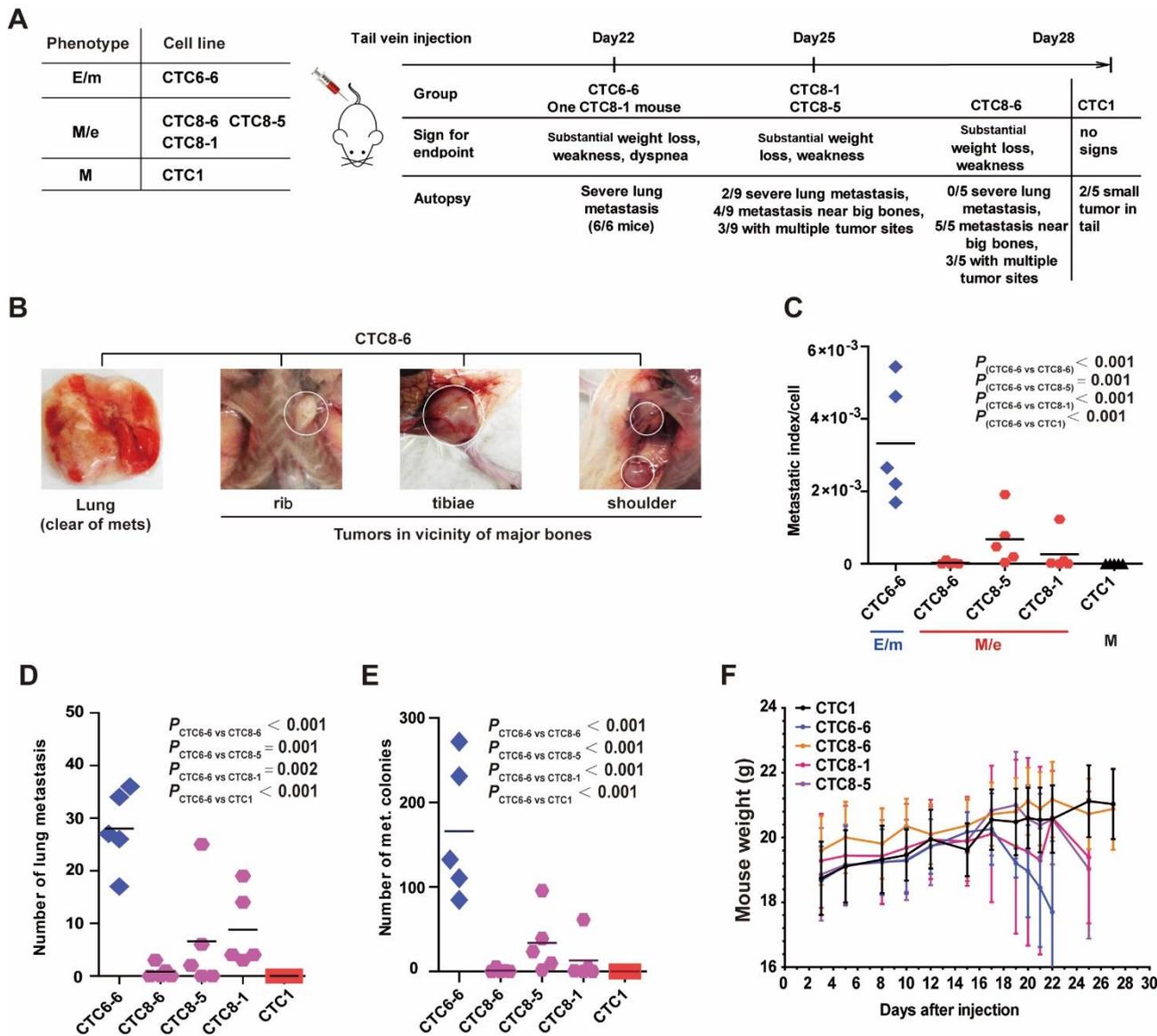


Figure 26: Metastasis formation by E/m-, M/e-, and M- type CTCs. (A) Scheme of *i.v.* injection experiment: E/m-type (CTC6-6), M/e-type (CTC8-6, CTC8-5, CTC8-1), and mesenchymal M-type (CTC1) cells were transplanted to BALB/c mice through intravenous tail vein injection. Experimental groups were ended at the day of the indicated signs for a pre-defined endpoint. Results from autopsy at the given time points are described in the table. (B) Shown are pictures of autopsy results from CTC8-6-injected mice displaying the lack of lung metastasis, tumors in the vicinity of the rib, tibiae, and shoulder blade. (C) Shown is the metastatic index per cell of E/m-type (CTC6-6), M/e-type (CTC8-6, CTC8-5, CTC8-1) and mesenchymal M-type (CTC1) cells with means (lines) and P values as indicated. (D) Dot plot graph shows numbers of metastases counted in the lungs with means (lines) and P values as indicated. (E) Dot plot graph shows numbers of colonies counted in metastasis colony formation assay with means (lines) and P values as indicated. (F) Shown is weight growth curves for each group as means \pm SD.

3.6 EMT in CTC lines is not a mere reflection of cell heterogeneity in 4T1 and DTC1 cells

4T1 cells and DTC1 cells are characterized by high genetic alterations, which could be a reason for the heterogeneous EMT phenotypes observed in CTC lines. In other words, the observed EMT heterogeneity in CTC lines might be the result of a clonal selection of pre-existing EMT phenotypes, rather than an actual induction of EMT during the metastatic cascade. To test this hypothesis, single cell clones (SCCs) of 4T1 (n = 30), CTC1 (n = 23), and DTC1 (n = 30) were generated *in vitro* and compared with *ex vivo* DTC1-derived CTC lines (n = 26). EMT scores were assessed and EpCAM expressions were measured by flow cytometry in 4T1-, CTC1-, and DTC1- derived SCCs. EMT scores of 4T1-SCCs were ranged from 0 to 80, reflecting a high proportion of cells with an epithelial phenotype. EMT scores of CTC1-SCCs were ranged from 380 to 400, which confirmed a predominant mesenchymal phenotype in CTC1-SCCs. The largest variation in EMT scores amongst single cell clones was observed in DTC1-SCCs, with EMT scores ranging from 0 to 150. However, DTC1-derived CTC lines had EMT scores ranging from 0 to 360 (**Figure 27A and B**). Therefore, the range of EMT in DTC1-derived CTC lines was substantially broader than in DTC1-SCCs, CTC1-SCCs, or 4T1-SCCs, and thus, it was not simply a result of the heterogeneity of 4T1 cells in general. Variances (squared standard deviations) of SCCs and DTC1-derived CTC lines were calculated and compared. Significant differences in EMT scores were found between DTC1-SCCs (range 90, variance 736.86) and DTC1-derived CTC lines (range 345, variance 14428.12) (**Figure 27B**).

EpCAM expressions were measured in 4T1-, CTC1-, and DTC1-SCCs and compared with DTC1-derived CTC lines. 4T1-SCCs had the largest variance (344.77), which might be attributed to a generally very high EpCAM level in 4T1 cells. Similar to EMT scores, DTC1-derived CTC lines had a wider distribution of mean EpCAM expression levels with increased variance and range compared to DTC1-SCCs (**Figure 27C**). In conclusion, these data demonstrated that EMT phenotypes in DTC1-derived CTC lines were not the sole result of the heterogeneity of parental 4T1 or DTC1 cells.

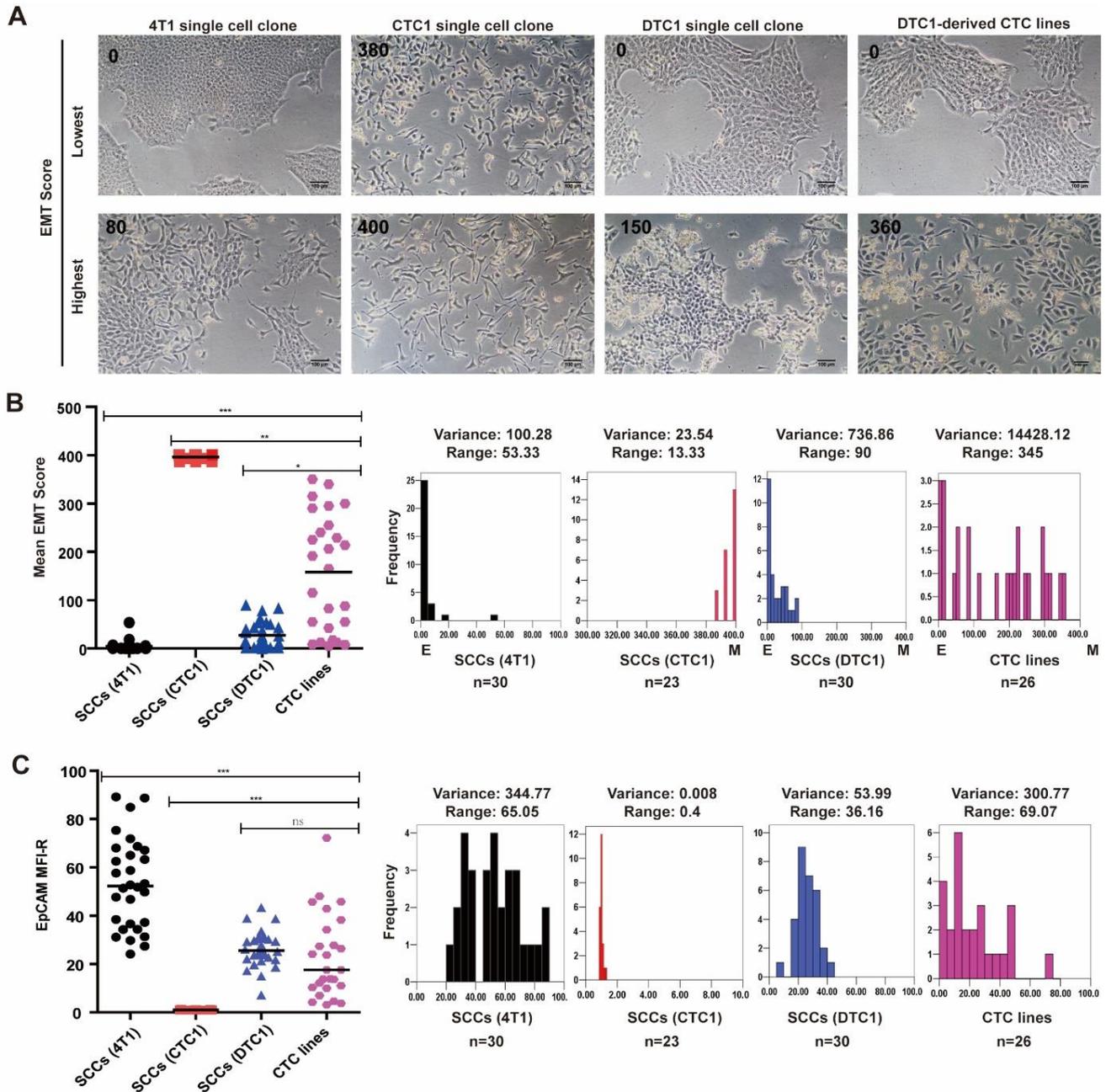


Figure 27: EMT scores and EpCAM expressions in 4T1-, CTC1- and DTC1-SCCs and in DTC1-derived CTC lines. (A) Shown are representative pictures of 4T1-, CTC1-, DTC1- SCCs, and DTC1-derived CTC lines ($n = 26$) for the highest and lowest EMT scores. Individual EMT scores are indicated within each picture. (B) Dot plots show mean values of EMT scores from $n = 3$ independent scores in 4T1-, CTC1-, and DTC1-SCCs, and DTC1-derived CTC lines. Black lines represent means of EMT scores for all 4T1-, CTC1-, and DTC1-SCCs, and DTC1-derived CTC lines. (C) Dot plot graph shows means of EpCAM expression from $n = 2$ independent measurement in 4T1-, CTC1-, and DTC1- SCCs, and DTC1-derived CTC lines. Black lines represent means of EpCAM expression for all 4T1-, CTC1-, and DTC1- SCCs, and DTC1-derived CTC lines. Data are additionally shown as frequency diagrams with variance (squared value of standard deviation) and range (difference between lowest and highest values) (right panel).

3.7 EpCAM expression is negatively correlated with EMT

EpCAM is an epithelial marker that is widely used to detect and isolate CTCs from the blood of cancer patients. However, cells can lose EpCAM expression during EMT (Gorges et al., 2012). To test whether EpCAM expression in DTC1-derived CTC lines was related to EMT, a Spearman's rank correlation test was performed to analyze a potential correlation between EpCAM expression and EMT score. The result proved that the expression of EpCAM was negatively correlated with the EMT score ($r = -0.728$, $P < 0.001$) (**Figure 28**). This finding further indicated that CTCs that undergo EMT would lose EpCAM and escape from EpCAM-dependent CTC isolation using enrichment techniques.

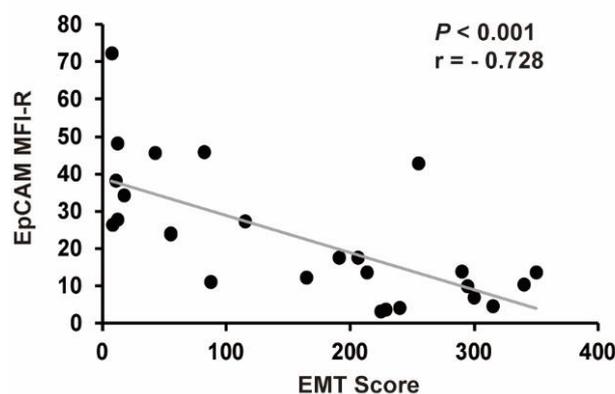


Figure 28: Cluster plot analysis of Spearman's rank correlation between EpCAM expression and EMT score in DTC1-derived CTC lines. P value and r value are indicated.

3.8 EpCAM expression is higher in metastatic sites compared with primary tumors

After subcutaneously transplanting 4T1, CTC1, and DTC1 cells in BALB/c mice, *ex vivo* cultured cell lines were established from primary tumors, blood, bones, and metastases. EpCAM expression in these cell lines was measured by flow cytometry (**Figure 29**).

In 4T1-derived cell lines, EpCAM was down-regulation in a proportion of the primary tumor cell sublines compared to parental 4T1 cells, while metastatic cell lines presented generally high EpCAM expression (**Figure 29A**). EpCAM was significantly higher in metastases comparing to primary tumors. Furthermore, EpCAM remained absent in CTC1-derived primary tumors and metastatic cell lines (**Figure 29B**). Former IHC staining of EpCAM in CTC1 primary tumors was also negative. Accordingly, CTC1-derived primary tumor and metastatic cells maintained a mesenchymal phenotype

in *ex vivo* culture conditions. This result further suggested that tumor cells can form primary tumors and metastases without considerable MET, although with decreased frequency. In DTC1-derived cell lines, EpCAM expression was higher in DTCs and metastasis-derived cell lines compared to CTCs and primary tumor cell lines (**Figure 29C**). Furthermore, DTC1-derived CTC lines displayed significant heterogeneity of EpCAM expression (**Figure 29C**).

Referring to the results of the former mouse experiment, showing that E/m-type cells possessed the highest metastatic ability, higher EpCAM expression in metastases could be the consequence of E/m-type CTCs being the primary source of metastasis.

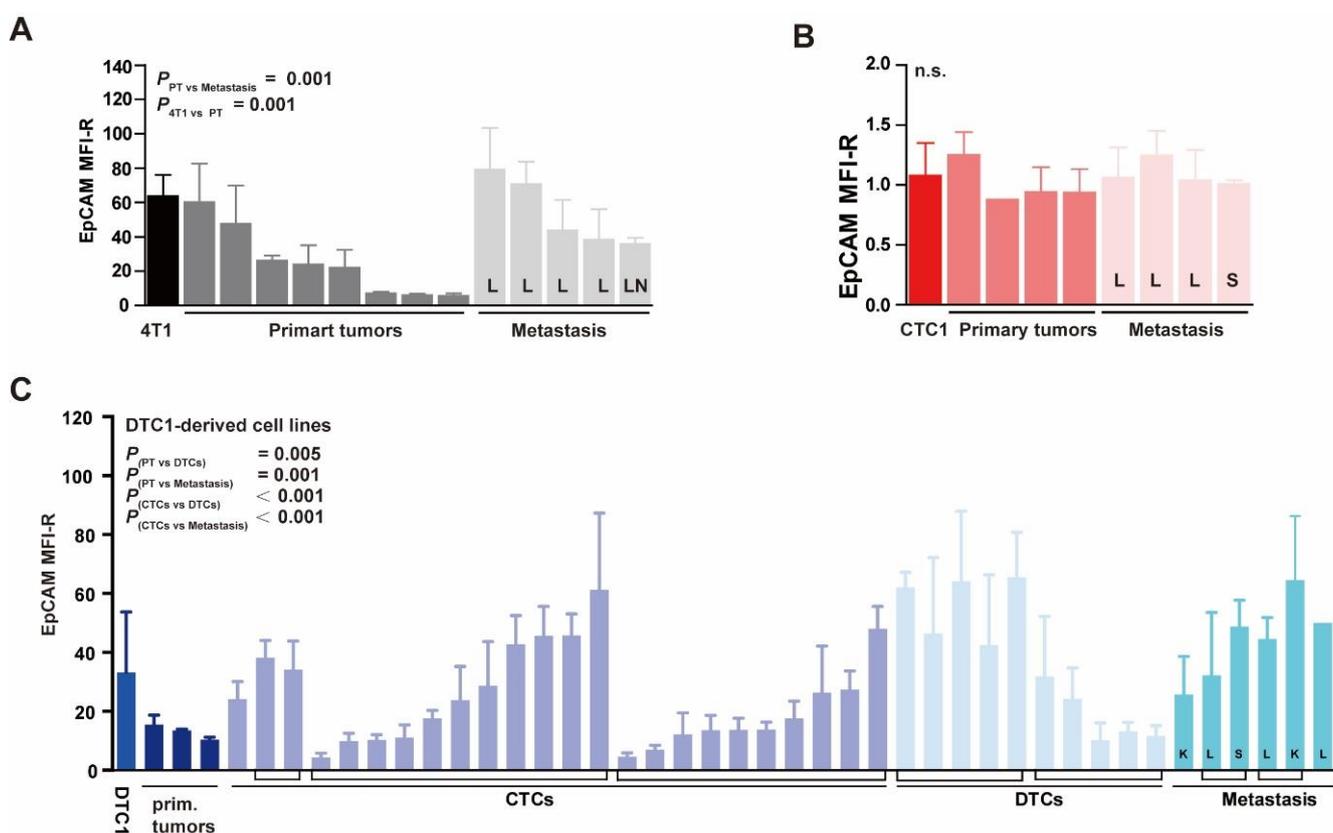
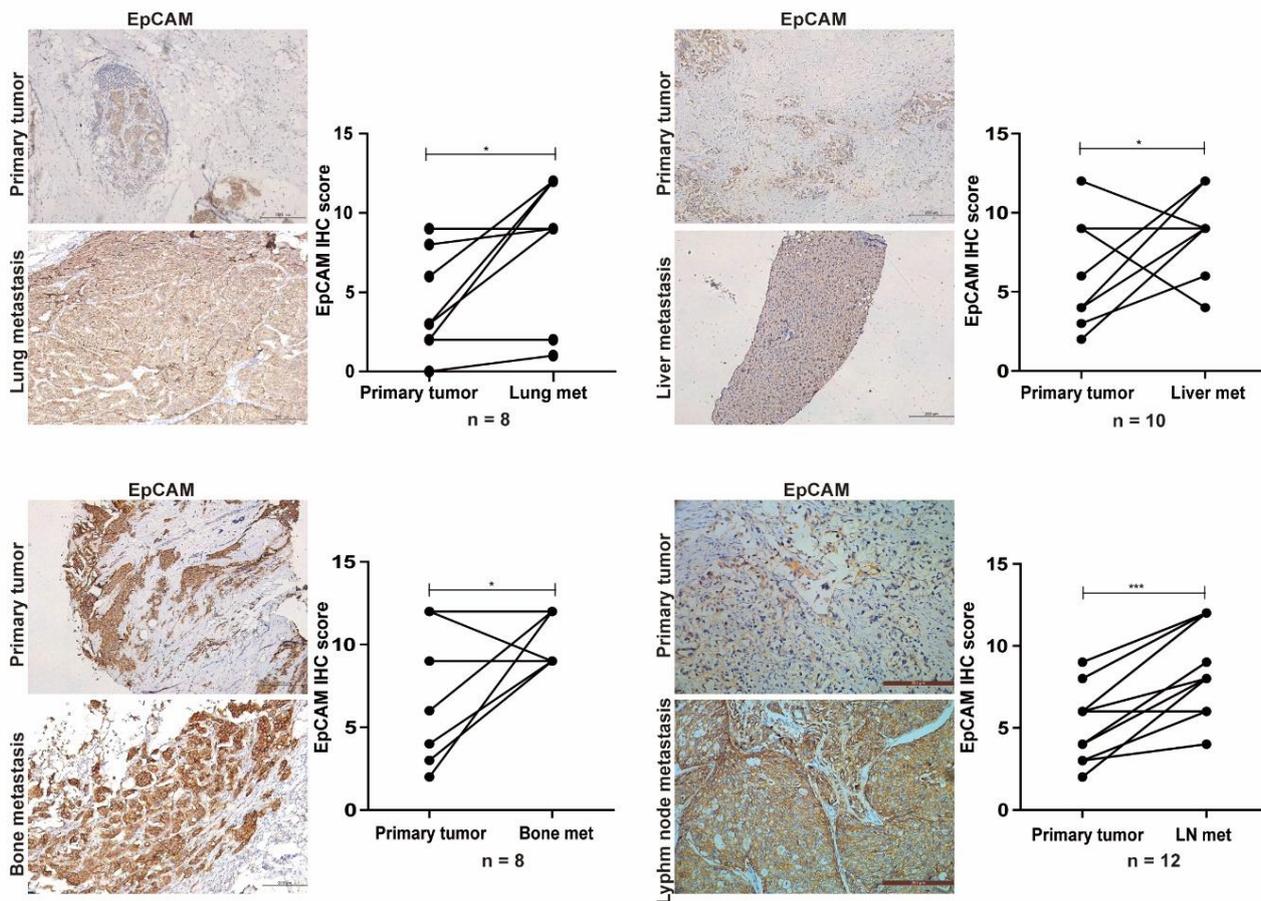


Figure 29: EpCAM expression of 4T1-, CTC1-, and DTC1-derived cell lines. EpCAM expression was tested by flow cytometry in cell lines from primary tumors (PT), CTCs, DTCs, and metastases isolated from 4T1-, CTC1-, and DTC1-injected BALB/c mice. (A) Bar chart shows EpCAM expression in 4T1-derived cell lines as the mean fluorescence intensity ratio (MFI-R) \pm SD from $n \geq 3$ independent measurements. (B) Bar chart shows EpCAM expression in CTC1-derived cell lines as the MFI-R \pm SD from $n \geq 3$ independent measurements. (C) Bar chart shows EpCAM expression in DTC1-derived cell lines as the MFI-R \pm SD from $n \geq 3$ independent measurements. Cell lines originating from the same mouse were marked with brackets. L: lung, S: spleen, K: kidney, LN: lymph node.

A



B

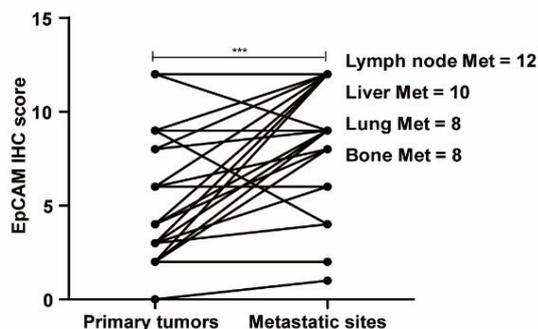


Figure 30: EpCAM expression in primary tumors and corresponding metastases in patients with MBC. (A-B) EpCAM expression was assessed in n = 38 breast cancer patients. Shown are representative IHC stainings of EpCAM in primary tumors and corresponding metastatic sites (lung, liver, bone marrow, lymph node), and the quantification of EpCAM IHC scores of paired tumor and metastases samples are shown as paired line charts.

In the 4T1 mouse model, EpCAM was expressed to higher levels in metastatic sites compared to primary tumors. To explore whether comparable observations could be done in patients with MBC, pairs of primary tumors and corresponding lymph node metastases (n = 12), liver metastases (n = 10), lung metastases (n = 8), and bone metastases (n = 8) were obtained from breast cancer patients after surgery. IHC staining of EpCAM expression was performed at the Shanghai General Hospital,

Shanghai, China. The results of EpCAM IHC scores confirmed that EpCAM levels were higher in metastases compared to primary tumors (**Figure 30 A and B**), which was similar to the findings in the 4T1 mouse model.

3.8 The EpCAM-positivity rates in DTCs is related to metastasis and predicts survival in patients with MBC

To extend our findings, an EpCAM-independent CTC isolation technology that integrates subtraction enrichment and immunostaining fluorescence *in situ* hybridization (SE-iFISH) was used to assess the EMT status of CTCs and DTCs in patients with MBC (Lin, 2015). In collaboration with Prof. Hongxia Wang's group (Shanghai General Hospital, China), CTCs and DTCs were isolated by multi-marker subtraction enrichment (Methods 2.2.16). In total, n = 34 patients with breast cancer stage III - IV were recruited from September 2015 to April 2017. The characteristics of patients are shown in **Figure 31**. Patients were newly diagnosed with MBC and have received standard care at the Shanghai General Hospital. The age of patients ranged from 28 to 76 years, with a median of 57 years. The majority of these patients were suffering from stage IV MBC (67.65%) and had the intrinsic subtype of luminal B breast cancer (44.12%). All patients were followed for a median of 11 months. Lungs and bones were the main sites of metastases (both represented 14.71% of the metastatic sites, each).

In the current study, CTCs and DTCs were isolated from blood and bone marrow, respectively. Blood and bone marrow samples were firstly deprived of red blood cells, and the remaining cells were detected with CD45-specific antibody to exclude white blood cells. *In situ* hybridization with chromosome enumeration probes hybridizing to human chromosome 8 (CEP8) was conducted to assess the ploidy status of each single cell (**Figure 32**). Aneuploidy on chromosome 8 was frequently observed using CEP8 in breast cancer, and thus the CEP8 status served to identify cancer cells in the blood and bone marrow (Lin, 2018). EpCAM expression was examined in the enriched cell population. As shown in **Figure 32**, CTCs and DTCs were identified as CD45-negative, DAPI-positive, CEP8-aneuploid and EpCAM positive or negative cells, or CEP8-diploid EpCAM positive cells.

Characteristic	No. (%)
Total no. of patients	34 (100)
Gender	
Female	34 (100)
Age (years)	
Median	57
Range	28-76
Tumor stage	
Stage IIIa	1 (2.94)
Stage IIIc	10 (29.41)
Stage IV	23 (67.65)
Intrinsic Subtype	
Luminal A	6 (17.65)
Luminal B	15 (44.12)
HER2 +	9 (26.47)
HER2 -	6 (17.65)
HER2 positive (non-luminal)	7 (20.59)
Triple negative (basal-like)	4 (11.76)
Normal-like	2 (5.88)
Follow-up (Months)	
Median	11
Range	1-28
Distant detectable metastases	
Lung	5 (14.71)
Bone	5 (14.71)
Liver	4 (11.76)
Others	2 (5.88)
Multiple	7 (20.59)
Numbers of detected CTC per patient	
Median	9
Range	1-185
Numbers of detected DTC per patient	
Median	413
Range	4-40240

HER2 = human epidermal growth factor receptor 2.

Figure 31: Characteristics of patients. Shown are clinical parameters for n = 34 patients with MBC, including gender, age, tumor stage, intrinsic subtype, follow-up time, metastatic status, and CTC and DTC numbers.

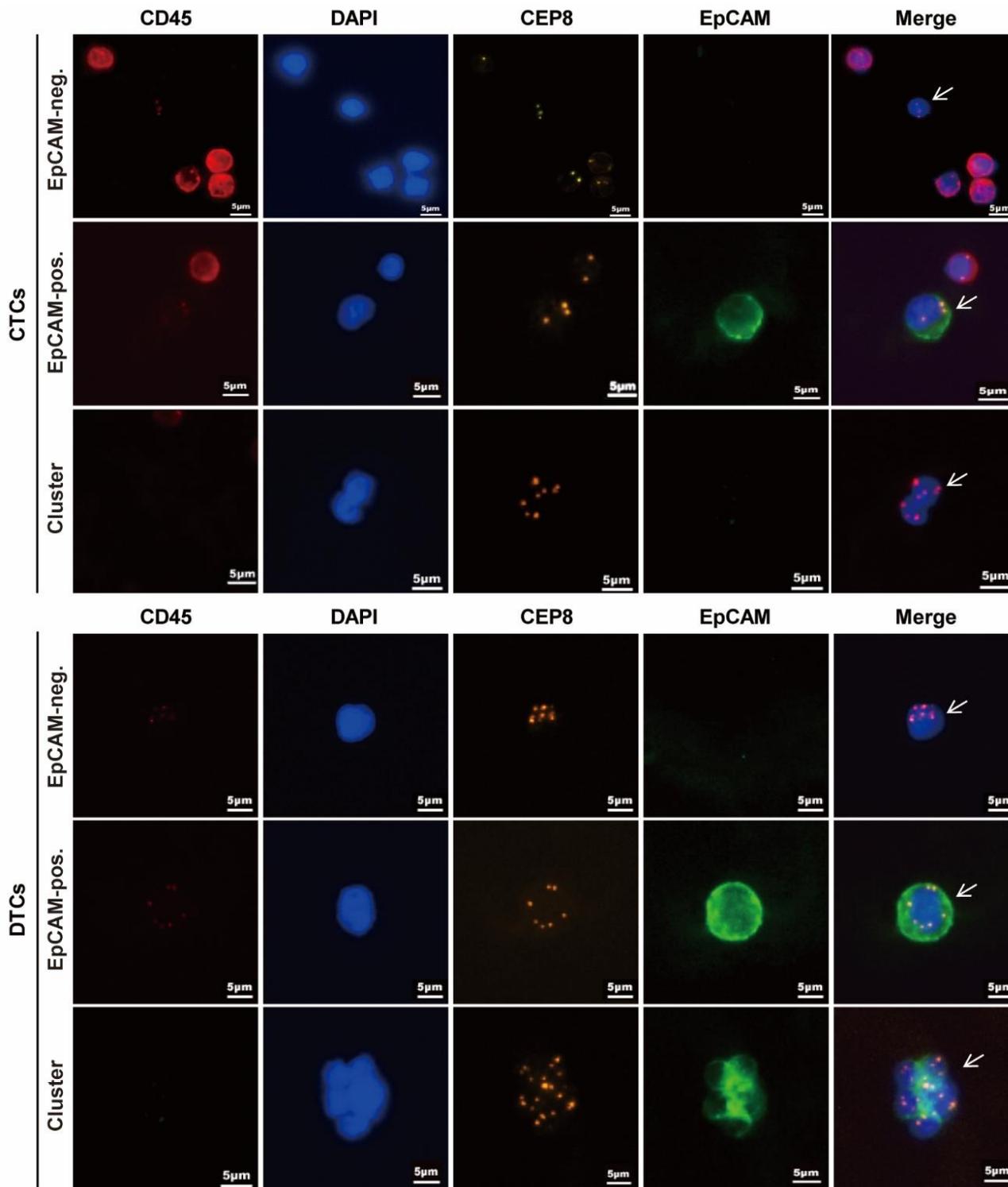


Figure 32: SE-iFISH enrichment and detection of CTCs and DTCs. Immunofluorescence staining of CD45 (red), DAPI (blue), EpCAM (green), and CEP8 (orange) are depicted as indicated. CTC and DTC, and tumor cell clusters are indicated by white arrows. Shown are representative examples as indicated. EpCAM-pos: EpCAM-positive; EpCAM-neg: EpCAM-negative.

A total number of 845 CTCs and 71,910 DTCs were detected from $n = 34$ patients with MBC. The median cell numbers detected per patient were 9 CTCs and 413 DTCs, respectively. As shown in

Figure 33 A and B, the overall numbers of CTCs were less than of DTCs per patient, and the numbers of cell clusters were also significantly higher in DTCs compared to CTCs.

The EpCAM-positivity rate was defined as the number of EpCAM-positive cells divided by the total number of detected cells in each patient. The EpCAM-positivity rate was found higher in DTCs compared to CTCs, which is in line with the result obtained in the 4T1 mouse experiment (**Figure 33C**).

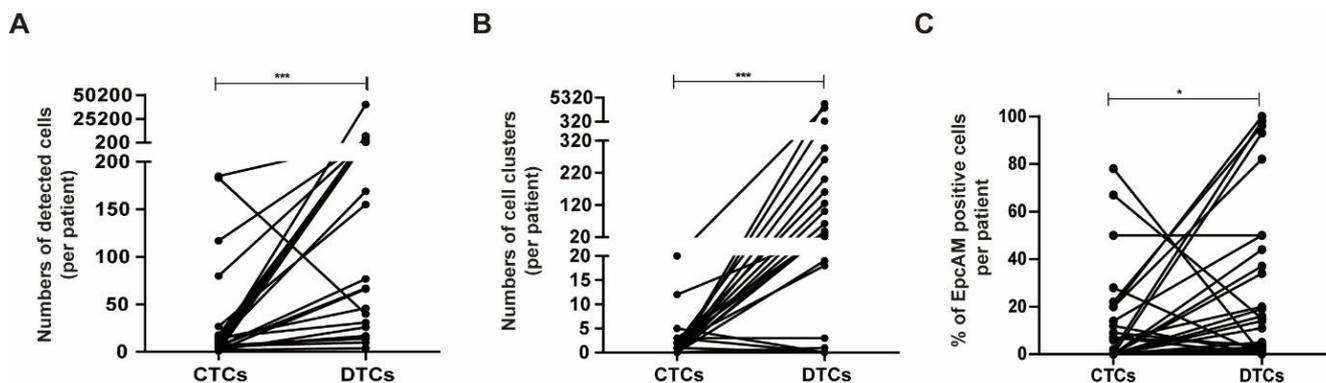


Figure 33: Paired CTCs and DTCs numbers. (A) Paired line chart shows the numbers of CTCs and DTC detected in $n = 34$ MBC patients. Each line represents one patient. (B) Paired line chart shows the numbers of cell clusters. Each line represents a patient. (C) Paired line chart shows the EpCAM positive rate in each patient. Each line represents a patient.

The proportion of patients containing EpCAM-positive cells is shown in **Figure 34**. Twenty of 34 patients (58.8%) did not contain detectable EpCAM-positive CTCs in their blood. In contrast, only 12 patients (35.3%) had no EpCAM-positive DTCs in their bone marrow (**Figure 34**). Referring to the entirety of detected CTCs and DTCs, the proportion of EpCAM-positive CTCs and DTCs was 22.4% and 65.9%, respectively (**Figure 34**). The ploidy statuses of CTCs and DTCs are shown in **Figure 35A**. Triploid (36.6%) and tetraploid (36.3%) represented the highest proportion of CTCs and DTCs, respectively. Cell size was compared referring to WBC, and the majority of cells had a size similar to WBC (50.7% in CTCs and 77.0% in DTCs, respectively) (**Figure 35B**). **Figure 36** further shows the ploidy status of CTCs and DTCs in relation to cell size. For cells larger than WBC, 64.6% of CTCs and 84.6% of DTCs had CEP 8 numbers larger than 3.

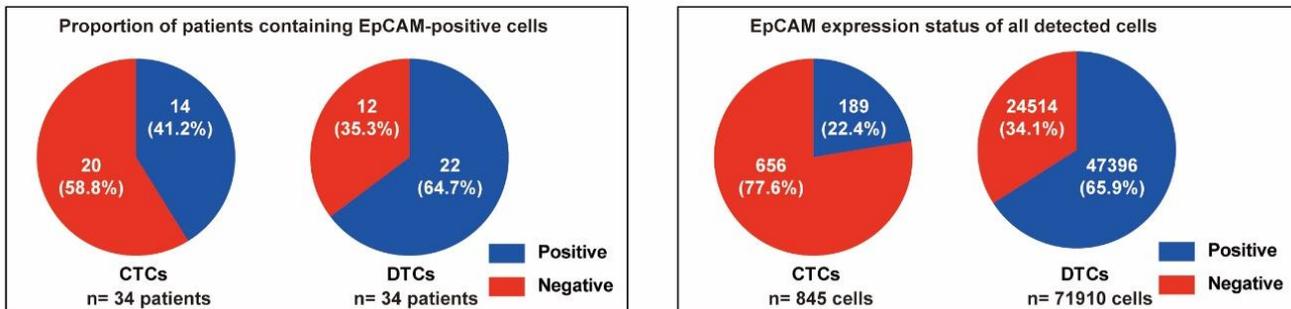


Figure 34: Proportion of EpCAM expression in CTCs and DTCs. Pie charts show the proportion of MBC patients containing EpCAM-positive cells (left) and the EpCAM expression status of all detected cells (right). EpCAM-positive and -negative proportions are depicted in blue and red, respectively.

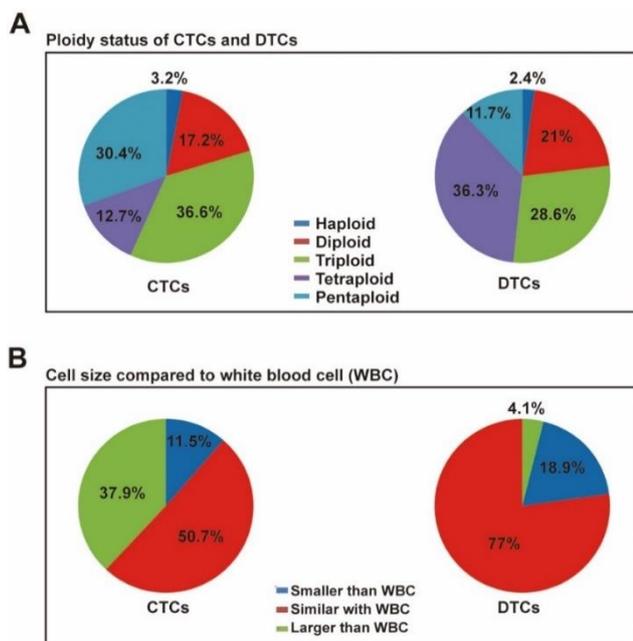
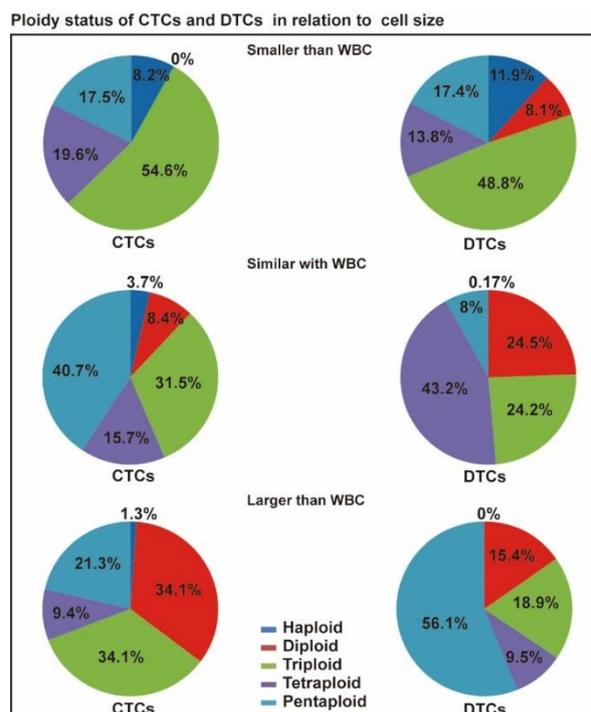


Figure 35: Pie charts show (A) the ploidy status of CTCs and DTCs, and (B) the cell size compared with WBC

Figure 36: Ploidy status of CTCs and DTCs in relation to the cell size, referring to smaller, similar, and larger than WBC. Ploidy statuses are colour-coded as indicated and are given in percentages.



The expression of EpCAM was negatively correlated to the EMT status of cells and the EpCAM-positivity rate reflected the proportion of epithelial tumor cells in CTCs and DTCs derived from the 4T1 mouse model. Therefore, the EpCAM-positivity rate was applied as a parameter to stratify $n = 34$ MBC patients and to predict the patient's metastatic status and clinical outcome. The results showed that the EpCAM-positivity rates of CTCs and DTCs were higher in patients with detectable organ metastasis (M_1), compared to patients without (M_0) (**Figure 37A**). In addition, higher EpCAM-positivity rates in DTCs were significantly associated with the occurrence of lung metastases (**Figure 37A**). All patients were followed for a median of 11 months. A receiver operating characteristic (ROC) curve analysis was conducted to assess the efficiency of the EpCAM-positivity rate to predict the 6-months survival of patients. The EpCAM-positivity rate of DTCs, but not CTCs, allowed to predict the risk of 6-months mortality (**Figure 37B**; AUC = 0.785, 95% CI 0.588 – 0.983; $P = 0.018$). A cut-off value of 19.78% EpCAM expression was further obtained, with a sensitivity and specificity of 75.0% and 82.6%, respectively. Based on the results of the ROC analysis, we used a 20% EpCAM-positivity rate as a cut-off value to predict the overall survival of patients with MBC. The Kaplan-Meier survival analyses demonstrated that patients with EpCAM-positivity rates of $\geq 20\%$ in DTCs had substantially poorer overall survival (**Figure 37C**).

Considering diploid EpCAM-positive cells could be normal epithelial cells that in the blood of patients, EpCAM-positivity rates of CTCs and DTCs were re-calculated by excluding diploid EpCAM-positive cells. Shown in **Figure 37 D, E, and F**, the results were similar to the results obtained when including diploid EpCAM-positive cells. However, a significant difference was observed in the EpCAM-positivity rates of aneuploid CTCs between patients with and without lung metastases. The proportions of EpCAM-positive aneuploid DTCs could also predict the 6-months mortality in patients with MBC (AUC = 0.793, 95% CI 0.599 – 0.988; $P = 0.015$), and the cut-off value decreased to 16.87% (**Figure 37E**). Hence, we applied a 15% EpCAM-positivity rate as a cut-off value to predict the patients' outcome. The Kaplan-Meier curves demonstrated that patient with EpCAM-positive aneuploid DTCs $\geq 15\%$ had a substantially poorer outcome (**Figure 37F**).

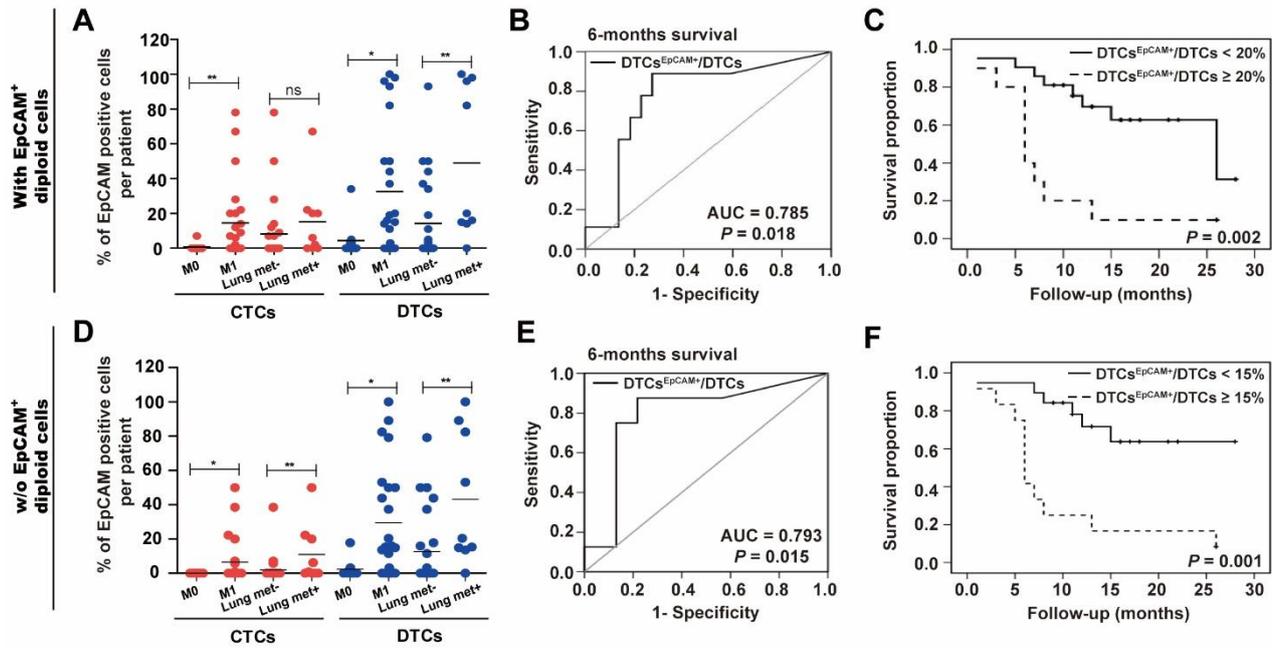


Figure 37: EpCAM-positivity rate as a clinical parameter is associated with the patient's metastatic status and can predict the clinical outcome. EpCAM⁺: EpCAM positive. (A and D) Dot plot graph shows the percentage of EpCAM-positive CTCs and DTCs per patient (n = 34) and is stratified according to distant metastases statuses M₀ and M₁, and the presence or absence of lung metastases ((A) includes EpCAM-positive diploid cells, (D) excludes EpCAM-positive diploid cells). (B and E) Receiver operating characteristic (ROC) curves predict 6 months survival of patients with MBC ((B) includes EpCAM-positive diploid cells, (E) excludes EpCAM-positive diploid cells). The area under the curve (AUC) and P value are presented. (C and F) Kaplan-Meier survival curves show overall survival of patients (n = 34) with EpCAM-positivity rate cutoff values of 20% (C) and 15% (F) in DTCs ((C) includes EpCAM-positive diploid cells, (F) excludes EpCAM-positive diploid cells).

In summary, the EpCAM-positivity rates of CTCs and DTCs are useful markers to predict the risk of metastasis, the 6-months survival, and the overall survival of patients. These clinical data strongly supported the results of the 4T1 mouse model, in which EpCAM-positive CTCs (E- and E/m- type cells) were correlated with the generation of lung metastasis and poorer outcome.

4. DISCUSSION

In the current study, the 4T1 MBC mouse model comprehensively recapitulated the different stages of tumor progression including the formation of primary tumors, the systemic dissemination of cancer cells, and, ultimately, the formation of organ metastases. Based on the in-depth analysis of the EMT heterogeneity of 4T1-derived CTC lines, we found that epithelial tumor cells with a limited mesenchymal transition were the main reservoir of metastasis-inducing systemic cancer cells. EpCAM, as a marker of epithelial cells, was reduced in expression or entirely lost in breast carcinoma cells undergoing EMT. On the basis of this regulation, we used the EpCAM-positivity rate in CTCs and DTCs as a surrogate marker for their EMT status, and could show an association with the metastatic status and clinical outcome in clinical samples of patients with MBC.

4.1 The 4T1 metastatic mouse model

The 4T1 breast cancer model is a well-established animal model that recapitulates the process of tumor progression, including the formation of primary tumors, CTCs, DTCs, and, eventually, distant metastases (Aslakson and Miller, 1992; Miller et al., 1987). 4T1 cells were first isolated from a lung metastasis in a breast cancer-bearing mouse through *in vivo* selection by 6-TG (Aslakson and Miller, 1992; Miller et al., 1981). As a breast cancer cell line, 4T1 cells have a typical epithelial phenotype with compact cell-cell connections and a cobblestone-like morphology. In the present study, 4T1 cells were subcutaneously transplanted in the flank of BALB/c mice to generate primary tumors. CTC lines were isolated at late stages of tumor formation (*i.e.* 3 weeks following transplantation) from the blood of tumor-bearing animals. CTCs were selected based on their resistance to 6-TG, as opposed to normal blood cells, which are vulnerable to 6-TG. Potential DTC lines were isolated in a similar manner from the bone marrow of transplanted mice at the same time points. By doing so, one CTC and one DTC line were generated following the transplantation of 4T1 cells. CTC1 had a strong mesenchymal morphology with loose cell-cell contact, complete loss of EpCAM expression, and distinct spindle shape. All these features reflect a pronounced EMT in CTC1. The DTC1 cells showed a hybrid EMT phenotype (E/m-type), with a majority of cells remaining in an epithelial state and a subpopulation of

cells that have acquired a more mesenchymal phenotype. This was accompanied by an overall 50% decrease of EpCAM expression in comparison with 4T1 cells. Hence, CTC1 and DTC1 cells presented two different statuses of EMT: CTC1 were in a complete mesenchymal transition, whereas DTC1 cells appeared in a stage of partial EMT. These results confirmed that EMT occurred during the metastatic cascade in the 4T1 model. Furthermore, the results are also in strong support of the notion that EMT represents a gradual process with different cellular stages in cancer, rather than an “all-or-nothing” process.

In vitro functional assays performed in the course of the study further show that CTC1 cells have increased invasion and anchorage-independent cell growth, while decreased proliferation and adhesion compared with parental 4T1 cells. DTC1 cells were characterized by a hybrid EMT type in which they have retained in a proliferative state, are more adhesive to endothelial cells, and possess enhanced invasion and anchorage-independent cell growth. Therefore, it can be suggested that DTC1 cells have made an evolution compared to 4T1 cells through a partial EMT, that allowed them to adapt to the tumor microenvironment. DTC1 cells seem to have undergone an EMT to a degree sufficient to provide them with the above mentioned mesenchymal traits (*i.e.* invasion, migration), without suffering the loss of proliferation and the ability to form cell contacts. Oppositely, full EMT as observed in CTC1 cells generates more quiescent cells with improved anchorage-independent growth. However, these features were gained at the expense of the strong capacity to form primary tumors and, even more so, metastases in the lungs. Recently, Kröger C *et al.* reported a hybrid E/M phenotype that is important for the tumorigenicity of carcinoma cells, which cannot be phenocopied by simply mixing E- and M-staged tumor cells together. These findings demonstrated that the hybrid E/M state of tumor cells is an independent cell phenotype that is highly relevant in tumor progression (Kroger et al., 2019). In order to isolate more CTC and DTC lines to study the effect of EMT on metastasis formation, 4T1, CTC1, and DTC1 cells were re-transplanted into BALB/c mice. It was frequently reported that CTCs in general are rare in the blood, with less than 1 CTC in ten million WBCs in patients with MBC (Cristofanilli, 2006; Janni et al., 2016a). In the 4T1 mouse model, only one CTC line (CTC1) was established after transplantation of parental 4T1 cells and no CTC line was retrieved from CTC1-

injected mice. Hence, these results reflected and further confirmed the extremely low numbers of CTCs reported in the blood of patients. An additional reason for the observed low incidence of *ex vivo* cultured CTCs and DTCs could be based on differences between *in vivo* and *ex vivo* culture conditions, where CTCs and DTCs may be not able to grow in an *ex vivo* culture environment. It is conceivable that soluble and cellular factors present in the microenvironment *in vivo* are essential for CTCs and DTCs, which will be lacking under *ex vivo* culture conditions. For example, Min Y *et al.* could culture MBC patient-derived CTC lines in ultralow attachment plates in RPMI-1640 medium supplemented with epidermal growth factor (20 ng/mL), basic fibroblast growth factor (20 ng/mL), B27 (10 mL), and antibiotic/antimycotic. However, their CTC lines were unable to grow in other culture media (*i.e.* MCF10A medium and Mammary Epithelial cell Growth Medium) (Yu *et al.*, 2014). Despite potential issues in the *ex vivo* culture of CTCs and DTCs from the 4T1 mouse model, we didn't observe any bias in EMT phenotypes of *ex vivo* established CTC lines, which demonstrated that the culture method applied in the current study did not select for particular EMT types. Selection with 6-TG did also not impact on the epithelial status of 4T1 in culture, as was shown in long-term cultures.

In summary, the 4T1 breast cancer mouse model allowed to recapitulate all major aspects of breast cancer generation and progression in the presence of an intact immune system. It was further confirmed that EMT occurs in systemic cancer cells, and that CTCs and DTCs represent rare cells in the blood and bone marrow of mice.

4.2 EMT phenotypes of DTC-derived CTC lines

In the group of DTC1-injected mice, 26 CTC lines were successfully established *ex vivo*. Given the low numbers of cancer cells appearing in the blood and because selection of CTCs with 6-TG occurred in serial dilutions in a 96-well format, the resulting CTC lines most likely originated from either mono- or oligoclones. The cell morphology differed considerably across these 26 CTC lines, but was however stable within cell lines, further supporting the notion that they represent different mono- or oligoclones. In order to compare the degree of EMT across these CTC lines, an EMT scoring system was introduced. This EMT score system was built by reference to the widely used IHC scoring system, as a product of two fundamental EMT parameter, namely the level of cell-cell connection and the

percentage of spindle-shaped cells within a given cell population. Owing to the heterogeneity of 4T1 carcinoma cells (Roulot et al., 2016; Wang et al., 2017b), the EMT phenotypes observed in *ex vivo* selected CTC lines could have been the result of a tumor heterogeneity that is intrinsic to the parental 4T1 population, rather than an EMT fulfilled during the metastatic cascade *in vivo*. Hence, it was important to provide evidence that differences in EMT phenotypes of 4T1 derivative cell lines were indeed the result of an EMT performed *in vivo*. To do so, single cell clones of 4T1 (n = 30), CTC1 (n = 23), and DTC1 (n = 30) cells were generated *in vitro*, in order to assess an EMT in these clones. Clearly, EMT shifts observed in 4T1- and DTC1-SCCs were significantly inferior to the EMT variations detected in DTC1-derived CTC lines. Hence, a substantial degree of the observed EMT measured in CTC lines must have occurred *in vivo* after transplantation of 4T1 and DTC1 cells.

Recent evidence demonstrated that EMT was not an “all-or-nothing” event and that cancer cells usually acquire hybrid EMT phenotypes, *i.e.* E/m- or M/e-types (Brabletz et al., 2018a; Huang et al., 2013; Jordan et al., 2011; Lambert et al., 2017; Pastushenko et al., 2018; Sikandar et al., 2017; Thompson and Nagaraj, 2018; Yu et al., 2013). As mentioned, DTC1-derived CTC lines showed a substantial diversity in EMT phenotypes, which allowed us to further subdivide CTCs into E/m- and M/e-phenotypes. Because our EMT scoring system was based on morphological criteria rather than on genetic criteria, we opted to divide CTCs only into two major types, *i.e.* an E/m- and an M/e-type, to reflect their stages of epithelial and mesenchymal transitions. These E/m- and M/e-phenotypes observed in the 4T1 model replicate the frequently discussed partial EMT in tumor cells (Chaffer et al., 2016; Nieto et al., 2016; Voon et al., 2017), and support the notion that tumor cells do undergo a limited EMT. We selected three E/m- and M/e-type CTC lines each for further *in vitro* functional characterization and *in vivo* metastasis formation analysis. Functional characterization showed that E/m-type cells have increased invasion, proliferation, and adhesion capacities compared to M/e-type CTCs. The mRNA transcript levels of the epithelial markers Epcam, E-cadherin, Rab25 were significantly higher in E/m-type cells, whereas the mesenchymal marker Vimentin, Slug, and Zeb2 were expressed to higher levels in M/e-type cells. No differences were observed in the expression levels of Ddr1, Grhl2, Krt19, N-cad, Vim, Slug, Zeb1, Erbb2/3, Snail, and Twist. Hybrid EMT

phenotypes were also addressed in ovarian and breast cancer cells (Huang et al., 2013; Yu et al., 2013). Huang *et al* have defined EMT groups similar to ours: epithelial, intermediate epithelial, intermediate mesenchymal, and mesenchymal, and have analyzed the correlation between EMT phenotypes and the outcome of patients suffering from a broad spectrum of carcinomas, comprising breast cancer (Huang et al., 2013; Tan et al., 2014). In the current study, we tested EMT-related genes, including Ddr1, Grhl2, Krt19, N-cad, Vim, Slug, Zeb1, Erbb2/3, Snail, and Twist, as referenced by Tan *et al.* (Tan et al., 2014), in nine cell lines with different EMT statuses (4T1, CTC1, DTC1, E/m-, and M/e- type cells). Although we observed significant differences across CTC lines, the expression pattern of some EMT genes revealed unexpected. For instance, EMT-inducing genes Twist and Snail were decreased in the mesenchymal CTC1 cell line, as compared to 4T1 and DTC1 cells. As the EMT program is controlled by complex pathways and since plenty of genes can be involved, the activation of EMT genes can be tumor-specific (Chaffer et al., 2016; Nieto et al., 2016; Saitoh, 2018; Zhang et al., 2016). For example, NF κ B can increase Snail expression in breast, colon, and pancreas cancer cells (Barbera et al., 2004; Baulida et al., 2019; Cichon and Radisky, 2014; Ordonez-Moreno et al., 2017), while in atypical teratoid/rhabdoid and liver tumor cells, the activation of Snail is mainly done by STAT3 (Liu et al., 2015; Saitoh et al., 2016; Zhou et al., 2017). Furthermore, it was shown in pancreatic adenocarcinomas that EMT is rather induced by an alternative program that is based on the internalization of essential proteins, rather than the induction of a genetic program. Importantly, tumor cells undergoing an EMT program based on post-translational processes displayed a different migration in cell clusters instead of single cells. Such behaviour was also observed in breast and colorectal cancers (Aiello et al., 2018).

4.3 E/m-type CTCs have the highest capacity to form lung metastasis

In order to assess the potential of CTCs to form metastases when they are in the blood stream, *i.v.* injections were performed in BALB/c mice. Nineteen days following *i.v.* injection of E-type (4T1), E/m-type (CTC6-6, CTC6-11, CTC8-12, DTC1), M/e-type (CTC8-6, CTC8-5, CTC8-1), and M-type (CTC1) cells, we found that E/m-type CTCs have the significantly highest capacity to generate lung metastases. These findings are in line with the emerging notion that carcinoma cells with a partial

rather than a full EMT comprise clinically relevant tumor cells, which contribute to metastases formation and relapse in general (Aiello and Kang, 2019; Aiello et al., 2018; Kroger et al., 2019; Li and Balazsi, 2018; Puram et al., 2018; Puram et al., 2017; Saitoh, 2018). However, since M/e- and M-type cells had decreased proliferation rates compared with E- and E/m-type cells, we assessed whether M/e- and M-type cells have identical lung metastases formation capacities as E- and E/m-type cells after longer incubation times *in vivo*. Interestingly, the second *i.v.* experiment confirmed that mice injected with M/e- and M-type CTCs had a prolonged disease-free survival in comparison with E/m-type cells (CTC6-6), based on pre-defined parameters including substantial weight loss, weakness, and dyspnoea. Importantly, the analysis of sacrificed mice in all injection groups confirmed that E/m-type CTCs induced severe lung metastases at early time points, whereas M/e- and M-type CTCs did not or barely induce any lung metastases, even at later disease time points. Noteworthy, M/e-type CTCs-injected mice bore metastases mainly in large bones, but not in the lungs. In breast cancer patients, tumor-related deaths are mainly attributed to metastases, especially in the important life-supporting organs, such as lung and liver. In contrast, metastases appearing in bones are correlated with relatively longer survival time (Siegel et al., 2019). A study analysed the survival of 9143 patients with MBC and showed that patients with bone metastasis only had a median overall survival of 38 months (95%CI 36 - 40). However, the median overall survival significantly decreased to 19 months (95%CI 17 - 20) and 17 months (95%CI 16 - 19) in patients with lung or liver metastasis, respectively (Leone et al., 2017).

Besides a higher proliferation, E/m-type CTCs have increased adhesion ability to endothelial cells, which can promote the cell-cell contact to the blood vessel endothelium and, subsequently, the extravasation into lung tissue and the formation of lung metastases (**Figure 38**). Oppositely, M/e-type cells had a reduced ability to attach to endothelium, which may restrain the extravasation and retention in lungs and thereby promote the observed formation of metastases in the major sites of large bones. The major findings and the deduced working hypothesis are schematically shown in **Figure 38**.

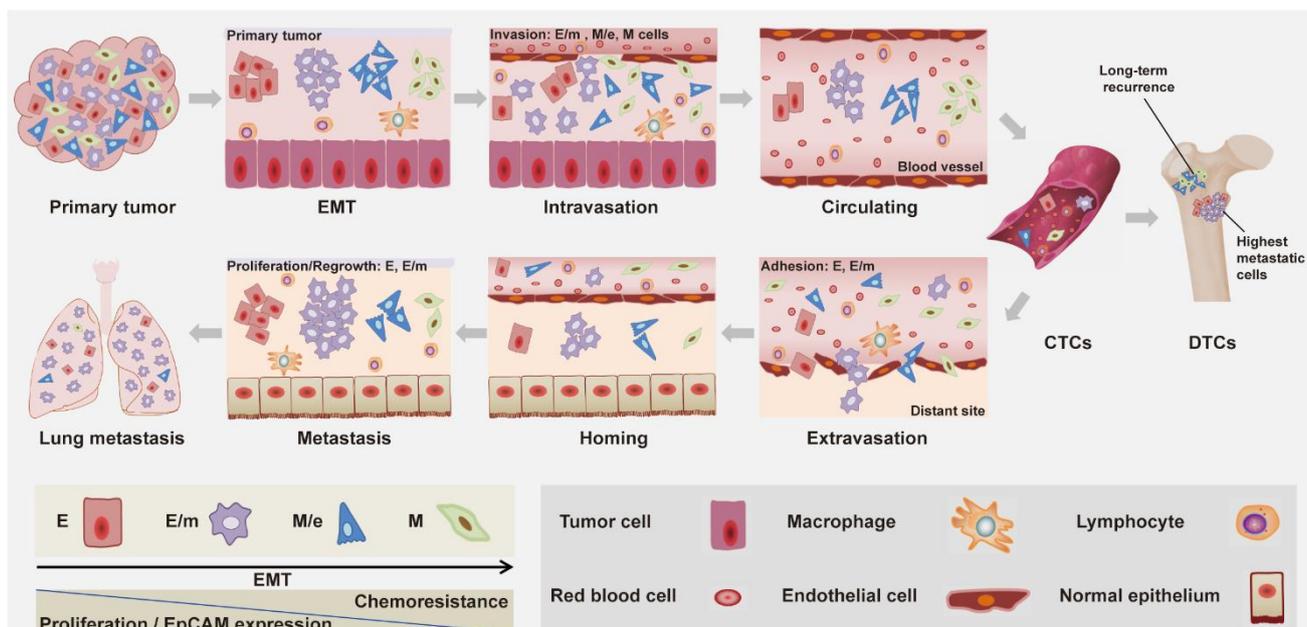


Figure 38: Schema of the influence of EMT on the metastatic process. Cancer cells undergo EMT to different degree (E-, E/m-, M/e-, M- type) with decreasing EpCAM expression and increasing chemoresistance. E/m-, M/e-, and M-type cells have improved ability to intravasate into blood vessels compared to strongly epithelial 4T1 parental cells. Tumor cells present in the blood stream are termed CTCs. E- and E/m-type CTCs have increased adhesion properties, which could allow for an improved capacity to extravasate out of blood vessels into distant sites. Tumor cells are termed DTCs when they are homing to distant organs (*e.g.* the bone marrow). In the metastatic sites, E- and E/m-type cells have a better ability to proliferate and to regrow as metastases, whereas M/e- and M-type cells are associated with long-term tumor recurrence.

4.4 The influence of EMT on metastasis

The EMT is the phenotypic change of epithelial cells to mesenchymal cells, and cells that undergo EMT will accordingly change the expression of markers, such as E-cadherin, EpCAM, Vimentin, and others (Brabletz et al., 2018b). Former studies demonstrated that EMT was correlated with therapy resistance and mesenchymal CTCs were the main cause of resistance towards chemotherapeutic drugs (Fischer et al., 2015; Yu et al., 2013; Zheng et al., 2015). The concept that EMT is mandatory for metastasis formation itself, is however still under debate. By using cell tracking techniques in an MBC and a pancreatic cancer mouse model, two publications proved that metastases-forming tumor cells had not undergone EMT and that the EMT transcription factors Snail and Twist were dispensable for metastases formation. However, mesenchymal tumor cells were responsible for chemo-resistance in both tumor models (Fischer et al., 2015; Zheng et al., 2015).

In our study, Vimentin was comparably expressed in 4T1 cells and all 4T1-derived cell lines; E-

cadherin levels differed amongst E-type (4T1), E/m-type (CTC6-6, CTC6-11, CTC8-12, DTC1), M/e-type (CTC8-6, CTC8-5, CTC8-1), and M-type (CTC1) cells. Furthermore, a comparison between 4T1, CTC1 and DTC1 cells showed that the EMT-TFs Twist and Snail were paradoxically higher in 4T1 cells. Despite new techniques that enable the tracking of EMT-related gene expressions during the metastatic process in mouse models (Fischer et al., 2015; Zheng et al., 2015), a comprehensive, all-in-one tracking of all EMT-related genes is not feasible. For example, using Twist and Snail as referential genes in the 4T1 mouse model would fail to recognise EMT-related changes in CTC1 cells. Therefore, even tracking the expression of several EMT-related genes, a contribution of EMT during metastases formation cannot be terminally excluded (Brabletz et al., 2018a; Ye et al., 2017).

In the current study, DTC1-driven CTC lines showed diverse EMT phenotypes (EMT scores range from 5 to 350), which is a reflection of phenotypic plasticity and an evidence that EMT occurs in CTCs and during tumor metastasis. Recently, many studies report the importance of a hybrid or partial EMT phenotype in cancer metastasis, which means epithelial tumor cells can benefit from a partial EMT program and are not necessary to undergo a complete change to a mesenchymal status (George et al., 2017; Hiew et al., 2018; Kai et al., 2018; Kroger et al., 2019; Saitoh, 2018). The definition of a partial or hybrid EMT phenotype is mostly based on the co-expression of epithelial and mesenchymal markers, or gene expression signatures in primary tumors (George et al., 2017; Hiew et al., 2018; Kroger et al., 2019; Tan et al., 2014). On the basis of EMT gene signatures, Tan *et al.* built up an EMT scoring system and categorized tumor samples to epithelial-, intermediate-, and mesenchymal-groups. From their findings, they concluded that higher EMT scores correlated with poorer survival in patients with ovarian and colorectal cancer, but not in patients with breast cancer (Tan et al., 2014). Similarly, George *et al.* developed an EMT scoring metric and classified primary tumors as epithelial-, E/M-, and mesenchymal-phenotype. Their result proved that patients with an epithelial-phenotype breast cancer had poorer survival in comparison with E/M- and mesenchymal-phenotype tumors, while patients with E/M-phenotype lung or ovarian tumors presented poorer outcome as compared to epithelial-phenotypes (George et al., 2017).

In cancer cells, EMT is a relative concept and how to define the edges of E and M statuses could

substantially influence the definition of EMT subtypes. Therefore, an optimal EMT scoring system is still under exploration and it may be tumor-, or even patient-specific.

Through *i.v.* injection of E/m-, M/e- and M-type CTCs, we found that the M/e-type cells can form metastasis mainly in the big bones instead of the lungs. Theoretically, bone marrow-resident M/e-type tumor cells could still release malignant cells in the blood, which could be capable of generating lung metastases. Although mice injected with M/e-type cells had tumors in multiple sites, most of them remained in the bones and lung metastases were few. This observation is a strong indication that EMT has a substantial impact on cancer organotropism. E/m-type cells preferentially metastasized in the lung, whereas M/e-type cells usually develop metastases in the bone marrow. Hence, studying functional interactions between EMT phenotypes and cancer organotropism should represent a future direction.

4.5 EpCAM expression during tumor metastasis

EpCAM is an epithelial marker on the cell surface that is widely applied for the capture of CTCs from the blood of tumor patients. In order to describe the expression of EpCAM during metastasis formation, EpCAM expression levels were tested in all 4T1-derived cell lines. Our results proved that the expression of EpCAM was higher in metastatic sites than primary tumors derived from 4T1 cells. Similar results on a higher expression level of EpCAM in metastatic sites in comparison with primary tumors were demonstrated in a clinical cohort of patients with MBC. Given that the EpCAM level is negatively correlated to the degree of EMT in CTCs, the conclusion can be made that compared to mesenchymal CTCs (EpCAM-low or -negative), epithelial CTCs (EpCAM-high) have significantly higher metastatic ability in the 4T1 mouse model. In accordance with this conclusion, the percentage of EpCAM-positive CTCs and, even more pronounced, of DTCs was associated with the presence of distant metastases, including lung metastases in patients with MBC.

In our study, dynamic EMT changes that might occur during the formation of metastases cannot be tracked in real-time in tumor cells. Furthermore, EpCAM expression and EMT phenotypes were observed in primary tumors, CTCs, DTCs, and metastases at the endpoint of animal experiments. Therefore, higher EpCAM levels in metastatic sites could be the result of M/e- or M- type CTCs that

have reverted their phenotype through the process of MET, following their homing to distant sites. Nevertheless, comparing to energy-consuming MET process (Chaffer et al., 2016; Jolly et al., 2017; Nieto et al., 2016), the outgrowth of metastasis directly from epithelial-type tumor cells (*i.e.* EpCAM-high, E-, and E/m-type cells) appears more probable.

The CELLSEARCH system is the most widely used technique for the detection of EpCAM-positive CTCs. It has been proved in a wide range of tumors containing breast, lung, colon, and esophageal carcinoma, that the numbers of CTCs are correlated to the patients' outcome (Arrazubi et al., 2019; Matsushita et al., 2015; Moussavi-Harami et al., 2014; Qi and Wang, 2017). In our DTC1-derived CTC lines, EpCAM represented a suitable surrogate marker for E-, E/m-type systemic cancer cells, which defined cells with higher metastatic potential. Hence, using the CELLSEARCH system to capture CTCs will certainly lose the information of EpCAM-negative EMT cells, but it can capture clinically relevant metastatic cells, *i.e.* EpCAM positive, E- and E/m-type cells.

Recently, more and more attention is paid to mesenchymal CTCs, and some studies proved that these cells are associated with chemo-resistance (Fischer et al., 2015; Zheng et al., 2015). In the chemoresistance assay, M/e- and M-type cells revealed more resistant to Cisplatin and Doxorubicin compared to E- and E/m-type cells. In addition to an important missing information concerning the total numbers of CTCs, the population of EpCAM-negative or mesenchymal-type CTCs can reflect the degree of the EMT process and be a marker for patient's therapeutic response. Hence, the application of EpCAM-independent CTCs/DTCs enrichment technologies is critical to collect the complete information of both EpCAM positive and negative systemic tumor cells, but addressed cells of utmost clinical relevance.

4.6 Characterization of CTCs and DTCs isolated by SE-iFISH

The SE-iFISH technology is based on the aneuploidy status of carcinoma cells and is an EpCAM-independent CTCs and DTCs isolation method (Lin et al., 2017). Aneuploidy refers to the abnormal alteration of chromosomes (including gains and losses) in a cell (Lin, 2018). Somatic aneuploidy is a common characteristic of human carcinoma cells (Gordon et al., 2012). Aneuploidy in chromosome 8 was reported to occur in various solid tumors, for example breast, pancreatic, lung, and colon cancer

(Gordon et al., 2012; Lin, 2015, 2018). Such chromosome 8 aneuploidy can be monitored using chromosome 8 enumeration probes (CEP8). Using the SE-iFISH technology (Lin, 2015; Lin et al., 2017), CTCs and DTCs were considered as CEP8-aneuploid cells with a DAPI-positive nucleus, a lack of CD45 expression, with or without EpCAM expression. Diploid EpCAM-positive cells can represent CTCs and DTCs, but could theoretically also represent normal epithelial cells aberrantly present in the blood or bone marrow. Therefore, the impact of the percentage of EpCAM-positive CTCs and DTCs on the clinical outcome of patients was assessed including and excluding CEP8-diploid EpCAM-positive cells. By doing so, it could be demonstrated that independently of the inclusion of CEP8-diploid EpCAM-positive cells, the percentage of EpCAM-positive cells was prognostic. However, following the exclusion of CEP8-diploid EpCAM-positive cells, EpCAM-positive CTCs significantly correlated with the presence of lung metastases. Diploid EpCAM-positive cells could be circulating endothelial or epithelial cells. Studies confirmed the prognostic value of circulating endothelial cells in lung, prostate, rectal, and gastric carcinoma, *etc* (Bertolini et al., 2013; Gu et al., 2015; Ha et al., 2013; Najjar et al., 2017). The invasion of surrounding normal vascular or epithelial tissues, and rebuilding tumor vessels are important steps for the metastatic cascade, and these events can cause increasing numbers of endothelial and epithelial cells in the blood (Garmy-Susini and Varner, 2005; Goon et al., 2006; Viallard and Larrivee, 2017). Therefore, including or excluding diploid EpCAM positive cells could have only minor impact on the prediction performance of the EpCAM-positivity rate. Also, using EpCAM as a surrogate marker for the EMT status, the analysis of systemic tumor cells revealed that CTCs contained proportionally more mesenchymal cells, whereas DTCs were comprised of more epithelial cells. These results were in accordance with our findings in the 4T1 mouse experiment, where a broader range of EpCAM expression was observed in CTCs, whereas DTCs generally expressed EpCAM to a high level.

As already mentioned, it remains under debate whether rather epithelial or mesenchymal CTCs comprise the major metastasis-inducing cells. Numerous studies support the notion that mesenchymal CTCs are correlated with distant metastasis and poorer prognosis (Guan et al., 2019; Li et al., 2017; Zhang et al., 2017). On the other hand, using the CELLSEARCH system, former studies confirmed

that the numbers of EpCAM-positive CTCs were correlated patients' outcomes in both non-metastatic and metastatic breast cancer (Janni et al., 2016a; Ye et al., 2019). Owing to the fact that current CTC enrichment techniques are mainly based on epithelial markers (*i.e.* EpCAM, EGFR, and Cytokeratins), and because it is uncertain which genes will be actually activated within the highly complex EMT program in tumors, EMT marker-depending CTC enrichment techniques can hardly reflect the complete and actual EMT statuses of CTCs. Furthermore, patients have usually received surgery, medication, and/or radiation interventions after the diagnosis of cancer. Epithelial tumor cells are more sensitive to chemotherapeutic drugs and could be largely eliminated after chemotherapy (Zhang et al., 2017). Accordingly, mesenchymal CTCs have been correlated to therapy resistance *in vivo*, poorer outcome, and might be linked to the incidence of long-term recurrences (Li et al., 2017; Yu et al., 2013; Zhang et al., 2017). Given epithelial CTCs are superior in proliferation and can easily form cell clusters and adhere to the endothelium, no matter before or after chemo- or radio-therapy, the proportion of epithelial CTCs can be a reliable marker for the acute metastatic risk. In our study, the EpCAM-positivity rates were used to reflect the proportions of epithelial cancer cells in CTCs and DTCs instead of absolute cell numbers, which represented better the dynamic balance of E and M status in CTCs and DTCs.

In summary, the SE-iFISH technique provides new insights in the study both EpCAM-positive and -negative systemic tumor cells, and the EpCAM-positivity rates of CTCs and DTCs were correlated with the lung metastasis status and overall survival of patients with MBC.

5. CONCLUSIONS

5.1. Summary

In summary, the current study facilitates the understanding of the influence of EMT on breast cancer progression and on its role in the formation of lung metastases. In the 4T1 mouse model, it was demonstrated that: (1) tumor cells could undergo EMT to variable degree and CTCs with different EMT types co-existed within one mouse; (2) E/m-type CTCs primarily contributed to the formation of lung metastases; (3) M/e-type CTCs were associated with long-time recurrence in bone marrow *in vivo* and resistance to treatment *in vitro*; (4) EpCAM expression was negatively correlated to EMT; (5) EpCAM expression was higher in metastatic sites (lungs, kidneys, spleens, and lymph nodes) compared to primary tumors.

These results were further verified in a clinical cohort of patients with MBC, where: (1) EpCAM expression was likewise higher in metastatic sites (lungs, livers, lymph nodes and bones) compared to primary tumors; (2) EpCAM was a useful marker for the epithelial status of systemic tumor cells, which allowed to demonstrate a higher proportion of mesenchymal CTCs and epithelial DTCs in patients; (3) the EpCAM-positivity rate of CTCs and DTCs (both including and excluding diploid EpCAM-positive cells) were associated with the metastatic status of patients and EpCAM-positive DTCs correlated with 6-months and overall survival. All in all, epithelial-type (EpCAM-positive) systemic tumor cells with a restricted mesenchymal transition are the major metastasis-initiating cells in breast cancer.

5.2. Zusammenfassung

Die Bildung von Metastasen bei Patienten mit soliden Tumoren stellt das größte Risiko für eine schlechte Prognose dar. Karzinomzellen durchlaufen eine epitheliale-mesenchymale Transition (EMT), welche für die Bildung von Metastasen verantwortlich gemacht wurde. Der exakte Beitrag einer EMT-induzierten Tumorzellheterogenität zum Krankheitsverlauf wird dennoch kontrovers diskutiert. Insbesondere der Einfluss der EMT auf das metastatische Potenzial systemischer Tumorzellen im Blut und im Knochenmark von Patienten ist bislang unzureichend untersucht. In der vorliegenden Arbeit wurde der EMT-Status bei *ex vivo* kultivierten zirkulierenden und disseminierten Tumorzellen (CTC/DTC) aus dem Blut bzw. Knochenmark in einem Mausmodell des metastasierten Brustkrebses untersucht. Epitheliale CTC mit einem unvollständigen mesenchymalen Übergang (E/m-Type) besaßen die stärkste Fähigkeit zur Bildung von Lungenmetastasen, wohingegen mesenchymale CTC mit einer geringen epithelialen Transition (M/e-Type) eine stark eingeschränkte Fähigkeit zur Metastasierung aufwiesen. Im Folgenden diente die Expression des epithelialen Zelladhäsionsmoleküls EpCAM als Surrogatmarker zur Bewertung der EMT-Heterogenität von klinischen Proben von Patienten mit metastasiertem Brustkrebs, einschließlich Metastasen, CTC und DTC. Die vermehrte Anwesenheit epithelialer CTC und insbesondere epithelialer DTC im Blut bzw. im Knochenmark dieser Patientinnen korrelierte mit dem Auftreten von Fernmetastasen und mit einem signifikant schlechteren klinischen Verlauf.

Diese Studie fördert grundlegend das Verständnis des Beitrags der EMT zur Ausbildung von Metastasen und untermauert den heterogenen EMT-Phänotypen von Tumorzellen als wichtigen Parameter für die Tumorprognose und -behandlung. Wir schlagen ferner vor, dass EpCAM-abhängige CTC-Isolationssysteme die CTC Zahlen im Blut von Patienten unterschätzen, jedoch klinisch relevante metastatische Zellen quantifizieren.

6. REFERENCES

- (2002). Breast Cancer Treatment (PDQ(R)): Health Professional Version. In PDQ Cancer Information Summaries (Bethesda (MD)).
- Aiello, N.M., Brabletz, T., Kang, Y., Nieto, M.A., Weinberg, R.A., and Stanger, B.Z. (2017). Upholding a role for EMT in pancreatic cancer metastasis. *Nature* 547, E7-E8.
- Aiello, N.M., and Kang, Y. (2019). Context-dependent EMT programs in cancer metastasis. *J Exp Med* 216, 1016-1026.
- Aiello, N.M., Maddipati, R., Norgard, R.J., Balli, D., Li, J., Yuan, S., Yamazoe, T., Black, T., Sahmoud, A., Furth, E.E., *et al.* (2018). EMT Subtype Influences Epithelial Plasticity and Mode of Cell Migration. *Dev Cell* 45, 681-695 e684.
- Akhtar, M., Haider, A., Rashid, S., and Al-Nabet, A. (2019). Paget's "Seed and Soil" Theory of Cancer Metastasis: An Idea Whose Time has Come. *Adv Anat Pathol* 26, 69-74.
- Allard, W.J., Matera, J., Miller, M.C., Repollet, M., Connelly, M.C., Rao, C., Tibbe, A.G., Uhr, J.W., and Terstappen, L.W. (2004). Tumor cells circulate in the peripheral blood of all major carcinomas but not in healthy subjects or patients with nonmalignant diseases. *Clin Cancer Res* 10, 6897-6904.
- Arrazubi, V., Mata, E., Antelo, M.L., Tarifa, A., Herrera, J., Zazpe, C., Teijeira, L., Viudez, A., Suarez, J., Hernandez, I., *et al.* (2019). Circulating Tumor Cells in Patients Undergoing Resection of Colorectal Cancer Liver Metastases. Clinical Utility for Long-Term Outcome: A Prospective Trial. *Ann Surg Oncol*.
- Aslakson, C.J., and Miller, F.R. (1992). Selective events in the metastatic process defined by analysis of the sequential dissemination of subpopulations of a mouse mammary tumor. *Cancer Res* 52, 1399-1405.
- Azevedo, A.S., Follain, G., Patthabhiraman, S., Harlepp, S., and Goetz, J.G. (2015). Metastasis of circulating tumor cells: favorable soil or suitable biomechanics, or both? *Cell Adh Migr* 9, 345-356.
- Balzar, M., Winter, M.J., de Boer, C.J., and Litvinov, S.V. (1999). The biology of the 17-1A antigen (Ep-CAM). *J Mol Med (Berl)* 77, 699-712.
- Banko, P., Lee, S.Y., Nagygyorgy, V., Zrinyi, M., Chae, C.H., Cho, D.H., and Telekes, A. (2019).

- Technologies for circulating tumor cell separation from whole blood. *J Hematol Oncol* 12, 48.
- Barbera, M.J., Puig, I., Dominguez, D., Julien-Grille, S., Guaita-Esteruelas, S., Peiro, S., Baulida, J., Franci, C., Dedhar, S., Larue, L., *et al.* (2004). Regulation of Snail transcription during epithelial to mesenchymal transition of tumor cells. *Oncogene* 23, 7345-7354.
- Baulida, J., Diaz, V.M., and Herreros, A.G. (2019). Snail1: A Transcriptional Factor Controlled at Multiple Levels. *J Clin Med* 8.
- Bednarz-Knoll, N., Alix-Panabieres, C., and Pantel, K. (2012). Plasticity of disseminating cancer cells in patients with epithelial malignancies. *Cancer Metastasis Rev* 31, 673-687.
- Bertolini, F., Shaked, Y., and Mancuso, P. (2013). On the clinical relevance of circulating endothelial cells and platelets in prostate cancer. *Br J Cancer* 108, 1387.
- Borowicz, S., Van Scoyk, M., Avasarala, S., Karuppusamy Rathinam, M.K., Tauler, J., Bikkavilli, R.K., and Winn, R.A. (2014). The soft agar colony formation assay. *J Vis Exp*, e51998.
- Brabletz, T., Kalluri, R., Nieto, M.A., and Weinberg, R.A. (2018a). EMT in cancer. *Nat Rev Cancer*.
- Brabletz, T., Kalluri, R., Nieto, M.A., and Weinberg, R.A. (2018b). EMT in cancer. *Nat Rev Cancer* 18, 128-134.
- Campbell, K. (2018). Contribution of epithelial-mesenchymal transitions to organogenesis and cancer metastasis. *Curr Opin Cell Biol* 55, 30-35.
- Chaffer, C.L., San Juan, B.P., Lim, E., and Weinberg, R.A. (2016). EMT, cell plasticity and metastasis. *Cancer Metastasis Rev* 35, 645-654.
- Chaffer, C.L., and Weinberg, R.A. (2011). A perspective on cancer cell metastasis. *Science* 331, 1559-1564.
- Chambers, A.F., Groom, A.C., and MacDonald, I.C. (2002). Dissemination and growth of cancer cells in metastatic sites. *Nat Rev Cancer* 2, 563-572.
- Chaves-Perez, A., Mack, B., Maetzel, D., Kremling, H., Eggert, C., Harreus, U., and Gires, O. (2013). EpCAM regulates cell cycle progression via control of cyclin D1 expression. *Oncogene* 32, 641-650.
- Chen, W., Hoffmann, A.D., Liu, H., and Liu, X. (2018). Organotropism: new insights into molecular

mechanisms of breast cancer metastasis. *NPJ Precis Oncol* 2, 4.

Chikaishi, Y., Yoneda, K., Ohnaga, T., and Tanaka, F. (2017). EpCAM-independent capture of circulating tumor cells with a 'universal CTC-chip'. *Oncol Rep* 37, 77-82.

Cichon, M.A., and Radisky, D.C. (2014). ROS-induced epithelial-mesenchymal transition in mammary epithelial cells is mediated by NF- κ B-dependent activation of Snail. *Oncotarget* 5, 2827-2838.

Cohen, S.J., Punt, C.J., Iannotti, N., Saidman, B.H., Sabbath, K.D., Gabrail, N.Y., Picus, J., Morse, M., Mitchell, E., Miller, M.C., *et al.* (2008). Relationship of circulating tumor cells to tumor response, progression-free survival, and overall survival in patients with metastatic colorectal cancer. *J Clin Oncol* 26, 3213-3221.

Cristofanilli, M. (2006). Circulating tumor cells, disease progression, and survival in metastatic breast cancer. *Semin Oncol* 33, S9-14.

Cristofanilli, M., Budd, G.T., Ellis, M.J., Stopeck, A., Matera, J., Miller, M.C., Reuben, J.M., Doyle, G.V., Allard, W.J., Terstappen, L.W., *et al.* (2004). Circulating tumor cells, disease progression, and survival in metastatic breast cancer. *N Engl J Med* 351, 781-791.

Dasgupta, A., Lim, A.R., and Ghajar, C.M. (2017). Circulating and disseminated tumor cells: harbingers or initiators of metastasis? *Mol Oncol* 11, 40-61.

de Bono, J.S., Scher, H.I., Montgomery, R.B., Parker, C., Miller, M.C., Tissing, H., Doyle, G.V., Terstappen, L.W., Pienta, K.J., and Raghavan, D. (2008). Circulating tumor cells predict survival benefit from treatment in metastatic castration-resistant prostate cancer. *Clin Cancer Res* 14, 6302-6309.

De Francesco, E.M., Maggiolini, M., and Musti, A.M. (2018). Crosstalk between Notch, HIF-1 α and GPER in Breast Cancer EMT. *Int J Mol Sci* 19.

Derynck, R., and Weinberg, R.A. (2019). EMT and Cancer: More Than Meets the Eye. *Dev Cell* 49, 313-316.

DeSantis, C.E., Ma, J., Goding Sauer, A., Newman, L.A., and Jemal, A. (2017). Breast cancer statistics, 2017, racial disparity in mortality by state. *CA Cancer J Clin* 67, 439-448.

- Dey, P., Rathod, M., and De, A. (2019). Targeting stem cells in the realm of drug-resistant breast cancer. *Breast Cancer (Dove Med Press)* *11*, 115-135.
- Dong, X. (2018). Current Strategies for Brain Drug Delivery. *Theranostics* *8*, 1481-1493.
- Einsele, H., Ehninger, G., Schneider, E.M., Kruger, G.F., Vallbracht, A., Dopfer, R., Schmidt, H., Waller, H.D., and Muller, C.A. (1988). High frequency of graft-versus-host-like syndromes following syngeneic bone marrow transplantation. *Transplantation* *45*, 579-585.
- Fischer, K.R., Durrans, A., Lee, S., Sheng, J., Li, F., Wong, S.T., Choi, H., El Rayes, T., Ryu, S., Troeger, J., *et al.* (2015). Epithelial-to-mesenchymal transition is not required for lung metastasis but contributes to chemoresistance. *Nature* *527*, 472-476.
- Franken, N.A., Rodermond, H.M., Stap, J., Haveman, J., and van Bree, C. (2006). Clonogenic assay of cells in vitro. *Nat Protoc* *1*, 2315-2319.
- Gabriel, M.T., Calleja, L.R., Chalopin, A., Ory, B., and Heymann, D. (2016). Circulating Tumor Cells: A Review of Non-EpCAM-Based Approaches for Cell Enrichment and Isolation. *Clin Chem* *62*, 571-581.
- Gao, Y., Bado, I., Wang, H., Zhang, W., Rosen, J.M., and Zhang, X.H. (2019). Metastasis Organotropism: Redefining the Congenial Soil. *Dev Cell* *49*, 375-391.
- Garmy-Susini, B., and Varner, J.A. (2005). Circulating endothelial progenitor cells. *Br J Cancer* *93*, 855-858.
- Ge, F., Zhang, H., Wang, D.D., Li, L., and Lin, P.P. (2015). Enhanced detection and comprehensive in situ phenotypic characterization of circulating and disseminated heteroploid epithelial and glioma tumor cells. *Oncotarget* *6*, 27049-27064.
- George, J.T., Jolly, M.K., Xu, S., Somarelli, J.A., and Levine, H. (2017). Survival Outcomes in Cancer Patients Predicted by a Partial EMT Gene Expression Scoring Metric. *Cancer Res* *77*, 6415-6428.
- Gires, O., and Stoecklein, N.H. (2014). Dynamic EpCAM expression on circulating and disseminating tumor cells: causes and consequences. *Cell Mol Life Sci* *71*, 4393-4402.
- Gloushankova, N.A., Rubtsova, S.N., and Zhitnyak, I.Y. (2017). Cadherin-mediated cell-cell

interactions in normal and cancer cells. *Tissue Barriers* 5, e1356900.

Goetz, M.P., Gradishar, W.J., Anderson, B.O., Abraham, J., Aft, R., Allison, K.H., Blair, S.L., Burstein, H.J., Dang, C., Elias, A.D., *et al.* (2019). NCCN Guidelines Insights: Breast Cancer, Version 3.2018. *J Natl Compr Canc Netw* 17, 118-126.

Gonzalez, B., Denzel, S., Mack, B., Conrad, M., and Gires, O. (2009). EpCAM is involved in maintenance of the murine embryonic stem cell phenotype. *Stem Cells* 27, 1782-1791.

Gonzalez, D.M., and Medici, D. (2014). Signaling mechanisms of the epithelial-mesenchymal transition. *Sci Signal* 7, re8.

Goon, P.K., Lip, G.Y., Boos, C.J., Stonelake, P.S., and Blann, A.D. (2006). Circulating endothelial cells, endothelial progenitor cells, and endothelial microparticles in cancer. *Neoplasia* 8, 79-88.

Gordon, D.J., Resio, B., and Pellman, D. (2012). Causes and consequences of aneuploidy in cancer. *Nat Rev Genet* 13, 189-203.

Gorges, T.M., Tinhofer, I., Drosch, M., Rose, L., Zollner, T.M., Krahn, T., and von Ahsen, O. (2012). Circulating tumour cells escape from EpCAM-based detection due to epithelial-to-mesenchymal transition. *BMC Cancer* 12, 178.

Gradilone, A., Naso, G., Raimondi, C., Cortesi, E., Gandini, O., Vincenzi, B., Saltarelli, R., Chiapparino, E., Spremberg, F., Cristofanilli, M., *et al.* (2011). Circulating tumor cells (CTCs) in metastatic breast cancer (MBC): prognosis, drug resistance and phenotypic characterization. *Ann Oncol* 22, 86-92.

Gu, W., Sun, W., Guo, C., Yan, Y., Liu, M., Yao, X., Yang, B., and Zheng, J. (2015). Culture and Characterization of Circulating Endothelial Progenitor Cells in Patients with Renal Cell Carcinoma. *J Urol* 194, 214-222.

Guan, X., Ma, F., Li, C., Wu, S., Hu, S., Huang, J., Sun, X., Wang, J., Luo, Y., Cai, R., *et al.* (2019). The prognostic and therapeutic implications of circulating tumor cell phenotype detection based on epithelial-mesenchymal transition markers in the first-line chemotherapy of HER2-negative metastatic breast cancer. *Cancer Commun (Lond)* 39, 1.

Ha, X., Zhao, M., Zhao, H., Peng, J., Deng, Z., Dong, J., Yang, X., Zhao, Y., and Ju, J. (2013).

- Identification and clinical significance of circulating endothelial progenitor cells in gastric cancer. *Biomarkers* 18, 487-492.
- Hay, E.D. (1995). An overview of epithelio-mesenchymal transformation. *Acta Anat (Basel)* 154, 8-20.
- Herlyn, D., Herlyn, M., Steplewski, Z., and Koprowski, H. (1979). Monoclonal antibodies in cell-mediated cytotoxicity against human melanoma and colorectal carcinoma. *Eur J Immunol* 9, 657-659.
- Hiew, M.S.Y., Cheng, H.P., Huang, C.J., Chong, K.Y., Cheong, S.K., Choo, K.B., and Kamarul, T. (2018). Incomplete cellular reprogramming of colorectal cancer cells elicits an epithelial/mesenchymal hybrid phenotype. *J Biomed Sci* 25, 57.
- Hsu, Y.T., Osmulski, P., Wang, Y., Huang, Y.W., Liu, L., Ruan, J., Jin, V.X., Kirma, N.B., Gaczynska, M.E., and Huang, T.H. (2016). EpCAM-Regulated Transcription Exerts Influences on Nanomechanical Properties of Endometrial Cancer Cells That Promote Epithelial-to-Mesenchymal Transition. *Cancer Res* 76, 6171-6182.
- Huang, H.P., Chen, P.H., Yu, C.Y., Chuang, C.Y., Stone, L., Hsiao, W.C., Li, C.L., Tsai, S.C., Chen, K.Y., Chen, H.F., *et al.* (2011). Epithelial cell adhesion molecule (EpCAM) complex proteins promote transcription factor-mediated pluripotency reprogramming. *J Biol Chem* 286, 33520-33532.
- Huang, R.Y., Wong, M.K., Tan, T.Z., Kuay, K.T., Ng, A.H., Chung, V.Y., Chu, Y.S., Matsumura, N., Lai, H.C., Lee, Y.F., *et al.* (2013). 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* 4, e915.
- Imrich, S., Hachmeister, M., and Gires, O. (2012). EpCAM and its potential role in tumor-initiating cells. *Cell Adh Migr* 6, 30-38.
- Janni, W., Rack, B., Terstappen, L.W., Pierga, J.Y., Taran, F.A., Fehm, T., Hall, C., de Groot, M., Bidard, F.C., Friedl, T.W., *et al.* (2016a). Pooled Analysis of the Prognostic Relevance of Circulating Tumor Cells in Primary Breast Cancer. *Clin Cancer Res*.
- Janni, W.J., Rack, B., Terstappen, L.W., Pierga, J.Y., Taran, F.A., Fehm, T., Hall, C., de Groot,

- M.R., Bidard, F.C., Friedl, T.W., *et al.* (2016b). Pooled Analysis of the Prognostic Relevance of Circulating Tumor Cells in Primary Breast Cancer. *Clin Cancer Res* 22, 2583-2593.
- Jolly, M.K., Ware, K.E., Gilja, S., Somarelli, J.A., and Levine, H. (2017). EMT and MET: necessary or permissive for metastasis? *Mol Oncol* 11, 755-769.
- Josse, S.A., Gorges, T.M., and Pantel, K. (2015). Biology, detection, and clinical implications of circulating tumor cells. *EMBO Mol Med* 7, 1-11.
- Jordan, N.V., Johnson, G.L., and Abell, A.N. (2011). Tracking the intermediate stages of epithelial-mesenchymal transition in epithelial stem cells and cancer. *Cell Cycle* 10, 2865-2873.
- Kai, K., Iwamoto, T., Zhang, D., Shen, L., Takahashi, Y., Rao, A., Thompson, A., Sen, S., and Ueno, N.T. (2018). CSF-1/CSF-1R axis is associated with epithelial/mesenchymal hybrid phenotype in epithelial-like inflammatory breast cancer. *Sci Rep* 8, 9427.
- Kalluri, R., and Weinberg, R.A. (2009). The basics of epithelial-mesenchymal transition. *J Clin Invest* 119, 1420-1428.
- Kang, Y., and Pantel, K. (2013). Tumor cell dissemination: emerging biological insights from animal models and cancer patients. *Cancer cell* 23, 573-581.
- Khalili, A.A., and Ahmad, M.R. (2015). A Review of Cell Adhesion Studies for Biomedical and Biological Applications. *Int J Mol Sci* 16, 18149-18184.
- Khetani, S., Mohammadi, M., and Nezhad, A.S. (2018). Filter-based isolation, enrichment, and characterization of circulating tumor cells. *Biotechnol Bioeng* 115, 2504-2529.
- Kim, D.H., Xing, T., Yang, Z., Dudek, R., Lu, Q., and Chen, Y.H. (2017). Epithelial Mesenchymal Transition in Embryonic Development, Tissue Repair and Cancer: A Comprehensive Overview. *J Clin Med* 7.
- Kim, M.Y., Oskarsson, T., Acharyya, S., Nguyen, D.X., Zhang, X.H., Norton, L., and Massague, J. (2009). Tumor self-seeding by circulating cancer cells. *Cell* 139, 1315-1326.
- Kroger, C., Afeyan, A., Mraz, J., Eaton, E.N., Reinhardt, F., Khodor, Y.L., Thiru, P., Bierie, B., Ye, X., Burge, C.B., *et al.* (2019). Acquisition of a hybrid E/M state is essential for tumorigenicity of basal breast cancer cells. *Proc Natl Acad Sci U S A* 116, 7353-7362.

- Ladwein, M., Pape, U.F., Schmidt, D.S., Schnolzer, M., Fiedler, S., Langbein, L., Franke, W.W., Moldenhauer, G., and Zoller, M. (2005). The cell-cell adhesion molecule EpCAM interacts directly with the tight junction protein claudin-7. *Exp Cell Res* 309, 345-357.
- Lambert, A.W., Pattabiraman, D.R., and Weinberg, R.A. (2017). Emerging Biological Principles of Metastasis. *Cell* 168, 670-691.
- Leal, F., Liutti, V.T., Antunes dos Santos, V.C., Novis de Figueiredo, M.A., Macedo, L.T., Rinck Junior, J.A., and Sasse, A.D. (2015). Neoadjuvant endocrine therapy for resectable breast cancer: A systematic review and meta-analysis. *Breast* 24, 406-412.
- Leone, B.A., Vallejo, C.T., Romero, A.O., Machiavelli, M.R., Perez, J.E., Leone, J., and Leone, J.P. (2017). Prognostic impact of metastatic pattern in stage IV breast cancer at initial diagnosis. *Breast Cancer Res Treat* 161, 537-548.
- Li, C., and Balazsi, G. (2018). A landscape view on the interplay between EMT and cancer metastasis. *NPJ Syst Biol Appl* 4, 34.
- Li, S., Chen, Q., Li, H., Wu, Y., Feng, J., and Yan, Y. (2017). Mesenchymal circulating tumor cells (CTCs) and OCT4 mRNA expression in CTCs for prognosis prediction in patients with non-small-cell lung cancer. *Clin Transl Oncol* 19, 1147-1153.
- Liljefors, M., Nilsson, B., Fagerberg, J., Ragnhammar, P., Mellstedt, H., and Frodin, J.E. (2005). Clinical effects of a chimeric anti-EpCAM monoclonal antibody in combination with granulocyte-macrophage colony-stimulating factor in patients with metastatic colorectal carcinoma. *Int J Oncol* 26, 1581-1589.
- Lim, B., and Hortobagyi, G.N. (2016). Current challenges of metastatic breast cancer. *Cancer Metastasis Rev* 35, 495-514.
- Lin, P.P. (2015). Integrated EpCAM-independent subtraction enrichment and iFISH strategies to detect and classify disseminated and circulating tumors cells. *Clin Transl Med* 4, 38.
- Lin, P.P. (2018). Aneuploid CTC and CEC. *Diagnostics (Basel)* 8.
- Lin, P.P., Gires, O., Wang, D.D., Li, L., and Wang, H. (2017). Comprehensive in situ co-detection of aneuploid circulating endothelial and tumor cells. *Sci Rep* 7, 9789.

- Liu, Q., Zhang, H., Jiang, X., Qian, C., Liu, Z., and Luo, D. (2017). Factors involved in cancer metastasis: a better understanding to "seed and soil" hypothesis. *Mol Cancer* 16, 176.
- Liu, W.H., Chen, M.T., Wang, M.L., Lee, Y.Y., Chiou, G.Y., Chien, C.S., Huang, P.I., Chen, Y.W., Huang, M.C., Chiou, S.H., *et al.* (2015). Cisplatin-selected resistance is associated with increased motility and stem-like properties via activation of STAT3/Snail axis in atypical teratoid/rhabdoid tumor cells. *Oncotarget* 6, 1750-1768.
- Lowe, D.J., and Raj, K. (2015). Quantitation of Endothelial Cell Adhesiveness In Vitro. *J Vis Exp*, e52924.
- Lu, T.Y., Lu, R.M., Liao, M.Y., Yu, J., Chung, C.H., Kao, C.F., and Wu, H.C. (2010). Epithelial cell adhesion molecule regulation is associated with the maintenance of the undifferentiated phenotype of human embryonic stem cells. *J Biol Chem* 285, 8719-8732.
- Maetzel, D., Denzel, S., Mack, B., Canis, M., Went, P., Benk, M., Kieu, C., Papior, P., Baeuerle, P.A., Munz, M., *et al.* (2009). Nuclear signalling by tumour-associated antigen EpCAM. *Nat Cell Biol* 11, 162-171.
- Marshall, J. (2011). Transwell((R)) invasion assays. *Methods Mol Biol* 769, 97-110.
- Massague, J., and Obenauf, A.C. (2016). Metastatic colonization by circulating tumour cells. *Nature* 529, 298-306.
- Matsushita, D., Uenosono, Y., Arigami, T., Yanagita, S., Nishizono, Y., Hagihara, T., Hirata, M., Haraguchi, N., Arima, H., Kijima, Y., *et al.* (2015). Clinical Significance of Circulating Tumor Cells in Peripheral Blood of Patients with Esophageal Squamous Cell Carcinoma. *Ann Surg Oncol* 22, 3674-3680.
- McDonald, E.S., Clark, A.S., Tchou, J., Zhang, P., and Freedman, G.M. (2016). Clinical Diagnosis and Management of Breast Cancer. *J Nucl Med* 57 *Suppl 1*, 9S-16S.
- Miller, B.E., Miller, F.R., and Heppner, G.H. (1981). Interactions between tumor subpopulations affecting their sensitivity to the antineoplastic agents cyclophosphamide and methotrexate. *Cancer Res* 41, 4378-4381.
- Miller, B.E., Miller, F.R., Wilburn, D.J., and Heppner, G.H. (1987). Analysis of tumour cell

composition in tumours composed of paired mixtures of mammary tumour cell lines. *Br J Cancer* 56, 561-569.

Mosmann, T. (1983). Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. *J Immunol Methods* 65, 55-63.

Moussavi-Harami, S.F., Wisinski, K.B., and Beebe, D.J. (2014). Circulating Tumor Cells in Metastatic Breast Cancer: A Prognostic and Predictive Marker. *J Patient Cent Res Rev* 1, 85-92.

Moustakas, A., and Heldin, C.H. (2016). Mechanisms of TGFbeta-Induced Epithelial-Mesenchymal Transition. *J Clin Med* 5.

Mrozik, K.M., Blaschuk, O.W., Cheong, C.M., Zannettino, A.C.W., and Vandyke, K. (2018). N-cadherin in cancer metastasis, its emerging role in haematological malignancies and potential as a therapeutic target in cancer. *BMC Cancer* 18, 939.

Munz, M., Baeuerle, P.A., and Gires, O. (2009). The emerging role of EpCAM in cancer and stem cell signaling. *Cancer Res* 69, 5627-5629.

Munz, M., Kieu, C., Mack, B., Schmitt, B., Zeidler, R., and Gires, O. (2004). The carcinoma-associated antigen EpCAM upregulates c-myc and induces cell proliferation. *Oncogene* 23, 5748-5758.

Nagrath, S., Sequist, L.V., Maheswaran, S., Bell, D.W., Irimia, D., Ulkus, L., Smith, M.R., Kwak, E.L., Digumarthy, S., Muzikansky, A., *et al.* (2007). Isolation of rare circulating tumour cells in cancer patients by microchip technology. *Nature* 450, 1235-1239.

Najjar, F., Alammari, M., Al-Massarani, G., Almalla, N., Aljapawe, A., and Ikhtiar, A. (2017). Circulating endothelial cells and microparticles for prediction of tumor progression and outcomes in advanced non-small cell lung cancer. *Cancer Biomark* 20, 333-343.

Nan, X., Wang, J., Liu, H.N., Wong, S.T.C., and Zhao, H. (2019). Epithelial-Mesenchymal Plasticity in Organotropism Metastasis and Tumor Immune Escape. *J Clin Med* 8.

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

Obenauf, A.C., Zou, Y., Ji, A.L., Vanharanta, S., Shu, W., Shi, H., Kong, X., Bosenberg, M.C., Wiesner, T., Rosen, N., *et al.* (2015). Therapy-induced tumour secretomes promote resistance and

tumour progression. *Nature* 520, 368-372.

Oeffinger, K.C., Fontham, E.T., Etzioni, R., Herzig, A., Michaelson, J.S., Shih, Y.C., Walter, L.C., Church, T.R., Flowers, C.R., LaMonte, S.J., *et al.* (2015). Breast Cancer Screening for Women at Average Risk: 2015 Guideline Update From the American Cancer Society. *JAMA* 314, 1599-1614.

Ordonez-Moreno, A., Rodriguez-Monteros, C., Cortes-Reynosa, P., Perez-Carreón, J.I., and Perez Salazar, E. (2017). Erythropoietin Induces an Epithelial to Mesenchymal Transition-Like Process in Mammary Epithelial Cells MCF10A. *J Cell Biochem* 118, 2983-2992.

Pace, L.E., and Keating, N.L. (2014). A systematic assessment of benefits and risks to guide breast cancer screening decisions. *JAMA* 311, 1327-1335.

Pan, M., Schinke, H., Luxemburger, E., Kranz, G., Shakhtour, J., Libl, D., Huang, Y., Gaber, A., Pavsic, M., Lenarcic, B., *et al.* (2018). EpCAM ectodomain EpEX is a ligand of EGFR that counteracts EGF-mediated epithelial-mesenchymal transition through modulation of phospho-ERK1/2 in head and neck cancers. *PLoS Biol* 16, e2006624.

Pastushenko, I., and Blanpain, C. (2019). EMT Transition States during Tumor Progression and Metastasis. *Trends Cell Biol* 29, 212-226.

Pastushenko, I., Brisebarre, A., Sifrim, A., Fioramonti, M., Revenco, T., Boumahdi, S., Van Keymeulen, A., Brown, D., Moers, V., Lemaire, S., *et al.* (2018). Identification of the tumour transition states occurring during EMT. *Nature* 556, 463-468.

Puram, S.V., Parikh, A.S., and Tirosh, I. (2018). Single cell RNA-seq highlights a role for a partial EMT in head and neck cancer. *Mol Cell Oncol* 5, e1448244.

Puram, S.V., Tirosh, I., Parikh, A.S., Patel, A.P., Yizhak, K., Gillespie, S., Rodman, C., Luo, C.L., Mroz, E.A., Emerick, K.S., *et al.* (2017). Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer. *Cell* 171, 1611-1624 e1624.

Qi, Y., and Wang, W. (2017). Clinical significance of circulating tumor cells in squamous cell lung cancer patients. *Cancer Biomark* 18, 161-167.

Rashid, O.M., and Takabe, K. (2014). Does removal of the primary tumor in metastatic breast cancer improve survival? *J Womens Health (Larchmt)* 23, 184-188.

- Roulot, A., Hequet, D., Guinebretiere, J.M., Vincent-Salomon, A., Lerebours, F., Dubot, C., and Rouzier, R. (2016). Tumoral heterogeneity of breast cancer. *Ann Biol Clin (Paris)* 74, 653-660.
- Saitoh, M. (2018). Involvement of partial EMT in cancer progression. *J Biochem* 164, 257-264.
- Saitoh, M., Endo, K., Furuya, S., Minami, M., Fukasawa, A., Imamura, T., and Miyazawa, K. (2016). STAT3 integrates cooperative Ras and TGF-beta signals that induce Snail expression. *Oncogene* 35, 1049-1057.
- Sarrach, S., Huang, Y., Niedermeyer, S., Hachmeister, M., Fischer, L., Gille, S., Pan, M., Mack, B., Kranz, G., Libl, D., *et al.* (2018). Spatiotemporal patterning of EpCAM is important for murine embryonic endo- and mesodermal differentiation. *Sci Rep* 8, 1801.
- Siegel, R.L., Miller, K.D., and Jemal, A. (2019). Cancer statistics, 2019. *CA Cancer J Clin* 69, 7-34.
- Sikandar, S.S., Kuo, A.H., Kalisky, T., Cai, S., Zabala, M., Hsieh, R.W., Lobo, N.A., Scheeren, F.A., Sim, S., Qian, D., *et al.* (2017). Role of epithelial to mesenchymal transition associated genes in mammary gland regeneration and breast tumorigenesis. *Nat Commun* 8, 1669.
- Smith, R.A., Andrews, K.S., Brooks, D., Fedewa, S.A., Manassaram-Baptiste, D., Saslow, D., and Wender, R.C. (2019). Cancer screening in the United States, 2019: A review of current American Cancer Society guidelines and current issues in cancer screening. *CA Cancer J Clin*.
- Sobin LH, G.M., Wittekind C. (2009). *TNM Classification of Malignant Tumours*. 7th ed. . American Joint Committee on Cancer.
- Stemmler, M.P., Eccles, R.L., Brabletz, S., and Brabletz, T. (2019). Non-redundant functions of EMT transcription factors. *Nat Cell Biol* 21, 102-112.
- Stockert, J.C., Horobin, R.W., Colombo, L.L., and Blazquez-Castro, A. (2018). Tetrazolium salts and formazan products in Cell Biology: Viability assessment, fluorescence imaging, and labeling perspectives. *Acta Histochem* 120, 159-167.
- Stott, S.L., Hsu, C.H., Tsukrov, D.I., Yu, M., Miyamoto, D.T., Waltman, B.A., Rothenberg, S.M., Shah, A.M., Smas, M.E., Korir, G.K., *et al.* (2010). Isolation of circulating tumor cells using a microvortex-generating herringbone-chip. *Proc Natl Acad Sci U S A* 107, 18392-18397.
- Tam, W.L., and Weinberg, R.A. (2013). The epigenetics of epithelial-mesenchymal plasticity in

cancer. *Nat Med* 19, 1438-1449.

Tan, T.Z., Miow, Q.H., Miki, Y., Noda, T., Mori, S., Huang, R.Y., and Thiery, J.P. (2014).

Epithelial-mesenchymal transition spectrum quantification and its efficacy in deciphering survival and drug responses of cancer patients. *EMBO Mol Med* 6, 1279-1293.

Thiery, J.P., and Lim, C.T. (2013). Tumor dissemination: an EMT affair. *Cancer cell* 23, 272-273.

Thompson, E.W., and Nagaraj, S.H. (2018). Transition states that allow cancer to spread. *Nature* 556, 442-444.

Viallard, C., and Larrivee, B. (2017). Tumor angiogenesis and vascular normalization: alternative therapeutic targets. *Angiogenesis* 20, 409-426.

Voon, D.C., Huang, R.Y., Jackson, R.A., and Thiery, J.P. (2017). The EMT spectrum and therapeutic opportunities. *Mol Oncol* 11, 878-891.

Wang, H., Stoecklein, N.H., Lin, P.P., and Gires, O. (2017a). Circulating and disseminated tumor cells: diagnostic tools and therapeutic targets in motion. *Oncotarget* 8, 1884-1912.

Wang, L., Balasubramanian, P., Chen, A.P., Kummar, S., Evrard, Y.A., and Kinders, R.J. (2016). Promise and limits of the CellSearch platform for evaluating pharmacodynamics in circulating tumor cells. *Semin Oncol* 43, 464-475.

Wang, L., Li, Y., Xu, J., Zhang, A., Wang, X., Tang, R., Zhang, X., Yin, H., Liu, M., Wang, D.D., *et al.* (2018a). Quantified postsurgical small cell size CTCs and EpCAM(+) circulating tumor stem cells with cytogenetic abnormalities in hepatocellular carcinoma patients determine cancer relapse. *Cancer Lett* 412, 99-107.

Wang, M.H., Sun, R., Zhou, X.M., Zhang, M.Y., Lu, J.B., Yang, Y., Zeng, L.S., Yang, X.Z., Shi, L., Xiao, R.W., *et al.* (2018b). Epithelial cell adhesion molecule overexpression regulates epithelial-mesenchymal transition, stemness and metastasis of nasopharyngeal carcinoma cells via the PTEN/AKT/mTOR pathway. *Cell Death Dis* 9, 2.

Wang, R., Jin, C., and Hu, X. (2017b). Evidence of drug-response heterogeneity rapidly generated from a single cancer cell. *Oncotarget* 8, 41113-41124.

Warrier, S., Tapia, G., Goltzman, D., and Beith, J. (2016). An update in breast cancer screening and

- management. *Womens Health (Lond)* 12, 229-239.
- Went, P.T., Lugli, A., Meier, S., Bundi, M., Mirlacher, M., Sauter, G., and Dirnhofer, S. (2004). Frequent EpCam protein expression in human carcinomas. *Hum Pathol* 35, 122-128.
- Wu, C., Zhuang, Y., Jiang, S., Liu, S., Zhou, J., Wu, J., Teng, Y., Xia, B., Wang, R., and Zou, X. (2016). Interaction between Wnt/beta-catenin pathway and microRNAs regulates epithelial-mesenchymal transition in gastric cancer (Review). *Int J Oncol* 48, 2236-2246.
- Wu, C.J., Mannan, P., Lu, M., and Udey, M.C. (2013). Epithelial cell adhesion molecule (EpCAM) regulates claudin dynamics and tight junctions. *J Biol Chem* 288, 12253-12268.
- Yan, J., Gumireddy, K., Li, A., and Huang, Q. (2013). Regulation of mesenchymal phenotype by MicroRNAs in cancer. *Curr Cancer Drug Targets* 13, 930-934.
- Ye, X., Brabletz, T., Kang, Y., Longmore, G.D., Nieto, M.A., Stanger, B.Z., Yang, J., and Weinberg, R.A. (2017). Upholding a role for EMT in breast cancer metastasis. *Nature* 547, E1-E3.
- Ye, Z., Wang, C., Wan, S., Mu, Z., Zhang, Z., Abu-Khalaf, M.M., Fellin, F.M., Silver, D.P., Neupane, M., Jaslow, R.J., *et al.* (2019). Association of clinical outcomes in metastatic breast cancer patients with circulating tumour cell and circulating cell-free DNA. *Eur J Cancer* 106, 133-143.
- Yu, M., Bardia, A., Aceto, N., Bersani, F., Madden, M.W., Donaldson, M.C., Desai, R., Zhu, H., Comaills, V., Zheng, Z., *et al.* (2014). Cancer therapy. Ex vivo culture of circulating breast tumor cells for individualized testing of drug susceptibility. *Science* 345, 216-220.
- Yu, M., Bardia, A., Wittner, B.S., Stott, S.L., Smas, M.E., Ting, D.T., Isakoff, S.J., Ciciliano, J.C., Wells, M.N., Shah, A.M., *et al.* (2013). Circulating breast tumor cells exhibit dynamic changes in epithelial and mesenchymal composition. *Science* 339, 580-584.
- Zhang, J., Tian, X.J., and Xing, J. (2016). Signal Transduction Pathways of EMT Induced by TGF-beta, SHH, and WNT and Their Crosstalks. *J Clin Med* 5.
- Zhang, S., Wu, T., Peng, X., Liu, J., Liu, F., Wu, S., Liu, S., Dong, Y., Xie, S., and Ma, S. (2017). Mesenchymal phenotype of circulating tumor cells is associated with distant metastasis in breast cancer patients. *Cancer Manag Res* 9, 691-700.
- Zheng, X., Carstens, J.L., Kim, J., Scheible, M., Kaye, J., Sugimoto, H., Wu, C.C., LeBleu, V.S., and

Kalluri, R. (2015). Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer. *Nature* 527, 525-530.

Zhou, J.J., Meng, Z., He, X.Y., Cheng, D., Ye, H.L., Deng, X.G., and Chen, R.F. (2017). Hepatitis C virus core protein increases Snail expression and induces epithelial-mesenchymal transition through the signal transducer and activator of transcription 3 pathway in hepatoma cells. *Hepatol Res* 47, 574-583

7. LIST OF FIGURES

Figure 1: Scheme of the 4T1 mouse model.....	36
Figure 2: EMT phenotypes of 4T1, CTC1, and DTC1 cells <i>in vitro</i>	37
Figure 3: CD45 expression in 4T1, CTC1, and DTC1 cells was tested by flow cytometry.....	37
Figure 4: 6-TG treatment of 4T1 cells.....	38
Figure 5: Immunocytochemistry staining of EpCAM, E-cadherin, Cytokeratin, and Vimentin in 4T1, CTC1, and DTC1 cell lines.....	38
Figure 6: Flow cytometry analysis of EpCAM expression in 4T1, CTC, and DTC1 cells.....	39
Figure 7: mRNA expression levels of epithelial and EMT markers in 4T1, CTC1, and DTC1 cells.	40
Figure 8: Comparison of cell metabolism and proliferation between 4T1, CTC1, and DTC1 cells..	41
Figure 9: 2D colony formation capacity of 4T1, CTC1, and DTC1 cells.....	42
Figure 10: 3D colony formation capacity of 4T1, CTC1, and DTC1 cells.....	43
Figure 11: Adhesion of 4T1, CTC1, and DTC1 cells to bEnd.3 endothelial cells, Matrigel, and Gelatin.....	43
Figure 12: Invasion capacity of 4T1, CTC1, and DTC1 cells.....	44
Figure 13: Tumorigenic ability of 4T1, CTC1, and DTC1 cells <i>in vivo</i>	45
Figure 14: Immunohistochemistry (IHC) staining of EpCAM and Vimentin in primary tumors of 4T1, CTC1, and DTC1 cells-injected groups of mice.....	46
Figure 15: Frequencies of <i>ex vivo</i> cultures of 4T1, CTC1, and DTC1 cells following syngeneic transplantation.....	46
Figure 16: Tumorigenic ability of 4T1, DTC1, and CTC1 cells with increased cell numbers in the CTC1-injected group of mice.....	47
Figure 17: Metastatic potential of 4T1, DTC1, and CTC1 cells.....	48
Figure 18: CD45 expression on DTC1-derived CTC lines.....	49
Figure 19: EMT scores of DTC1-derived CTC lines.....	50
Figure 20: Immunocytochemistry staining of control, Vimentin, EpCAM, and E-cadherin in E/m	

type (CTC6-6, CTC6-11, CTC8-12) and M/e type (CTC8-6, CTC8-5, CTC8-1) CTCs	51
Figure 21: Expression of EMT-related genes in E/m- and M/e-type CTCs	52
Figure 22: Cell metabolism and proliferation of E/m- and M/e- type DTC1-derived CTC lines	53
Figure 23: Adhesion and invasion assay of E/m- and M/e- type DTC1-derived CTC lines	53
Figure 24: Chemoresistance of E-, E/m-, M/e-, and M- type cells.....	54
Figure 25: Lung metastasis formation by E-, E/m-, M/e-, M- type cells.....	57
Figure 26: Metastasis formation by E/m-, M/e-, and M- type CTCs.....	58
Figure 27: EMT scores and EpCAM expressions in 4T1-, CTC1- and DTC1-SCCs and in DTC1- derived CTC lines	60
Figure 28: Cluster plot analysis of Spearman's rank correlation between EpCAM expression and EMT score in DTC1-derived CTC lines	61
Figure 29: EpCAM expression of 4T1-, CTC1-, and DTC1-derived cell lines	62
Figure 30: EpCAM expression in primary tumors and corresponding metastases in patients with MBC.....	63
Figure 31: Characteristics of patients	65
Figure 32: SE-iFISH enrichment and detection of CTCs and DTCs.....	66
Figure 33: Paired CTCs and DTCs numbers.....	67
Figure 34: Proportion of EpCAM expression in CTCs and DTCs.....	68
Figure 35: Pie charts show (A) the ploidy status of CTCs and DTCs, and (B) the cell size compared with WBC.....	68
Figure 36: Ploidy status of CTCs and DTCs in relation to the cell size, referring to smaller, similar, and larger than WBC.....	68
Figure 37: EpCAM-positivity rate as a clinical parameter is associated with the patient's metastatic status and can predict the clinical outcome.	70
Figure 38: Schema of the influence of EMT on the metastatic process	77

8. TABLE LIST

Table 1 List of chemicals	12
Table 2 List of experimental kits	13
Table 3 List of antibodies	13
Table 4 List of qRT-PCR primers.....	15
Table 5 List of cell lines	18
Table 6 List of equipment	19
Table 7 List of consumables.....	21
Table 8 List of software in present project.....	21

9. ABBREVIATIONS

Abbreviation	Annotation
°C	Celsius degree
6-TG	6-Thioguanine
bp	Base pairs
BSA	Bovine serum albumin
CD	Cluster of differentiation
cDNA	Complementary DNA
CEP8	Chromosome 8
CK	Cytokeratin
CT	Computed tomography
CTCs	Circulating tumor cells
ddH ₂ O	Double distilled water
del	Deletion
der	Derived
DMEM	Dulbecco`s Modified Eagle Medium
dmin	Double minute chromosomes
DMSO	Dimethylsulfoxid
DNA	Desoxyribonucleic acid
DNase	Deoxyribonuclease
dNTP	Desoxyribonucleotidtriphosphate
DTCs	Disseminated tumor cells
E	Epithelial
E/m	Epithelial/mesenchymal
EGFR	Epidermal growth factor receptor
EMT	Epithelial to mesenchymal transition
EMT-TFs	EMT-transcription factors
EpCAM	Epithelial cell adhesion molecule
EpEX	Extracellular domain
EpICD	Intracellular domain of EpCAM
ERK	Extracellular signal regulated kinases
FACS	Fluorescence activated cell sorting
FBS	Fetal bovine serum
FDA	Drug Administration

Abbreviation	Annotation
FITC	Fluorescein isothiocyanate
g	Gravity
gDNA	Genomic DNA
GO	Gene Ontology
h	Hour
H ₂ O	Water
Her2	Human epidermal growth factor receptor 2
IC ₅₀	The half maximal inhibitory concentration
IHC	Immunohistochemistry
<i>i.v</i>	Intravenous injection
KCl	Potassium chloride
kDa	Kilo Dalton
L	Litre
LMP Agarose	Low Melting Point Agarose
M	Mesenchymal
M/e	Mesenchymal/epithelial
mA	Milli ampere
mar	Marker chromosome
max	Maximal
MBC	Metastatic breast cancer
MET	Mesenchymal to epithelial transition
MFI-R	Mean fluorescence intensity ratio
mg	Milligram
MIC	Metastases-inducing cell
min	Minute
mL	Millilitre
mM	Millimolar
MRI	Magnetic resonance imaging
MTT	Thiazolyl Blue Tetrazolium Bromide
NaCl	Modium chloride
nm	Nanometer
OD	Optical density
PBS	Phosphate buffered saline

Abbreviation	Annotation
PCR	Polymerase chain reaction
PI	Propidium iodide
qRT-PCR	Quantitative Real Time PCR
RBC	Red blood cell
RNA	Ribonucleic acid
RNase	Ribonuclease
rob	Robertsonian translocation
ROC curve	Receiver operating characteristic curve
rpm	Revolutions per minute
RT	Reverse transcriptase
RT-PCR	Reverse transcription PCR
SCC	Single cell clone
SE-iFISH	Subtraction enrichment and immunostaining-fluorescence in situ hybridization
sec	Seconds
siRNA	Small interfering RNA
T	Thymidine
t	Translocation
TBS	Tris buffered saline
TIC	Tumor-initiating cells
TNF	Tumor Necrosis Factor
UICC	Union for International Cancer Control
V	Volt
μg	Microgram
μL	Microlitre
μm	Micrometer
μM	Micromolar
WB	Western blot
WBC	White blood cell

10. PUBLICATIONS

This thesis is part work of the following publication:

X. Liu, J. Li, B. L. Cadilha, A. Markota, C. Voigt, Z. Huang, P. P. Lin, D. D. Wang, J. Dai, G. Kranz, A. Krandick, D. Libl, H. Zitzelsberger, I. Zagorski, H. Braselmann, M. Pan, S. Zhu, Y. Huang, S. Niedermeyer, C. A. Reichel, B. Uhl, D. Briukhovetska, J. Suárez, S. Kobold, O. Gires, H. Wang. Epithelial-type systemic breast carcinoma cells with a restricted mesenchymal transition are a major source of metastasis. **Sci. Adv.** 5, eaav4275 (2019).

11. ACKNOWLEDGEMENTS

Foremost, I would like to thank my supervisor Prof. Olivier Gires for giving me this interesting project and guiding me throughout my doctoral study. In the past 4 years, I learned a lot from him. Without his advice and constant help, my research results would not be achievable.

My sincere gratitude also goes to my labmates: Zhe Huang, Yuanchi Huang, Min Pan, Sebastian Niedermeyer, Anna Krandick, Julius Shakhtour, Henrik Schinke, Vera Kohlbauer, Alexandra Blancke-Soares, and Natasha Mahanta. Thanks for their support, inspiration, and companion.

I would like to express my appreciation to various people for their contribution to our common project: Thank Prof. Hongxia Wang and her group for conducting the clinical study and Prof. Sebastian Kobold and his group for helping to perform the animal experiment. Technique support given by Peter P. Lin and his company has been a great help too. Thank Junjian Li for collecting and analyzing the clinical data and thank Bruno Loureiro Cadilha, Anamarija Markota, and Cornelia Voigt for their assistance with the mouse experiment.

I would also like to thank the technicians Gisela Kranz, Darko Libl, Elke Luxenburger, Sabina Schwenk-Zieger for their help in offering me the resources in running the program.

I sincerely appreciate the Guangzhou government for the financial support of my oversea study.

My special thanks are to the experimental mice, they suffered a lot for bearing tumors and had contributed their lives to our research. I would like to express my great respect and appreciation to them.

Finally, I would like to thank my parents for their support and encouragement.